General Psychiatry

Sample sizes based on three popular indices of risks

Hongyue Wang,[1] Bokai Wang,[1] Xin M Tu,[2] Jinyuan Liu,[2] Changyong Feng[1,3]

To cite: Wang H, Wang B, Tu XM, *et al*. Sample sizes based on three popular indices of risks. *General Psychiatry* 2018;**31**:e100011. doi:10.1136/gpsych-2018-100011

Received 11 April 2018
Revised 28 May 2018
Accepted 01 June 2018

[1]Department of Biostatistics and Computational Biology, University of Rochester, Rochester, New York, USA
[2]Division of Biostatistics and Bioinformatics, University of California San Diego, La Jolla, California, USA
[3]Department of Anesthesiology, University of Rochester, Rochester, New York, USA

Correspondence to
Professor Changyong Feng; Changyong_Feng@URMC.Rochester.edu

SUMMARY

Sample size justification is a very crucial part in the design of clinical trials. In this paper, the authors derive a new formula to calculate the sample size for a binary outcome given one of the three popular indices of risk difference. The sample size based on the absolute difference is the fundamental one, which can be easily used to derive sample size given the risk ratio or OR.

INTRODUCTION

Sample size calculation is an essential part in the design of clinical trials. In many cases, a primary outcome of interest is compared between two groups (namely control and treatments groups) in a trial. Usually, the sample size calculation is related to the test statistic used in such comparisons. However, as discussed in section 2, more assumptions are needed to uniquely determine the sample size.

Suppose we want to design a clinical trial to determine if the treatment effect of a new drug is better than that of the current one. The most popular design is to assign patients randomly to the treatment group (the new drug) and the control group (current drug). If the outcome is a success or a failure, it is a binary variable. Generally, for binary outcomes, there are three popular measures of treatment difference: risk difference, relative risk and OR.[1] See Feng and colleagues[2] for relationships among these three measures of effect size.

Formulas for sample size estimation for the binary outcome has been well developed and incorporated in many statistical software packages. See, for example, the formula of Chow and colleagues[3] (4.2.2). In this paper, the authors derive another formula using more information in the hypotheses. Once the sample size formula based on risk difference is obtained, the sample size formulas for the other two indices can be acquired easily.

In this paper, the authors consider the sample size calculation in the parallel design. This paper is organised as follows. In section 2, the authors derive a new formula based on the null hypothesis of the success rate in the control group and the proposed difference of success rates of two groups under the alternative hypothesis. The authors compare it with the formula in Chow and colleagues.[3]. Sections 3 and 4 derive a formula for OR and relative risk. The conclusion and discussion are in section 5.

SAMPLE SIZE CALCULATION BASED ON DIFFERENCE OF SUCCESS RATES

Suppose the true success rates of two groups are p_1 and p_2, respectively. Since p_1, $p_2 \in (0,1)$, let $\Theta=(0,1)\times(0,1)$ be the parameter space and $\Theta_0=\{\theta \in \Theta: p1=p2\}$. Usually, given the significance level α and power $1-\beta$, the sample size calculation depends on the null and alternative hypotheses about the parameters. In the current context, we want to see if the success rates of two groups are the same. Here we consider several scenarios of the hypotheses.

Scenario 1

The null and alternative hypotheses are specified as:

$$H_0: p_1 = p_2 \text{ and } H_1: p_1 \neq p_2 \quad (1)$$

The specification in[1] is the same as:

$$H_0: \theta \in \Theta_0 \text{ and } H_1: \theta \in \Theta \backslash \Theta_0.$$

Although we can test H_0 with data in data analysis, the specification in equation 1 does not offer us enough information to calculate the sample size, as the alternative hypothesis lacks specific details about the treatment effect. Both the null and alternative hypotheses in equation 1 are composite. In fact, any p_1 and p_2 are potential candidates for H_1 as long as they are not equal to each other. However, the power to reject the null hypothesis depends on the true success rates in two groups if they are different. For example, $p_1 = 0.1$, $p_2 = 0.2$ and $p_1 = 0.2$, $p_2 = 0.3$ both satisfy the alternative hypothesis. We have different powers to reject the hypothesis that they have the same success rates in these two cases.

