Introduction

This document provides guidance for RDS advisors supporting applications for feasibility studies and stand-alone pilot studies. Issues specific to internal pilot studies are not addressed in this guide. It is informed by the experiences and reflections of the author, by relevant research literature, by the observations of more than sixty RDS advisors across nine regions who completed a 2015 survey, and by their comments on early and subsequent drafts.

Before attempting to describe what is meant by the terms ‘feasibility study’ or ‘pilot study’, let’s first consider the circumstances in which research teams might want to undertake preparatory research before conducting full-scale clinical research.

When planning large clinical studies, research teams will draw heavily on their knowledge and experiences, findings from other studies, routinely collected clinical data and the views of patients and other investigators. Beyond expertise and information, investigators require access to clinical sites and health data and the cooperation of a range of people, not least the prospective study participants. If important information is missing or equivocal, or the extent of access or cooperation is in doubt, preparatory studies may be needed to resolve planning uncertainties. Preparatory studies can also help investigators to decide whether full-scale studies are viable and, if they are, how large they need to be to produce reliable findings. Furthermore, when investigators believe they have sufficient information to deliver main studies, preparatory studies allow plans and assumptions to be tested before resources are committed to fully powered trials or other large-scale research.
NIHR and MRC definitions and differences

The most common labels associated with preparatory studies are ‘feasibility’ and ‘pilot’, although other terms, such as ‘vanguard’, ‘exploratory’, and ‘preliminary’ are sometimes used. Until quite recently, most investigators used the terms ‘feasibility study’ and ‘pilot study’ interchangeably. The MRC guidance on developing and evaluating complex interventions (1) includes a ‘feasibility and piloting’ stage:

“The feasibility and piloting stage includes testing procedures for their acceptability, estimating the likely rates of recruitment and retention of subjects, and the calculation of appropriate sample sizes. Methodological research suggests that this vital preparatory work is often skimmed. Evaluations are often undermined by problems of acceptability, compliance, delivery of the intervention, recruitment and retention, smaller-than-expected effect sizes, and so on, that could be anticipated by thorough piloting. A pilot study need not be a ‘scale model’ of the planned mainstage evaluation, but should address the main uncertainties that have been identified in the development work […]. Pilot study results should be interpreted cautiously when making assumptions about the required sample size, likely response rates, etc., when the evaluation is scaled up. Effects may be smaller or more variable and response rates lower when the intervention is rolled out across a wider range of settings. A mixture of qualitative and quantitative methods is likely to be needed, for example to understand barriers to participation and to estimate response rates. Depending on the results, a series of studies may be required to progressively refine the design, before embarking on a full-scale evaluation.”

In contrast, some NIHR programmes (EME, PHR, HTA and RfPB) have separate definitions for feasibility and pilot studies. (2) They describe feasibility studies as:

“.. research done before a main study in order to answer the question ‘Can this study be done?’. They are used to estimate important parameters that are needed to design the main study. For instance:

• standard deviation of the outcome measure, which is needed in some cases to estimate sample size;
• willingness of participants to be randomised;
• willingness of clinicians to recruit participants;
• number of eligible patients; carers or other appropriate participants;
• characteristics of the proposed outcome measure and in some cases feasibility studies might involve designing a suitable outcome measure;
• follow-up rates, response rates to questionnaires, adherence/compliance rates, ICCs in cluster trials, etc;
• availability of data needed or the usefulness and limitations of a particular database;
• time needed to collect and analyse data.

Feasibility studies for randomised controlled trials may not themselves be randomised. Crucially, feasibility studies do not evaluate the outcome of interest; that is left to the main study. If a feasibility study is a small randomised controlled trial, it need not have a primary outcome and the usual sort of power calculation is not normally undertaken. Instead the sample size should be adequate to estimate the critical parameters (e.g. recruitment rate) to the necessary degree of precision.”

Pilot studies are defined as:

“.. a version of the main study that is run in miniature to test whether the components of the main study can all work together. It is focused on the processes of the main study, for example to ensure recruitment, randomisation, treatment, and follow-up assessments all run smoothly. It will therefore resemble the main study in many respects, including an assessment of the primary outcome. In some cases this will be the first phase of the substantive study and data from the pilot phase may contribute to the final analysis; this can be referred to as an internal pilot. Or at the end of the pilot study the data may be analysed and set aside, a so-called external pilot.”

