

Relevance of Sample Size Determination in Medical Research

Sathian B¹, Sreedharan J², Baboo N S³, Sharan K⁴, Abhilash E S⁵, Rajesh E⁶

¹ Assistant Professor, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal.

² Assistant Director, Research Division, Gulf Medical University, Ajman, UAE.

³ Professor, Department of Physiology, Manipal College of Medical Sciences, Pokhara, Nepal.

⁴ Assistant Professor, Department of Radiotherapy, Manipal University, India.

⁵ Lecturer, Department of Biological Sciences, Grammar School, UAE.

⁶ Assistant Professor, Department of Health and Behavioural Sciences, Mahatma Gandhi University, India.

Review Article

Corresponding Author:

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

Email: drsathian@gmail.com

Abstract

Sample size determination is one of the central tenets of medical research. If the sample size is inadequate, then the study will fail to detect a real difference between the effects of two clinical approaches. On the contrary, if the sample size is larger than what is needed, the study will become cumbersome and ethically prohibitive. Apart from this, the study will become expensive, time consuming and will have no added advantages. A study which needs a large sample size to prove any significant difference in two treatments must ensure the appropriate sample size. It is better to terminate such a study when the required sample size cannot be attained so that the funds and manpower can be conserved. When dealing with multiple sub-groups in a population the sample size should be increased the adequate level for each sub-group. To ensure the reliability of final comparison of the result, the significant level and power must be fixed before the sample size determination. Sample size determination is very important and always a difficult process to handle. It requires the collaboration of a

specialist who has good scientific knowledge in the art and practice of medical statistics. A few suggestions are made in this paper regarding the methods to determine an optimum sample size in descriptive and analytical studies.

Key Words

Sample size, Power analysis, Medical research

Background

In Medical research, it is important to determine a sample size sufficient enough to ensure reliable conclusions. 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. Every researcher must strive for the proper sample size and the protocol should contain its details.

Inferential statistics has two parts: estimation of population parameter and testing of hypothesis. According to the type of medical research, any one of them can be adopted. The estimation method is used in prevalence/descriptive studies and the testing of hypothesis is used for cohort/case control/clinical trials.

Using estimation method, the best estimates for population characteristics such as prevalence, incidence, mean, standard deviation, etc. can be found out.

By testing the hypothesis, correctness of whatever values or any relationship or association between variables derived from estimation can be verified.

These are the two requirements for the analysis of data in medical research. Before the testing of the hypothesis, one

must confirm the type of normality of the data so that the type of the test (parametric or non parametric) can be decided. Violation of this rule will result in wrong conclusion. Once the correct test is selected, the next important step is to determine the sample size. If proper attention is not given to the determination of the sample size, a real difference will become statistically insignificant. Thus, the study has to be repeated on a larger sample so that the real difference can be statistically proved. A randomly decided sample will invite non sampling errors in the study. An under sized sample will not give correct information and that will turn into a waste of time and resources. An over sized sample will end up with loss of resources with respect to money, man power and time. As a result, both errors will entail even unethical outcomes. Therefore, sample size determination is an important issue in medical research but availability of literature in this topic is scanty. On a recent survey a few of them were located¹⁻²³. Referring to the available literature and from the personal experience in this important topic, the authors would like to suggest a few methods for the determination of relevant sample size in various situations in medical research. The authors believe that this brief discourse will be of help to all personnel involved in medical research.

Sample size determination

Choosing a sample size is to be done by combination of logistical and pragmatic considerations which include:
 (a) The number of subjects who can be recruited in the given time period within available resources, and
 (b) The final figure calculated must represent the minimum number of subjects required to procure a reliable answer to the research question.

Factors that affect the sample size

1. Power of the study

The power of a study is the probability of detecting a real difference. In other words, it is same as the probability of making a positive diagnosis when the disease is present. For a good study one should have to take at least 80% power. An increase in the sample size or a greater difference between the control and the test groups leads to an increase in the power of the test, while an increasing standard deviation of the characteristic and an increase in significance level lead to a fall in the power of the study.

2. Level of significance

Level of significance is the probability of rejecting null hypothesis when it is true, in other words, detecting a significant difference when it is absent. This is the most important factor in determination of sample size. Therefore, the level of significance must be fixed before the testing of hypothesis, estimation and sample size calculation. In a standard situation, the probability can be taken as 0.05 or 0.01. Of late, researchers have also been taking the probability up to 0.2 (20%).