Scenario 2

The hypotheses are specified as:

$$H_0: p_1 = p_2 \text{ and } H_1: p_2 = p_1 + \Delta. \quad (2)$$

where Δ is a prespecified known constant. Although both null and alternative hypotheses are still composite, the alternative hypothesis in equation 2 is much simpler than that in equation 1. It turns out that we still do not have sufficient information to determine the sample size. For example, consider the following two special cases:

Case 1. The hypotheses are:

$$H_0: p_1 = p_2 = 0.1 \text{ and } H_1: p_1 = 0.1,\ p_2 = 0.3.$$

Case 2. The hypotheses are:

$$H_0: p_1 = p_2 = 0.4 \text{ and } H_1: p_1 = 0.4,\ p_2 = 0.6.$$

In both cases, $p_2 = p_1 + 0.2$ under the alternative hypothesis. However, in the following sections, we will show that the sample sizes in these two cases are different. Given the difference of success rates, usually it is much easier to reject the null hypothesis in case 1 than in case 2.

Scenario 3

The null and alternative hypotheses are:

$$H_0: p_1 = p_2 = p_0 \text{ and } H_1: p_1 = p_0,\ p_2 = p_0 + \Delta, \quad (3)$$

where p_0 and Δ are prespecified constants. Without the loss of generality, we assume that $\Delta>0$ in the following discussion. It turns out that we can uniquely determine the sample size in this case.

Sample size formula

We derive a sample size formula based on the hypotheses specified in[3] using the large sample theory.[4] The typical way is to first derive the asymptotic distribution of a test statistic under the null and alternative hypothesis followed by solving an equation to obtain the sample size formula (with the given significance level and power) (see, eg, Tu *et al*[5]).

Although the treatment and control groups have the same sample size in many studies, it is unnecessary in practice. Some studies intentionally assign more patients in one group. Suppose the sample size in groups 1 and 2 are n and $n\kappa$, respectively, where κ is a prespecified positive constant. Group 2 has more (less) subjects than group 1 depending on if $\kappa > 1\ (0 \le \kappa \le 1)$. If $\kappa = 1$, the two groups have an equal sample size.

Let $\hat{p}_1 = \frac{m_1}{n}$ and $\hat{p}_2 = \frac{m_2}{n\kappa}$ denote the estimates of p_1 and p_2, where m_1 (m_2) denote the number of events of success in group 1 (equation 2). According to the central limit theorem,[6]

$$\sqrt{n}\left(\hat{p}_1 - p_1\right) \to N\left(0,\ p_1\left(1 - p_1\right)\right),$$

$$\sqrt{n\kappa}\left(\hat{p}_2 - p_2\right) \to N\left(0,\ p_2\left(1 - p_2\right)\right),$$

as n is large enough.

Under the hypothesis of $p_1 = p_2 = p_0$, the variances of $\hat{p}_1$ and $\hat{p}_2$ are $p_0\left(1 - p_0\right)/n$ and $p_0\left(1 - p_0\right)/\left(n\kappa\right)$, respectively. To test the null hypothesis that $p_1 = p_2$, we consider the following test statistics:

$$T = \frac{\hat{p}_2 - \hat{p}_1}{\sqrt{\frac{p_0(1-p_0)}{n} + \frac{p_0(1-p_0)}{n\kappa}}}.$$

Then $T \to N\left(0,\ 1\right)$ as n grows unbounded.