The NIHR Research for Patient Benefit Programme elaborates further on the NIHR descriptions in their Guidance Information for Applicants, (3) which attempts to clarify why feasibility studies are funded by the programme while generally pilot studies are not:

“Pilot trials are therefore usually funded as part of the phased development of a full trial. In effect, they offer the researchers and funders a stopping point should the trial not prove viable. It is therefore unlikely that RfPB would fund a pilot trial as continuation would involve a separate application to another funder.

Feasibility studies, on the other hand, attempt to derive more precise estimates of various parameters which will be required for a full trial. Guidance on what these parameters might be has been provided on the website […]. The design of a feasibility study involves listing those parameters (recruitment rates, retention rate, variability of primary endpoint, willingness to be randomised, etc) which are uncertain and describing the methods for improving their precision so that a full trial will have a better chance of being successfully funded.”

The guidance adds that pilot studies should include an assessment of the viability of the main trial and provide a stopping point should the main study be considered not viable.

Differences between pilot and feasibility studies are explored further in the question-and-answer section later in this guide.
Table 1. Comparison of NIHR and MRC definitions of feasibility and pilot studies.

<table>
<thead>
<tr>
<th>NIHR definition</th>
<th>MRC definition (in relation to complex interventions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasibility and pilot studies are considered distinct study types.</td>
<td>Piloting and feasibility assessment are portrayed as an integrated activity.</td>
</tr>
<tr>
<td>Feasibility studies are principally conducted to establish whether large studies can be delivered. Parameter estimation is emphasised. Pilot studies are conducted to assess whether key study elements run &quot;smoothly&quot; - there is no explicit mention of feasibility assessment in the definition of pilot studies.</td>
<td>Pilot studies are conducted to inform the design of a main study. Feasibility assessment is an implied activity within pilot studies.</td>
</tr>
<tr>
<td>Little is said about the design of feasibility studies, other than that they need not follow the design of the main study, and may or may not be randomised. Pilot studies should be miniature versions of the main study, sharing &quot;many&quot; of its features.</td>
<td>Pilot studies need not be &quot;scale models&quot; of the main study but, presumably, can be if necessary. Data collection activities should address the main uncertainties identified in the development work.</td>
</tr>
<tr>
<td>The focus of the analysis of feasibility studies is largely quantitative (parameter estimation) although some suggested activities will involve qualitative judgements (&quot;usefulness and limitations of a particular database&quot;) or developmental work (&quot;designing a suitable outcome measure&quot;). No analysis methods are mentioned in relation to pilot studies.</td>
<td>Both qualitative and quantitative methods are likely to be needed in pilot studies.</td>
</tr>
<tr>
<td>A feasibility study should not include evaluation of the main outcome that will be assessed in the full-scale study, and may not necessarily even measure that outcome. The outcomes of interest should be measured in a pilot study, although any analysis should be &quot;set aside&quot;. The size of the feasibility study should be determined by the degree of accuracy needed to estimate parameters. No mention is made of how large pilot studies need to be. No mention is made of the possibility of multiple related feasibility or pilot studies.</td>
<td>Estimates derived from pilot studies &quot;should be interpreted cautiously&quot; when used to inform the sample size and other aspects of the intended larger study. No mention is made of how large pilot studies need to be. More than one study may be needed to refine the design of the main study.</td>
</tr>
</tbody>
</table>

Issues observed by RDS advisors supporting feasibility and pilot studies

In a 2015 survey, RDS advisors reported correcting a range of mistakes and misunderstandings when supporting applications for feasibility and pilot studies. Table 2 lists the six most common issues found in early consultations with clients, ordered by how often they were experienced. For each, more than half of respondents encountered the issue ‘often’.

Table 2. Common mistakes and misunderstandings encountered by RDS advisors when supporting feasibility and pilot studies.

<table>
<thead>
<tr>
<th>Issue</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicants are uncertain about whether the research should be described as a ‘pilot’ or a ‘feasibility’ study.</td>
<td></td>
</tr>
<tr>
<td>It is not clear what criteria will be used to decide whether to continue to a full-scale study.</td>
<td></td>
</tr>
<tr>
<td>Potentially important objectives are missing from proposals.</td>
<td></td>
</tr>
<tr>
<td>The number of participants to be recruited is not properly justified.</td>
<td></td>
</tr>
<tr>
<td>Not all outcome measures are clearly justified.</td>
<td></td>
</tr>
<tr>
<td>Objectives for the pilot/feasibility study include establishing the effectiveness of the intervention.</td>
<td></td>
</tr>
</tbody>
</table>

These common difficulties are addressed in the question and answer section on the next page.
Supporting applications to NIHR programmes: Answers to common questions

The advice set out in this guide purposely aligns to the NIHR descriptions of feasibility and pilot studies. Whilst these definitions are not universally recognised, much of the advice below remains pertinent to applications to other funders.

