

Medical statistics online help:

Sample size & power for clinical trials

[Clinical Trials: an overview](#)

[Sample size calculation](#)

[Dedicated software](#)

[Reference texts](#)

It is very important that the design of a clinical trial has been thoroughly considered before it is undertaken. A crucial component of the design is the number of participants required, in order to reliably answer the clinical question. The aim of these pages is to clarify some of the key issues regarding sample size and power.

Clinical Trials: an overview

A Clinical Trial is a carefully planned experiment involving human participants, which attempts to answer a pre-defined set of research questions with respect to an intervention. Types of clinical trials include:

- Intervention or therapeutic - e.g. a drug treatment for stroke; a surgical intervention such as simple mastectomy for breast cancer
- Preventative - e.g. screening for cervical cancer; educational methods and lifestyle interventions; vaccine trials for TB

[top](#)

Sample size

When planning a clinical trial, it is very important to consider how many participants you will need to reliably answer the clinical question. Too many participants is a needless waste of resources (and possibly lives), which could result in a beneficial treatment being denied to patients unnecessarily. Too few participants will not produce a precise, reliable and definitive answer, which can also be considered unethical. Under this common scenario, patients might be denied a useful treatment because the trials were frequently underpowered (i.e. too small to detect a treatment effect) – this can also result in further studies being cancelled without good reason. Choosing a sample size is a combination of logistical and pragmatic considerations. These

include (a) the number of participants you can reasonably expect to recruit in the given time period within available resources, and (b) mathematical calculation. The final figure calculated indicates the minimum number of participants required to reliably answer the research question.

There are many packages dedicated to performing sample size/power calculations. It is often possible to perform such calculations within standard statistical software, and stand-alone packages are available for this purpose (for details of these, see sample size software). However, if you do not have access to any dedicated packages, the examples listed below will possibly help you to do your own calculations by hand. We have included the most commonly encountered clinical trial scenarios - the comparison of two equal-sized, concurrently studied groups where the primary outcomes are either proportions (e.g. percentage responding to treatment) or means (e.g. average blood pressure). The methods listed below are extracts from Mark Woodward's book (see reference texts), which includes a whole chapter on sample size determination.

One of the most important decisions to make before calculating a sample size is to define a clinically important treatment effect, δ (or delta), which should not be confused with a statistically significant treatment effect – neither one implies the other and the distinction between them is important. Of course, since you are embarking on a study, you would expect your new intervention to be an improvement, but can you estimate by how much reliably or realistically? If not, another way to come up with an estimate is to consider what observed treatment effect would make you change your current clinical practice. For example, if your trial was looking at a treatment to lower blood pressure, you might argue that an average lowering of systolic BP of 5mm Hg is clinically important; however, you might decide that an average lowering of systolic BP of 10mm Hg would be clinically important and necessary before you would possibly think about prescribing this treatment.

Before proceeding, you also need to set two parameters:

1. the statistical significance level, alpha (α), typically 5% (sometimes written $\alpha = 0.05$); also known as the false-positive rate.
2. the power – adequate power for a trial is widely accepted as 0.8 (or 80%). Power is defined as $1 - \beta$, where β , the false-negative rate, in this case would be 0.2 (or 20%).

Where the primary outcome is a (continuous) measurement such as blood pressure, you also need an estimate of the natural variability in the population, usually the standard deviation, s .

[top](#)

Basic formula

The number of participants required in each intervention group, m , is given by:

$$m = \frac{2 \times [z_{(1-\alpha/2)} + z_{(1-\beta)}]^2}{\Delta^2}$$

where $z_{(1-\alpha/2)}$ and $z_{(1-\beta)}$ represent percentage points of the normal distribution for statistical significance level and power, respectively (see Table 1 for typical values), and Δ represents the standardised difference (i.e. the treatment difference divided by its standard deviation).

