Are pilot trials useful for predicting randomisation and attrition rates in definitive studies: A review of publicly funded trials

Cindy L Cooper¹, Amy Whitehead²,³, Edward Pottrill², Steven A Julious² and Stephen J Walters²

Abstract

Background/aims: External pilot trials are recommended for testing the feasibility of main or confirmatory trials. However, there is little evidence that progress in external pilot trials actually predicts randomisation and attrition rates in the main trial. To assess the use of external pilot trials in trial design, we compared randomisation and attrition rates in publicly funded randomised controlled trials with rates in their pilots.

Methods: Randomised controlled trials for which there was an external pilot trial were identified from reports published between 2004 and 2013 in the Health Technology Assessment Journal. Data were extracted from published papers, protocols and reports. Bland–Altman plots and descriptive statistics were used to investigate the agreement of randomisation and attrition rates between the full and external pilot trials.

Results: Of 561 reports, 41 were randomised controlled trials with pilot trials and 16 met criteria for a pilot trial with sufficient data. Mean attrition and randomisation rates were 21.1% and 50.4%, respectively, in the pilot trials and 16.8% and 65.2% in the main. There was minimal bias in the pilot trial when predicting the main trial attrition and randomisation rate. However, the variation was large: the mean difference in the attrition rate between the pilot and main trial was 2.4% with limits of agreement of 237.1% to 28.2%. Limits of agreement for randomisation rates were 247.8% to 77.5%.

Conclusion: Results from external pilot trials to estimate randomisation and attrition rates should be used with caution as comparison of the difference in the rates between pilots and their associated full trial demonstrates high variability. We suggest using internal pilot trials wherever appropriate.

Keywords
Pilot trial, randomisation, attrition, dropout, recruitment, feasibility, trial design

Background

In the United Kingdom 2014/2015, the National Institute for Health Research in England invested £237.6 million to assess new health technologies.¹ A large proportion of this research has been randomised controlled trials (RCTs).

A common problem with publicly funded RCTs is that the recruitment in the trial is not as good as anticipated with many trials failing to reach their target sample size. A review of a cohort of trials funded by the UK Medical Research Council and the Health Technology Assessment Programme between 2002 and 2008 demonstrated that of 73 funded studies, 55% (40/73) of the trials achieved their original patient recruitment target, 16/73 (22%) achieved <80% of their original target and 45% (33/73) were awarded an extension.²

One way to mitigate the risks of conducting a large clinical trial is to undertake small preliminary or pilot trials first. This may help facilitate the design of the

¹Clinical Trials Research Unit, School of Health and Related Research (ScHARR), The University of Sheffield, Sheffield, UK
²School of Health and Related Research (ScHARR), The University of Sheffield, Sheffield, UK
³University of Southampton, Southampton, UK

Corresponding author:
Cindy L Cooper, Clinical Trials Research Unit, School of Health and Related Research (ScHARR), The University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK.
Email: c.l.cooper@Sheffield.ac.uk
main trial by enabling an estimate assessment of the recruitment and attrition rates for the trial. A recent review of Health Technology Assessment–funded trials found that the average attrition rate was 11%, while the average proportion of eligible patients who are randomised was 70%. Accurate estimates of both these rates are important a priori for the design of future studies.

It should be highlighted though that pilot studies are not just designed to inform on recruitment. They can also give confidence that a clinically meaningful effect is likely to be observed and inform sample size calculations. When they are successful, they can provide sufficient confidence that a larger trial can be conducted. Even when the results are not as expected, the experiences from the trial can still be useful.

The sample size calculation is an essential step in the design of a RCT. An important aspect of a sample size calculation is that it provides the researcher with an estimate of the number of evaluable participants required in the dataset at the end of the study in order to investigate the study hypothesis with the required level of power for a given level of statistical significance.

An important consideration when estimating the evaluable sample size and designing the trial are the facts that not all patients who are approached will consent to be in the trial, and of those who consent, not all will complete the trial and have evaluable outcome data and information for the statistical analysis. Participants with non-evaluable data are defined as the participants lost during the course of the study due to withdrawal or dropout. Participants can withdraw from a study for any number of reasons such as discontinued participation in treatment, missing study visits at outcome measure time points or non-completion of study data collection forms. Therefore, to assist in the design of a RCT, we need to have estimates of the proportion of eligible participants who will enter the study and are randomised as well as the proportion of randomised participants who will have evaluable data. The proportion randomised and the proportion of patients with evaluable data can be estimated from a pilot trial.