This Q&A is a work in progress and some may find aspects contentious. If you believe you can improve any of the answers, or would like to suggest additional questions for this guide, please contact Martin Williams at mjw48@le.ac.uk.

Q1 What is the difference between feasibility studies and stand-alone pilot studies and why are they often confused?

Answer 1: The NIHR guidance (2) describes the focus of feasibility studies as the estimation of parameters to inform whether main studies are possible and pilot studies as involving small-scale testing of the “smooth” delivery of research plans of larger studies. However, whilst these definitions appear straightforward, investigators often have difficulty interpreting and applying them to their own research needs.

Investigator confusion over the differences between feasibility and pilot studies is understandable as many authors and institutions use the terms to mean different things. Even within the NIHR descriptions there is scope for overlap in their designs and objectives: parameter estimation is characteristic of feasibility studies but often happens in pilot studies too; pilot studies look like small-scale versions of main studies, but some feasibility studies can also resemble main studies, including mini-trials; and both typically lead to decisions about whether to proceed to full-scale studies. Investigator confusion is a particular problem in applications to the NIHR Research for Patient Benefit programme where proposals for feasibility studies are encouraged but pilot studies are generally considered out-of-scope. (3) Despite their similarities, there are features that distinguish the two study types.

Pilot studies
Pilot studies purposefully test putative plans/protocols for main studies; feasibility studies do not. To plan and deliver “a version of the main study [...] in miniature” investigators will need to have settled on the design and organisation of their main study before delivering their pilot. Pilot studies can assess the coherence of main study plans/protocols and test key assumptions, such as recruitment rates and the estimates underpinning sample size calculations. They may also examine a range of organisational and operational issues, such as resource requirements, training needs, activity timelines, the fidelity of delivery of the intervention, data quality and the acceptability of aspects of the study protocol to clinicians and participants. These considerations will lead to judgements about the viability of the main study. Information gained from pilot studies can also help to refine the plans for main studies (e.g. by making modifications that will improve their efficiency or minimise potential sources of bias).

As previously mentioned, the designs of pilot studies will be similar to the designs of intended main studies, albeit with fewer participants. However, the designs of external pilots can sometimes deviate from those of their definitive equivalents in aspects other than scale. For example, external pilot studies may sometimes have shorter follow-up periods or fewer measures or measurement points than their full-scale counterparts. Deviations of this kind can make the pilots quicker and less expensive, but care should be taken so that they do not invalidate important inferences made about the main study. Pilot studies may also involve data collection that will not be part of the main study, such as interviews with participants to improve recruitment procedures or to reduce participant burden for better retention.

Feasibility studies
Feasibility studies are used to collect information to establish the feasibility of main studies and to inform their designs and processes. When feasibility studies are planned, important aspects of the larger follow-on studies will be unknown or undecided. They can take many forms, including those that resemble small-scale versions of trials (e.g. with randomisation to treatment arms and follow-up). It is important, however, that all of the design features of feasibility studies are driven by information needs. For example, a need to measure willingness to be randomised and to establish treatment adherence rates in the two study arms might justify randomisation to new and standard interventions.

O’Cathain and colleagues (4) provide useful guidance on maximising the impact of qualitative research in feasibility studies for randomised trials. They highlight the value of qualitative information to help to: refine the intervention content and delivery; improve the trial design, conduct and processes; inform the selection of appropriate outcomes; and enhance measurement completion and accuracy. The Association of Medical Research Charities (AMRC) and NIHR Medicines for Children Research Network (MCRN) have produced a guide to feasibility assessment for research funders that emphasises the importance of assessing recruitment issues, clinical relevance and site logistics. (5)
Q2 What criteria should be used to assess the feasibility of the future trial in a feasibility or pilot study?

**Answer 2:** There are a number of ways in which findings from a feasibility or pilot study might indicate that a main study could not or should not be attempted. For example, they may reveal problems with: participant recruitment and retention; the delivery of an intervention; missing data; participant compliance; or acceptability of trial procedures to patients or clinical staff. Where such issues cannot be resolved, the investigators may conclude that the main trial is not viable or not worthwhile.