3. Event rate

If an event studied occurs more commonly among the study population, naturally, the power can be expected to be higher. Even though the expected event rate can be found out from previous studies, there is a possibility that the event rate can be estimated wrongly because of the background of the referred study, like differences in place, time, population etc. If the overall event rate falls to an unexpectedly low level, the sample size must be re-estimated by adopting the new (currently observed) event rate.

4. Effect of compliance

Compliance is another factor that directly affects the sample size. So, it should be calculated correctly. The compliance adjustment formula is as follows:

Adjusted sample size n_1 per group equals

$$n_1 = \frac{n}{[c_1 + c_2 - 1]^2}$$

where n is the original sample size, and c_1, c_2 are the average compliance rates per group.

In addition to the above factors, other factors that affect the sample size include consideration for unequal allocation, effect of important clinical treatment, etc. One of the most important decisions to make before calculating a sample size is to define the effect of important clinical treatment, δ (delta), which should not be confused with a statistical significance of the treatment effect – neither one implies the other and the distinction between them is important.

Table 1- Typical values for significance level and power

Significance level			Power			
5%	1%	0.1%	80%	85%	90%	95%
1.96	2.58	3.29	0.84	1.04	1.29	1.64

Descriptive Study

Descriptive studies are designed to describe the occurrence of disease by time, place and person. Descriptive study deals with estimation of population parameters. Two commonly used parameters are the mean (measure of central tendency) and the proportion.

Sample size calculation, when mean is the parameter of study

The confidence interval contains an estimate, above or below a margin of error. The margin of error for a 95% confidence interval is 1.96 times the standard error. It shows the accuracy of the guess and is based on the variability of the estimate.

Let E denote the margin of error. Then

$$E = \frac{1.96\sigma}{\sqrt{n}} \quad n = \frac{1.96^2 \sigma^2}{E^2}$$

Similarly for 99% confidence,

$$n = \frac{2.58^2 \sigma^2}{E^2} \quad E = \frac{2.58\sigma}{\sqrt{n}}$$

e.g. The mean pulse rate of a population is believed to be 70 per minute with a standard deviation of 8 beats. Calculate the minimum sample size required to verify this if allowable error is 1 beat at 1% risk.

$$n = \frac{2.58^2 \sigma^2}{E^2} \quad n = \frac{2.58^2 \times 8^2}{1^2}$$

$$n = 426$$

Sample Size calculation, when proportion is the parameter of study

For 95% confidence

$$E = 1.96\sqrt{\frac{PQ}{n}} \quad n = \frac{1.96^2 PQ}{E^2}$$

For 99% confidence

$$E = 2.58\sqrt{\frac{PQ}{n}} \quad n = \frac{2.58^2 PQ}{E^2}$$

Where P is the population proportion and Q= 1 – P

If E is given as a percentage, then it is to be taken as a percentage of P.

e.g. Hookworm prevalence rate was 30% before the specific treatment and adoption of other measures. Find out the size of sample required to find the prevalence rate now if available error is 10% at 5% risk.

$$n = \frac{1.96^2 PQ}{E^2} \quad E = 30 \times \frac{10}{100} = 3$$

$$n = \frac{1.96^2 \times 30 \times 70}{3^2}$$

$$n = 896$$

Calculate sample size for a sensitivity of a Test

$$\text{Sensitivity} = \frac{\# + \text{ve on new test \& gold standard}}{\# + \text{ve on gold standard}}$$

e.g. Sensitivity (P) =75%, Q= 100- P = 100-75 =25%, Precision (E) = 10%, Formula is

$$n = \frac{1.96^2 PQ}{E^2}$$

$$n = \frac{1.96^2 \times 75 \times 25}{10^2}$$

n=72

Analytical study

Analytical studies are designed to examine etiology and causal associations. Testing of hypothesis is the statistical method in analytical study. Analytic studies can be divided into two main types: Observational Studies and Experimental studies - Clinical trials.

Observational Studies

Calculating sample size for a case-control study: binary exposure

Use difference in proportions formula

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

n = **Sample size** in the case group
r = **ratio** of controls to cases

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

Z_β = Represents the desired power (typically 0.84 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)

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

e.g. For 80% power, you want to detect an odds ratio (OR) of 2.0 or greater, $Z_\beta=0.84$, for 0.05 significance level, $Z_\alpha=1.96$, r=1 (equal number of cases and controls), the proportion exposed in the control group is 20%, to get proportion of cases exposed:

$$p_{caseexp} = \frac{OR p_{controlsexp}}{p_{controlsexp}(OR - 1) + 1}$$

$$p_{caseexp} = \frac{2.0(.20)}{(.20)(2.0-1)+1} = \frac{.40}{1.20} = .33$$

Average proportion exposed = (.33+.20)/2=0.265

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

$$n = 2 \frac{(.265)(1-.265)(.84+1.96)^2}{(.33-.20)^2} = 181$$

Therefore, n=362 (181 cases, 181 controls).

