

Standard statistical methods and calculation of sample size in Medical education research

Dr . Gopinath Agnihotram, Biochemist, Institute for Post Graduate Teaching and Research in Ayurveda, Jamnagar, Gujarat

Address for correspondence: Dr.Gopinath Agnihotram, Biochemist, Institute for Post Graduate Teaching and Research in Ayurveda,Gujarat Ayurved University, Jamnagar, Gujarat, India-361008

E-mail: agnbio1985@gmail.com

ABSTRACT

After confirmation of any line and area of Research work, the Basic Question starts with Collection of Sample size determination, but before that in a health science research "why statistical methods are so important?" Simple answer is that, statistical methods are used throughout a study that includes planning, designing, collecting data, analyzing and drawing meaningful interpretation and report the findings. Hence, it is important that a researcher knows the concepts of at least basic statistical methods used at various stages of a research study. This helps the researcher in the conduct of an appropriately well-designed study leading to valid and reliable results that can be generalized to the population. A well-designed study possesses fewer biases, which intern gives precise, valid and reliable results. There are many statistical methods and tests that are used at various stages of a research. In this communication, we discuss the overall importance of statistical considerations in medical research with the main emphasis on estimating minimum sample size for different study objectives

Keywords: Biostatistics, Medical Research, Confidence Interval, Probability, Sample size

INTRODUCTION

Statistics is a branch of science concerned with (i) The collection, organization, summarization and analysis of data and (ii) the drawing of inferences about whole body of data when only a part of the data is observed.¹ Statistics has a role in a study starting from planning, designing, collecting data, analyzing until drawing meaningful interpretation from it. Many health science researchers may wonder around the question "why statistical methods are so important in research?" The answer is "bad statistics lead to bad research and bad research is unethical."² Statistical methods revolutionized science in 20 th century³ with a vast majority of advanced methodological developments taking place during this period. Further, the invention of high end computers toward the end of the last century enabled the researchers to apply advanced statistical methods in their research with ease and comfort in computations.

The ease in computations has, at least to some extent; changed the course of analytical considerations. For example, sophisticated multivariate methods are no more considered as difficult to apply from computations point of view. Today, statistics is an indispensable tool in each and every field of health science research, whether it is Medicine, Ayurved, Pharmacy, Dental or other allied health sciences. Statistics helps even clinicians in extracting vital information from the empirical data that ultimately lead to improved patient care. Statistical concepts are required to be considered throughout a study, from planning to the final reporting stage. This article provides a brief overview of statistical methods used at various stages of a research study with the main emphasis on estimation of minimum sample size for various types of objectives.

STATISTICS IN RESEARCH STUDIES

The first step in any research study is to clearly state aims of the study followed by its objectives, which are focused at achieving the aim. Many use the terms aim and objective interchangeably, whereas the two terms have their own meaning with clear-cut distinction. Aim of the study is a general statement about main and broad study question, meaning that it is not measurable. On the other hand, objectives of a research study should be (Specific, Measureable, Achievable, Relevant and Time-bound) and limited in number. Before stating the objectives, the researcher should know about the types of variables and/or attributes being assessed in the study, especially the exposure and outcome measurements and distinction between the two in the analytical studies. This is required in order to ensure the characteristics of the objectives stated above. The second step is to choose the appropriate study design for meeting the stated objectives, which depends on many factors related to study, such as, study question/aim, availability of the fund and other logistic resources, type of study subjects/variables, etc., After selecting the study design, the third step is to define the target population and also the method of selecting study participants from this population.

Statistical techniques called sampling types/methods are used for selecting subjects from a study population to form the sample. There are broadly two types of sampling methods namely probability (random) sampling and non-probability (non-random) sampling methods. Statistical methods pertaining to inference have been developed based on assumptions that study participants are randomly selected, i.e., the sampling method used should be a probability sampling. Even within probability sampling, there are a number of ways of selecting a sample randomly. The choice of probability sampling method to be used in a study depends on the objective of the study, heterogeneity of characteristics of individuals in the study population and also the feasibility of getting study participants in terms of cost, time and manpower. Once the study design and sampling technique has been finalized the fourth step in a study is to estimate the sample size. This aspect is being dealt in detail in the present communication.

The fifth step in a research study is to decide on data collection tools used for collecting relevant information from the selected participants. Some studies require the researcher to develop new tools for data collection. Statistical methods are used for the development and validation of research tools. Reliability coefficients such as Cronbach's alpha, intra-class correlation coefficient, split-half coefficient etc., are used for measuring the reliability of a questionnaire. Another aspect of tool preparation is validity, which requires looking for content, criterion and face validity. Once data collection is over it has to undergo quality checking, coding and computer entry.

