CLINICAL TRIALS

SESSION 3

THE SIZE OF TRIALS

1 Objectives

The aim of this session is to understand the importance of recruiting sufficient numbers into a trial and be able to calculate the required sample size when designing a trial.

By the end of this session students will be able to:

- Understand the importance of the size of a trial to provide reliable results
- Explain the key factors in determining the required size for a particular trial
- Discuss the importance of power for sample size calculations
- Calculate a sample size for a trial based on two percentages (or proportions) or means
- Adjust sample sizes to allow for factors such as losses to follow-up
- Discuss the problems of trials which are too small

2 Introduction

In previous sessions the focus has been on the importance of randomisation, control groups and blinding (where appropriate) to avoid bias in the results from a clinical trial. A further key issue in the design of a clinical trial is ensuring there are sufficient numbers of patients.

The aim of a trial should be to provide reliable evidence of the effectiveness and safety of a treatment. To achieve this, enough patients are needed in our trial to give a good chance of detecting a clinically important treatment difference if such a difference exists, while being able to reasonably conclude that no such difference exists if our results do not show it.

Given the importance of sample size there are several approaches that can be taken to answer the question "How many patients do we need?". For example:

(1) **Statistical/scientific approach.** This approach is concerned with determining how many patients are needed to get firm evidence of a treatment difference if it exists, and to estimate any difference precisely (i.e. within a sufficiently small confidence interval (CI)).

(2) **Economic/pragmatic approach.** In reality, other constraints will impact on the ability to recruit the numbers required from a purely statistical approach. For example, how many patients are available, and how much time, effort and cost is involved?

(3) **Ethical approach.** In many clinical trials decisions need to be made as to how long it is ethical to continue a trial. The question that might be asked here is "How soon can we stop a trial to avoid some patients getting inferior treatment?". This will be dealt with in detail in session 4: "Data Monitoring".

(4) **Credibility.** If a trial is very small, it's results may not be seen as providing believable evidence.

In determining the required size of a trial all of the above approaches are usually considered, but this session is mainly concerned with (1) the statistical approach. The compromises to be made between purely scientific objectives and practical constraints, and the implications of trials which are too small will also be considered.

In order to appreciate the impact of sample size for interpreting trial results consider the table below showing the results from two placebo-controlled trials assessing the impact of streptokinase for reducing mortality in patients who have experienced a myocardial infarction (MI).

Table 1: Results from the 1st Australian and ISIS-2 trials for reducing mortality post-MI

	Streptokinase	Placebo	P-value
1st Australian	**n=264**	**n=253**	
Deaths	26 (9.8%)	32 (12.6%)	p=0.32
	Risk ratio = 0.78 (95% CI 0.48 to 1.27)		
ISIS-2	**n=8592**	**n=8595**	
Deaths	791 (9.2%)	1029 (12.0%)	p<0.0001
	Risk ratio = 0.77 (95% CI 0.70 to 0.84)		

The estimated treatment effect is very similar in the two trials. Further, the proportions of patients who die in the two trials are also very similar. However, while the ISIS-2 trial provides very strong evidence of a treatment effect with p<0.0001 and a tight confidence interval, the 1st Australian trial is unconvincing with p=0.32 and a wide confidence interval. Even though more than 500 patients were recruited, the trial was too small in these circumstances.

3 Key factors in determining the required sample size

The previous section indicated the impact the size of a trial can have on the interpretation of results. In particular, trials which are too small lead to results which are inconclusive. There are a number of key factors that determine the required size of a trial and in particular there are five key questions that need to be asked. In order to illustrate these factors, the UK PACE study in fitting a dual chamber vs. single chamber pacemaker for patients with atrioventricular block is used.

Note that while questions (2) and (5) are more technical these are fairly standard whereas questions (3) and (4) are perhaps the most important and difficult to address. The answers to these two questions will have the most impact on the required size as will be shown later in the session.

(1) What is the principal outcome measure of the trial?

In designing a trial, a suitable primary outcome measure is chosen to reflect the main purpose of conducting the trial. The sample size requirements need to be based on this outcome measure in order to reliably address the prime purpose of the trial.

For example, in the UK PACE trial the purpose was to assess whether dual chamber pacing could reduce mortality compared to the more usual single chamber by comparing mortality at four years.

(2) How will the data be analysed to detect a treatment difference?

