

How to calculate the sample size for animal studies?

Hitesh Shah

Department of Community Medicine
Govt. Medical College, Surat (India)

Correspondence to:

Dr. Hitesh Shah
Department of Community Medicine
Govt. Medical College, Surat (India)
395001
E Mail – drhitesh12@rediffmail.com

Received – 19-3-2011

Accepted - 23-3-2011

Researcher must calculate the sample size before starting of any animal study. It should be adequate enough to detect a small significant difference between the groups. Small sample size is not only responsible for the insignificant result but also for the less power of the study.

Calculation of sample size involves complex statistics but it can be simplified to help the researchers who are not from statistical background. Calculation of sample size depends on the type of study. Mainly animal studies are of two types-

1. Pilot and Exploratory experiments
2. Experiments to test a formal hypothesis

Pilot and Exploratory experiments

Pilot studies are performed to know about the precision of measurement of variables which is to be measured in main study and to check the logistics of the proposed experiments. The sample size in pilot studies is based on the researcher's previous experiences or guess work as no prior data is available. Pilot studies are useful for the calculation of sample size in main study as factors which needed for calculation of sample size like standard deviation and effect size can be measured by this.

Exploratory studies are performed to generate new hypothesis. Aim of these studies is to see the trend or pattern of responses. There is no need to use the test of significance. Sample size of these studies is calculated on the basis of guesswork which may be based

on previous studies. Data obtained (SD, difference between means of two samples) from these studies are helpful in calculation of sample size in main study.

Experiments to test a formal hypothesis

There are three type of variable a researcher may measure.

1. Dichotomous variables (nominal)
e.g - alive/dead, blood groups
2. Continuous variables (ratio, interval, ordinal)
e.g- Weight, volume (ratio)
Temperature (interval)
Any scale (ordinal)
3. Time to occurrence of events.
e.g. - Mortality after 6 months

A. Sample size for dichotomous variables.

In these types of data variables are measured in rates or proportions. The aim of experiment is to see the difference between proportions in two groups. Prerequisite for sample size calculation for this type of data is-

1. Power of the study ($1 - \beta$) (β is type II error)
(Usually $1 - \beta$ is fixed at $> 80\%$)
2. Significance level (α) (α is type I error)
(Usually α is fixed at 5% i. e. $p < 0.05$)
3. Difference of outcome between two group, researcher wishes to detect.

Power of study and significant levels are fixed values. The difference of outcome between two groups fixed on the basis of previous studies or pilot studies. If researcher is interested in multiple end points than he/she should calculated the sample size for most important primary end point. If all the end points are important than sample size for all the endpoints should be calculated and biggest size should be used in experiment. Formula for calculation of sample size for dichotomous variables are-
Sample size for one group is –

$$n = C \frac{p_c q_c + p_e q_e}{d^2} + \frac{2}{d} + 2$$

p_c - Proportion of events in control group

q_c - $1 - p_c$

p_e - Proportion of events in experimental group

$q_e = 1 - p_e$

d - Difference of outcome between two group, researcher wishes to detect.

C - Constant. It measured from the table given below. One should know α and $1 - \beta$ to calculate the constant.

Table for calculation of C when α and $1 - \beta$ are known

$\alpha \rightarrow$	0.05	0.01
$1 - \beta \downarrow$		
0.8	7.85	11.68
0.9	10.51	14.88

So if α is 0.01 and $1 - \beta$ is 0.9 than C will be 14.88.

B. Sample size for continuous variables

1. Studies comparing the unpaired data

Here the mean of variables of two groups are measured and compared.

Prerequisites of calculation of sample size in unpaired data are-

1. Standard deviation (SD) - It can be calculated by previous studies or pilot study.
2. Power of the study ($1 - \beta$) (β is type II error)
(Usually $1 - \beta$ is fixed at $> 80\%$)
3. Significance level (α) (α is type I error)
(Usually α is fixed at 5% i. e $p < 0.05$)
4. Difference of outcome between two group, researcher wishes to detect.

Standard deviation and effect size should be calculated from previous studies or pilot study (if former not available). α and β are arbitrary values which are fixed at the start of study.

Sample size per group in this condition is-

$$n = 1 + 2C \left(\frac{s}{d} \right)^2$$

s - Standard deviation

d - Expected difference between two means

C - Constant

Constant should be measured from table given above (see sample size calculation for dichotomous variables).

2. Studies comparing the paired data

In paired data, there is only one group and variables are measured before intervention and after intervention. Mean value of variable/s are compared. In paired data sample size per group is-

$$n = 2 + C \left(\frac{s}{d} \right)^2$$

s- Standard deviation

d- Expected difference between two means

C- Constant

Constant should be measured from table given above (see sample size calculation for dichotomous variables).

3. Correlation coefficient between two continuous variables in single group

In a group of animal two variables may be correlated. Researcher uses correlation coefficient r to see the strength of association.

Sample size formula for the calculation of correlation coefficient is-

$$n = 3 + \frac{4C}{\left[1n \left(\frac{1+r}{1-r} \times \frac{1-r_o}{1+r_o} \right) \right]^2}$$

r- Postulated correlation coefficient

r^o – specified correlation coefficient

C- Constant

4. Single group experiments

If aim is to determine whether an event has occurred like presence of particular pathogen in colony of animals, then sample size is-

$$= \frac{\log \alpha}{\log p}$$

α- Significance level- fixed by researcher

p- Proportion of colony which is not affected.

C. Sample size for time to occurrence of events

This is a simplified approach. The first approach is to estimate sample size using the proportions in the two experimental groups exhibiting the event by a certain time. This method converts time to an event into a dichotomous variable. The second approach is to treat time to occurrence as a continuous variable. This approach is applicable only if all animals are followed to event occurrence (e.g., until death or time to exhibit a disease such as cancer), but it cannot be used if some animals do not reach the event during the study.

This is the simplified approach. For more elaborate discussion readers are requested to consult some standard books of biostatistics.

References:

1. Kirkpatrick LA, Feeney BC. 2000. A Simple Guide to SPSS for Windows: Versions 8, 9, and 10. 4th Ed. Wadsworth Publishing Co. <www.wadsworth.com>.
2. SAS. 2000. Step-by-Step Programming with Base SAS Software. Cary NC: SAS Institute Inc.
3. Fleiss JL. 1981. Statistical Methods for Rates and Proportions. 2nd Ed. New York: Wiley.
4. Snedecor GW, Cochran WG. 1989. Statistical Methods. 8th Ed. Ames: Iowa State Press.
5. Cohen J. 1988. Statistical Power Analysis for the Behavioral Sciences. 2nd Ed. Mahwah NJ: Lawrence Erlbaum Associates.