Calculating sample size for a case-control study: continuous exposure

Use difference in means formula

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

n = **Sample size** in the case group

r = **ratio** of controls to cases

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)

e.g. For 80% power, $Z_\beta = .84$, For 0.05 significance level, $Z_\alpha = 1.96$, $r=1$ (equal number of cases and controls), $\sigma=10.0$, Difference = 5.0.

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

$$n = (2) \frac{10^2 (7.84)}{(5)^2} = (2) 2^2 (7.84) = 63$$

Therefore, $n=126$ (63 cases, 63 controls)

Sample size for independent cohort studies

This function gives the minimum number of case subjects required to detect a true relative risk or experimental event rate with power and two sided type I error probability alpha. This sample size is also given as a continuity-corrected value intended for use with corrected chi-square and Fisher's exact tests.

Information required

- Power
- alpha
- p_0 : probability of event in controls (can be estimated as the population prevalence of the event under investigation)
- p_1 : probability of event in experimental subjects
- **RR** : relative risk of events between experimental subjects and controls
- input either P_1 or RR, where $RR=P_1/P_0$
- **m** : number of control subjects per experimental subject

Practical issues

- Usual values for power are 80%, 85% and 90%; try several in order to explore/scope.
- 5% is the usual choice for alpha.
- p_0 can be estimated as the population prevalence of the event under investigation.

- If possible, choose a range of relative risks that you want have the statistical power to detect.

Technical validation

The estimated sample size n is calculated as:

Where,

$$n = \frac{[Z_1 + Z_2]^2}{(p_0 - p_1)^2}$$

$$Z_1 = Z_\alpha \sqrt{(1+1/m)(\bar{p})(1-\bar{p})}$$

$$Z_2 = Z_\beta \sqrt{p_0(1-p_0)/m + p_1(1-p_1)}$$

$$\bar{p} = \frac{p_1 + mp_0}{m+1}$$

$$n_c = \frac{n}{4} \left[1 + \sqrt{1 + \frac{2(m+1)}{nm|p_0 - p_1|}} \right]^2$$

where α = alpha, β = 1 - power, n_c is the continuity corrected sample size. n is rounded up to the closest integer.

Experimental studies

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)

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

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)

If unequal numbers in each group

- Ratio of cases to controls
- Use if want λ patients randomized to the placebo arm for every patient randomized to the treatment arm
- Take no more than 4-5 controls/case

$$n_1 = \frac{(Z_\beta + Z_{\alpha/2})^2 \left(\frac{\sigma_1^2 + \sigma_2^2}{\lambda} \right)}{\delta^2}$$

$n_2 = \lambda n_1 \rightarrow \lambda$ controls for every case

K:1 Sample Size Shortcut

- Use equal variance sample size formula: total sample size increases by a factor of

$$\frac{(k+1)^2}{4k}$$

- Total sample size for two equal groups = 26; want 2:1 ratio
- $26 * (2+1) / (2) = 26 * 3 / 2 = 39$
- 20 in one group and 10 in the other

Unequal numbers in Each Group: Fixed number of Cases

- Case-Control Study
- Only so many new devices
- Sample size calculation says $n=13$ cases and controls are needed
- Only have 11 cases!
- Want the same precision
- $n_0 = 11$ cases
- $kn_0 =$ number of controls

$$k = \frac{n}{2n_0 - n} \quad k = \frac{13}{(2 \times 11 - 13)}$$

$= 13 / 9 = 1.44$

$kn_0 = 1.44 * 11 \approx 16$ controls (and 11 cases) this will give the same precision as 13 controls and 13 cases

If number of events are important

- Cohort of exposed and unexposed people
- Relative Risk = R
- Prevalence in the unexposed population = π_1

$$R = \frac{\text{Risk of event in exposed group}}{\text{Risk of event in unexposed group}}$$

$$n_1 = \frac{(Z_\beta + Z_{\alpha/2})^2}{2(\sqrt{R} - 1)^2}$$

$n_1 =$ number of events in unexposed group

$n_2 = Rn_1 =$ number of events in exposed group

n_1 and n_2 are the number of events in the two groups required to detect a relative risk of R with power $1-\beta$

$$N = \frac{n_1}{\pi_1}$$

N = number of subjects per group.

