Chapter 27
Pilot Studies

Ryan E. Ferguson and Mary T. Brophy

Introduction

Pilot studies are fundamental components of the research process that are conducted to examine the feasibility of an experimental approach for a subsequent larger study. The high cost of traditional explanatory clinical research (i.e., randomized controlled trials) and the restriction of financial support for funding agencies accounts for an increasing interest in pilot studies and the demand for pilot data prior to full-scale funding of a trial. Despite the increasing demand for pilot studies from funding agencies and the relative ubiquity of the design in academic settings, training on the design, planning, and execution of pilot studies is often missing in formal training programs for clinical science researchers [1]. To our knowledge, very few epidemiology or biostatistics text books cover the material in the necessary detail. Some texts mention design in passing and few provide more than cursory details, and relatively few textbooks dedicate an entire chapter to the topic. The objective of this chapter is to provide a detailed examination of the key issues of the design and conduct of pilot studies done in a clinical research setting.

R.E. Ferguson (✉) · M.T. Brophy
Massachusetts Veterans Epidemiology Research and Information Center, Boston Cooperative Studies Program Coordinating Center, VA Boston Healthcare System, 150 South Huntington Ave (151-MAV), Boston 02130, MA, USA
e-mail: ryan.ferguson@va.gov

M.T. Brophy
e-mail: mary.brophy@va.gov

K.M.F. Itani and D.J. Reda (eds.), *Clinical Trials Design in Operative and Non Operative Invasive Procedures*, DOI 10.1007/978-3-319-53877-8_27

What Is a Pilot Study?

Pilot studies are preparatory studies that are designed to "test the performance characteristics and capabilities of study designs, measures, procedures, recruitment criteria, and operational strategies that are under consideration for use in a subsequent, often larger, study" [1]. Pilot studies, then, are the vanguard for a full-scale clinical research study. Table 27.1 provides a description of types of clinical trials by 'phase' of drug development. Traditionally, pilot studies are restricted to inform Phase III or IV studies. Pilot studies are **not** first-in-human studies, early phase safety trials, or Phase 1–2 studies.

We will focus our discussion on pilot studies that are being conducted for Phase III clinical investigations as defined in Table 27.1. This restriction in scope is consistent with the recent recommendation from the British Medical Research Council which explicitly recommends the use of feasibility studies prior to the conduct of Phase III trials, especially those that include complex interventions [2]. Restriction of the discussion in this chapter to Phase III pilot studies is not meant to imply that pilot and feasibility studies cannot be done in other settings. In fact, they can be done in a variety of research areas (i.e., drug development, population science, genomic analysis, etc.) and across multiple study designs (i.e., randomized trial, prospective cohort studies, etc.) and are routinely used in qualitative research as well.

Classification of Pilot Studies

Pilot studies can be broadly categorized into four classifications: Process, Resources, Management, and Scientific [3].

Table 27.1 Phases of clinical investigation

Phase	Objective
I	To investigate the pharmacokinetics of a drug and to identify a dose that can be tolerated with minimal toxicity. Usually not randomized; small sample size
II	To assess preliminary evidence on clinical efficacy. Can be randomized or non-randomized; usually small sample size
III	To compare the efficacy and safety of two (or more) interventions, usually the investigational agent and a placebo. Studies are usually randomized; very large sample sizes
IV	To assess the post-marketing experience of the drug (e.g., long-term safety profile, drug–drug interactions, etc.). Studies are often non-randomized; often very large sample sizes

Process—this rationale applies when pilot studies assess the feasibility of key procedures that will take place in the main trial such as consent refusal rates given different types of informed consent documents or procedures as well as the overall accrual rates for the protocol. Draft case report forms (e.g., data collection forms) are often piloted prior to implementation to assess ease of completion, skip pattern and conditionality of questions. The intent is to iteratively improve quality through revisions.
Resource—this rationale applies to pilot studies that assess time and resource issues that are important to the main trial. For example, time for the completion of a subject interview can be assessed to understand how this will impact workload requirements of the study staff and ultimately factor into recruitment potential for the site. Resource piloting is also helpful to assess the availability and use of equipment needed for the trials, especially if the equipment is shared with the clinical staff.
Management—this rationale applies when the objective is to assess the potential human and data management issues that may arise in the main trial and provide opportunity to maximize data integrity and use of human resources. Pilot studies focused on management issues will often assess the challenges that study personnel will encounter when conducting different aspects of the main trial. Examples of key questions addressed include Are participating centers able to see patients within expected visit time interval? Are sites able to collect and capture the data?
Scientific—this rationale assesses study outcomes such as treatment effect size and variance around the estimate allowing for 'fine tuning' of the research hypothesis. In limited circumstances additional parameters such as drug safety and dosing can be assessed. Often, important rates associated with the analysis will be estimated in these pilot studies (i.e., missing data rates and participant attrition rates) and used to inform the analytic plan of the main trial.

