

REVIEW ARTICLE



Estimation of sample size in dental research

Sandhya Jain¹, Alpana Gupta¹, Deshraj Jain²

¹Department of Orthodontics, Government College of Dentistry, Indore, Madhya Pradesh, India, ²Department of Prosthodontics, Government College of Dentistry, Indore, Madhya Pradesh, India

Keywords

Level of significance, power, sample size, Type I error, Type II error

Correspondence

Dr. Sandhya Jain, Department of Orthodontics, Government College of Dentistry, Indore, Madhya Pradesh, India. Phone: +91-9425045455, Email: researchorthodontics@gmail.com

Received 16 March 2015;

Accepted 19 April 2015

doi: 10.15713/ins.idmjar.9

Abstract

Sample size determination is an important step while planning a statistical study. The determination of minimum required sample size is extremely important not only for ethical and economic purposes but also to achieve scientifically and statistically sound results. The objective of this article is to emphasize the importance and general principles of sample size calculation for different study designs to produce quality results. This article emphasizes various methods of sample size estimation commonly used in dental research especially formulae method used for simple random sampling. Different formulae are described for comparing different outcome measure (means or proportions) by taking dental data as examples.

Introduction

The sample size estimation is the mathematical procedure for deciding how many individuals or specimen should be included in the investigation from the population. It must be carried out before collecting the data.^[1] Inappropriate sample size cannot produce a useful result and expose the participants to unnecessary risk. The determination of optimum sample size or minimum required sample size is extremely important not only for ethical and economic purposes but also to achieve scientifically and statistically sound results.^[2]

The sample size largely depends on the study design. The study design can be either descriptive type and/or analytical type. The descriptive type study focuses on the entire population describing the pattern of occurrence of disease, for e.g., prevalence of dental caries in the school district relative to age, gender, socioeconomic status, etc.

Analytical studies, ascertain statistical association between two things by the test of significance. Analytical studies can be either observational type and/or experimental type. In observational type, we simply observe either retrospectively (case-control study) or prospectively (most of cohort study).

In experimental studies, we test the efficacy of a new drug or a new treatment with the conventional method.

Steps in designing a study

Thus, following are the steps^[3] for quality dental research:

1. Formation of the research question with its aim and objectives
2. Hypothesis formation - e.g., null hypothesis for comparative study
3. Decide type of study with its outcome measures
4. Data-qualitative and quantitative
5. Deciding about study population with sampling method and sample size estimation
6. Randomization and blinding in a comparative study
7. Check validity and reliability of the method
8. Selection of proper statistical tests (parametric and non-parametric).

Factors Affecting Sample Size [Table 1]

Study design and sampling method

Descriptive study needs hundreds of subjects to give acceptable confidence interval (CI) (amount of error that is allowed in the statistical data and analysis) while experimental study generally need lesser number of subjects. For example, a descriptive study conducted on 600 samples to evaluate the relationship between patient's education level and knowledge on dental

Table 1: Factors affecting sample size^[3]

Factor	Magnitude	Required sample size
SD	Large	Large
	Small	Small
Effect size	Large	Small
	Small	Large
Power	More	Large
	Less	Small
Level of significance	Large	Small
	Small	Large

SD: Standard deviation

treatment of caries/periodontal disease and oral hygiene through a specific questionnaire.^[4] If multiple subgroups are to be analyzed, the sample should be increased to ensure adequate number of subjects in each group. The cross-over control trials need one-quarter of a number required compared to control group because every subject (both cases and control) gets the experimental treatment in cross-over study. A study design with one-tailed hypothesis (detect the difference or effect or association in one direction only) requires 20% lesser subjects as compared to two-tailed studies (detect the difference or effect or association in either direction). Non randomized studies need 20% more subjects compared to randomized studies in order to accommodate the confounding factors.^[5]

Simple formulae can be used to calculate the sample size for simple random sampling but for more complex designs like stratified random samples, several other factors are taken into account e.g., variances of subpopulation, strata, or clusters before estimation of variability in the entire population. Cluster or multistage sampling methods require a larger sample size to achieve the same precision. Therefore, the calculated sample sizes using the formulae need to be multiplied by the design effect^[6] (The design effect is basically the ratio of the actual variance, under the sampling method actually used, to the variance computed under the assumption of simple random sampling). The design effect for cluster random sampling is taken as 1.5-2. For the purposive sampling, convenience or judgment sampling, design effect may cross 10.

