



Simple sample size calculator for medical research: a necessary tool for the researchers

Sathian B, Sreedharan J, Banerjee I, Roy B

Correspondence to:

brijeshstat@gmail.com

Dr. Brijesh Sathian, PhD, Assistant Professor, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal.

Dr. Jayadevan Sreedharan, PhD, Professor & Asst Director, Research Division, Gulf Medical University.

Dr. Indrajit Banerjee, Lecturer, Department of Pharmacology, Manipal College of Medical Sciences, Pokhara, Nepal.

Bedanta Roy, Assistant Professor, Department of Physiology, Manipal College of Medical Sciences, Pokhara, Nepal.

Editors for this Article:

Dr. A.K. Pradhan, MBBS, MD. Professor, KIMS, Amalapuram, Editor-In-Chief, Medical Science.

Dr. I. A. Khan, MBBS, MD, former Professor, Physiology, MCOMS, Editorial board member, Medical Science.

Cite this article:

Sathian B, Sreedharan J, Banerjee I, Roy B Simple sample size calculator for medical research: a necessary tool for the researchers. *Medical Science*. 2014, 2(3):141-144.

Information about the article

Received: March. 16, 2014

Revised: May. 25, 2014

Accepted: June. 12, 2014

Published online: Sep 30, 2014

Abstract

It is mandatory to calculate sample size in any type of research in medical science to get a generalized conclusion for the population under study. If the study is well designed with a desired sample size then the standard error will be less and the power and precision will be good. All statistical procedures become valid in this context.

Key words

Sample size, mean, sample size calculator, medical research, proportion.



Background

It is mandatory to calculate sample size in any type of research in medical science to get a generalized conclusion for the population under study. If the study is well designed with a desired sample size then the standard error will be less and the power and precision will be good. All statistical procedures become valid in this context. It is the responsibility of the principal researcher/ author to include the optimum sample size in the protocol and the final report/ thesis/ manuscript [1-10].

Most of the good journals and universities made sample size calculation mandatory. Some of the medical researchers face severe problems in calculating sample size because of the mathematical formula.

Sample size calculation problems in medical research can be classified generally as two, one is proportion problem and the other is mean problem. Researcher will be focusing on the objective as to find out prevalence, incidence, and percentage. Then it is proportion problem.

Ex: Prevalence of HBV, HIV, TB

If one should focus on average, then it will be mean problem.

Ex: Lipid Profiles like TC, TG, LDL, VLDL

We can again sub classify this into another two groups so we will be having four type of sample size calculations in general medical research. Single mean, double mean, single proportion, double proportion.

1. Single mean

Average pulse rate of 240 medical students of Manipal College of Medical Sciences were 70/mts with a Standard Deviation of 9 beats. Determine the sample size to verify the same research in Gandaki Medical College if the allowable error is 2 beats at 5% risk.

Given,

Mean (\bar{x}) = 70/ min
Standard deviation (σ) = 9 beats
Error (E) = 2
Sample (n) = ?

$$\begin{aligned} \text{We know that for 95\% CI} \\ n &= 1.96^2 \sigma^2 / E^2 \\ &= 1.96^2 (9)^2 / 2^2 \\ &= 78 \end{aligned}$$

Therefore, the minimum sample size is 78

How to use the Excel Spread Sheet Programme (Supplementary file)

1st Step: the values should be entered to the Yellow highlights

Researcher should enter the sample size and standard deviation of the main study variable of the pilot study, which was done prior to the study or any of the study from the review of literature. Researcher will get the allowable error in the right end.

2nd Step Enter the values to the Blue Highlights

Researcher should enter the allowable error which he thinks better for the study [It can be the allowable error calculated early or researcher can even reduce the value if he feels it necessary]. Standard Deviation should be entered. Researcher will get the sample size in the right end.

2. Difference between 2 means

Simplified formula for difference in means

$$n = \frac{2\sigma^2 (Z_{\beta} + Z_{\alpha/2})^2}{\text{difference}^2}$$

n = Sample size in each group (assumes equal sized groups)

σ = Standard deviation of the outcome variable

Z_{β} = Represents the desired power (typically .84 for 80% power).

$Z_{\alpha/2}$ = Represents the desired level of statistical significance (typically 1.96).

difference = Effect Size (the difference in means)

How to use the Excel Spread Sheet Programme (Supplementary file)

1st Step Enter the values to the Yellow highlights

Researcher should enter the sample size and standard deviation of the main study variable of the pilot study which was done prior to the study or any of the study from the review of literature. Researcher will get the Allowable Error in the right end.

2nd Step Enter the values to the Blue Highlights

Researcher should enter the allowable error which he think better for the study [It can be the allowable error calculated earlier or researcher can even reduce the value if he feel it is necessary]. Enter the Standard Deviation. Sample size will be displayed in the right end.

3. Single proportion or percentage

A team of mental health professionals wanted to conduct a survey to evaluate the illegal drug use among medical students in their city. In a Pilot study of 100 showed percentages of drug users among medical students is known



to be 23%. How many medical students should be interviewed so that the investigators are 95% confidence?

Answer:

Given, (At 95% confidence)

$P=23$

$Q=100-p = 100-23= 77$

Margin of error $E= \sqrt{(4 * P * (1 - P))/100}$

Margin of error $E= \sqrt{(4 * 0.23 * 0.77)/100}$
 $= 0.08$

Therefore,

$N(\text{sample})=(1.96)^2 \times P \times Q$
 $(E)^2$

$N= \frac{(1.96)^2 \times 0.23 \times 0.77}{(0.08)^2}$

$n= 111$

Hence the no of sample required is 111

How to use the Excel Spread Sheet Programme (Supplementary file)

1st Step Enter the values to the Yellow highlights

Researcher should enter the sample size and the proportion [Percentage divided by 100] of the main study variable in the pilot study you have done prior to the study or any of the study from the review of literature. Allowable Error will be in the in the right end.