The sixth step in a research study is to perform appropriate statistical analysis. The first step in data analysis is to compute descriptive statistics for all important variables in the study. The descriptive statistics summarize various aspects about the data, giving details about the selected sample.⁴ To summarize the data, we use several statistical summary measures such as mean, median, standard deviation, inter quartile range, percentages etc., depending upon the type of variable, measurement scale used and the variability in each of the variables. Descriptive statistics is followed by inferential statistics dealing with generalization of the sample results to the population. Confidence intervals (CIs) are presented to take account of sampling error involved in the estimates and tests of significance are applied to find whether the results observed from a sample is due to chance or not. There are many statistical tests, but the choice depends on many factors such as the objective of the study, type of variables and measurement scales used, sample size, number of groups to be compared, number of variables, distribution of the outcome variable etc.⁵

Statistical hypothesis tests are broadly classified into two categories; based on whether they mandate the assumption of the distribution of variable under study as normal or not. Parametric tests have been built under the assumption that the variable under consideration follows the normal distribution. Non-parametric tests on the other hand, do not require any distributional assumption about the variable under consideration in the study population. Hence, non-parametric tests are called distribution free tests. However, one should know that the non-parametric tests bring the ease in their application on the cost of sacrificing the power of the study, a main concern in clinical studies.

The final step in a research study is to communicate the results and interpretations using the appropriate figures and tables. A table of appropriate summary measures for variables in the study with the number of subjects in each category is a necessary one to brief the sampled data. Reporting a CI for every population parameter became a mandatory pre-requisite in modern research. Several authors have discussed about the principles and reporting of CIs.⁶⁻¹⁴ It is recommended to report both CI and exact P value instead of reporting only $P < 0.05$ or as $P < 0.01$ etc., The next section discuss about the determination of minimum sample size required for studies with a common type of objectives.

ESTIMATION OF MINIMUM SAMPLE SIZE REQUIRED FOR A STUDY

The problem of sample size estimation can be broadly of two types namely (a) Sample size for an estimation study and (b) sample size for a hypothesis testing study, i.e., comparison study. In an estimation study, the researcher is interested in estimating the quantum of one or more characteristics of the population called parameter(s), for example, mean hemoglobin level or prevalence of arthritis, etc., In hypothesis testing studies, the investigators are interested in comparing a characteristic of the population for one or more time points or a characteristic of two or more populations, for example, comparison of prevalence of arthritis before and after administration of some intervention or between two populations. The objective of calculating sample size in an estimation study is to estimate the value of the parameter under study for a prefixed precision and level of confidence.

If a researcher wants the estimate to be more precise in his study then he should select a large number of subjects, i.e., as precision increases (or margin of error decreases) the minimum sample size required increases. Similarly, the sample size increases with the increase in level of confidence. For example, the sample size required for estimating a parameter with 99% confidence level is more than that required for 95%

confidence level. In testing of hypothesis studies, the objective of sample size calculation is to achieve a desired power for detecting a clinically or scientifically meaningful difference at a prefixed level of significance.¹⁵ Power of a study is the probability of rejecting a null hypothesis, which is false. Level of significance is the threshold set on the probability of rejecting a null hypothesis, which is true. In this section, we see how one can find the minimum sample size for (a) Estimating a population mean, (b) estimating a population proportion, (c) testing the equality of two means and (d) testing the equality of two proportions.

The sample size formula for estimating a population mean is given by

$$n = \left(\frac{Z_{1-\alpha/2} \sigma}{d} \right)^2$$

where, $Z_{1-\alpha/2}$ is the value obtained from the standard normal distribution table for 100 (1- α)% confidence level. Its value for 95% and 99% confidence level is 1.96 and 2.58 respectively.

σ is the population standard deviation, which is majority of times unknown. The value of σ can be taken from a similar published study or based on a pilot study. For the pilot study, there is no need of sample size estimation; one can do a pilot study according to the number available. d is the absolute allowable error (or precision) in the estimation. Thus, the minimum sample size (n) required for estimating a population mean is directly proportional to the level of confidence and the standard deviation, whereas it is inversely proportional to the absolute error that is allowed in the estimation. For example, a physician wants to estimate the average age of patients visiting his clinic. If the allowable error in estimating the average age is within ± 2 years, confidence level of 95% and assumed the standard deviation of age of patients visiting the clinic is 8 years, the minimum sample size required for his study as

$$n = \left(\frac{Z_{1-\alpha/2} \sigma}{d} \right)^2 = \left(\frac{1.96 \times 8}{2} \right)^2 = 62$$

In the above example if the researcher wants 99% confidence level, the sample size becomes

$$n = \left(\frac{Z_{1-\alpha/2} \sigma}{d} \right)^2 = \left(\frac{2.58 \times 8}{2} \right)^2 = 107$$

Similarly, the minimum number of subjects required for the above study at 95% confidence level and allowable error of ± 1 year is 246.