Thabane and colleagues (6) they are designed to assess the safety of treatment or interventions; to assess recruitment potential; to assess the feasibility of international collaboration or coordination for multicentre trials; to increase clinical experience with the study medication or intervention for the phase III trials. They are the best way to assess feasibility of a large, expensive full-scale study, and in fact are an almost essential pre-requisite. Conducting a pilot prior to the main study can enhance the likelihood of success of the main study and potentially help to avoid doomed main studies. The objective of this paper is to provide a detailed examination of the key aspects of pilot studies for phase III trials including: 1 discuss how feasibility criteria can be used to assess how and whether to deliver the main study. They propose that a pilot study can lead to one of four recommendations: (i) Stop – main study not feasible; (ii) Continue, but modify protocol – feasible with modifications; (iii) Continue without modifications, but monitor closely – feasible with close monitoring; and (iv) Continue without modifications – feasible as is. These recommendations cannot, however, be applied directly to feasibility studies as the final protocol of the main study will not be fully defined. Instead, a feasibility study might simply recommend that a main study is feasible or not.

The Thabane paper provides an example of feasibility criteria from the PeriOperative Epidural Trial (POET) (7) but recent larger trials have been inconclusive. We conducted a pilot study to assess the feasibility of conducting a large multicentre randomized controlled trial in Canada.

**FINDINGS:** After Research Ethics Board approvals from the participating institutions, subjects were recruited if they were > or = 45 years old, had an expected hospital stay > or = 48 hours, were undergoing a noncardiothoracic procedure amenable to epidural analgesia, met one of six risk criteria, and did not have contraindications to neuraxial blockade. After informed consent, subjects were randomly allocated to combined epidural analgesia (epidural group pilot study):

- one subject per centre per week (i.e., 200 subjects from four centres over 50 weeks) can be recruited;
- at least 70% of all eligible patients can be recruited;
- no more than 5% of all recruited subjects crossed over from one modality to the other; and
- complete follow-up in at least 95% of all recruited subjects.

Criteria, such as those used in the POET study, can help to provide clarity for how continuation recommendations will be made. Investigators should, however, be mindful of a number of issues when setting feasibility criteria:

- It can be tempting to use numerical targets as cut-points such that, if any value observed in a preparatory study should fall short, the conclusion would be that the main study could not be delivered. This, however, is an inflexible and restrictive approach. For example, it would seem overly prescriptive to decide in advance that a main study would not be feasible if x-1% of eligible patients could be recruited in a pilot, but feasible if x% of patients could be recruited.

- A more flexible approach would involve judgements about both the implications of the observed data for the main study and whether any identified deficiencies can be corrected. For example, poor patient recruitment in a pilot study, with no obvious way of improving the situation, might lead the investigators to conclude that the main study is not viable. However, had a solution to the recruitment issues been identified, then the recommendation might be different.

- Target values can be more useful when they clearly represent desirable values for the main study and are justified in terms of their implications for that study. For example, an estimate of the number of eligible patients will have important implications for the design, scale, cost and viability of a main study. Lower than expected numbers of eligible patients might imply a need for a longer recruitment period and/or additional study sites and/or changes to the eligibility criteria and it may or may not be possible to accommodate these changes.

- Feasibility criteria can involve assessment of qualitative information as well numerical data. (4)

- Not all information collected in feasibility or pilot studies will inform continuation decisions; some findings will be used to shape or refine the processes and design of the future main study should it prove feasible.

- Allowance should be made for unforeseen events and accumulating external evidence to be considered in final feasibility recommendations.
For pilot studies, an estimate of the desired sample size of the future trial can be very helpful when considering viability. One parameter that can potentially be estimated in pilot studies is the size of the difference between treatment outcomes. It may be tempting, therefore, to consider using this information to decide whether a main study should be conducted. For example, if an estimate of a treatment difference observed in a pilot study were small, it would seem sensible to assess how likely it would be that the main study could reach a significant result. Unfortunately, pilot studies are generally too small to determine ‘futility’ in this way and so having this as an objective would rarely be a realistic proposition.

Are there important aims and objectives that feasibility and pilot studies should include?

**Answer 3:** The aims and objectives of feasibility and pilot studies should directly address gaps in current knowledge, trial planning needs and viability assessment. However, there are certain objectives that many funding panels will expect to see. First, funding panels will anticipate data collection leading to a decision about whether the future main study can be delivered. This objective is explicit in the NIHR description of feasibility studies, but viability assessment will also be an expectation of most panels considering pilot studies.

A second objective that should be given strong consideration in both feasibility and pilot studies is the estimation of likely recruitment rates. This information is useful for assessing the future trial’s feasibility, but will also support study planning. Failure to recruit to target is a well-known risk to clinical research. (8–10) Even if a research team is confident of recruitment, research funders may seek reassurance from smaller scale preliminary research before committing resources to a full-scale study. Feasibility and pilot studies can be used to explore related issues, such as: how study sites will be selected and their willingness to take part; the numbers of eligible patients at study sites; numbers likely to agree to participate or to decline; how potential participants should best be approached; likely rates of participant dropout; and completion rates of follow-up measures.