Pilot studies have become increasingly more common in recent years and are often required by some sponsors to secure funding [4]. The National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health has an established funding mechanism (R34) specifically for the conduct of pilot studies in preparation of a larger, more robust clinical trial. The guidance from the NHLBI suggests that pilot studies should only be done to address gaps in knowledge that are required for the conduct or the design of the main trial.

In general, pilot studies present opportunities to clarify and sharpen the research hypotheses to be studied; identify potential barriers to study completion; evaluate performance of the trial systems and their acceptability to trial participants and providers; and enhance data integrity and human subjects' protections.

Internal Versus External Pilot Studies

Pilot studies may be classified as *internal* or *external*, each with its advantages and disadvantages. External pilot studies are those that are done completely independently of the main trial to assess its feasibility. They have their own specific aims, data collection procedures, and analytical plan. Data from external pilot studies should not be merged with data from the main trial [5]. Merging the data may create a selection bias and will inflate the type 1 error for the study. Figure 27.1 shows a graphical representation of the abbreviated workflow for an external pilot study.

Internal pilot studies are adaptive trials that are primarily designed to allow for re-estimation of sample size calculations of the main trial [3]. In this type of study, the main trial is planned using the best available data and is initiated on a pre-specified number of trial participants. Sample size is recalculated using the observed outcome rate and effect size seen in the pilot sample. If the originally calculated sample is large enough (or too big) then the original estimate will stand [5]. The principal advantage of internal pilot studies is that the design allows for sample size estimation without increasing the time for the conduct of the full trial [5]. All data collected from the initial patients can be used in the main trial and no effort (or data) is lost. Figure 27.2 shows a graphical representation of an internal pilot study. A major disadvantage of the internal pilot design is that other feasibility factors cannot be assessed as the pilot phase is, in fact, part of the main trial. In addition, the type 1 error will be slightly inflated as the pilot subjects and the main trial participants are considered to be independent when combined in the final analysis [5]. As long as the alpha level is controlled, internal pilot study designs offer flexibility and power [6].