Sample size for estimating a population proportion (or prevalence)

The sample size formula for estimating a population proportion is given by

$$n = \frac{Z_{1-\alpha/2}^2 (1-P)}{\epsilon^2 P}$$

where, $Z_{1-\alpha/2}$ is as defined earlier, P is the anticipated proportion of the condition in the study population and ϵ is the relative precision. An estimate for the anticipated proportion in the study population can be obtained from previous ϵ studies conducted in the same population or from a pilot study.

Thus, the minimum sample size n required for estimating the proportion of a rare condition in the study population is large compared with that of a common condition. Furthermore, the sample size decreases as the value of relative precision increases. For example, a survey was planned to assess the prevalence of current Medicine use among an urban adult population in India. It was anticipated that in the study population the current prevalence of Medicine use was 30%. The minimum sample size required for this study for relative precision of 10% and confidence level of 95% is given by

$$n = \frac{Z_{1-\alpha/2}^2 (1-P)}{\alpha^2 P} = \frac{1.96^2 \times 0.7}{0.1^2 \times 0.3} = 897$$

Hence, the above study requires a minimum of 897 participants to estimate the prevalence of Medicine use in the study population with relative precision of 10% and confidence level of 95%.

Sample size for testing equality of two population means

If the primary objective of a study is to test the null hypothesis of equality of means of two independent populations then the formula for estimating the sample size in each of the study group is given by

$$n = 2 \left[\frac{(z_{1-\alpha/2} + z_{1-\beta}) S}{d} \right]^2$$

where, n is the minimum sample size required in each group (equal allocation and hence total sample size is $2n$), $Z_{1-\alpha/2}$ is the two-tailed standard normal distribution table value for 100 α % level of significance, $Z_{1-\beta}$ is the value of standard normal distribution table value for 100 β % Type II error or 100 \times α % power of the study. Usually, the power of a study is selected as 80% ($\beta = 0.2$) or 90% ($\beta = 0.1$) and the value of $Z_{1-\beta}$ is respectively 0.842 and 1.282.

S is the pooled standard deviation (or average of two population standard deviations), which can be obtained from

previous literature or from a pilot study. The most important component in sample size estimation for hypothesis testing is fixing the minimum clinically or scientifically meaningful difference between two population means, which is given by d in the above equation. The researcher has to fix the value of d based on clinical or scientific judgment and not based on results of the pilot study or previous literature. The clinically significant difference should be a difference that makes some real difference. Thus, the minimum sample size required for testing the equality of two population means depends on

- The level of significance (inversely proportional)
- Power of the study (directly proportional)
- Pooled standard deviation of the variable under study (directly proportional)
- The minimum clinically significant difference (inversely proportional)

The above formula is used for selecting equal number of subjects in each of the study group. Sometimes the investigators want higher number of subjects in control/standard drug group. For example, in a clinical trial the research team wants more number of subjects to be recruited in the standard treatment arm than in the new treatment arm. If the sample size n_c in the standard treatment group is to be k times the size n_t of new treatment group, i.e., $n_c = kn_t$, then n_t is given by

$$n_t = \left(1 + \frac{1}{k}\right) \left[\frac{(z_{1-\alpha/2} + z_{1-\beta}) S}{d} \right]^2$$

Thus, the total number of subjects required for a clinical trial in which the allocation to the treatment and controlled group is in the ratio of 1:2 is 12.5% more than that required for a 1:1 allocation. Similarly, a study which allocates 1:3 requires a sample size of 33% more than that required for 1:1 allocation.¹⁶ The sample size required for comparison of more than two population means depends on the number of groups to be compared in addition to all the above mentioned factors. In this case, the value of $Z_{1-\alpha/2}$ has to be adjusted for the number of comparisons made between the study populations. Hence, sample size in this situation increases with the increase in number of group comparisons.

For example, a randomized controlled trial has been planned to compare the efficacy of an Ayurvedic medication with a placebo in controlling the thyroid stimulating hormone (TSH) levels among subjects with hypothyroidism. The researchers consider a difference of 3 mIU/L in TSH level between the two groups as clinically important and they want 80% power to detect this difference at 5% level of significance.