Other assumptions made here –

- equal allocation of participants to each of two treatment groups
- 2-sided statistical tests are to be used using a normal assumption or approximation

Table 1 – typical values for significance level and power

Significance level			Power			
5%	1%	0.1%	80%	85%	90%	95%
1.96	2.5758	3.2905	0.8416	1.0364	1.2816	1.6449

[top](#)

When your primary comparison is a proportion (or a percentage)

Intervention trial example

Your research question is “does a new treatment work better than the existing one?”. With standard therapy, 40% of patients, on average, achieve a favourable outcome (e.g. single layer compression stockings for the treatment of venous leg ulcers). It is anticipated that a new treatment (e.g. multi-layer compression stockings) will increase the ‘cure’ rate to 50%. What sample size would be required in order to detect such a treatment effect with 80% power at a 5% level of statistical significance?

First, calculate Δ the standardised difference. In the case of two proportions, p_1 and p_2 ,

the standardised difference $\Delta = \frac{p_1 - p_2}{\sqrt{\bar{p} \times (1 - \bar{p})}}$ where $\bar{p} = \frac{(p_1 + p_2)}{2}$

In our example, $p_1=0.50$ (or 50%), $p_2=0.40$ (or 40%) and so $\bar{p} = \frac{(0.50+0.40)}{2} = 0.45$

hence $\Delta = \frac{0.50 - 0.40}{\sqrt{0.45 \times (1 - 0.45)}} = \frac{0.10}{\sqrt{0.45 \times 0.55}} = \frac{0.10}{\sqrt{0.2475}} = \frac{0.10}{0.46749} = 0.201$

Using the values from the table for a significance level of 5%, $z_{(1-\alpha/2)} = 1.96$, and a power of 80%, $z_{(1-\beta)} = 0.8416$,

$$m = \frac{2 \times [1.96 + 0.8416]^2}{0.201^2} = \frac{2 \times [2.8016]^2}{0.201 \times 0.201} = \frac{2 \times [2.8016 \times 2.8016]}{0.0404} = \frac{2 \times [7.84896]}{0.0404} = \frac{15.6979}{0.0404} = 388.5$$

Round up to the nearest whole number and say that we require **389** participants per treatment group or **778** in total.

NB. Always work with plenty of decimal places, at least 6, in your calculations – only round up for the final figure. I have only rounded here for aesthetic reasons.

[top](#)

When your primary comparison is a mean

Prevention trial example

Your research question is “does a new treatment work when compared to placebo?”. You have been asked to find out how many patients with mild hypertension would need to be recruited into a trial, in order to detect an average difference of 5mm Hg in systolic blood pressure, between an intervention group who receive a new anti-hypertensive and a control group (who effectively receive delayed intervention). Assume the standard deviation (a measure of the patient-to-patient variability) of systolic blood pressure is 10mm Hg, 90% power and a significance level of 5%.

First, calculate Δ , the standardised difference, sometimes called the effect size. In the case of two means, μ_1 and μ_2 , with a common standard deviation ‘s’, the

standardised difference $\Delta = \frac{m_1 - m_2}{s}$. Alternatively, it can be written as $\Delta = \frac{d}{s}$ where d is the clinically important difference.

In our example, $\Delta = \frac{5}{10} = 0.5$

Using the values from the table for a significance level of 5%, $z_{(1-\alpha/2)} = 1.96$, and a power of 90%, $z_{(1-\beta)} = 1.2816$,

$$m = \frac{2 \times [1.96 + 1.2816]^2}{0.5^2} = \frac{2 \times [3.2416]^2}{0.5 \times 0.5} = \frac{2 \times [3.2416 \times 3.2416]}{0.25} = \frac{2 \times [10.50797]}{0.25} = \frac{21.0159}{0.25} = 84.1$$

Rounding up to the nearest whole number indicates that we require **85** participants per treatment group or **170** in total.

Points to remember

Sample size increases:

- when a small treatment difference is expected rather than a large one e.g. when comparing

two active treatments rather than an active treatment versus placebo

- with higher power – the higher the power, the more likely, on average, you are to detect a treatment effect if it exists e.g. more participants will be needed for a trial with 90% power rather than 80%
- with a lower significance level – the lower the significance level e.g. 1% (or $\alpha=0.01$) rather than the typical 5%, the less likely you are to get a chance (but spurious) treatment effect i.e. a false-positive result
- when measurements are highly variable – natural variability can be thought of as ‘noise’ and makes the ‘signal’ more difficult to hear e.g. measurements such as blood pressure and peak flow are highly variable – a simple way to alleviate this problem is to take a few repeated measurements and use the average (or the maximum for the latter)

[top](#)

References

No official endorsement by the NHS/ICRF of any of the following material is intended or should be inferred. This list is by no means exhaustive, but we, the authors, have used part or all of this material at some time and found it useful.