In applications to NIHR funding programmes, it can be helpful to encourage applicants to mirror the language used by the NIHR to describe feasibility and pilot studies. (2) That is, aside from establishing the feasibility of the main study, feasibility studies will “estimate […] parameters” that will inform the design of the main study, and pilot studies will assess whether “the components of the main study can all work together” and ensure that “processes of the main study […] all run smoothly”.
Q4 What advice should I give regarding the number of participants in feasibility and pilot studies?

Answer 4: The size of a feasibility or pilot study will be influenced by a range of factors including the type of information that will be collected, the purpose for which it is needed, the desired accuracy of any parameter estimates, and pragmatic considerations such as time and resource requirements. When there is no need to estimate parameters with particular accuracy, the size of feasibility and pilot studies is often a pragmatic balance between information needs and resources.

When pilot studies are used to estimate parameters that will underpin sample size calculations for a main study, such as outcome event rates or the distribution of an outcome variable, guidance is available on how large they should be. In a simulation study, (11) there is little consensus on how large pilot studies need to be, and some suggest inflating estimates to adjust for the lack of precision when planning the definitive RCT.

METHODS: We use a simulation approach to illustrate the sampling distribution of the standard deviation for continuous outcomes and the event rate for binary outcomes. We present the impact of increasing the pilot sample size on the precision and bias of these estimates, and predicted power under three realistic scenarios. We also illustrate the consequences of using a confidence interval argument to inflate estimates so the required power is achieved with a pre-specified level of confidence. We limit our attention to external pilot and feasibility studies prior to a two-parallel-balanced-group superiority RCT.

RESULTS: For normally distributed outcomes, the relative gain in precision of the pooled standard deviation (SD) that Teare and colleagues found that the optimal size for pilot studies used to inform sample size calculations was between 70 and 120 participants, depending on the nature of the outcome data. They concluded that pilot studies of this size provide efficiency gains by reducing the risks of running definitive trials that are under-powered or over-powered.

Feasibility studies and, to some extent, pilot studies may contain multiple sub-studies of varying sizes. For example, a particular feasibility study might involve: a survey of up to 40 clinical staff to assess their theoretical willingness to recruit patients and to follow trial procedures; interviews with up to 20 patients to assess their theoretical willingness to participate in a future trial and their views on options for intervention delivery and outcome measurement; and a review of clinical data, with no participants, to assess data accuracy and completeness.

Q5 How should outcome measures in feasibility and pilot studies be justified?

Answer 5: To assess the processes of the main study, a pilot will typically include all or most of the clinical and cost outcome measures anticipated for that study. Justification of measures will, therefore, relate to their importance to the main study. Separate justification will be needed for any additional information collected in the pilot that will not be part of the main study, such as data that will be used to refine study processes or to undertake feasibility assessment.

In feasibility studies, all data collection should be justified in terms of its value for informing the design of the main study and/or for assessing the feasibility and value of that main study. It can be useful to describe whether this information is available from other sources, such as published studies in related areas and, if it is, why further data collection is needed.
Should feasibility and pilot studies include an assessment of the effectiveness of the intervention?

**Answer 6:** The simple answer is “no”; whilst pilot studies and some feasibility studies may measure participant outcomes, they are by definition underpowered to make useful assessments of effectiveness.

The NIHR description of feasibility studies (2) is unequivocal in stating that they “do not evaluate the outcome of interest”. The description of pilot studies is more nuanced, stating that they will:

“...resemble the main study in many respects, including an assessment of the primary outcome. [...] In some cases this will be the first phase of the substantive study and data from the pilot phase may contribute to the final analysis, this can be referred to as an internal pilot. Or at the end of the pilot study the data may be analysed and set aside, a so-called external pilot.”

The NIHR distinguishes between how outcome data may be used in internal and in external pilot studies. Since the focus of this guidance is purely on external (stand-alone) pilot studies, then the advice relating to feasibility and pilot studies is the same; that any outcome measurements should not be used to assess the effectiveness of the intervention.

It should also be noted that most journals do not expect to see an assessment of the effectiveness of interventions in articles reporting on feasibility or stand-alone pilot studies.

References


Martin Williams, May 2016

© This guide may be freely distributed to advisors within the NIHR Research Design Service. It should not be circulated outside of the RDS or altered in any way without the prior permission of the author.