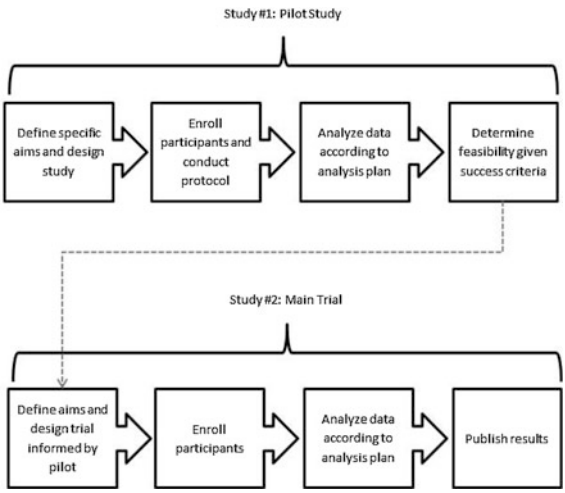

Fig. 27.1 Abbreviated workflow for an external pilot study

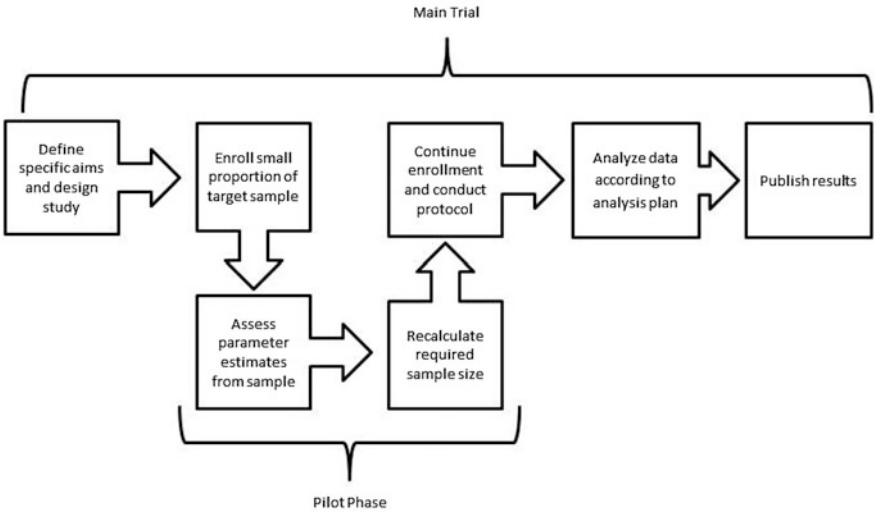

Fig. 27.2 Abbreviated workflow for an internal pilot study

Statistical Considerations

Design and Analytic Plan

The design of pilot studies should be guided by the same principles as the parent clinical trial particularly when feasibility of the parent study is the central issue. Pilot studies must have a well-elucidated statistical analysis plan with carefully constructed strategies for achieving each of its aims. The analysis plan should clearly identify the outcomes, the measures, and the acceptance criteria for each critical element. This axiom is true for all pilot studies regardless of classification and does not imply that formal analysis with inferential statistics is needed. Consider for example, a study investigating proper hydration levels to prevent contrast-induced nephropathy in diabetics receiving an angiography that underwent feasibility piloting within a large healthcare institution. Table 27.2 provides example questions that each category of pilot study may ask as well as sample outcomes that should be included in an analytic plan. As stated above, pilot studies will focus on feasibility; consequently, hypothesis testing on efficacy and safety endpoints are inappropriate analytic procedures for a pilot study analytic plan. As a result of the focus on feasibility, the analysis plan for a pilot will rely heavily on point and interval estimation and should only involve limited, if any, hypothesis testing (more on this below).

Table 27.2 Example analysis plans and acceptance criteria for each pilot classification

Category of pilot study	Example aim	Possible outcome	Sample acceptance criteria or analytic plan
Process	To assess the feasibility of the enrollment	% of eligible patients receiving an angiography that are consented	$\geq 20\%$ of eligible patients are consented into the protocol
Resource	To assess the resource requirements for the baseline participant interview	Time to completion	Successful interviews will be those conducted in less than 20 min (on average) across a sample of 20 patients
Management	To test whether the post-procedure hydration protocol can be implemented within the clinical care workflow	Proportion of consented patients receiving the post-procedure hydration protocol	$\geq 90\%$ of consented patients receive 100% of the hydration protocol within 3 h of the procedure
Scientific	To estimate the variance around the event rate within the healthcare system	Outcome (event) rate	Occurrence and distribution of the event rate with in 96 h of the procedure

Sample Size

Sample size estimation is often incorrectly viewed as 'not essential' for pilot studies because there will be limited hypothesis testing and restricted used of inferential statistics. However, this is a misconception that focuses only on the use of inferential statistics. Instead, sample size should be sufficiently large to obtain precise point estimates and confidence interval estimates for the parent study. Therefore, there is a very real need to have a clear and well-reasoned rationale for the number of participants to be included in the pilot study. The justification must be deeply rooted in the analytic plan and aligned with each of the aims of the pilot study. The choice of the appropriate sample size, then, will be driven by sound judgement and the aims of the pilot with specific consideration to the issues of practical feasibility and not by considerations related to power.

In 2005, Cook et al. [7] reported the results of a pilot study done in preparation for a large-scale study on the prophylaxis of thromboembolism. The pilot focused entirely on feasibility and reported recruitment rates, rates of protocol adherence, and an assessment of workload. The total sample of 120 participants from 16 intensive care units was selected (1) in order to obtain an estimate (with confidence intervals) of the proportion of people that would meet eligibility criteria and; and (2) to allow for an adequate sample (with at least 3 from each ICU) to refine protocol and screening procedures prior to full-scale deployment. All rates observed in the pilot study were then compared to rates that were specified a priori and "feasibility" of the larger trial was determined based on these "acceptance criteria."

There are no explicit rules or guidelines for the appropriate sample size of a pilot study. It should be large enough to provide point estimates and confidence intervals with sufficient precision to reduce statistical uncertainty but in practice they are typically too small to achieve this goal. A recent report from Billingham et al. looked at sample sizes in 79 funded trials recorded within the United Kingdom Clinical Research Network database and found that among pilot and feasibility studies the mean sample size for studies with dichotomous and continuous endpoints was only 36 (range: 10–300) and 30 (range: 8–114) per arm, respectively [8].

Power Calculations and Hypothesis Testing

While pilot studies are underpowered for testing of parent study hypothesis, they should be adequately sized to test operational issues and guide decisions about how the parent study will be conducted. Examples include the following: Is the RNA assay more accurate and more precise than the antigen assay? Is the taste of a particular dietary supplement acceptable to at least 95% of the target population? [1]. In these cases the power of the hypothesis test will depend on the choice of the sample and will be a function of the hypothesized parameter values. It is therefore very useful to calculate power with different sample sizes and to present power curves in the analytic plan. The biostatistical and hypothesis testing literature is rife with examples and formulae to guide calculations of the appropriate sample or power for given parameter estimates.

In summary, pilot studies are an important preparatory step in the progression of research that is hypothesis driven, but the studies themselves may not test a hypothesis. It is appropriate to focus on the level of precision for a given estimate (i.e., the statistical uncertainty and confidence interval) and not necessarily on the power level of a testing procedure.