Let Φ be the distribution of standard normal distribution. For each $\eta \in \left(0,\ 1\right)$, let z_η be such that $\Phi\left(z_\eta\right) = \eta$, that is, z_η is the $(100 \times \eta)$th percentile of the standard normal distribution. Given the significance level α, we reject the hypothesis of $p_1 = p_2$ if $\left|T\right| > z_{1-\alpha/2}$. Note that:

$$\Pr\left\{\left|T\right| > z_{1-\alpha/2}\right\} = \Pr\left\{T > z_{1-\alpha/2}\right\} + \Pr\left\{T < -z_{1-\alpha/2}\right\}.$$

Let

$$c_0 = \sqrt{p_0\left(1 - p_0\right)\left(1 + \frac{1}{\kappa}\right)}, \quad (4)$$

$$c_1 = \sqrt{p_0\left(1 - p_0\right) + \frac{\left(p_0+\Delta\right)\left(1-p_0-\Delta\right)}{\kappa}}, \quad (5)$$

We have

$$\Pr\left\{T > z_{1-\frac{\alpha}{2}}\right\} = \Pr\left\{\frac{\hat{p}_2 - \hat{p}_1 - \Delta}{\frac{c_1}{\sqrt{n}}} > \frac{c_0 z_{1-\frac{\alpha}{2}} - \sqrt{n}\Delta}{c_1}\right\},$$

$$\Pr\left\{T < z_{1-\alpha/2}\right\} = \Pr\left\{\frac{\hat{p}_2 - \hat{p}_1 - \Delta}{\frac{c_1}{\sqrt{n}}} < -\frac{c_0 z_{1-\frac{\alpha}{2}} + \sqrt{n}\Delta}{c_1}\right\}.$$

In most studies, $\alpha = 0.05$ or $\alpha = 0.01$ for a large sample size. Since $\Delta > 0, \Pr\left\{T < -z_{1-\alpha/2}\right\} \approx 0$ under H_1. Under the hypothesis that $p_2 = p_1 + \Delta$, to make the test statistic have power $1 - \beta$, we let:

$$\frac{c_0 z_{1-\alpha/2} - \sqrt{n}\Delta}{c_1} = z_\beta = z_{1-\beta}.$$

Solving this equation, we obtained the required sample size in group 1:

$$n = \left[\frac{c_0 z_{1-\alpha/2} + c_1 z_{1-\beta}}{\Delta}\right]^2. \quad (6)$$

This formula is the basis of sample size calculation based on other indices (see the next two sections).

Note that formula (4.2.2) in Chow and colleagues[3] is:

$$n = \left[\frac{c_1\left(z_{1-\alpha/2} + z_{1-\beta}\right)}{\Delta}\right]^2. \quad (7)$$

The sample sizes in equations 6 and 7 are equal if and only if $p_0 = \left(1 - \Delta\right)/2$. If $p_0 > (<) \left(1 - \Delta\right)/2$, then n in equation 6 is larger (smaller) than that in equation 7.

Figure 1 shows the sample size formulas equations 6 and 7 for different p_0 with $\Delta = 0.1$. Note that in the sample size calculation of Chow and colleagues [3] they did not use the fact that $p_1 = p_2 = p_0$ in calculating the variance of $\hat{p}_2 - \hat{p}_1$ under the null hypothesis.

SAMPLE SIZE BASED ON RELATIVE RISK

The null and alternative hypotheses are:

$$H_0: p_2 = p_1 = p_0 \text{ and } H_1: \frac{p_2}{p_1} = r,$$

where r is a known constant. Without loss of generality, assume $r > 1$.

From 6 if we only need new Δ and c_1 to obtain the sample size formula. Given relative risk and p_0, the proposed risk difference is:

$$\Delta = \left(r - 1\right) p_0.$$

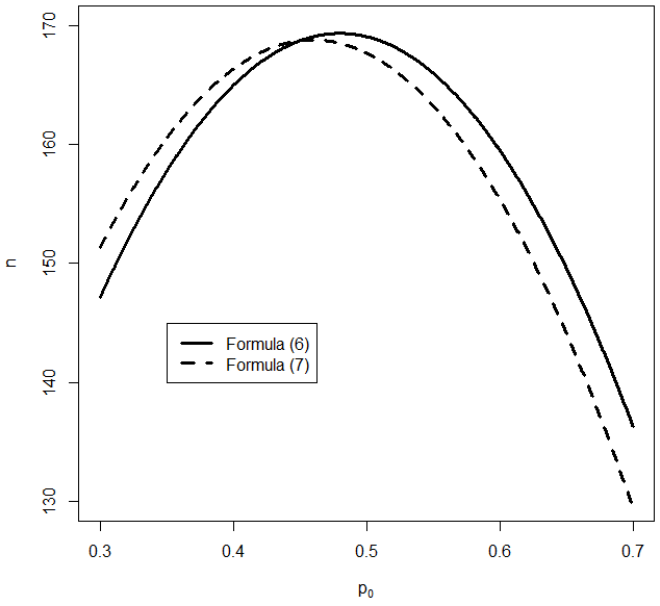

Figure 1 Sample sizes based formulas (6) and (7).

