Sample Size Calculation for Comparative Trials

Slobodan M. Jankovic

University of Kragujevac, Faculty of Medical Sciences, Kragujevac, Serbia

Corresponding author: Professor Slobodan M. Jankovic, MD, DSc, MSc, Prim. Faculty of Medical Science, University of Kragujevac. Svetozara Markovica Street, 69. 34000 Kragujevac, Serbia. E-mail: slobnera@gmail.com. ORCID ID: http//www.orcid.org/0000-0002-1519-8828.

Background: In order to find out truth about the target population size of the study sample has to be large enough to contain the characteristic (variable) we are seeking with variability sufficient to reliably reflect its variability in the population. **Objective**: The aim of this article was to explain logic of sample size calculation in comparative studies, and shed some light on key assumptions of the calculation. **Methods**: This article is a review of methodology used for estimating appropriate size of a study sample. **Results:** True difference in target parameter among the populations that are studied, and its variability (usually expressed as standard deviation from the mean) could not be changed according to our preferences; also maximum acceptable levels of probability if type one and type two errors cannot be further increased without compromising ability of the study to give us reliable information about the populations. What we can change is number of patients within the study groups, which if increased, will decrease variability of the results, and make distribution of the difference beteween the groups (if the study is hypothetically repeated many times) around true value of difference between the populations ***more narrow***. Through narrowing of the distributions we will decrease number of cases when the difference among the group (type one error) or lack of difference (type two error) happens by chance, i.e. put probabilities of these errors below limits of acceptability. **Conclusion**: Careful calculation of sample size is necessary to minimize probability of type one and type two errors and therefore obtain reliable answer to a research question.

Keywords: sample size; statistical power; type one error; type two error.

Editorial, Received: Dec 02, 2018, Accepted: Dec 25, 2018, doi: 10.5455/ijbh.2018.6.62-64, Int J Biomed Healthc. 2018; 6(2): 62-64

1. INTRODUCTION

Any kind of quantitative research, whether pre-clinical or clinical, is performed on a sample withdrawn from the target population, whose characteristics are objective of the research. Whether we will find out truth about the target population depends on the size of the sample: it has to be large enough to contain the characteristic (variable) we are seeking with variability sufficient to reliably reflect its variability in the population (1). Only if size of the sample is sufficient, we can be sure that results of our study are not like this just by chance, if positive, and are not false, if negative. Therefore, one of key elements of appropriate study design is planning sample of sufficient size (2).

If one is planning comparative study, main outcome of the study could be either continuous or categorical variable. If continuous, necessary sample size will depend on extent of the difference between a measure of central tendency in both groups, and on variability of the variable in both groups (3). If categorical, the sample size will depend on difference between the proportions of „positive" patients in the study groups. Although there are numerous calculators for estimation of minimum sample size, researchers should understand basic principles of the calculation in order to use them properly.

2. OBJECTIVE

The aim of this article was to explain logic of sample size calculation in comparative studies, and shed some light on key assumptions of the calculation.

3. LOGIC BEHIND THE SAMPLE SIZE CALCULATION

It is always possible that certain result of a research study, which is the most often difference between two or more treatment groups, is reached only by a chance, and not because there is true difference between the populations that were sampled. Such result is actually an error, occured by a chance, designated in statistics as „type one error". When designing a research study, our aim is to minimize this error, i.e. to decrease its probability as much as possible. The highest probabilty of the type one error that is acceptable is 0.05. On the other hand, if a research study shows that there is no difference among the treatment groups, such result may also happen just by chance, although there is true difference among the population sampled. This type of error is called „type two error", and its highest probability that is still acceptable for us is 0.2 (4).

It is obvious that we cannot change true difference in target parameter among the populations that are studied,

and its variability (usually expressed as standard deviation from the mean); also maximum acceptable levels of probability if type one and type two errors cannot be further increased without compromising ability of the study to give us reliable information about the populations. What we can change is number of patients within the study groups, which if increased, will decrease variability of the results, and make distribution of the difference beteween the groups (if the study is hypothetically repeated many times) around true value of difference between the populations ***more narrow***. Through narrowing of the distributions we will decrease number of cases when the difference among the group (type one error) or lack of difference (type two error) happens by chance, i.e. put probabilities of these errors below limits of acceptability. Therefore, minimal number of the study participants in the study groups should be calculated based on the following criteria: probability of type one error below 0.05, probability of type two error below 0.2 (or, what is complementary, statistical power of the study higher than 0.8), hypothetical difference between members of populations that will be sampled and variability of that difference (the ratio of the two is usually named in the literature as „effect size"), and type of statistical test that will be used for comparison of the study groups.

It is clear that concrete methods of calculating sample size will differ among themselves according to type of statistical test, nature of a study outcome (whether continuous or categorical variable) and number of the study groups. Here only two basic cases will be elaborated: comparison between two independent groups with normal distribution of continuous dependent variable (e.g. serum level of creatinine), and comparison between two independent groups with categorical, binary dependent variable (e.g. a patient survived or not) (5, 6).

