NIHR | National Institute for Health Research

Justifying sample size for a feasibility study

How do I calculate the sample size I need for a clinical trial?
The terms “pilot” and “feasibility” study are often used interchangeably in published research to indicate a study done in anticipation of a full-scale clinical trial, to test out different components of the methods or to provide information that will help with the trial’s design.

The National Institute for Health Research (NIHR) draws a distinction between the meanings of “pilot” study – a full trial in miniature which offers researchers a stopping point should the full trial not prove viable – and “feasibility” study – a catch-all term for research designed to investigate whether a trial will be feasible.[1] There is still some ambiguity and overlap in these definitions, but the latest guidance from the Research for Patient Benefit (RfPB) funding programme indicates they are unwilling to fund what they see as pilot studies (by NIHR’s definition), explaining that these are more appropriately funded as part of a programme of work which also includes the remainder of the full-scale trial.[2]

Pilot studies whose data are ultimately included in the analysis of the full trial (so-called “internal pilots”), and in particular those where there is an interim data analysis at the end of the pilot stage (“group sequential” and “adaptive” trial designs) need specialist statistical input to the design, and do need to be considered as a part of a whole. I will not discuss these here, but instead I give some brief advice on justifying sample size for feasibility studies – which *are* within the remit of the RfPB programme.

For a feasibility study, as with any piece of quantitative research, you need to justify the proposed sample size with reference to your stated aims and objectives – your sample size justification should, in fact, clearly match your objectives – otherwise the funder might wonder if the research couldn’t be done more quickly and cheaply while still achieving those objectives. Examples of objectives for feasibility studies are given in the NIHR glossary, and are varied,[1] but might include, for instance:

i) estimating a parameter such as a standard deviation which will be used in a sample size calculation for the full-scale trial;
ii) estimating the rate (proportion) of eligible people who are willing to participate, of participants who drop out of the trial, or of participants who comply with their allocated intervention.

A useful introduction to feasibility and pilot studies is given by Lancaster, Dodd & Williamson.[3] On the question of estimating a parameter such as a standard deviation for use in a sample size calculation, they recommend an overall sample size of 30. This reference is often cited as a justification for a figure of 30, though it unravels

slightly when we trace its provenance: Lancaster *et al* cite a paper by Browne, but Browne himself writes "the proverbial rule of thumb of 'use an *n* of 30 or greater to estimate a parameter' will not alleviate that problem [of the full-scale trial ending up under-powered]".[4]

Two other authoritative and contradictory papers that discuss how big a study you need to estimate a standard deviation for a sample size calculation are by Sim & Lewis, who recommend at least 50 participants,[5] and Julious, who recommends 24.[6] Thus a researcher who wants to address this particular objective, and would like to justify a feasibility study of 40-50 participants in total, might reasonably say in justification that "sample sizes between 24 and 50 have been recommended", and give these references.

If your objective is to estimate a rate (in the sense of a proportion of people), such as the rate of participation, drop-out, or compliance, then a simple approach is to relate the proposed sample size to the width of the 95% confidence interval for the rate – this relationship being determined by standard statistical methods for dealing with proportions. For example, you might say "with a sample size of 50, we will be able to estimate a drop-out rate of 80% to within a 95% confidence interval of +/- 11%", or "if we identify 100 eligible subjects we will be able to estimate a participation rate of 50% to within a 95% confidence interval of +/-10%". In these examples I calculated the width of the confidence interval (in %) as $1.96 \times \sqrt{(p \times (1-p) / n}$, where p is the percentage you expect to see, and n is the intended sample size. (Then, as with any sample size justification based on confidence intervals of estimates, expert reviewers must decide whether this achievement is good enough.)

Other feasibility-style objectives may require different approaches again to justifying sample size. You will want to consider these carefully with your statistician co-applicant(s) or discuss options with a specialist RDS London advisor.

Richard Hooper
RDS London
Queen Mary University of London

References

1. NIHR. *Feasibility and pilot studies*, in NIHR NETSCC glossary.
2. www.netscc.ac.uk/glossary/#glos6
3. NIHR. *Research for Patient Benefit (RfPB) programme: guidance information for*
4. *applicants* (competition number 22).
5. http://www.ccf.nihr.ac.uk/RfPB/Documents/RfPB%20-%20Guidance%20for%20Applicants%20Competition%2022.pdf
6. Lancaster GA, Dodd S, Williamson PR. *Design and analysis of pilot studies:*
7. *recommendations for good practice. J Eval Clin Practice* 2004;10:307-312
8. Browne RH. *On the use of a pilot sample for sample size determination. Stat Med*
9. 1995;14:1933-1940
1. Sim J, Lewis M. *The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol* 2012;65:301-308
2. Julious SA. *Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat* 2005;4:287-291