Calculating Sample Size

How do I calculate the sample size I need for a clinical trial?

Justifying the number of participants – or the “sample size” – is a vital part of planning a clinical trial for ethical reasons. The trial should be large enough to answer the research question, but not so large that participants are involved in medical research needlessly. Any research funder will look carefully at how you have justified the sample size for your trial. The aim of most trials is to show that one treatment is better than another: these are known as superiority trials. (Less commonly, the aim might be to show that one treatment is no worse than another – a non-inferiority trial.) The sample size for a superiority trial is usually justified in terms of the trial’s statistical “power”, which is the chance of finding evidence for superiority when a clinically important benefit is present. “Finding evidence” is generally taken to mean obtaining a result that is significant at the two-sided 5% significance level. Trials are usually planned to have 80% or 90% power to detect the smallest effect that would be considered clinically important. The precise standard can vary, so it helps to know your funder: the National Institute for Health Research’s Efficacy and Mechanisms Evaluation (EME) funding programme, for example, routinely expects researchers to plan for 90% power.

In order to work out the sample size to achieve given power, you need to know several things, including what your outcome measure is, and how big the improvement in this outcome has to be to be considered clinically important. The latter is a clinical issue, not a statistical one, and as an expert in your field you will be in a better position to answer this than a statistician. Depending on the method and the type of data, you may also need to specify some additional information: e.g. for a numerical outcome the standard deviation in each group, or for a dichotomous outcome the risk of an adverse outcome in controls. The existing research literature may help quantify these things.

There are a variety of resources available – papers, textbooks, websites, and software – that can help you perform sample size calculations. A good, basic reference is by Campbell, Julious & Altman, which describes the theory and methods behind sample size calculation for comparing two means or two proportions.

1. Schoenfeld & Richter describe how to calculate sample size for a trial with a time-to-event outcome.
2. A method for correcting for a cluster-randomised design (for example where you randomise entire general practices rather than individual patients) is described by Kerry & Bland.
3. Machin et al have produced a comprehensive textbook on sample size calculation that comes with free software.
4. Other proprietary sample size software is available. Some websites offer free sample size calculators: although these can be dangerous toys in the wrong hands, they can still be a good way to have a go yourself. One site with a range of options and some responsibly written accompanying text is Russ Lenth’s power and sample size page.
5. An RDS London statistical advisor will be happy to point you to further resources, interpret results you have obtained, and work through example calculations with you.

Remember finally that all funders want to be convinced that you have all the necessary expertise within your research team to analyse and interpret your data, so there should be someone within the team of applicants (it might be a statistical co-applicant, or it might have to be you) who can understand and take responsibility for whatever appears in the sample size justification section of your application. The EME programme provides some helpful advice to researchers on sample size as part of their general advice on getting the appropriate methodological input to your proposal.6 This useful resource describes the elements of a sample size calculation, and gives examples of a poor sample size justification as well as (from EME’s point of view) a more convincing one.

Richard Hooper
RDS London
Queen Mary, University of London

References

2. Schoenfeld & Richter. *Nomograms for calculating the number of patients needed for a clinical trial with survival as an endpoint*. Biometrics 1982;38:163-170