Acknowledgements – these pages were adapted by Ed Juszcak, based on an idea by Mike Bradburn and Sharon Love. Many thanks go to Lesley Smith for her very useful comments. Words of caution - there are many web-based calculators freely available over the Internet – be careful, since these are not always accurate.

Sample size book which comes with dedicated sample size software:

Sample Size Tables for Clinical Studies (2nd Edition, 1997) by D Machin, M Campbell, P Fayers & A Pinol - published by Blackwell Science Ltd [ISBN 0-86542-870-0]. Contact <http://www.blackwell-science.com>

A disk accompanies the book and provides a small, menu-driven package dedicated to sample size and power calculations.

Software packages with sample size capability:

StatsDirect – statistical software which links to Microsoft Excel: for a **free** trial visit: <http://statsdirect.com/>

Epi Info – **free** epidemiology/statistics software from the CDC, Atlanta, “designed for the global community of public health practitioners and researchers”. For your free download visit: <http://www.cdc.gov/epiinfo/>

Electronic articles on medical statistics and clinical trials:

For short (no more than 1 page), easy to understand articles aimed at non-statisticians try the BMJ series of Statistics notes written in the main by Martin Bland & Douglas Altman

visit – <http://www.sghms.ac.uk/depts/phs/staff/jmb/pbstnote.htm>

For an overview of clinical trials visit – <http://www.wiley.co.uk/eob/sample1.pdf>

Classic papers on clinical trials:

Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer*. 1976;**34**(6):585-612.

Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG. "Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer* 1977;**35**(1):1-39.

Medical statistics / Epidemiology books:

Practical Statistics for Medical Research by Douglas G Altman published by Chapman & Hall (1990) [ISBN 0-412-27630-5]

Epidemiology: Study Design and Data Analysis by Mark Woodward published by Chapman & Hall/CRC (1999) [ISBN 1-58488-009-0]

Statistics from Scratch - An Introduction for Health Care Professionals by David Bowers published by Wiley (1996) [ISBN 0-471-96325-9]

Further Statistics from Scratch - For Health Care Professionals by David Bowers published by Wiley (1997) [ISBN 0-471-97136-7]

General clinical trials books:

Clinical Trials: A Practical Approach by Stuart J Pocock - published by John Wiley & Sons, 1984 [ISBN: 0471901555]

Fundamentals of Clinical Trials by LM Friedman, C Furberg, DL Demets published by Springer Verlag, 1998 [ISBN: 0387985867]

Randomised Controlled Trials : A User's Guide by Alejandro R. Jadad - published by BMJ Books, 1998 [ISBN: 0727912089]

Clinical Trials: A Methodologic Perspective by Steven Piantadosi - published by John Wiley & Sons, 1997 [ISBN: 0471163937]

Clinical Trials: Design, Conduct, and Analysis (Monographs in Epidemiology and Biostatistics, Vol 8) by Curtis L. Meinert - published by Oxford University Press, 1986 [ISBN: 0195035682]

Statistical Issues in Drug Development by Stephen Senn published by John Wiley &

Sons, 1997 [ISBN: 0471974889]

An Introduction to Randomised Controlled Clinical Trials by J.N.S Matthews published by Edward Arnold, 2000 [ISBN: 0340761431]

Specialised clinical trials books:

Cross-over Trials in Clinical Research by Stephen Senn published by John Wiley & Sons, 1993 [ISBN: 0471934933]

Design and Analysis of Cluster Randomisation Trials: In Health Research by Allan Donner, Neil S. Klar published by Edward Arnold, 2000 [ISBN: 0340691530]

[top](#)

(29/11/2001)