Power Analysis and Sample Size, When and Why?

Editorial

Özgür Kemal
Department of Otorhinolaryngology, Ondokuz Mayıs University School of Medicine, Samsun, Turkey

Clinical or experimental study results need to be processed precisely to lead to development and advances in medicine. At this stage, biostatistics plays an important role in collecting healthy data, making unbiased comparisons and interpreting the findings correctly. In order to interpret the findings correctly and to adapt this to the diagnosis or treatment of patients, it is very important to conduct power analysis in scientific research. By determining the number of samples to be included in the study by power analysis, it can be demonstrated that the results obtained are really significant or not (1, 2).

Rosenfeld and Rockette (1) showed in their study that only 1% of 541 original research articles published in four prestigious otolaryngology journals of 1989 studied sample size or power analysis.

Today, the first step of a clinical or experimental study is design. Before beginning the study, one should determine the study population, than find a sample which is considered to represent the population, and it is clear that, the most important part of a study design is the sample size (2). A small sample size might lead to failure of the study and statistical analysis will be ineffective; on the other hand, a big sample size might lead statistically significant results with unnecessary numbers of subjects and cost (2, 3). Also, including more participants than needed is an ethical problem (2). For both statistical adequacy and unnecessary cost avoidance we have to find the exact number of patients, subjects or laboratory animals (3, 4).

The power analysis is performed by some specific tests and they aim to find the exact number of population for a clinical or experimental study (5).

In fact, there are two situations while testing the hypothesis in a clinical trial. These are null hypothesis (H_0) and alternative hypothesis (H_1). Null hypothesis always argues that there is no difference between groups. The opposite of this is called as alternative hypothesis. Other than the hypothesis types, there are two types of error in biostatistics. Type I error means that incorrectly rejection of the hypothesis where the hypotesis is true. Type II error is an error that we accept the hypothesis, when the hypothesis is false, but we incorrectly do not have the ability to reject it (Table 1) (6).

Type I Error value is predetermined by the researchers and usually set at 0.05 or 0.01. If authors define type I error as 0.05 and if the result is found as no difference, that is 95% true (1). Type II error is defined as the power of the study. It is usually set at 0.20, sometimes 0.10. If it is set to 0.20, the power of the study is 80%. In other words, the probability of not detecting the difference between two groups is considered as 20% (5-7).

Table 1. Type I, Type II errors and their relationship

	The real situation	Hypothesis is true	Hypothesis is false
The result of research	Hypothesis is true	Correct decision	Type I Error (α)
	Hypothesis is false	Type II error (β)	Correct decision

ORCID ID of the author:
Ö.K. 0000-0002-6419-6204.

Cite this article as: *Kemal Ö. Power Analysis and Sample Size, When and Why?. Turk Arch Otorhinolaryngol 2020; 58(1): 3-4.*

Corresponding Author:
Özgür Kemal, drozgurkemal@gmail.com

DOI: 10.5152/tao.2020.0330

The other two parameters which affect the sample size are minimal clinically relevant difference and variance. The minimal clinically relevant difference is the smallest difference outcome between the study groups. It can be called as minimal scientific outcome that is significant for the investigator. So, this difference should be determined by the author. For example, if you are working on a treatment for sudden hearing loss. The level for the outcome should be determined as 20 or 30 dB by the authors.

The last important parameter is the variance of the outcome. This outcome is usually obtained from the clinical knowledge or previous data (6).

After determining these parameters, calculations can be done easily by using different software by a biostatistician.

As a conclusion, at the beginning of a clinical or experimental study, the researcher should determine the type I and type II error values, minimal clinically relevant difference and the variance of their own study. Than by using these parameters, biostatisticians can be able to help us find the most appropriate number of samples that will obtain effective and qualified scientific values.

References

1. Rosenfeld RM, Rockette HE. Biostatistics in otolaryngology journals. Arch Otolaryngol Head Neck Surg 1991; 117: 1172-6. [Crossref]
2. Hickey GL, Grant SW, Dunning J, Siepe M. Statistical primer: Sample size and power calculations-why, when and how? Eur J Cardiothorac Surg 2018; 54: 4-9. [Crossref]
3. Cohen J. Statistical power analysis. Current directions in psychological science 1992; 1: 98. [Crossref]
4. Dell RB, Holleran S, Ramakrishnan R, Sample size determination. ILAR 2002; 43: 207-13. [Crossref]
5. Fitzner K, Heckinger E. Sample size calculation and power analysis: a quick review. Diabetes Educ 2010; 36: 701-7. [Crossref]
6. Kul S. Sample size determination for clinical research. Pleura Bulletin 2011; 2: 129-32. [Crossref]
7. Lovell DP. Null hypothesis significance testing and effect sizes: Can we 'effect' everything … or … anything? [published online ahead of print, 2020 Jan 13]. Curr Opin Pharmacol. 2020;S1471-4892(19)30117-1. doi:10.1016/j.coph.2019.12.001. [Crossref]