The new c_1 is:

$$c_1 = \sqrt{p_0\left(1 - p_0\right) + \frac{rp_0\left(1 - rp_0\right)}{\kappa}}.$$

Substituting (Δ, c_1) into equation 6, we obtain the required sample size based on relative risk.

SAMPLE SIZE BASED ON ORS

The null and alternative hypotheses are

$$\mathrm{H}_0: \; p_2 = p_1 = p_0 \text{ and } \mathrm{H}_1: \frac{p_2/(1-p_2)}{p_1(1-p_1)} = \theta,$$

where θ is a known constant. Without loss of generality, assume $\theta > 1$. The risk difference is

$$\Delta = \frac{(\theta-1)p_0(1-p_0)}{1+(\theta-1)p_0}.$$

The new c_1 is:

$$c_1 = \sqrt{p_0\left(1 - p_0\right)\left[1 + \frac{\theta}{\kappa\left(1+(\theta-1)p_0\right)^2}\right]}.$$

Substituting (Δ, c_1) into equation 6, we obtain the required sample size based on OR.

CONCLUSION

In this paper, we derive new formulas to calculate sample size based on three popular indices of difference of success rates in two treatment groups. Generally, we cannot uniquely determine the sample size only given one of the risk difference indices. We need the success rates of both groups under the alternative hypothesis. This is usually done through the specification of the success rate of the control group and the difference of two groups. The easiest one is to specify the absolute risk difference Δ between the two groups. If the difference between the two groups is specified by the risk ratio or OR, we can easily transfer them to risk difference Δ and use formula 6 to calculate the sample size.

Figure 1 shows that the sample size calculated based on our formula is generally different than that reported in Chow and colleagues. [3] We compare the accuracies of those two formulas by comparing the powers under different situations. This work is on-going.

Contributors CF and HW derived the theoretical results; BW and JL constructed the examples and graphs; and XMT drafted the manuscript.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Commissioned; internally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1 Fleiss JL, Levin B, Paik MC. *Statistical methods for rates and proportions*. 3rd edn. Hoboken: Wiley, 2003.
2 Feng C, Wang H, Wang B, *et al*. Relationships among three popular measures of differential risks: relative risk, risk difference, and odds ratio. *Shanghai Arch Psychiatry* 2016;28:56–60.
3 Chow SC, Shao J, Wang H. *Sample size calculations in in clinical research*. 2nd edn. Boca Raton: Chapman & Hall/ CRC, 2008.
4 van der Vaart W. *Asymptotic statistics*. New York: Cambridge University Press, 1998.
5 Tu XM, Kowalski J, Zhang J, *et al*. Power analyses for longitudinal trials and other clustered designs. *Stat Med* 2004;23:2799–815.
6 Kowalski J, XM T. *Modern applied U statistics*. New York: Wiley, 2007.

Hongyue Wang obtained her BS in Scientific English from the University of Science and Technology of China (USTC) in 1995, and PhD in Statistics from the University of Rochester in 2007. She is a Research Associate Professor in the Department of Biostatistics and Computational Biology at the University of Rochester Medical Center. Her research interests include longitudinal data analysis, missing data, survival data analysis, and design and analysis of clinical trials. She has extensive and successful collaboration with investigators from various areas, including Infectious Disease, Nephrology, Neonatology, Cardiology, Neurodevelopmental and Behavioral Science, Radiation Oncology, Pediatric Surgery, and Dentistry. She has published more than 80 statistical methodology and collaborative research papers in peer-reviewed journals.