4. SAMPLE SIZE FOR COMPARING TWO INDEPENDENT GROUPS WITH NORMALLY DISTRIBUTED CONTINUOUS VARIABLE

Formula for calculating sample size for comparing two independent groups with normal distribution is based on setting the point of the difference between the study groups where probabilities of type one and type two errors are those that we want (7):

$$z_{1-\alpha/2} * \sqrt{\sigma^2} = d - z_{1-\beta} * \sqrt{\sigma^2} \quad \text{(I)}$$

In the formula I: $z_{1-\alpha/2}$ is value of difference between the means divided by standard deviation around zero that corresponds to half of probability of type one error; $z_{1-\beta}$ is value of difference between the means divided by standard deviation around mean difference that corresponds to probability of type two error; d is supposed true difference between the populations from which the groups will be sampled; is square root of variance around mean difference between the study groups .

The formula I is then solved for σ^2:

$$\sigma^2 = \frac{d^2}{\left(z_{1-\beta} + z_{1-\alpha/2}\right)^2} \quad \text{(II)}$$

The variance around difference between the two study groups is equal to sum of variances around mean in each of the groups divided by the number of participants in that group (8):

$$\sigma^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \quad \text{(III)}$$

If now we assume that number of participants is the same in both groups (n1 = n2 = n), we can combine the equations II and III:

$$\sigma^2 = \frac{\sigma_1^2 + \sigma_2^2}{n} = \frac{d^2}{\left(z_{1-\beta} + z_{1-\alpha/2}\right)^2} \quad \text{(IV)}$$

Finally, if we solve the equation IV for n, we can calculate minimum sample size of each of the two groups ($z_{1-\alpha/2}$ is 1.96 for probability of type two error of 0.05, and $z_{1-\beta}$ is 0.84 for probability of type two error of 0.2):

$$n = \frac{(\sigma_1^2 + \sigma_2^2) * \left(z_{1-\beta} + z_{1-\alpha/2}\right)^2}{d^2} \quad \text{(V)}$$

5. SAMPLE SIZE FOR COMPARING TWO INDEPENDENT GROUPS WITH BINARY CATEGORICAL VARIABLE

Formula for calculating sample size for comparing two independent groups with categorical variable (i.e. percent of certain outcome in each of the groups) is also based on setting the point of the difference between the study groups where probabilities of type one and type two errors are those that we want (9):

$$z_{1-\alpha/2} * \sqrt{\sigma^2} = \varepsilon - z_{1-\beta} * \sqrt{\sigma^2} \quad \text{(VI)}$$

The formula VI could be rearranged as:

$$z_{1-\alpha/2} + z_{1-\beta} = \frac{\varepsilon}{\sqrt{\sigma^2}} \quad \text{(VII)}$$

Since variance around difference in proportions () is:

$$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad \text{(VIII)}$$

when equations VII and VIII are combined, and solved for number of participants in one of the groups (assuming that both groups have the same number of participants: n1 = n2 = n), we will have the minimum sample size ($z_{1-\alpha/2}$ is 1.96 for probability of type two error of 0.05, and $z_{1-\beta}$ is 0.84 for probability of type two error of 0.2):

$$n = \frac{\left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)^2}{\varepsilon^2} * [p_1(1-p_1) + p_2(1-p_2)] \quad \text{(IX)}$$

6. CONCLUSION

Careful calculation of sample size is necessary to minimize probability of type one and type two errors and therefore obtain reliable answer to a research question.

Young researchers should become familiar with principles of calculating sample size, and then regularly conduct this excercise when planning future studies.

- *Author's contribution:* ***Athor was involved in all steps of preparation this paper and also made final proofreading.***
- *Conflict of interest:* ***None declared***
- *Financial support and sponsorship:* ***Nil***

REFERENCES

1. Noordzij M, Dekker FW, Zoccali C, Jager KJ. Sample size calculations. Nephron Clin Pract. 2011; 118(4): c319-323.
2. Röhrig B, du Prel JB, Wachtlin D, Kwiecien R, Blettner M. Sample size calculation in clinical trials: part 13 of a series on evaluation of scientific publications. Dtsch Arzteblatt Int. 2010; 107(31-32): 552–556.
3. Sami W, Alrukban MO, Waqas T, Asad MR, Afzal K. Sample Size Determination In Health Research. J Ayub Med Coll Abbottabad JAMC. 2018; 30(2): 308-311.
4. Flight L, Julious SA. Practical guide to sample size calculations: an introduction. Pharm Stat. 2016; 15(1): 68-74.
5. Flight L, Julious SA. Practical guide to sample size calculations: superiority trials. Pharm Stat. 2016; 15(1): 75-79.
6. Julious SA. Sample sizes for clinical trials with normal data. Stat Med. 2004; 23(12): 1921-1986.
7. Kamangar F, Islami F. Sample size calculation for epidemiologic studies: principles and methods. Arch Iran Med. 2013; 16(5): 295-300.
8. Lane DM. Sampling Distribution of Difference Between Means. In: Online Statistics Education: An Interactive Multimedia Course of Study.
9. Wang H, Chow SC. Sample Size Calculation for Comparing Proportions. In: Wiley Encyclopedia of Clinical Trials [Internet]. American Cancer Society; 2007 [cited 2020 Dec 5]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780471462422.eoct005