A pilot study showed the pooled standard deviation of TSH levels to be 5 mIU/L. Then, the minimum number of subjects required in each group is computed as

$$n = 2 \left[\frac{(z_{1-\alpha/2} + z_{1-\beta}) S}{d} \right]^2 = 2 \left[\frac{(1.96 + 0.84) 5}{3} \right]^2 = 44$$

In the above study if the researchers want 90% power then the sample size in each group is

$$n = 2 \left[\frac{(z_{1-\alpha/2} + z_{1-\beta}) S}{d} \right]^2 = 2 \left[\frac{(1.96 + 1.28) 5}{3} \right]^2 = 51$$

Similarly, in the above example if the level of significance is fixed at 1% and power at 80%, the minimum sample size required is

$$n = 2 \left[\frac{(z_{1-\alpha/2} + z_{1-\beta}) S}{d} \right]^2 = 2 \left[\frac{(2.58 + 0.84) 5}{3} \right]^2 = 65$$

Sample size for testing equality of two population proportions

The sample size required for testing the null hypothesis of equal proportion of a particular condition under study in two independent populations is given

$$n = \frac{(Z_{1-\alpha/2} \sqrt{2PQ} + Z_{1-\beta} \sqrt{P_1Q_1 + P_2Q_2})^2}{(P_1 - P_2)^2}$$

where $P = \frac{P_1 + P_2}{2}$, $Q = 1 - P$,

$Q_1 = 1 - P_1$ and $Q_2 = 1 - P_2$

n is the minimum number of subjects required in each group, P_1 and P_2 are proportions of the disease or condition expected in two independent populations. In the above equation, the denominator term $P_1 - P_2$ is the minimum clinically significant difference between the two proportions and as mentioned earlier the researcher has to fix it. Once the researcher fixes the minimum clinically significant difference and also he knows either P_1 or P_2 , then he can obtain the proportion in the other group.

For example it was planned to conduct a randomized controlled trial to compare the percentage survival 5 years of diagnosis of breast cancer between two treatment groups. One half of the study participants with confirmed breast cancer will be randomly allocated to intervention (a new treatment) group and the other half to the control group (standard treatment). The investigators consider a difference of 6% between the two groups as clinically significant. It is known that the percentage of survival after 5 years in the standard treatment group is 70%. Then one would expect a percentage survival of 76% (or 64%) in the intervention group. Then, the minimum sample size required in each group at 80% power and 5% level of significance is obtained as

(assuming no loss to follow-up).

$$n = \frac{(Z_{1-\alpha/2}\sqrt{2PQ} + Z_{1-\beta}\sqrt{P_1Q_1 + P_2Q_2})^2}{(P_1 - P_2)^2}$$

$$= \frac{(1.96\sqrt{2 \times 0.73 \times 0.27} + 0.84\sqrt{0.70 \times 0.30 + 0.76 \times 0.24})^2}{(0.76 - 0.70)^2} = 851$$

Approximately, 860 breast cancer patients are required in each group to detect a clinically significant difference of 6% between the two treatment groups at 80% power and 5% level of significance. In the above example if the clinically significant difference is fixed at 5% (i.e., in the intervention group we expect 75% survival after 5 years) then the minimum number of subjects required in each group is 1251.

The above mentioned four sample size formulas assume that the sample being selected using simple random sampling technique and the sampling distribution follows normal. There are many more sample size formulae available in the literature depending on the situations; for example, if one wants to estimate a correlation coefficient or test a regression coefficient then one needs to use a different set of formulas. Thus, the sample size formula to be used in a study depends on:

- The main objective of the study
- The primary outcome variable
- The study design used
- The sampling technique used and
- The summary statistics and the type of statistical analysis (estimation or testing of hypothesis) used for the main objective in the study.

There is an abundance of literature available regarding the choice and principles of sample size estimation.¹⁷⁻³² Furthermore, there are many statistical packages that provide estimates of minimum sample size. However, the researchers must provide the relevant information in order to get the correct number and to do that a researcher should have complete conceptual understanding of determining minimum sample size.

CONCLUSION

Statistical methods are applied in almost all stages of a research study from the planning and design stage to the final reporting of findings. Utmost care must be taken in choosing the right statistical method at each stage of a study. Sample size estimation is very important and choosing the right formula is crucial in all types of research studies. The selection of appropriate sample size formula depends on the primary

objective of the study, the type of outcome variable, study design used, the statistical analysis planned, number of groups in the study and sampling technique to be adopted. The number of subjects to be recruited in a study depends on the power of the study, the precision of the estimated value, level of significance or confidence level, clinically significant difference and also other constraints such as money, manpower, availability of subjects and time in particular and feasibility in general. There are many statistical tests and the choice of appropriate analysis depends on the objective of the study, types of the outcome variable, number of variables, sample size, number of groups in the study, whether the groups are related or not and also on distributional assumptions.