2nd Step Enter the values to the Blue Highlights

Researcher will enter the allowable error which he thinks better for the study [It can be the allowable error calculated earlier or researcher can even reduce the value if he feels it necessary]. The proportion should be entered. The Sample size in the right end will appear.

4. Difference between 2 proportions or percentages

Simplified formula for difference in proportion

$$n = \frac{2(\bar{p})(1 - \bar{p})(Z_{\beta} + Z_{\alpha/2})^2}{(p_1 - p_2)^2}$$

n = Sample size in each group (assumes equal sized groups)

$(\bar{p})(1 - \bar{p})$ = A measure of variability (similar to standard deviation)

Z_{β} = Represents the desired power (typically .84 for 80% power).

$Z_{\alpha/2}$ = Represents the desired level of statistical significance (typically 1.96).

$P_1 - P_2$ = Effect Size (the difference in proportions)

$\bar{p} = (p_1 + p_2) / 2$

How to use the Excel Spread Sheet Programme (Supplementary file)

1st Step Enter the values to the Yellow highlights

Researcher should enter the sample size and the proportion [Percentage divided by 100] of the main study variable in the pilot study, the researcher has done prior to the study or any of the study from the review of literature. Allowable Error will be available in the right end.

2nd Step Enter the values to the Blue Highlights

Researcher should enter the allowable error which he thinks better for his study [It can be the allowable error he calculated early or he can even reduce the value if he feel it necessary]. The proportion should be entered. Sample size in the right end will appear.

Conclusion

Sample size calculators are simple and very useful to the common researchers who don't have strong mathematical background. Properly designed medical research will end up with a useful outcome. Researcher should calculate the sample size for their study rather than taking any fixed and assumed sample size.

Competing interests

None declared.

Authors' contribution

BS, JS drafted the manuscript and revised. IB and BR critically revised the manuscript, edited. Final manuscript is approved by all authors.

Authors' information

Dr. Brijesh Sathian, PhD, Assistant Professor, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal.

Dr. Jayadevan Sreedharan, PhD, Professor and Asst Director, Research Division, Gulf Medical University.



Dr. Indrajit Banerjee, Assistant Professor, Department of Pharmacology, Manipal College of Medical Sciences, Pokhara, Nepal.

Bedanta Roy, Assistant Professor, Department of Physiology, Manipal College of Medical Sciences, Pokhara, Nepal.

References

1. Sathian B, Sreedharan J, Baboo NS, Sharan K, Abhilash E S, Rajesh E. Relevance of Sample Size Determination in Medical Research. *Nepal Journal of Epidemiology* 2010; 1(1): 4-10.
2. Roy B, Banerjee I, Sathian B, Mondal M, Saha CG. Blood Group Distribution and Its Relationship with Bleeding Time and Clotting Time: A Medical School Based Observational Study among Nepali, Indian and Srilankan Students. *Nepal Journal of Epidemiology* 2011;1(4):135-40.
3. Roy B, Banerjee I, Sathian B, Mondal M, Kumar SS, Saha CG. Attitude of Basic Science Medical Students towards Post Graduation in Medicine and Surgery: A Questionnaire based Cross-sectional Study from Western Region of Nepal. *Nepal Journal of Epidemiology* 2010; 1(4):126-34.
4. Banerjee I, Roy B, Sathian B, Banerjee I, Kumar SS, Saha A. Medications for Anxiety: A Drug utilization study in Psychiatry Inpatients from a Tertiary Care Centre of Western Nepal. *Nepal Journal of Epidemiology* 2010; 1(4):119-25.
5. Mittal A, Sathian B, Kumar A, Chandrasekharan N, Farooqui MS, Singh S, Yadav KS. Hyperuricemia as an Additional Risk Factor for Coronary Artery Disease: A Hospital Based Case Control Study in Western Region of Nepal. *Nepal Journal of Epidemiology* 2011;1(3):81-5.
6. Banerjee I, Jauhari AC, Bista D, Johorey AC, Roy B, Sathian B. Medical Students View about the Integrated MBBS Course: A Questionnaire Based Cross-sectional Survey from a Medical College of Kathmandu Valley. *Nepal Journal of Epidemiology* 2011;1(3): 95-100.
7. Mittal A, Sathian B, Poudel B, Farooqui MS, Chandrasekharan N, Yadav KS. The Significance of Hepatobiliary Enzymes for Differentiating Liver and Bone Diseases: A Case Control Study from Manipal Teaching Hospital of Pokhara Valley. *Nepal Journal of Epidemiology* 2011;1(5): 153-9.
8. Poudel B, Mittal A, Yadav BK, Sharma P, Jha B, Raut KB. Estimation and Comparison of Serum Levels of Sodium, Potassium, Calcium and Phosphorus in Different Stages of Chronic Kidney Disease. *Nepal Journal of Epidemiology* 2011;1 (5): 160-7.
9. Mittal A, Sathian B, Chandrasekharan N, Lekhi A, Rahib R, Dwedi S. Hepatic Steatosis and Diabetes Mellitus: Risk Factors, Pathophysiology and with its Clinical Implications: A Hospital Based Case Control Study in Western Region of Nepal. *Nepal Journal of Epidemiology* 2011;1(2):51-56.
10. Banerjee I, Roy B, Banerjee I, Sathian B, Mondol M, Saha A. Depression and its Cure : A Drug Utilization Study from a Tertiary Care Centre of Western Nepal. *Nepal Journal of Epidemiology* 2011;1 (5):144-52.