To account for the attrition (also referred to as withdrawal or dropout) of participants during the trial, the calculated sample size must be increased to deliver a sufficient number of evaluable participants at the end of the trial when the outcome data are analysed to ensure the appropriate power. Hence, if the number of withdrawals is underestimated, the trial will not have the appropriate evaluable sample size and necessary power. Conversely, overestimating the number of withdrawals could lead to over-recruitment in the trial with more patients potentially exposed to inferior treatments. This can occur, for example, where the follow-up phase is long (e.g. 2 years) in relation to the recruitment phase and recruitment finishes before the extent of the lower than anticipated attrition is known.

Once we have calculated the number of participants that we need to randomise in order to have the required number of evaluable participants, we need to have some estimate of the proportion of eligible patients who will be randomised. This is required for planning purposes to ascertain the recruitment period for the trial.

Using a pilot trial to estimate parameters for the main RCT has a number of advantages over using data from published literature. The main advantage is that the pilot trial can be designed to be a miniature version of the main trial – for example, in the same trial population with similar inclusion and exclusion criteria. In contrast, previously reported trials may differ from the trial being planned in ways which impact on the accuracy of the estimates.

Eldridge et al. propose that a pilot study should be considered to be a subtype of feasibility study with both feasibility and pilot studies aimed at answering whether and how something can be done, a pilot study being designed specifically as the full study but on a smaller scale. Guidance on good reporting of pilot and feasibility trials has been published.

The National Institute for Health Research defines a pilot study as a miniature version of the main study to determine whether the components of the main study can work together. Their guidance distinguishes between internal and external pilots, the former being the first phase of a substantive study which contributes data to the main analysis. The latter are undertaken prior to the substantive study with the data being analysed separately from the main study.

Lancaster et al. have outlined the possible objectives that an external pilot study can achieve including testing the integrity of the study protocol, testing the randomisation procedure, finding the rates of recruitment and consent, assessing the acceptability of the intervention and identifying the most appropriate primary outcome measure.

The aims of this article are to describe and compare the attrition rates and the rates of eligible to randomised patients of external pilot trials with those of their respective main RCTs in order to ascertain whether external pilot trials usefully predict these attributes in the main trial.

**Method**

**Trial identification**

To fulfil the aims of this study, an audit of RCTs with pilot trials was carried out in two phases. An initial set of RCTs published between 2004 and 2013 was collected from reports published online in the *Health Technology Assessment Journal*. The criteria for
inclusion of trials in this first stage of the audit were single or multi-centre RCTs, fully or partially randomised main trials and trials reporting early termination. The following were excluded: cluster randomised trials, trials that used adaptive designs, trials of influenza (because they recruit very quickly over a short period) and external pilot trials. Cluster trials were excluded as the unit of randomisation is different from that of individually randomised trials. In a cluster trial, one may not be obtaining individual consent from patients, and what is of interest in cluster trials is how many eligible clusters when approached agree to be randomised and what is the attrition of the clusters.

In this initial audit, information was collected about whether or not the RCT had a pilot trial. Trials which had been identified as having a pilot trial then went forward into the second phase of the audit. The second phase of the audit assessed each of the pilot trials to see whether they matched our definition of a pilot trial and whether they were therefore eligible to be entered into our study.

Although there are a number of definitions of pilot studies, we based our definition on that of the National Institute for Health Research at the time of the study which is that a pilot study is a smaller version of the main study undertaken to test whether the components of the main study can all work together. This definition has been recommended for use as pilot studies have been found to be poorly reported, in part due to variation in definitions and lack of distinction between pilot and feasibility studies. To qualify the study for inclusion, in line with this definition, the pilot trial had to be a version of the main study run in miniature and include assessment of some processes of the main study, for example, recruitment and randomisation rates and/or outcome measure collection. Pilot studies expressly aimed at assessing factors other than recruitment and attrition rates (e.g. sample size and outcome measures) were included if data were available or could be obtained from authors to allow assessment of recruitment or attrition rates or both.

**Eligibility criteria**

Trials were eligible for analysis if they were eligible in the initial phase of the audit and had an external pilot trial which conformed to our definition of a pilot trial.

**Data extraction**

For the main trials, data were extracted using the individual trial papers, trial protocols and *Health Technology Assessment* Journal reports. For the pilot trials where published papers were available, these were used to extract the required data. Some information was also gained by consulting the main trial study protocols. In some circumstances, not all of the required information was available from these sources. There was no published paper for most of the pilot trials, so limited information was available in some cases. For trials with missing data, the information was requested directly from the corresponding author of the trial report. The data extracted were entered and analysed in Microsoft Excel.