Some thought should be given to the appropriate analysis for the outcome being measured. For example, is the outcome the occurrence of some event (e.g. death) or a continuous outcome (e.g. blood pressure). In addition, the level of statistical significance that will be accepted as indicating evidence of a treatment difference should also be considered. This is usually set at 5% i.e. if the final p-value <0.05 then this would be accepted as there being some evidence for a real treatment effect.

For, example, the simplest analysis for the UK PACE trial is a comparison between the groups of the percentage of patients who die within four years of randomisation using a χ^2 test. A 5% significance level was used.

(3) What results are expected in the control group?

Clearly the results that will be observed in the control arm will not be known. However, there may well be information from previous trials or observational studies as to what might be expected.

For example, in the UK PACE trial it was expected that 24% of patients would die by four years in the single chamber group.

(4) How small a treatment difference, if it exists, is important to detect?

This is perhaps the key decision that needs to be taken in determining sample size. If a very large treatment difference exists, then this can be detected with relatively few numbers. For example, in the UK PACE trial if the dual chamber pacing reduced mortality by three-quarters (i.e. from 24% to 6%) then relatively few patients are needed (in fact less than 200). Of course, such a difference is unlikely and, more importantly, a difference much smaller than this would be clinically very important to detect if it were true.

It could be argued that any difference would be important to detect but this is unrealistic as huge trials would be required (an example of this will be given later). What needs to be decided upon is the smallest clinically relevant difference that would be important to detect if it were true.

For example, in the UK PACE trial it was decided that a one quarter reduction (i.e. from 24% to 18%) would be clinically important to detect.

(5) What degree of certainty is needed to be able to detect the treatment difference in (4)?

In a clinical trial conclusions are based on <u>observed</u> treatment effects rather than the <u>true</u>, unknown effects. These observed effects could be larger (i.e. further away from the null hypothesis of no effect) <u>but</u> could also be smaller (i.e. closer to the null hypothesis of no effect) than the true effects. The trials need to give a good chance of providing evidence for a treatment difference even if the observed effects are smaller than the true effects. The chances of this can be improved by increasing the numbers in our trial.

For example, the investigators in the UK PACE trial wanted to be 90% sure of picking up a statistically significant (at $p<0.05$) result if a true difference of 24% vs. 18% existed.

In total, the UK PACE trial needed to recruit approximately 2000 patients to satisfy the above criteria as will be seen in later sections.

4 Type I and type II errors

Note the true population values in each treatment group are not known but in calculating the sample size some estimate of these values are needed based on the alternative hypothesis. In the following sections the observed and anticipated percentages will be distinguished using the following notation:

p_1 = the observed percentage in those on standard treatment
p_2 = the observed percentage in those on 'new' treatment

i.e. $p_1 - p_2$ is the observed treatment effect (observed absolute percentage reduction on 'new' treatment)

π_1 = the anticipated percentage in those on standard treatment
π_2 = the anticipated percentage in those on 'new' treatment

i.e. $\pi_1 - \pi_2$ is the true difference which has been decided it is important to detect

Further notations are as follows:

α = significance level for detecting a difference (usually set at 0.05 (or 5%))
$1-\beta$ = degree of certainty that a true difference of $\pi_1 - \pi_2$ would be detected i.e. the power (often aim for $1- \beta = 0.9$ (or 90% power))

Consider the table below. Ideally, the results of the trial should fall either:

(i) in the top left hand cell of the table, i.e. if, in truth, no difference exists (the null hypothesis), then no evidence for a difference is observed in the trial, or
(ii) in the bottom right hand cell of the table, i.e. if, in truth, a difference of $\pi_1 - \pi_2$ does exist (the alternative hypothesis), then evidence for a difference is observed in the trial.

These are marked by 'X' in the table.

Table 2: Observed trial results compared to the 'truth' of (i) no difference (ii) a true $\pi_1 - \pi_2$ difference

	Truth	
	No difference exists	**Difference $\pi_1 - \pi_2$ exists**
No significant difference observed	**X**	
Significant difference observed		**X**

However, there are two particular types of errors that can be made in interpreting trial results. These are referred to as type I and type II errors.

- **Type I error.** A type I error is when a treatment difference is claimed based on a statistically significant observed result when in truth no such difference exists i.e. a false positive result.