Number of Covariates and number of Subjects

- At least 10 subjects for every variable investigated
 - In logistic regression
 - No general justification
 - This is stability, not power
- Principle component analysis (PCA) $N \geq 10m + 50$ or even $N \geq m^2 + 50$

One-sample t-test and Paired t-test

For testing the hypothesis:

$H_0: \mu = k$ vs. $H_1: \mu \neq k$

With a two-tailed test, the formula is:

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

$Z_\beta =$ Represents the desired power (typically .84 for 80% power).

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

Note: this formula is used even though the test statistic could be a t-test.

Lehr's formula

Lehr's formula can be used for determining the sample size for the studies expected to be verified by paired or unpaired t-tests or Chi-squared test.

It is a very simple formula. In a standard study where the power is 80% and a two-sided significance level of 0.05, the required sample size in each group will be:

$$n = \frac{16}{(\text{Standardized difference})^2}$$

If the standardized difference is small, this overestimates the sample size. We have to overcome this by using a numerator of 21 (instead of 16) relates to a power of 90%.

Unpaired t-test is to be applied in a research when both or one samples size will be less than 30 then Standardized difference is δ / σ and $N/2$ observation in each group σ the smallest difference in means that is clinically important. The assumed equal standard deviation (σ) of the observations in each of the two groups. You can estimate it using results from a similar study conducted previously or from published information. Alternatively, you can perform a pilot study to estimate it. Another approach is to express σ as a multiple of the standard deviation (e.g. the ability to detect a difference of two standard deviations).

If the trial is to conduct a before and after treatment comparison study (e.g. efficacy of drug, surgery etc.) in same subjects then you should use Paired t-test. In this case Standardized difference is $2\delta/\sigma$ and N pairs of observations. δ is the smallest difference in means that is clinically important. σ is the standard deviation of the **differences** in response, usually estimated from a pilot study.

To find out the relationship or association between exposure variable and outcome variable, one should use the Chi-square test (e.g. Smoking [exposure variable] and Cancer [outcome variable]). In this case Standardized difference is $p_1-p_2/(\bar{p}(1-\bar{p}))$ and N/2 observation in each group. p_1-p_2 the smallest difference in the proportions of 'success' in the two groups that is clinically important. One of these proportions is often known, and the relevant difference evaluated by considering what value the other proportion must take in order to constitute a noteworthy change.

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

Cluster Randomised Trial

When our research funding, manpower, time etc. are low and the disease is concentrated in a particular area, a cluster randomised trial may be carried out to estimate the sample size. Notations are given below

k- Number of clusters (clusters – villages, communities, households, schools, class rooms etc.)

m – Size of the cluster (5 members HH)

σ_A^2 - Variance among the clusters;

σ_W^2 - Within cluster variability

ρ - ICC (Intracluster correlation coefficient)

d - precision; i - intervention groups

(i = 1 treatment and i = 2 control)

To take account of the clustered nature of the data, or estimate a quantity with a given precision, the total overall sample size to detect a given difference will need to be multiplied by an amount known as the 'design factor' which is calculated as:

$$\text{Design effect} = 1 + ((\text{Number per cluster} - 1) \times \text{ICC})$$

where ICC is the intraclass correlation. This formula assumes a constant number per cluster, if this is variable then an average can be taken.

Estimation (Prevalence Study)

Continuous outcome:

$$n = \frac{z_{\alpha/2}^2 \sigma^2 [1 + (m-1)\rho]}{d^2}$$

σ - standard deviation of the outcome

where,

$$\rho = \sigma_A^2 / (\sigma_A^2 + \sigma_W^2)$$

Equivalently, the number of clusters required is given by

$$k = \frac{n}{m}$$

Binary outcome:

$$n = \frac{z_{\alpha/2}^2 pq [1 + (m-1)\rho]}{d^2}$$

p - prevalence (percentage or proportion)

where,

$$\rho = \sigma_A^2 / (\sigma_A^2 + \sigma_W^2)$$

Equivalently, the number of clusters required is given by

$$k = \frac{n}{m}$$

Testing of Hypothesis (RCT)

Comparison of means:

Equal cluster size:

The number of subjects required per intervention group to test the hypothesis $H_0: \mu_1 = \mu_2$ is given by

$$n = \frac{(z_{\alpha/2} + z_{\beta})^2 (2\sigma^2) [1 + (m-1)\rho]}{(\mu_1 - \mu_2)^2}$$

μ_1 - mean in the intervention group

μ_2 - mean in the control group

Note: For unequal cluster size, replace m by

$\frac{m}{m_{\max}}$ or more conservatively m_{\max} .