Data extracted included number of patients eligible, randomised and available for analysis; number of centres in the main trial; reporting of involvement of a Clinical Trials Unit in the main trial and length of follow-up period in pilot and main trial.

**Analysis**

The analysis is mainly descriptive. Bland–Altman plots were used to test the level of agreement and to look at the bias of the results from the pilot trial in predicting the main trial outcomes. The primary outcome variables were

- The attrition rate – derived by dividing the number of withdrawn participants (calculated as the number of people retained in the study subtracted from the total number randomised) by the number originally randomised.
- The rate of eligible patients who are randomised – derived by dividing the number of patients randomised into the trial by the number originally identified as eligible.

The difference was calculated as pilot minus main trial results.

Secondary exploratory analysis was undertaken to investigate a number of factors which could potentially explain the differences in the estimates including number of centres, number of participants and involvement of a Clinical Trials Unit.

**Results**

In total, 561 reports published in the *Health Technology Assessment* Journal were available from trials completed between 2004 and 2013. From which, 459 of the reports were excluded for not being RCTs or for being cluster RCTs leaving 102 RCTs. A further 3 RCTs were excluded from the initial audit in the first phase, leaving 99 RCTs. Of these trials, 58 trials had no pilot trial. The remaining 41 studies were reviewed and further studies were excluded as: did not have sufficient data (n = 2), was a cluster RCT (n = 1), did not conform to our definition of a pilot trial, that is, did not include assessment of some processes of the main study, for example, recruitment and randomisation rates and/or outcome measure collection (n = 17); the pilot trial was an internal pilot (n = 5 studies). The studies which did not provide sufficient data were excluded after
emails were sent, without a sufficient response, to the trial investigators requesting the missing information. A total of 16 external pilot trials were analysed. Figure 1 summarises the flow of the trials throughout the review.

Primarily, we were interested in how well the pilot trial attrition rate predicts the attrition rate in the main trial. The mean attrition rate in the pilot trials was 21.1% (standard deviation = 16.0%; median = 23.5%, interquartile range = 5.9%, 31.9%) and the mean attrition rate in the main trials was 16.8% (standard deviation = 11.66%; median = 10.8%, interquartile range = 5.8%, 28.0%). Figure 2 compares the attrition rates of the pilot trials with those of the main trials. The difference between the two rates is plotted on the vertical axis against the mean attrition rate for the two trials (main and pilot) on the horizontal axis. This graph shows minimal bias in the pilot trial when predicting the main trial attrition rate, and the average difference in the attrition rate between the pilot and the main trials is -4.4%. Thus, on average, the attrition difference (depicted by the red line) is 4.4% percentage points less in the main trial compared to the pilot trial. However, the standard deviation of the differences is large (16.3%). The 95% limits of agreement are -37.1% to 28.2%.

To investigate the predictability of the pilot studies for the main trial retention rates, a regression analysis was undertaken. With just the mean pilot retention in the model for a mean pilot retention rate of 21.1%, the model would predict the main trial retention to be 16.7%, which is 4.4% lower than in the pilot. The standard deviation of the difference from the same model is 16.2%. These results are consistent with the results above. If in the same model we had also the pilot sample size, the mean predicted retention (for an average-size pilot trial of 62 subjects) is 16.6%. The mean difference is thus similar to before at -4.5%. However, the standard deviation of this difference is reduced to 13%. The implication of this result is that the bigger the pilot study, the less variation there would be between the observed retention rates in the pilot and main trials.

Figure 3 compares the rates of patients who were eligible to be in the trials who were randomised for the pilot trial and main trial. For the limited data in the analysis, the graph shows minimal bias.

The mean percentage of eligible patients who were randomised for pilot trials was 50.4% (standard deviation = 22.4%; median = 48.9%, interquartile range = 30.3%, 61.3%) and for main trials the average was 65.2% (standard deviation = 23.2%; median = 60.9%, interquartile range = 45.7%, 92.5%). The mean of the differences between the pilot and the main trial is 14.9. This shows that the bias for pilot trials predicting the main trial rate of eligible patients who are randomised is small. However, again the standard deviation of this is large, 31.3%. The limits of agreement are -47.8% to 77.5% implying that the pilot could underestimate the rate of eligible patients who are randomised by as much as 77.5% and overestimate it by as much as 47.8%.