Outcome measure of the study (categorical variable or continuous)

Larger sample size is required to assess the categorical variable (qualitative) compared to continuous outcome variable (quantitative).

Effect size

It is the minimum clinically important difference between cases and control. It is inversely proportional to sample size. We can estimate the effect size based on previously reported or preclinical studies.

Standard deviation (SD)

SD for sample size calculation is obtained either from previous studies or from the pilot study. Larger the SD, larger is the sample size required in a study.

Significance level (α)

Sample size is determined according to ' α ' or Type I error - how much error is allowable in a study. Type I error is the probability of falsely claiming the difference in reading but actually there is no difference (false positive), and the null hypothesis is rejected erroneously. Type I error is fixed in advance, and its upper limit of tolerance is known as level of significance. The alpha level used in determining the sample size in most of the academic research studies are either 0.05 or 0.01.^[7] For critical results, the study should be more precise hence α - error is set at lower level. Lower the alpha level, larger is the sample size and more precise will be the study.

Power

Power of the study is determined before collecting the data, as it helps in determining the sample size. Power is the probability that the test will correctly identify the difference if it is there. Usually, most study accepts power of 80%^[8] i.e., 20% chance of missing the real difference. Sometimes a larger study power is set at 90% i.e., 10% possibility of false negative results due to β error. Type II error is falsely stating that the two variables are equivalent when they are actually different. Power proportionality increases as the sample size for study increases.

There are several approaches to determining the sample size. These include using:

- A. Using published tables
- B. Applying formulas to calculate a sample size (for simple random sampling)
- C. Nomogram
- D. Computer software

Table method

There are some published tables, which provide the sample size for a given set of criteria. Tables present sample sizes based on combinations of precision, confidence levels, and variability. The sample sizes presume that the data being measured is distributed normally in a population. Additional 10-20% subjects are required to allow adjustment of other factors such as withdrawals, missing data, lost to follow-up etc.

Example: How large a sample of patients should be followed up if an investigator wishes to estimate the incidence rate of Class II malocclusion in a particular area within 10% of its true value with 95% confidence? If the level of precision is $e = 0.10$ (Precision is the degree to which a variable has the same value when measured several times, It is a measure of consistency) and confidence level of 95%, from the given table a total sample size of 385 would be needed.

- Table 2 given below can be used to calculate the sample size making the desired changes in the relative precision and confidence level e.g. if the level of confidence is reduced to 90%, then the sample size would be 271.

Table 2: Sample size according to relative precision and confidence level

Relative precision (e)	Confidence level		
	99%	95%	90%
0.01	66358	38417	27061
0.02	16590	9605	6766
0.03	7374	4269	3007
0.04	4148	2402	1692
0.05	2655	1537	1083
0.06	1844	1068	752
0.07	1355	785	553
0.08	1037	601	423
0.09	820	475	335
0.10	664	385	271

Such table that give readymade sample sizes are available for different designs and situation.

Formula method

Several formulas are used to calculate sample size.^[9,10]

There are two approaches to sample size calculations:

1. Precision-based
2. Power-based

It is very important to understand that the method of sample size calculation is different for different study designs and one formula for sample size calculation cannot be used for all study designs.

For descriptive studies: Descriptive study deals with estimation of population parameters. Two commonly used parameters are the proportion and the mean (measure of central tendency).

Formula calculating a sample for proportions:

- For population that are large, Cochran (1963) developed the Equation to yield a representative sample for proportions.^[9]

$$n_0 = \frac{Z^2 pq}{e^2}$$

n_0 = desired sample size (when the population > 10,000)

Z = standard normal deviate; usually set at 1.96 which corresponds to 95% confidence level (Different values of Z_α and Z_β is given in Table 3)

p = proportion in the target population. If there is no reasonable estimate, use 50% (i.e. 0.5)

$q = 1 - p$ (proportion in the target population)

e = precision level (degree of accuracy required usually set at 0.05 level, occasionally at 2.0)

- Simple formula can be used instead of above formulae, but it will give a sample size slightly more, whenever precise estimate is required then above formulae should be used.

$$n = 4pq / e^2$$

Examples:

1. From a pilot study, it is reported that the patients having malocclusion coming to orthodontic department has dental

Table 3: Standard normal deviates Z_α and Z_β , for selected values (Z_α and Z_β values are obtained from standard normal distribution, for the predetermined values of α and β)^[11]

α and β	One-sided Z_α and Z_β	Two-sided Z_α
0.05	1.64	1.96
0.10	1.28	1.64
0.20	0.84	1.28

Z_β is the same for one-sided and two-sided tests

midline deviation in 28% of cases. It is decided to have 95% CI and 10% variability in the estimated 28%. How many patients are necessary to conduct this type of prevalence study?

