Pilot *and* Feasibility Studies

Research Design & Analysis Core

Version: 2.0 • 02 May 2022

Duke GLOBAL HEALTH INSTITUTE | Research Design & Analysis Core

Contents

Introduction

Definition of a Feasibility Study

A feasibility study is performed to assess whether *some aspect* of a proposed project or study will work [1]. They may also be used to estimate important parameters that are needed to design a larger study. The following are examples of values an investigator may be interested in measuring [2]:

- Time commitment needed by interventionists and study participants
- Willingness of participants to be randomized
- Willingness of clinicians to recruit participants
- Number of eligible patients, caretakers or other appropriate participants
- Characteristics of the proposed outcome measure (in some cases feasibility studies might involve designing a suitable outcome measure)
- Follow-up rates, response rates to questionnaires, adherence/compliance rates
- Availability of data needed or the usefulness and limitations of a particular database
- Ability to validly and precisely measure variables of interest
- Time needed to collect and analyze data

Definition of a Feasibility Study that is a Pilot:

A pilot study is a type of feasibility study in which an investigator is testing a potential future study protocol *as a whole* to see if it will work [1]. Pilot studies are used to test whether all components of a study can work together, with the intent that findings from the pilot will lead to a larger full-scale study in the future.

In some cases, this will be the first phase of the larger substantive study and data from the pilot phase may contribute to the final analysis; this can be referred to as an *internal pilot*. Internal pilot studies are often incorporated into the main study design of a larger randomized controlled trial. In contrast, an *external pilot* is a stand-alone piece of work planned and carried out independently of and prior to the main study [2, 3].

What is the difference between a pilot study and a feasibility study?

In the literature, it is not uncommon to see "pilot" and "feasibility" used interchangeably. In fact, all pilot studies are a type of feasibility study, though the contrary may not be true. The distinguishing feature of a pilot study is that it tests the feasibility of all aspects of the protocol as whole and how the components of the protocol work together [4]. On the other hand, a feasibility study that is NOT a pilot study might test various subsets of future full-scale study protocol but not the protocol as a whole.

Some simple definitions have been proposed to distinguish between three main types of feasibility studies [5]:

1. Randomized pilot studies: Studies in which the future RCT, including the randomization of participants, is conducted on a smaller scale to see if it can be done.
2. Non-randomized pilot studies: Similar to randomized pilot studies, these are studies in which all or part of the intervention to be evaluated and other processes to be undertaken in a future trial is/are carried out but without randomization of participants.
3. Feasibility studies that are not pilot studies: Studies in which investigators attempt to answer a question about whether some element of the future trial can be done but do not implement the intervention to be evaluated or other processes to be undertaken in a future trial, though they may be addressing intervention development in some way.

It is important to note that exact definitions and distinctions between randomized *pilot studies, non-randomized pilot studies,* and *feasibility studies that are not pilot studies* are not necessarily consistent in the literature. See Eldridge et al [5], for examples of different classifications and proposed definitions for the three types of studies.

Objectives of Interval Pilot Studies

Internal pilot studies are usually planned simultaneously with a larger study and conducted using a pre-specified number of the initial participants in the full trial. The main purpose of an internal pilot study is to provide a check on the adequacy of the sample size calculation [6]. One drawback to such a design is that analyzing the pilot participants along with participants in the rest of the study may lead to increased Type I error due to non-independence of samples (i.e. sample size in the second stage is allowed to depend on observed responses in the pilot phase). However, the inflation is likely to be small in all but very small pilot sample sizes and can be remedied by incorporating the non-independence into the sample size calculation [7].

Though both internal pilot studies and interim analyses allow for sample size reviews for clinical trials, there is an important distinction between them. An interim analysis involves a formal calculation of the treatment effect and a corresponding hypothesis test for the purposes of determining futility or success of the treatment in order to determine whether early stopping is appropriate. In contrast, internal pilot studies involve calculation or re-calculation of nuisance parameters[1] (e.g. sample variance) for the purpose of determining whether the original sample size calculation was appropriate [7].

Use of an internal pilot study ultimately depends on the goals of investigators. For instance, internal pilot studies do NOT allow for the pre-testing of the feasibility, acceptability, or many other components of the larger study since the internal pilot is already a part of the larger study [7], thus establishment of the feasibility of the study must occur before the internal pilot.

[1] A nuisance parameter is any parameter that is NOT of primary interest in an analysis but must be accounted for in order to make inference about parameters that ARE of primary interest (e.g. variance is often considered a nuisance parameter when the mean of a distribution is the parameter of interest).