We repeated the regression analysis to see whether the size of the pilot improved predictability of the main trials randomisation rates. With just the pilot randomisation rates in the model for a mean pilot rate of 50.6%, the prediction was that the mean randomisation rate in the main trial would be 65.2%. This gives an average difference of 14.6% which is comparable to before. The standard deviation for this difference is a little bigger than before at 35.9%. Having the pilot sample size in the model reduces the mean predicted randomisation rate to 38.1% for the main trial and increases the standard deviation to 40.0%. On the face of it, this is a wide change in the prediction, but in actuality, it is confirming that the pilot trials in our study did not predict the randomisation rates in the main trials.

The trial that gave the largest difference (67.1%) between the rates from the pilot and the main trial noted in the *Health Technology Assessment* Journal.
reports that there were fewer eligible patients presenting than projected from the pilot trial, and trial staff took action to maximise the recruitment.

When making any assessment, it is not just the point estimate which is of importance but also the accompanying confidence interval which gives a range of plausible responses for the true effect. When we investigated the confidence intervals for the dropout, we found that for 10 of the 13 trials, the dropout rates in the main trials were contained within the 95% confidence intervals for the pilot estimate dropout rates. For the 3 where the main trial rates were not within the confidence interval, 2 had lower dropout rates compared to the pilot and 1 had higher. For the eligible to randomised rate of the 7 trials where we had the information, only 2 of the main trials had rates which fell within the 95% confidence interval for the pilot trial.

Secondary results

Table 1 shows the results from a secondary analysis investigating some features of a trial which could cause the differences in the attrition rates and the randomisation rates between the pilot and the main trials, including number of centres, number of participants and involvement of a Clinical Trials Unit in the main trials.

Figure 2. Bland–Altman plot looking at the difference in attrition rates between the pilot and main trial.

Figure 3. Bland–Altman plot looking at the difference in percentage of eligible patients randomised between the pilot and main trials.
and stated differences in length of follow-up period between the pilot and main trial.

For the mean difference between attrition rates, a negative value implies that the withdrawal rate was less in the main trials than in the pilot trial. For the difference between the rates of eligible patients randomised, a negative value depicts a situation where the conversion of eligible patients into randomised patients was better in the pilot trial than in the main trial.

Table 1 demonstrates that for trials where it was stated that a Clinical Trials Unit was involved, the dropout in the main trial was less than that in the pilot, and the randomisation rate was also greater in the main trial. Where there was no indication that a Clinical Trials Unit was involved in the main trial, attrition rates were higher and randomisation rates lower in the main trial compared with the pilot. These results must be viewed with caution though as no reporting of Clinical Trials Unit involvement does not necessarily mean that one was not involved.

### Discussion

We found in our investigation that in terms of predicting the proportion of participants who withdraw from a trial and the rate of eligible patients who were randomised, external pilot trials were consistent in terms of the mean difference and provided unbiased estimates for these values in the main trial. However, we also found evidence of large variation in the randomisation and attrition rates between pilot trials and their associated main trial.

The mean attrition rate in the main trials was 16.8% compared with 11% in a recent review of Health Technology Assessment-funded trials (with and without pilot trials). The mean randomisation rate in main trials in this study was 65.8% compared with 70% in the recent review.

There is a perception in the clinical research community that external pilot studies are likely to overestimate recruitment rates. Avery identified the cause of this as pilot centres being chosen because of where co-applicants work or have previous collaborations or through selecting safe or enthusiastic centres. Their concern was that recruitment rates in pilot studies may be difficult to replicate when widening out to include more centres in the full study.21 It is interesting that this study demonstrated no consistent overestimate of recruitment rates by the external pilot studies, neither did it find consistent underestimation of attrition rates.

The results need to be interpreted with caution as there were outliers in the analysis, which contributed to wide limits of agreement. Consideration also needs to be given to the fact that a study team is only likely to proceed to a main trial if the pilot trial gives sufficient confidence that the main trial is plausible. Pilot trials which do not provide sufficient confidence in the value of proceeding, for example, where randomisation or recruitment rates are very low or attrition is very high, are less likely to progress to a main trial. As only pilot studies which progressed to a main study have been included here, there may be some bias in studies available for assessment.

A limitation to this study is that the differences in the data between the pilot and the full trial could have been due to remedial action taken in the full trial to address problems with recruitment and attrition identified in the pilot, for example, changes to recruitment processes. Such changes would have been intended to improve the randomisation and attrition rates from the pilot to the full trial. This could explain the finding that involving a Clinical Trials Unit in the main trial was associated with an increase in randomisation rates and reduction in dropout rates in the main trial compared with the pilot. However, there was insufficient information available to assess whether this was the case or whether differences were due to non-systematic factors.