The resulting sample size is:

Here p - 28%

q i.e. $1 - p = 72\%$

$Z = 1.96$

$d = 10\%$ of 28% = 2.8

$$n_0 = Z^2 pq / e^2 = (1.96)^2 \times 28 \times 72 / (2.8)^2 = 987 \text{ patients}$$

2. From a pilot study conducted on 100 patients coming to the prosthodontic department. The prevalence of headache in patients with occlusal disturbance is 67.4%. Now we wish to calculate the prevalence of same in a large population. How much sample is required to conduct this type of prevalence study? Assuming 90% CI and 5% variability in estimated prevalence. The resulting sample size is:

Here p - 67.4%

q i.e. $1 - p = 32.6\%$

$Z = 1.64$ (For 90% CI)

$d = 5\%$ of 67.4% = 3.37

$$n_0 = Z^2 pq / e^2 = (1.64)^2 \times 67.4 \times 32.6 / (3.37)^2 = 521 \text{ patients}$$

Formula for sample size for the mean

- The formula of the sample size for the mean is similar to that of the proportion, except for the measure of variability.
- The formula for the mean employs σ^2 instead of $(p \times q)$, as shown in equation

$$n_0 = \frac{Z^2 \sigma^2}{e^2}$$

Where n_0 is the sample size,

- Z is standard normal deviate
- e is the desired level of precision
- σ^2 is the variance of an attribute in the population

Similarly simple formula for the mean is:

$$n = 4 \sigma^2 / e^2$$

Example: From previous studies it is reported that Facial angle in Downs analysis has a mean value 87.8 degree with SD of 3.6, consider the precision as 0.8. Assume 95 % CI. The required sample size to conduct such type of study in a different population.

- The resulting sample size is:

$Z = 1.96$ which correspond to 95% confidence level

$\sigma = 3.6$

$e = 0.8$
 $n = 1.96 \times 1.96 \times 3.6 \times 3.6 / 0.8 \times 0.8 = 78$ patients

For analytic study

Formulae:
 When proportion is the parameter of our study

$$n = \left[\frac{r+1}{r} \right] \frac{p \times q (Z_\beta + Z_\alpha)^2}{d^2}$$

Where
 Z_α (two sided test) = Z value for α level
 Z_β = Z value for β level
 p = average percentage between two groups
 $q = 1 - p$
 d = Clinically meaningful difference between two groups
 r = ratio of controls to cases
 If, cases are equal to controls
 Then

$$n = \frac{2p \times q (Z_\alpha + Z_\beta)^2}{d^2}$$

Suppose we are planning a case-control study to evaluate the association between the extent and severity of gingival recession in orthodontically treated and untreated adult patients considering confidence level of 95%. From the pilot study on eight patients in each group, the data are:

Exposure	Gingival recession (%)	No gingival recession	Total patients
Ortho t/t	6 (75)	2	8
No Ortho t/t	3 (37)	5	8

From the formula:
 $p = 75 + 37 / 2 = 56\%$
 $q = 100 - 56 = 44\%$
 $Z_\alpha = 1.96$ for α at 0.05, $Z_\beta = 1.282$ for β at 0.10
 $d = 75 - 37 = 38$
 Sample size = $(1.96 + 1.282)^2 \times 56 \times 44 \times 2 / (38)^2 = 36$ /group

When mean is the parameter of our study:

$$n = \left[\frac{r+1}{r} \right] \frac{\sigma^2 (Z_\beta + Z_\alpha)^2}{d^2}$$

Where
 Z_α (two sided test) = Z value for α level
 Z_β = Z value for β level
 σ = Common SD between two groups
 d = clinically meaningful difference of mean between two groups
 r = ratio of controls to cases
 If cases are equal to controls
 Then

$$n = \frac{2 \times \sigma^2 (Z_\alpha + Z_\beta)^2}{d^2}$$

Example: Suppose in above example mean value of gingival recession in case group is 0.36 mm and SD is 0.42 and mean values of gingival recession in control group are 0.22 and SD is 0.32. To conduct such type of study, the required sample size is:

From above formula:
 $Z_\alpha = 1.96$ for α at 0.05, $Z_\beta = 1.282$ for β at 0.10
 $S = 0.42 + 0.32 / 2 = 0.37$
 $d = 0.36 - 0.22 = 0.14$
 $n = (1.96 + 1.282)^2 \times (0.37)^2 \times 2 / (0.14)^2 = 147$ /group

There are certain formulae which are power based (preferred for clinical trial study) given by Pocock. Suppose we will conduct a sample calculation for a trial with qualitative outcomes, 1:1 allocation ratio, and a 2-sided test we will use the formula given below.^[12,13]

Formula for proportion

$$n = f(\alpha, \beta) \times \frac{\pi_1 \times (1 - \pi_1) + \pi_2 \times (1 - \pi_2)}{(\pi_1 - \pi_2)^2}$$

- n is the required sample size per group
- π_1 is the expected percentage of proportion in one group
- π_2 is the expected proportion in other group
- $f(\alpha, \beta)$ is a function of α and β derived from the standard normal distribution (Table 4)

For example: Our aim is to compare failures of bonded molar tube with chemically cured or a light-cured adhesive. Let us assume that we have found a study that reported on the proportion of failures for the chemically cured adhesive is 34% and proportion of failures for light-cured adhesive is (20%), and the desired significance is set at 5% and power is at 90%. The required sample size is:

$$10.5 \times 0.34 \times (1 - 0.34) + 0.2 \times (1 - 0.2) / (0.34 - 0.2)^2 = 128$$
/group

Formula for mean

To compare the mean in one group to the mean in another. The number required in each group is given by

$$n = f(\alpha, \beta) \times \frac{2sd^2}{(\mu_1 - \mu_2)^2}$$

$\mu_1 - \mu_2$ is the smallest difference in means
 sd - Is the standard deviation (estimated from the previous study)

For example: Suppose we want to compare the mean value of tooth separation with two different separator. From previous study, it is found that the mean value of tooth separation between 2nd PM and 1st molar by using elastomeric separator is 0.14 mm and SD is 0.03 on day 1 and mean value of tooth separation between 2nd PM and 1st molar by using spring separator is 0.12 mm and SD is 0.03 on day 1, desired significance level is set at 5% (0.05) and power is at 90%.

$$n = 10.5 \times 2 \times (0.03)^2 / (0.02)^2 = 47$$
/group

Nomogram

Altman^[16] created a quick method for sample size calculation. To calculate sample size with this method, one needs to specify the study (Group 1) and the control group (Group 2).

In Figure 1, on the right vertical line, is power level and on left vertical line there is the effect size, required sample size is shown on the diagonal line in the center. The effect size in terms of the standardized difference is defined as the ratio of the minimum difference to be detected to the SD.

Table 4: Values of function $f(\alpha, \beta)$ for different values of α and β ⁽¹²⁻¹⁵⁾

α	β			
	0.05 (95% power)	0.1 (90% power)	0.2 (80% power)	0.5 (50% power)
0.05	13.0	10.5	7.85	3.84
0.01	17.8	14.9	11.7	6.63

By joining with a straight line the specific values for two of the variables the third value can easily be known from the nomogram. For example, when using quantitative outcomes, if the difference to be detected is 30 units, and the SD is 50 units, then the standardized difference is $30/50 = 0.6$.

If we draw a line from the left vertical line starting at 0.6 and connect it with the value of 0.90 on the right vertical line representing power, the point where the connecting line crosses the diagonal sample line indicates the total required sample size for both arms. In this example, the line crosses between 100 and 120.

Software

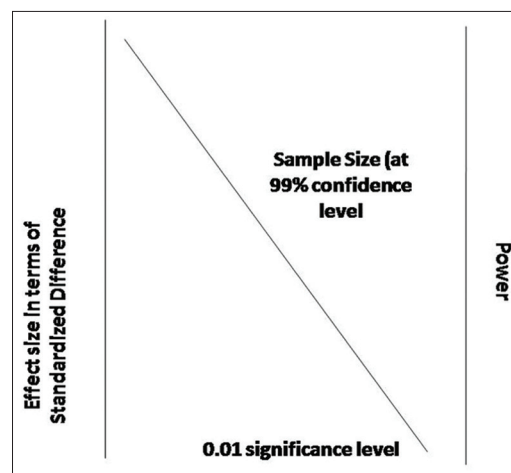
Numerous programs are available for performing power and sample size calculations. These include nQuery Advisor, PASS, PS, Russ Lenth's power, minitab, Stata, SAS Power, etc. PROC POWER and GLMPower are new additions to SAS version 9.0.