Objectives of External Pilot Studies

The main objective of an external pilot study is to test all aspects of the integrity of a study protocol and feasibility of the intervention AND of the trial design. The following are a list of objectives of a pilot study, which include objectives meant to establish feasibility of the intervention itself, as well as objectives meant to test the feasibility of the future trial designed to test the efficacy or effectiveness of the intervention.

In order to test the feasibility of the intervention itself, the pilot study may establish, for example:

1. Resources needs, such as:
    - Intervention and administrative staffing needs
    - Training needed for intervention and administrative staff
    - Mobilization of equipment/materials and other logistics involved in the roll-out of the intervention
    - Establishing and/or testing ongoing regulatory and reporting procedures

      Refining or establishing monitoring/oversight procedures (especially in cases of multiple sites)

2. Acceptability of the intervention:
    - Is the intervention appealing to participants given any known side effects?
    - Are there any difficulties with administration of intervention to participants?
    - How long does the intervention take to administer, and is the time commitment and length of time the intervention takes acceptable to participants?
    - What are the rates of compliance to the intervention?
    - What is the retention rate of participants in the intervention?
    - Are the inclusion/exclusion criteria for the intervention acceptable/appropriate?

In order to test the feasibility of the full-scale trial, the pilot study may be used to, for example:

1. Test appropriateness of collection forms, questionnaires:
    - Ensure the form is comprehensible and appropriate, and that questions are well defined, clearly understood and presented in a consistent manner.

2. Test the randomization procedure:
    - Testing the logistics of the implementation of randomization procedure
    - Testing the success of the randomization procedure
    - Testing the acceptability of the randomization procedure for participants (i.e. does knowing that they may be randomized to control affect recruitment?)

3. Test recruitment and consent:
    - Testing the informed consent procedures
    - Estimating recruitment, consent, and retention rates
    - Identifying barriers to recruitment
    - Identify issues with treatment cross-over and/or potential contamination (i.e. are control participants inadvertently exposed to the treatment?)

4. Assist in the selection of most appropriate primary outcome measure:
    - Is it feasible to measure the primary outcome?
    - Can the chosen outcome be reliably measured?

Because the goal of the pilot is to identify issues with implementation of BOTH the intervention and the trial, investigators should strongly consider collecting qualitative data to supplement quantitative data collection [2].

Using Pilot Studies to Inform Sample Size Calculations

A common goal of pilot studies is to inform the sample size calculation for the future full-scale trial. There are conflicting opinions about whether and how pilot results should be used for this purpose.

Some feel that pilot study results should NOT be used for sample size determination due to the inherent imprecision in between group effect size estimates and elevated levels of Type I and II error [8]. For example, a pilot trial that finds an overly large effect may lead to underpowering of the full scale trial. A pilot trial that finds an overly small effect could lead to termination of an intervention that may in fact be effective OR the overpowering of the larger trial. Instead, it is recommended that sample size for larger trials be based on the smallest difference that is clinically meaningful.

Others feel that providing data for the sample size calculation of a larger trial is a major objective of conducting a pilot study [1]. In such cases, one might use the pilot sample to estimate central tendency (e.g. mean or proportion) and/or variability (e.g. standard deviation) in the population of interest. When a major objective of the pilot/feasibility study is to provide data for a sample size calculation, investigators must factor this into the proposed sample size for the pilot study itself, as discussed in the next section.

There is some consensus in the fact that the results of pilot/feasibility studies can be used for sample size determination if investigators proceed with extreme caution, preferably basing such calculation on a range of possible parameter values (i.e. not just those obtained from the pilot/feasibility study) [1, 2].

Determining Samples Size for a Pilot/Feasibility Study

The size of the sample needed for a pilot/feasibility study is determined by the precision by which investigators wish to estimate various end targets, which may range from participant adherence to study protocol (which may require few participants) to having estimates for the length of time needed to fill out a questionnaire, determine response rates, estimate adherence and attrition [9], or to estimate parameters needed for calculation of sample size of a larger study (which may require more participants).

For example, if an investigator wants to estimate potential dropout, expected to be around 20%, within 10 percentage points, with a confidence level of 90%, then s/he will need N=52 participants total to estimate that quantity. If the investigator wishes to be conservative in planning for dropout in the larger trial, he/she might choose to use the upper end of the confidence interval of that estimate (i.e. if the realized dropout rate was 20% with a 90% CI of [10%, 30%] then 30% may be assumed for the purposes of sample size calculation for the larger trial).

If one of the goals of the pilot study is to determine whether an instrument is appropriate for the population that will be studied in the larger trial, then he/she may wish to use the pilot to calculate statistics such as Cronbach's alpha or test-retest reliability, then calculate the sample needed for a desired level of precision at an acceptable confidence level.