- **Type II error.** A type II error is when in truth there exists a difference of $\pi_1 - \pi_2$ but the observed results fail to reach statistical significance, i.e. a false negative result.

Alternative ways of describing α and β are as follows and represented in the table below:

α is the risk of a type I error
β is the risk of a type II error. As mentioned previously, 1- β is termed statistical power.

Table 3: Risk of type I and type II errors based on (i) the null hypothesis being true (ii) the alternative hypothesis being true

	Truth	
	(i) No difference exists	**(ii) Difference $\pi_1 - \pi_2$ exists**
No significant difference observed	1-α	β
Significant difference observed	α	1-β

It needs to be decided in advance what levels will be set for α and 1-β and in the following sections the influences of the choice of these on the required sample size will be demonstrated.

5 Comparing two percentages (or proportions) – power calculation

5.1 Sample size formula using 5% significance and 90% power

Up to now the key factors needed in calculating a sample size have been considered and the concept of power in this process has been introduced.

In the UK PACE example the intended power was set at 90%. This means that if the true underlying percentages are 24% and 18%, there is a 10% probability that the observed results will give $p>0.05$. This will occur if the observed difference is less than the true difference to the extent that it is more compatible with the null hypothesis of no difference. Some risk of a false negative result has to be accepted (otherwise an infinitely large study is needed).

Note that some risk of a false positive result also has to be accepted: it is quite possible that $p<0.05$ might be observed even if there is no true treatment difference.

Common choices for α (significance level) and 1-β (power) are 0.05 and 0.90 respectively as already stated. Based on these choices, the equation to calculate the required sample size in each group for the comparison of two proportions is given below:

$$n = 10.5 \times \frac{[\pi_1 \times (100 - \pi_1) + \pi_2 \times (100 - \pi_2)]}{(\pi_1 - \pi_2)^2} \text{ in each group}$$

There are two key notes that you need to be aware of given below.

Note 1: The above equation calculates the required number of patients required in each group. Therefore for a trial with two treatments (with equal numbers planned for each group), the required number of patients in total is double this amount.

Note 2: Up to this point the outcomes have been considered as percentages. If proportions are preferred in the calculation of the sample size then the above equation must be adapted by substituting '1' in place of '100' throughout.

By using the sample size formula above, the total sample size required for the UK PACE trial is 1926 (963 in each group). In practice the UK PACE trial aimed to recruit 2000 patients.

Note that the aim of a sample size calculation is to approximate the number of patients required rather than provide exact figures. There are several assumptions that need to be made including the anticipated true values which of course cannot be known.

5.2 More general formula for any combination of significance and power

The formula given for the comparison for two proportions is based on a type I error of α=0.05 (equating to a significance level of 5%) and a type II error of β=0.1 (equating to 90% power). However, a more general formula for different values of α and β is needed.

A more general formula is as follows:

$$n = f(\alpha,\beta) \times \frac{[\pi_1 \times (100 - \pi_1) + \pi_2 \times (100 - \pi_2)]}{(\pi_1 - \pi_2)^2} \text{ in each group}$$

where f(α,β) is a function of α and β. The table below provides the value of f(α,β) for different values of α and β.

Table 4: Values of f(α,β) for different levels of α and β

		β			
		0.05 (95% power)	0.1 (90% power)	0.2 (80% power)	0.5 (50% power)
α	0.05	13.0	10.5	7.85	3.84
	0.01	17.8	14.9	11.7	6.63

Note that the value of $f(\alpha,\beta)$ can be calculated for any level of α and β and this is given by the formula below (you will need to understand what $z_{1-\alpha/2}$ and $z_{1-\beta}$ represent: the z-distribution has a mean of 0 with a standard deviation of 1).

$f(\alpha,\beta) = (z_{1-\alpha/2} + z_{1-\beta})^2$

For example, for α=0.05 and β=0.1: $f(\alpha,\beta) = (1.96 + 1.28)^2 = 10.5$
For example, for α=0.05 and β=0.2: $f(\alpha,\beta) = (1.96 + 0.84)^2 = 7.85$

6 Impact of choice of scientific criteria on sample size

The key factors required in order to calculate the required sample size and needed for a trial comparing two percentages have been introduced. Assuming the primary outcome variable has been identified these are (1) the expected percentage or risk in the control arm (2) the smallest clinically important difference that needs detecting (3) the level of significance, and (4) the power to detect the difference specified in (2) if such a true difference exists.