Suggestions

- ✓ If the effect of a clinical treatment is not marked when compared to a placebo, or power of the study is low, or a lower significance level (lower 'p' value) is expected, the sample size should be increased.
- ✓ If the measurements are highly varying, use the average of repeated measurements.
- ✓ Determine the scientifically acceptable power and level of significance.
- ✓ Estimate the event rate from similar population.
- ✓ In research protocols, statistically determined sample size, power of the study, significance level, event rate, duration of the study, and compliance should be mentioned.
- ✓ The sample size should be increased to adequate level for each sub-group when dealing with multiple sub-groups in a population.
- ✓ Always aim for a cost effective sample size.
- ✓ In small negative trials, meta analysis can be tried.
- ✓ When a study requires very large sample size net working with other researchers engaged in similar projects and Multi-centre trials will be beneficial.
- ✓ A study which needs a large sample size to prove any significant difference in two treatments must ensure the required sample size. Otherwise such studies may not

provide much information by any method and are better terminated so that the money and time are at least saved.

Conclusion

Carefully and well planned Medical research will result in relevant and socially useful results. Planning has several parts, such as well defined relevant research hypothesis, objectives, subjects must be selected from appropriate population, and instruments should be reliable, carefully undergone through best possible procedures and other guidelines. Sample size determination is very important and always difficult process to handle. It requires the collaboration of a specialist who has a good scientific knowledge in the art and practice of medical statistics.

Conflict of Interests

The authors do not have any conflict of interest arising from the study.

References

- Mace AE. Sample-Size Determination. New York: Reinhold Publishing Corporation, 1964.
- Cochran WG. Sampling techniques. New York, NY: John Wiley & Sons, 1977.
- Altman DG. Statistics and ethics in medical research: III How large a sample?. *Br Med J* 1980;281(6251): 1336-8.
- Gore SM. Statistics in question. Assessing clinical trials--trial size. *Br Med J (Clin Res Ed)* 1981;282(6277):1687-9.
- Boen JR, Zahn DA. The Human Side of Statistical Consulting. Belmont, CA: Lifetime Learning Publications, 1982.
- Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. *N Engl J Med* 1978;299(13): 690-4.
- Kraemer HC, Thiemann S. How many subjects? Statistical power analysis in research. Newbury Park, CA: Sage Publications, 1987.
- Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. 1988.
- Desu MM, Raghavarao D. Sample Size Methodology. Boston, MA: Academic Press, Inc. 1990.
- Lipsey MW. Design Sensitivity: Statistical Power and Experimental Research. Newbury Park, CA: Sage Publications, 1990.
- Shuster JJ. CRC Handbook of Sample Size Guidelines for Clinical Trials. Boca Raton, FL: CRC Press, 1990.
- Odeh RE, Fox M. Sample Size Choice: Charts for Experiments with Linear Models. 2nd ed. New York: Marcel Dekker, 1991.
- Muller KE, Benignus VA. Increasing scientific power with statistical power. *Neurotoxicol Teratol* 1992;14(3):211-9.
- Murray DM, Rooney BL, Hannan PJ, Peterson AV, Ary DV, Biglan A et al. Intraclass correlation among common measures of adolescent smoking: estimates, correlates, and applications in smoking prevention studies. *Am J Epidemiol.* 1994;140(11): 1038-50.
- Murray DM, Short B. Intraclass correlation among measures related to alcohol use by young adults: estimates, correlates and applications in intervention studies. *J Stud Alcohol.* 1995;56(6):681-94.
- Murray DM, Short B. Intraclass correlation among measures related to alcohol use by school aged adolescents: estimates, correlates and applications in intervention studies. *J Drug Educ* 1996;26(3):207-30.
- Friedman L, Furberg C, DeMets D. Fundamentals of Clinical Trials. 3rd ed. New York, NY: Springer-Verlag, 1998.
- Kerry SM, Bland JM. Analysis of a trial randomised in clusters. *BMJ* 1998;316(7124):54.
- Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ* 1998;317(7167):1181-4.
- Donner A, Klar N. Design and analysis of cluster randomization trials in health research. London, UK: Arnold Publishing Co, 2000.
- Hoenig JM, Heisey DM. The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis. *Am Stat* 2001;55(1):19-24.
- Lui KJ, Cumberland WG. Sample size determination for equivalence test using rate ratio of sensitivity and specificity in paired sample data. *Control Clin Trials* 2001;22(4):373-89.
- Gebski V, Marschner I, Keech AC. Specifying objectives and outcomes for clinical trials. *Med J Aust* 2002; 176(10): 491-2.