A Cautionary Tale on the Use of Pilot Study Data to Guide Power Calculations

Kraemer et al. (2006) have shown that pilot studies can generate unreliable, unrealistic, and biased sample sizes for the larger parent trials because they are limited by small samples themselves [3, 9]. As a result, the parameter estimates generated by pilot studies should be used with extreme caution when estimating effect size within a larger population. Parameters generated from pilot studies may not have been estimated with sufficient accuracy to serve as the basis of power calculations or to serve as a basis for decision making on whether the main trial should be funded [9]. The authors report that the two likely outcomes of using pilot study data to drive power computation are as follows:

1. The study proposal will be aborted even when the actual effect is clinically significant.
2. If not aborted, the study sample estimated from the pilot data will be too small and will result in a study that is underpowered to detect the effect sizes of clinical significance.

In short, studies that calculate sample size that are based on effect sizes that are estimated from pilot studies will "likely" end in failed clinical trials and result in wasted resources. Therefore, the results of pilot studies should be used with caution as the data can potentially mislead sample size calculations.

Ethical Considerations

Informed Consent

There is a long standing history of debate on the ethical considerations of conducting underpowered research. In particular, underpowered studies are considered unethical because such studies will not adequately test their underlying hypotheses and they will be "scientifically useless" [10] yet will expose participants to both risks and burdens. However, similar discussion for pilot studies is lacking in this literature [10]. While pilot studies primarily address study feasibility with much less emphasis on statistical power, consideration of the same principles of informed consent is appropriate. Specifically, the consent process for pilot studies must convey the limited scope of the pilot to the subject [10].

Thabane et al. [3] investigated the obligation that researchers have, to patients or to participants in a trial, to disclose the feasibility nature and, hence the "limited" scientific value, of pilot studies. The authors reviewed the most cited research guidelines in the literature (e.g., the Nuremburg code, the Belmont Report, ICH GCP, etc.) and found that pilot studies are not addressed in any of the guidelines [3]. Thabane et al. [3] conclude that "given the special nature of feasibility or pilot studies, the disclosure of their purpose to study participants requires special wording—that informs them of the definition of a pilot study, the feasibility objectives of the study, and also clearly defines the criteria for success of feasibility". In order to fully inform participants, the authors have suggested template language for informed consent documents [3].

Publication

Although pilot studies can be very informative, few are ever published, perhaps because undue emphasis is placed on the statistical significance of findings rather than on feasibility issues that were the primary focus of the pilot study [3].

Underreporting of pilot study data results in publication bias [5] and further compounds the ethical considerations of the conduct of the pilot.

Recommendations for the Conduct of Pilot Studies

1. Keep the next study in mind!
   The pilot should be designed to maximize the information needed for the main trial. The design of the pilot should mimic the main trial as should the study procedures.
2. Maintain methodological rigor.
   The same principles that guide the design of the main study should be followed for a pilot study. The small size and feasibility focus does not remove the obligation to generate accurate and precise data.
3. Clearly define aims, objectives, and the definitions of success.
   The aims, the objectives, and the design should all be aligned. Acceptance criteria and definitions of success should be clearly articulated a priori as should a clear plan to use the data generated by the pilot study.
4. Align analysis plan with objectives and design of study.
   The analysis should be mainly descriptive and contain very limited hypothesis testing. If hypothesis testing is used, results should be treated as preliminary and not definitive. Sample size must be justified in the analysis plan.
5. Must convey limited value to participants.
   Ethical principles demand informed consent and notification of the limited value of the pilot study.
6. Publish the results.
   Results from all pilot studies should be reported. Reporting of results should follow the guidelines adapted from the CONSORT Statement by Thabane et al. in 2010. Reporting will help to reduce the impact of publication bias and will contribute by advancing the scientific community.

References

1. Moore CG, Carter RE, Nietert PJ, Stewart PW. Recommendations for planning pilot studies in clinical and translational research. Clin Transl Sci. 2011;4(5):332–7.
2. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new medical research council guidance. BMJ. 2008;29 (337):a1655.
3. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, Robson R, Thabane M, Giangregorio L, Goldsmith CH. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol. 2010;10(1):1.
4. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res. 2011;45(5):626–9.

5. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval clin Pract. 2004;10(2):307–12.
6. Wittes J, Brittain E. The role of internal pilot studies in increasing the efficiency of clinical trials. Stat Med. 1990;9(1–2):65–72.
7. Cook DJ, Rocker G, Meade M, Guyatt G, Geerts W, Anderson D, Skrobik Y, Hebert P, Albert M, Cooper J, Bates S. Prophylaxis of thromboembolism in critical care (PROTECT) trial: a pilot study. J Crit Care. 2005;20(4):364–72.
8. Billingham SA, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom clinical research network database. BMC Med Res Methodol. 2013;13(1):1.
9. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry. 2006;63 (5):484–9.
10. Halpern SD, Karlawish JH, Berlin JA. The continuing unethical conduct of underpowered clinical trials. JAMA. 2002;288(3):358–62.