If a major goal of the pilot study is to estimate parameters related to treatment effect size for the full study (e.g. means, proportions, standard deviations), then the investigator will need to factor this into planned sample size for the pilot (note: pilots are not meant to be powered to find the true effect size as this is the goal of the larger study). Some authors have advocated simple rules of thumb, such as having N=30 per parameter estimated (e.g. mean or proportion) for the pilot study, then to use at least an 80% upper one-sided confidence limit of

the estimated treatment effect to estimate power for the larger trial [10]. Overall, precision needed to estimate feasibility targets will need to be balanced between what is realistic financially and logistically and what the consequences of imprecision of that estimate will be as it applies to the larger trial.

Overall, precision needed to estimate feasibility targets will need to be balanced between what is realistic financially and logistically and what the consequences of imprecision of that estimate will be as it applies to the larger trial.

WORDS OF CAUTION

- A pilot should never be conducted simply because of small available sample size.
- A pilot should NOT be considered a preliminary test of the intervention hypothesis. Despite the view that pilot studies should NOT be used to assess treatment effectiveness, Arain et al (2010) [13] found that 81% of studies (in a meta-analysis of pilot/feasibility studies) incorporated hypothesis testing and some included tests of treatment effectiveness [11].
- Pilot studies are NOT, in general, appropriate for determining safety due to small sample size, except in extreme, unfortunate cases where a death occurs or repeated serious adverse events occur [8]. A pilot/feasibility study COULD, however, be used to examine feasibility of adverse event reporting system.

Recommendations for Analysis of Pilot Studies

Analysis of Feasibility

Criteria for success of a pilot study should be stated clearly and should be based primarily on feasibility objectives [2]. Feasibility objectives are determined by the investigator and study team. For example, the study team may decide that feasibility is defined as having at least 70% sample retention, having a recruitment rate of at least 10 participants per month, and average satisfaction rating by participants of “very good” or “excellent,” and no more than 10% missing data for outcome measures.

The end result of a pilot study should be one of the following:

1. Stop: The main study is not feasible
2. Continue, but modify the protocol: The study would be feasible with modifications
3. Continue without modifications, but monitor closely: The study would be feasible with modification and close monitoring
4. Continue without modifications: The study is feasible as is.

Analysis of Estimated Treatment Effects

A pilot study should NOT be considered a preliminary test of the intervention hypothesis. Analysis of pilot/feasibility studies should be mainly descriptive. Any estimation or description of treatment outcomes should focus on confidence interval estimation. Inferential statistical tests should NOT be proposed as part of the pilot proposal/protocol. Any hypothesis testing undertaken should be done with extreme caution as covariates are likely to be imbalanced due to low sample size and confidence intervals are likely to be imprecise even when significant differences do exist. All results should be treated as preliminary and interpreted with caution.

Some may wonder whether a control arm is needed at all if inferential comparisons are discouraged. For the purposes of a study pilot, having a control arm provides a more realistic examination of recruitment, randomization, implementation, blinding procedures, and differences in loss to follow-up by treatment arm. It is important to understand how feasibility, consistency and acceptability will occur in a control arm.

It is also important to determine the possible differences in recruitment, retention, and acceptability that occur when there is a control arm, e.g. if a participant knows that they may receive placebo OR they know that they will be part of the intervention arm, it may change their willingness to participate, acceptability of the intervention or reported outcomes [8].

Investigators should be compelled to publish the results of a pilot study. There is a publication bias against pilot/ feasibility studies, particularly those with negative or null results. It is important to the research community to have access to the results of pilot/feasibility studies to save resources from being unnecessarily spent on studies that are NOT feasible. Publishing the results of a pilot/feasibility study also helps avoid duplication of effort when assessing feasibility [2]. Examples of types of pilot/feasibility study findings potentially suitable for publication can be found in Table 1 on the next page.

Authors should state in the conclusion of a manuscript whether the aims and objectives of pilot/feasibility work have been met and whether the results obtained from the pilot/feasibility study will lead to a future large-scale study [11].

Reporting guidelines from CONSORT's extension to Pilot/Feasibility Studies [12] can be found in Supplemental Material 1 (note: guidelines are ONLY for pilot/feasibility studies for randomized controlled trials). Supplemental Material 2 provides examples of pilot and feasibility study abstracts.