While a set significance (usually 5%) and power (usually 90% or 80%) can be specified perhaps the most important assumptions that need to be considered are for the unknown underlying risk in the control group and the smallest clinically important decision that needed to be detected.

Some flexibility is usually needed in assessing sample size requirements and it is a useful exercise to calculate the numbers needed for several different scenarios. In this section the impact of altering each of the four criteria given above on the sample size for comparing two percentages will be demonstrated using the UK PACE trial. Note that previous calculations showed 1926 patients were required based on the following criteria:

Expected mortality in the single chamber arm (π_1) = 24%
Expected mortality in the dual chamber arm (π_2) = 18% (i.e. one quarter reduction)
α=0.05, β=0.1 (power = 90%) i.e. $f(\alpha,\beta) = 10.5$

(1) Impact of reducing the difference to detect.

Suppose one quarter reduction is considered unrealistically high and decide an absolute difference of 3% from 24% would clinically important to detect (i.e. $\pi_1 - \pi_2 = 3\%$). The total sample size required is 8128. Halving the absolute difference to detect for a given expected percentage outcome in the control arm resulted in more than quadrupling the sample size.

(2) Impact of overestimating the event rate in each group.

Suppose the true mortality is half that expected in both groups i.e. $\pi_1 = 12\%$, $\pi_2 = 9\%$. Note that this is still a one quarter reduction in mortality. The total sample size required is 4376, more than double the initial estimate.

Note however that this assumes the relative effect remains the same i.e. a quarter reduction. For a given absolute difference to detect, the total number decreases as the event rate decreases (for event rates below 50%). For example, suppose it is required to detect a difference of 12% vs. 6% then the total sample size would be 946 but this represents a halving in mortality.

(3) Impact of reducing the significance level (i.e. reducing risk of a type I error).

Suppose more stringent evidence for a treatment difference is required and the level of significance is set to be 1% i.e. require $p<0.01$ to claim a treatment difference. $f(\alpha,\beta)$ now becomes 14.9 and the total sample size needs to be increased to 2732.

(4) Impact of increasing the power (i.e. the risk of a type II error).

Suppose it is required for the study to give more chance of finding a treatment difference and 95% power is used. $f(\alpha,\beta)$ now becomes 13.0 and the total sample size needs to be increased to 2384.

In practice, these calculations are usually done using computer packages, such as Stata, R or Epi-Info. The number of patients which different packages arrive at may differ slightly from those above as slightly different approximations are used.

In deciding upon the required size of a trial the numbers are often disappointingly large in respect of investigators expectations and available patients. For example reducing the absolute required difference to detect by half meant increasing the sample size more than fourfold.

Sample sizes entail some judgement, and some compromise between using stringent statistical criteria and what can practically be achieved is required. However, the choices to be made need to be realistic.

7 Comparing two means

In some clinical trials the primary endpoint is a continuous outcome, e.g. kidney functioning as measured by glomerular filtration rate or systolic blood pressure, rather than the occurrence of some binary event, e.g. mortality. The principles behind calculating the required sample size when the outcome is a comparison of two means are the same as when comparing two percentages.

7.1 Sample size formula

However, one additional factor that needs to be considered is the standard deviation of the outcome in each treatment group. In this section it is assumed that the standard deviation is the same for each treatment group although alternative formulae are available when this is clearly not the case. The standard deviation can be estimated from previous observational studies or in some cases a pilot study will be required. However, as when considering two percentages, it is advisable to calculate the required sample size for a number of different scenarios.

To distinguish between means and percentages μ_1 and μ_2 will be used to refer to the anticipated means in the two treatment groups.

As for the comparison of percentages, there are certain assumptions and decisions that need to be made. These are as follows:

μ_1 = anticipated mean response on the standard treatment
μ_2 = anticipated mean response on the alternative treatment
σ = standard deviation of response
(assumed the same on both treatments)
α = risk of a type I error (significance level)
β = risk of a type II error (1-β = power)

Once these values have been decided upon the required number per group can be calculated as follows:

$$n = f(\alpha,\beta) \times \frac{2\sigma^2}{(\mu_1 - \mu_2)^2} \quad \text{in each group}$$

Note that the actual values of μ_1 and μ_2 are less important than the treatment difference to be detected. Suppose δ is the anticipated treatment difference then the above formula can be written as:

$$n = f(\alpha,\beta) \times \frac{2\sigma^2}{\delta^2} \quad \text{in each group}$$

7.2 Example in kidney transplantation

The REPAIR trial was designed to assess the impact of remote ischaemic preconditioning (RIPC) using repeated inflations of a blood pressure cuff on kidney function in patients undergoing kidney transplantation. The outcome of interest is a glomerular filtration rate at 1 year.