Conclusion

Sample size determination is one of the most essential parts of any research/study. Sample size calculation provides you with the appropriate number of participants you will need to address your research question properly. Determining sample size is a very important issue because the larger the sample size, the higher the degree of accuracy but samples that are too large may waste time, resources and money while samples that are too small may lead to inaccurate results.

The approach for the calculation of sample size for our study is first of all our research question must be clear that means what we are going to evaluate and the outcome measure of the study (continuous or categorical.) Then specify level of significance (α) and power ($1-\beta$) and effect size (from the previous study or from the pilot study). Use the appropriate formula according to what we want to measure i.e., mean or proportion (mean or proportion is either obtained from pilot study or from previously conducted similar type of study).

Increasing power or precision (without having to increase sample size): In this situation, we can change the design of the study. Using a paired design can be more efficient because here we are considering individual (within-subject) differences rather than between-subject differences; the former are usually less variable than the latter (i.e., the standard error is lower) and, therefore, the sample size required to detect a given difference will be lower. To decrease inter subject variation, the same subject can be examined on right and left side rather taking

**Figure 1:** Diagrammatic representation of nomogram

more subjects. eg., For comparing the effect of new treatment modality in comparison to the conventional method, one can use right and left segment within the same individual like comparison of two different methods of canine retraction or molar distalization.

Finally, one must need to consult a statistician, who can usually make precise calculations to determine the desirable sample size.

References

1. Last JM. Making the Dictionary of Epidemiology. *Int J Epidemiol* 1996;25:1098-101.
2. MacFarlane TV. Sample size determination for research projects. *J Orthod* 2003;30:99-100.
3. Jain S, Sharma N, Jain D. Basic fundamentals of designing a quality research. *J Adv Med Dent Sci Res* 2015;3:1-8.
4. Gomes AP, da Silva EG, Gonçalves SH, Huhtala ME, Martinho FC, Gonçalves SE, *et al.* Relationship between patient's education level and knowledge on oral health preventive measures. *Int Dent Med J Adv Res* 2015;1:1-7.
5. Habib A, Johargy A, Mahmood K, Humma. Design and determination of the sample size in medical research. *J Dent Med Sci* 2014;13:21-31.
6. Naing L, Win T, Rusli BN. Practical issues in calculating the sample size for prevalence studies. *Arch Orofac Sci* 2006;1:9-14.
7. Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd ed. New York, NY: Wiley; 1981. p. 45.
8. Phillips C. Sample size and power. What is enough? *Semin Orthod* 2002;8:67-76.
9. Cochran WG. *Sampling Techniques*. 3rd ed. New York: John Wiley & Son; 1977.
10. Sathian B, Sreedharan J, Baboo N, Sharan K, Abhilash ES, Rajesh E. Relevance of sample size determination in medical research. *Nepal J Epidemiol* 2010;1:4-10.
11. Suresh K, Chandrashekara S. Sample size estimation and power analysis for clinical research studies. *J Hum Reprod Sci* 2012;5:7-13.

12. Pocock SJ. *Clinical Trials: A Practical Approach*. Chichester, United Kingdom: Wiley; 1983. p. 125-9.
13. Pandis N, Polychronopoulou A, Eliades T. Sample size estimation: An overview with applications to orthodontic clinical trial designs. *Am J Orthod Dentofac Orthop* 2011;140:e141-6.
14. Pandis N. Sample calculations for comparing proportions. *Am J Orthod Dentofacial Orthop* 2012;141:666-7.
15. Pandis N. Sample calculations for comparison of 2 means. *Am J Orthod Dentofacial Orthop* 2012;141:519-21.
16. Altman DG. Statistics and ethics in medical research: III How large a sample? *Br Med J* 1980;281:1336-8.

How to cite this article: Jain S, Gupta A, Jain D. Estimation of sample size in dental research. *Int Dent Med J Adv Res* 2015;1:1-6.