Table 1. Pilot Study Findings Potentially Suitable for Publication

Sampling Information
- Number of eligible subjects in research sites
- Proportion of eligible persons who consent
- Reasons for non-consent of eligible potential subjects
- Flow of eligible subjects over time
- Relative efficacy of different recruitment approaches or locations
- Differences in subjects recruited in different sites
- Setting practice and organizational features that affect recruitment
- Sample characteristics compared to intended population
- Sample attributes that might be potential confounding variables
- Attrition rates and patterns over time
- Differential attrition by subject attributes or arm assignment
- Causes of attrition

Intervention delivery
- Intervention content integrity
- Intervention purity
- Intervention dose integrity
- Interventionist training adequacy and requirements
- Interventionist competence
- Intervention reliability between interventionists
- Intervention effectiveness between interventionists
- Participant responses to interventions beyond outcome measures

Measures in pilot studies
- Respondent burden
- Participant difficulties with particular measures or parts of measures
- Appropriateness of order of measures
- Instrument reliability and validity estimates
- Missing data rates and patterns

Table 1. Pilot Study Findings Potentially Suitable for Publication (continued)

Study implementation	
	Protocol integrity: Interventions and measures delivered consistent with protocol
	Adequacy of randomization procedures
	Success of masking arm assignment from subjects and data collectors
	Identification of potential site or sample extraneous variables
	Anticipated and unanticipated human subjects protection issues
	Management of sensitive or legal issues
	Personnel time for recruitment, retention, intervention delivery, and measurement of study variables
Pilot study outcomes	
	Measures of central tendency and variability
	Effect size estimates
	Estimated sample size for parent study to detect clinically and statistically meaningful findings
	Inferential tests when appropriate
	Characteristics of data
	Patterns of findings over time
	Comparisons of different measures of the same construct
	Results for subsamples with particular characteristics or individual subjects
	Potential mediator variable findings
	Safety and unanticipated outcomes
	Lessons learned when predicted outcomes are not achieved
	Intervention effectiveness between interventionists

*Source: Conn et al., 2010 [14]

References

1 Lancaster, G.A., S. Dodd, and P.R. Williamson, Design and analysis of pilot studies: recommendations for good practice. Journal of Evaluation in Clinical Practice, 2004. 10(2): p. 307-312.

2 Thabane, L., et al., A tutorial on pilot studies: the what, why and how. BMC medical research methodology, 2010. 10(1): p. 1.

3 Eldridge, S., et al., Definition and reporting of pilot and feasibility studies. Trials, 2013. 14 (1): p. 1-1.

4 Whitehead, A.L., B.G.O. Sully, and M.J. Campbell, Pilot and feasibility studies: Is there a difference from each other and from a randomised controlled trial? Contemporary Clinical Trials, 2014. 38(1): p. 130-133.

5 Eldridge, S.M., et al., Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. PLOS ONE, 2016. 11(3): p. e0150205.

6 Wittes, J. and E. Brittain, The role of internal pilot studies in increasing the efficiency of clinical trials. Statistics in Medicine, 1990. 9(1-2): p. 65-72.

7 Friede, T. and M. Kieser, Sample Size Recalculation in Internal Pilot Study Designs: A Review. Biometrical Journal, 2006. 48(4): p. 537-555.

8 Leon, A.C., L.L. Davis, and H.C. Kraemer, The role and interpretation of pilot studies in clinical research. Journal of Psychiatric Research, 2011. 45(5): p. 626-629.

9 Hertzog, M.A., Considerations in determining sample size for pilot studies. Research in Nursing & Health, 2008. 31(2): p. 180-191.

10 Browne, R.H., On the use of a pilot sample for sample size determination. Statistics in Medicine, 1995. 14(17): p. 1933-1940.

11 Lancaster, G.A., Pilot and feasibility studies come of age! Pilot and Feasibility Studies, 2015. 1(1): p. 1-4.

12 Eldridge, S.M., et al., CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot and feasibility studies, 2016. 2(1): p. 64.

13 Arain, M., et al., What is a pilot or feasibility study? A review of current practice and editorial policy. BMC medical research methodology, 2010. 10(1): p. 67.

14 Conn, V.S., et al., Publishing Pilot Intervention Work. Western Journal of Nursing Research, 2010. 32(8): p. 994-1010.

15 Barber, S.E., et al., Assessing the feasibility of evaluating and delivering a physical activity intervention for pre-school children: a pilot randomised controlled trial. Pilot and Feasibility Studies, 2016. 2(1): p. 1.

16 White, A., et al., Western medical acupuncture in a group setting for knee osteoarthritis: results of a pilot randomised controlled trial. Pilot and Feasibility Studies, 2016. 2(1): p. 1.

17 Tailor, V.K., et al., Prescribed computer games in addition to occlusion versus standard occlusion treatment for childhood amblyopia: a pilot randomised controlled trial. Pilot and Feasibility Studies, 2015. 1(1): p. 1.

Prepared by: Alyssa Platt and Rachel D'Anna
Reviewed by: Duke Global Health Institute | Research Design & Analysis Core
Photo by: Sergi Ferrete @sergiferrete via Unsplash