REFERENCES

1. Daniel WW. Biostatistics: A Foundation for Analysis in Health Sciences. 7 th ed. Singapore, Asia: John Wiley and Sons Pte. Ltd.; 2004
2. Bland M. An Introduction to Medical Statistics. 2 nd ed. England: Oxford University Press; 1995.
3. Applegate KE, Crewson PE. Statistical literacy. Radiology 2004;230:613-4.
4. Larson MG. Descriptive statistics and graphical displays. Circulation 2006;114:76-81.
5. McCluskey A, Lalkhen GA. Statistics III: Probability and statistical tests. Contin Educ Anaest Crit Care Pain 2007;7:167-70.
6. Montori VM, Kleinbart J, Newman TB, Keitz S, Wyer PC, Moyer V, et al. Tips for learners of evidence-based medicine: 2. Measures of precision (confidence intervals). CMAJ 2004;171:611-5.
7. Curran-Everett D. Explorations in statistics: Confidence intervals. Adv Physiol Educ 2009;33:87-90.
8. du Prel JB, Hommel G, Röhrig B, Blettner M. Confidence interval or P-value?: Part 4 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2009;106:335-9.
9. Gardner MJ, Altman DG. Confidence intervals rather than P-values: Estimation rather than hypothesis testing. Br Med J (Clin Res Ed) 1986;292:746-50.
10. Poole C. Low P-values or narrow confidence intervals: Which are more durable? Epidemiology 2001;12:291-4.
11. Rigby AS. Getting past the statistical referee: Moving away from P-values and towards interval estimation. Health Educ Res 1999;14:713-5.
12. Potter RH. Significance level and confidence interval. J Dent Res 1994;73:494-6.

13. Evans SJ, Mills P, Dawson J. The end of the P value? *Br Heart J* 1988;60:177-80.
14. Attia A. Why should researchers report the confidence interval in modern research? *Middle East Fertil Soc J* 2005;10:78-81.
15. Chow CS, Wang H, Shao J. *Sample Size Calculation in Clinical Research*. USA: Chapman and Hall/CRC Press; 2003.
16. Wittes J. Sample size calculations for randomized controlled trials. *Epidemiol Rev* 2002;24:39-53.
17. Endacott R, Botti M. Clinical research 3: Sample selection. *Accid Emerg Nurs* 2007;15:234-8.
18. Neely JG, Karni RJ, Engel SH, Fraley PL, Nussenbaum B, Paniello RC. Practical guides to understanding sample size and minimal clinically important difference (MCID). *Otolaryngol Head Neck Surg* 2007;136:14-8.
19. Devane D, Begley CM, Clarke M. How many do I need? Basic principles of sample size estimation. *J Adv Nurs* 2004;47:297-302.
20. Columb OM, Stevens A. Power analysis and sample size calculations. *Curr Anaesth Crit Care* 2008;19:12-4.
21. Brasher PM, Brant RF. Sample size calculations in randomized trials: Common pitfalls. *Can J Anaesth* 2007;54:103-6.
22. Schulz KF, Grimes DA. Sample size calculations in randomised trials: Mandatory and mystical. *Lancet* 2005;365:1348-53.
23. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. *Emerg Med J* 2003;20:453-8.
24. Donner A. Approaches to sample size estimation in the design of clinical trials - A review. *Stat Med* 1984;3:199-214.
25. Lwanga KS, Lemeshow S. *Sample Size Determination in Health Studies: A Practical Manual*. Geneva: WHO; 1992.
26. Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *BMJ* 1995;311:1145-8.
27. Parker AR, Bermman GN. Sample size: More than calculations. *Am Stat* 2003;57:166-70.
28. Lenth VR. Some practical guidelines for effective sample size determination. *Am Stat* 2001;55:187-93.
29. Eng J. Sample size estimation: How many individuals should be studied? *Radiology* 2003;227:309-13.
30. Florey CD. Sample size for beginners. *BMJ* 1993;306:1181-4.
31. Woodward M. Sample size, power and minimum detectable relative risk in medical studies. *J R Stat Soc Series DStat* 1992;41:185-96.
32. Bacchetti P, Wolf LE, Segal MR, McCulloch CE. Ethics and sample size. *Am J Epidemiol* 2005;161:105-10.

Please cite this article as: Agnihotram G. Standard statistical methods and calculation of sample size in Medical education research. *Perspectives in Medical Research* 2018;6(1):3-8.

Sources of Support: Nil, Conflict of interest: None declared