It is well established that a significant proportion of trials fail to meet their recruitment targets,22 and this has led to increased interest in undertaking external pilot trials. However, external pilot trials are not without cost in terms of time, effort, money and delay in

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean percentage difference between the attrition rates (standard deviation; n)</th>
<th>Mean % difference between the randomised/eligible (standard deviation; n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greater than 5× more centres in the main trial</td>
<td>Yes –4.3 (17.6; 8)</td>
<td>20.5 (65.9; 2)</td>
</tr>
<tr>
<td></td>
<td>No –0.18 (14.6; 4)</td>
<td>10.3 (21.2; 4)</td>
</tr>
<tr>
<td>Stated involvement of a Clinical Trials Unit in the main trial</td>
<td>Yes –7.8 (16.0; 9)</td>
<td>25.6 (29.5; 5)</td>
</tr>
<tr>
<td></td>
<td>No 3.3 (16.4; 4)</td>
<td>–12.0 (20.9; 2)</td>
</tr>
<tr>
<td>Change in the length of follow-up in the main trial</td>
<td>Yes –5.7 (17.4; 11)</td>
<td>19.3 (31.9; 6)</td>
</tr>
<tr>
<td></td>
<td>No 2.8 (6.6; 2)</td>
<td>–11.5 (N/A; 1)</td>
</tr>
<tr>
<td>Greater than 10× more patients in the main trial</td>
<td>Yes –12.3 (17.8; 7)</td>
<td>33.6 (29.4; 3)</td>
</tr>
<tr>
<td></td>
<td>No 4.8 (8.4; 6)</td>
<td>0.77 (27.8; 4)</td>
</tr>
</tbody>
</table>
delivering the definitive results about whether a potentially valuable health intervention is effective or not.

External pilot trials have demonstrated clear value where there are many unknown factors including whether the intervention or outcome measures are acceptable to participants and where uncertainties exist around the implementation of the intervention.\(^{5,23–26}\) In such cases, an external pilot trial is likely to be more appropriate than an internal pilot. However, based on the findings of this study, we would recommend that where the key unknown factors are randomisation and retention rates alone, investigators should consider using an internal pilot rather than an external pilot. An internal pilot as the first part of the main trial has a number of advantages over an external pilot.\(^{27}\) Undertaking an external pilot trial can delay the main study by 3–4 years by the time the pilot study has completed, reported and further funding has been applied for. In contrast, if randomisation and attrition rates within an internal pilot are on target and progression review is included within the trial protocol,\(^{8,28}\) progression to the full trial can be rapid. The use of internal pilot trials is reported as having the potential to optimise trial design, and recommendations have been produced on setting progression criteria to inform the decision to continue to the full study. Avery suggests that rather than employing dichotomous stop/go criteria, investigators should consider a traffic light red/amber/green system which identifies the level of risk to proceeding.\(^{21}\) In addition, data from the internal pilot are used in the main analysis which is not the case for external pilots and is therefore a more efficient use of data. This also means an internal pilot requires fewer total participants than using an external pilot which may be crucial where the availability of eligible patients is limited, for example, in rare conditions. Recognition of the importance of using the pilot data in the final analysis has lead to the development of a checklist to guide investigators about when this can be done appropriately.\(^{27}\)

It should be noted that the results in this article are based on a relatively small sample size. After exclusion of ineligible studies and those for which insufficient data were available, only 16 trials were included for analysis. Repeating this audit with more cases may produce more accurate estimates. There may be value in repeating this including subsequent reports.

**Conclusion**

This study demonstrates evidence of large variation in the randomisation and attrition rates between external pilot trials and their associated main trial in the published literature. We recommend that in circumstances where the intervention is developed and stable and the appropriate outcome measures are established, that is, where the main unknowns are recruitment rate and attrition, an internal pilot should be considered over an external pilot. Even if an external pilot had been undertaken, we would also recommend an internal pilot as the early phase of the main trial. The focus of the internal pilot would be to monitor randomisation and attrition rates within the early stages and then to modify processes to facilitate recruitment and retention or increase the numbers of participants recruited if necessary.

**Acknowledgements**

E.P. undertook this work as a Wellcome Summer Intern at The University of Sheffield.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**ORCID iD**

Cindy L Cooper https://orcid.org/0000-0002-2995-5447

**References**