The assumptions are as follows:

δ = 4.73 ml/min/1.73m^2 superiority with RIPC (= μ_1 - μ_2 equivalent to a 10% improvement)
σ = 13.9 ml/min/1.73m^2 from a previous study
α = 0.05
β = 0.2 i.e. power = 0.8 or 80%

The number required in each group is $7.85 \times \frac{2 \times 13.9^2}{4.73^2}$ = 136 i.e. 272 in total.

To allow for losses to follow-up the trial requires approximately 320 patients.

It has been seen already when estimating the sample size for comparing percentages that as the absolute difference to detect decreases then the number required increases. In addition, as the significance level decreases or as the power increases this also leads to an increase in the numbers required (by increasing f(α,β)). The same principles apply when comparing means as can be seen from the sample size formula.

The impact of the standard deviation (σ) also needs to be considered when comparing means. As the standard deviation increases so the number required in each group also increases. For example, suppose that the standard deviation was estimated to be 14.9 instead of 13.9 in the example given previously.

The number required in each group becomes $7.85 \times \frac{2 \times 14.9^2}{4.73^2}$ = 156 i.e. 312 in total.

Conversely, as the estimated standard deviation decreases so does the number of patients required. For example, suppose that the standard deviation was estimated to be 12.9%.

The number required in each group would be $7.85 \times \frac{2 \times 12.9^2}{4.73^2}$ = 117 i.e. 234 in total.

7.3 Potential of adjustment in the analysis on the sample size for comparing two means

It has already been stated that sample sizes are often disappointingly large even before making any adjustments. However, there is one scenario where it may be possible to reduce the required size if the primary outcome is a continuous measure which is also measured at baseline.

A detailed discussion is beyond the scope of this session and requires an understanding of an analysis technique call 'analysis of co-variance' (ANCOVA). Basically this method takes advantage of the fact that the outcome measure for individual patients is likely to be strongly correlated to the baseline value. For example, in the REPAIR trial an estimate of the glomerular filtration rate of the transplanted kidney from the living donor is available and is expected to be correlated to the kidney function at 1 year following transplantation. This information was used in the analysis of the data.

8 Adjustments to sample size calculations

In the calculations so far it has been assumed that full information is available on everyone (i.e. no losses to follow-up) and also that every patient receives the treatment they are allocated. In reality this is rarely achieved and some adjustments to the sample size needs to be made to account for this.

In this section the implications of these two scenarios will briefly be considered including what adjustments can be made. Note that these adjustments make certain assumptions but provide useful guidance to the number required. In addition, the situation where more patients will be randomised to one group than the other will also be discussed.

8.1 Losses to follow-up

If patients are lost to follow-up for some endpoint then this reduces the effective sample size. For example, if 2% of patients in the UK PACE trial were expected to be lost to follow-up for mortality then the effective sample size of 1926 would be reduced by 2% (although in the UK, the Office of National Statistics can notify the trial when a patient dies).

Suppose the proportion expected to be lost to follow-up is 'Q' then, in order to adjust for this, the total sample size needs to be multiplied by 1/(1-Q).

For example, the UK PACE trial would need to be increased to 1926 x (1/0.98) = 1966.

8.2 Patients who receive the alternative treatment

In many trials some patients will receive the alternative treatment. The impact of this is to make the groups more similar and therefore potentially reduce the observed treatment effect and therefore power is lost. In order to allow for this a correction formula is given below.

Let Q_1 = proportion in group 1 getting the treatment for group 2
Let Q_2 = proportion in group 2 getting the treatment for group 1

Correction formula: multiply the required sample size by $\frac{1}{(1-Q_1-Q_2)^2}$

For example, suppose in the UK PACE trial:

1% of patients randomised to the single chamber group received the dual chamber treatment
2% of patients randomised to the dual chamber group received the single chamber treatment

The required size = $\frac{1926}{(1-0.01-0.02)^2}$ = 2048

Alternatively the treatment effect to be detected might already account for the impact of patients receiving the alternative treatment.

Clearly, it is desirable to limit the number of patients who switch treatments and if the percentages become too large then interpretation becomes increasingly difficult. This is particularly important in non-inferiority trials (see session 7) where the aim is to show treatments are similar.

8.3 Patients who stop treatment

One simple adaptation to the formula introduced on the previous page relates to placebo controlled trials where a proportion, Q, on the active therapy stop taking their medication. For the purposes of the sample size adjustment, the proportion who stop their placebo medication is less important (as the randomised treatment contains no active ingredient expected to affect outcome).

The impact will be to make the active treatment group more similar to the placebo group. The correct formula used is to multiply the sample size by $1/(1-Q)^2$.

As mentioned, certain assumptions are made (for example, that those stopping active therapy do so early in the trial) but the adjustments do provide useful guidance.

8.4 Other adjustments

A trial which randomises patients to two treatment groups of equal size is usually the most efficient design in terms of total numbers required. However, it is sometimes the case that more patients are randomised to one group than the other.

Suppose patients are to be randomised to two groups in the ratio r:1. Then the total sample size required for equal numbers needs to be increased as follows:

Multiply the sample size by $(r+1)^2/4r$

For example, in the UK PACE trial, suppose the investigators wanted to randomise patients in the ratio 2:1 to the dual chamber treatment. The total sample size would need to be increased as follows:

$1926 \times (2+1)^2/8 = 2166.$

Since the required ratio is 2:1, the number required in the dual chamber group would be $(2166 \times {}^2/_3) = 1444$, and the number required in the single chamber group 722.

9 Trials which are too small and other issues

There are a number of important problems associated with trials which are too small. Such trials will be underpowered to detect realistic treatment differences, i.e. increasing the chance of a false negative result. In addition, the observed treatment effects could be far from the true values (whether the null hypothesis or some alternative hypothesis is true). These observed effects will also be measured imprecisely.

Unfortunately many trials are too small and as a result will be unable to reliably answer clinically meaningful questions. In addition they use valuable resources and it could be considered unethical to expose patients in a trial which has little chance of being able to address a clinically meaningful question.

In addition the results can be misleading and potentially lead to publication bias. Suppose a number of small trials are conducted addressing a similar question. It is likely that some small trials will produce results showing a large, statistically significant effect which might be a chance finding or at best an exaggeration of the truth. The trials showing a significant effect are more likely to be written up and accepted for publication.

Trials which are too small are a hindrance to medical progress but finding adequate numbers of patients is not easy. In addition investigators may be over-optimistic in the ability of a trial to recruit eligible patients and recruitment often takes longer than expected. Therefore a realistic appraisal of resources and potential patient accrual is needed.

However, there are a number of ways in which numbers can be potentially increased.

- The trial may need to be conducted as a multi-centre collaboration. Patient accrual will be quicker and more eligible patients available.
- Avoid being too restrictive in respect of the patient entry criteria. This will also have the advantage of making the results more broadly applicable.
- Trials of more than two treatments require more patients so some thought is necessary as to any benefits of including more than two treatment groups. One possible approach would be through the use of a factorial design (see session 8).

In this session the main techniques have been covered to calculate sample size when comparing percentages and means for a clinical trial where each individual is randomised. Similar considerations are needed when estimating the required sample size for other measures of effects, e.g. rates which account for varying follow-up times, however a detailed discussion of these formulae is beyond the scope of this module but included in the articles in the references.

In addition, other sample size considerations are required depending on the trial design, e.g. when the aim is to show one treatment is equivalent (or non-inferior) to another treatment or when groups (or clusters) are randomised rather than individuals. These will be discussed in more detail later in the module (sessions 7 and 8).

10 Summary

In this session the importance of estimating the required number of patients needed to produce a clinically meaningful result has been discussed. Sample size formulae have been introduced for comparing percentages (or proportions) and means together with adjustments which can be made under different conditions. More complex formulae exist but give similar results.

Trials which are too small are highly unlikely to produce conclusive results particularly for realistic moderate effects. Such trials could be considered unethical by exposing patients to a trial which has little chance of being able to address a clinically meaningful question.

It is helpful to produce tabulations under different scenarios. Realistic evaluation is needed to as to whether a trial is feasible given the resources both in terms of costs and patient availability.