5.3: Size to give adequate precision

This section describes how the trial size is determined if the aim is to obtain an estimate of the outcome of an intervention with a specified level of precision. The simplest case to consider is where just two groups of about the same size are to be compared (for example, the outcome of an intervention compared with that of a control group, or the comparison of outcomes of two interventions). More complex designs are discussed in Section 5. The methodology varies according to the type of outcome measure; the comparison of proportions, incidence rates, and means are considered in Sections 3.1 to 3.3.

3.1 Comparison of proportions

In this section, outcomes are considered that are binary (yes or no) variables. This includes cumulative incidence or risk, for example, the proportion of children experiencing at least one episode of clinical malaria during the follow-up period. It also includes examination of the prevalence of some characteristic, for example, the presence of a palpable spleen in a survey conducted at the end of the trial.

Suppose the true proportions in groups 1 and 2 are \( p_1 \) and \( p_2 \), respectively, giving a risk ratio (relative risk) of \( R = p_1/p_2 \). The approximate 95% CI for \( R \) extends from \( R/f \) to \( Rf \) where, in this case, the factor \( f \) is given by:

\[
f = \exp\left\{1.96\sqrt{\frac{1-p_1}{np_1}+\frac{1-p_2}{np_2}}\right\}
\]

where \( n \) is the number of children in each group, and \( f \) is commonly called the error factor. The required value of \( f \) is chosen, and rough estimates are made of the values of \( p_2 \) and \( R \) to enable the number required in each group \( n \) to be calculated as:

\[
n = \frac{(1.96/\log _ef)^2+((R+1)/(Rp_2)-2)}
\]
where \(\log_e f\) is the natural logarithm of \(f\).

For example, in the mosquito-net trial, one of the outcomes of interest is the prevalence of splenomegaly (the proportion of children with enlarged spleens) at the end of the trial. Prior data from the trial area suggest that, in the control group, a prevalence of approximately 40% would be expected. Suppose the intervention is expected to roughly halve the prevalence, so that \(\log(R=0.5)\) and an estimate of \(R\) is wanted to within about \(\pm 0.15\) This suggests setting \(f\) to about 1.3 (because then the upper 95% confidence limit on \(R\) is \(\log(R=-0.5\times1.3=0.65)\) which is 0.15 above \(\log(R=0.5)\)), and thus \(n=(1.96/\log_e1.3)^2 (1.5/(0.5\times0.4)-2)=307\) so that around 300 children would need to be studied in each group.

3.2 Comparison of incidence rates

Suppose a comparison of two groups is required, with respect to the rate of occurrence of some defined event over the trial period. Suppose the true incidence rates are \(r_1\) and \(r_2\) in groups 1 and 2, respectively, where each rate represents the number of events per person-year of observation. The rate ratio \(R\) (sometimes called incorrectly the relative risk, instead of the relative rate) of the incidence rate in group 1, compared to the incidence rate in group 2, is given by \(R=r_1/r_2\) (see Chapter 21, Section 5 for methods of analysis for the comparison of rates). If the total follow-up time for those in each group is \(y\) years (for example, \(y\) persons are each followed for 1 year, or \(y/2\) are each followed for 2 years), each group is said to experience \(y\) person-years of observation. The expected numbers of events in the two groups will be \(e_1=yr_1\) and \(e_2=yr_2\) respectively. When the results are analysed, the approximate 95% CI for \(R\) is expected to extend from \(R/f\) to \(Rf\) where:

\[f=\exp\{1.96\sqrt{1/e_1+1/e_2}\}\]

To decide on the necessary size of the trial, make a rough estimate of the likely value of \(R\), select the precision that is required by specifying a value for \(f\), the error factor, and calculate:

\[e_2=(1.96/\log_e f)^2 [(R+1)/R]\]

The trial size is then fixed so that the expected number of events in group 2 during the trial period is equal to the calculated value \(e_2\) The expected number of events in group 1 will be \(Re_2\)

It should be noted that these methods are only appropriate in the situation where each individual can experience only one event during the trial period or where the number of individuals experiencing multiple events is very small. If most individuals experience at least one event and many experience two or more, it is preferable to define a quantitative outcome for each individual, representing the number of events experienced during the trial period, and to use the methods described in Section 3.3.

Example: in the mosquito-net trial, suppose the trial groups are to consist of children aged 0–4 years and that the death rate associated with malaria in the trial area for that age group is estimated to be roughly 10 per 1000 child-years. If group 1 is the intervention group (treated bed-nets) and group 2 is the control group (no protection), \(R\) represents the ratio of the intervention and control death rates. Suppose \(R\) is expected to be about 0.4, corresponding to a reduction in the death rate of 60%. Suppose also that \(f\) is selected to be equal to 1.25, so that the 95% CI for \(R\) is expected to extend from \((0.4/1.25=0.32)\) to \((0.4\times1.25=0.50)\). In other words, it is desired to estimate the protective efficacy to within about 10% of the true value (i.e. 50–70% around the estimated efficacy of 60%). Then:
\[e_2=[1.96/\log_e(1.25)]^2(1.4/0.4)=270.\]

To expect 270 deaths in the control group, it would be necessary to observe an estimated 27 000 child-years
(=270/(10/1000)). This could be achieved by following 54 000 children for 6 months, or 27 000 children for 1 year, or
13 500 for 2 years, and so on, assuming an expected death rate of ten per 1000 child-years in each of these scenarios.
The magnitude of the required trial size (27 000 child-years of observation in each group) illustrates that, when rare
events are being studied, very large samples are needed to obtain a precise estimate of the impact of an intervention.

### 3.3 Comparison of means

Quantitative outcomes may be analysed by comparing the means of the relevant variable in the intervention and control
groups. This could be the mean of the values recorded at a cross-sectional survey, for example, the mean weight of
children in the trial at the end of the trial. Alternatively, it could be the mean of the changes recorded between baseline
and follow-up surveys, for example, the mean change in weight (or weight velocity, i.e. the change in weight divided by
the time between the two measurements) among the children in the trial.

Suppose the true means in groups 1 and 2 are \(\mu_1\) and \(\mu_2\). These would generally be compared in terms of the
difference in the means, \(D=\mu_1−\mu_2\). The 95% CI for \(D\) is given by \(D±f\), where:
\[f=1.96\sqrt{\left(\frac{\sigma^2_1+\sigma^2_2}{n}\right)}\]
where \(\sigma_1\) and \(\sigma_2\) are the standard deviations of the outcome variable in the two groups.

An acceptable value of \(f\) is chosen; values of \(\sigma_1\) and \(\sigma_2\) are selected, and the required number in each group is
calculated as:
\[n=(1.96/f)^2(\sigma^2_1+\sigma^2_2)\]

An estimate of the standard deviation of the outcome variable is often available from other studies. It is usually
reasonable to assume that the standard deviation will be roughly similar in the two trial groups. If no other estimate is
available, a rough approximation can be obtained by taking one-quarter of the likely range of the variable.

Example: In the mosquito-net trial, another outcome of interest is the PCV, or haematocrit, measured in blood samples
taken from the children at the end of the trial. From previous data, the mean PCV in the control group is expected to be
about 33.0, with a standard deviation of about 5.0 (the normal range is about \(33±10\)), and it has been assumed that
the normal range covers four standard deviations (i.e. \(±2\)). An increase in mean PCV in the intervention group of
between 2.0 and 3.0 is expected, and it is required to estimate the difference \(D\) between the two groups to within about
0.5, so that \(f=0.5\). Assuming that the standard deviation is about 5.0 in both groups:
\[n=(1.96/0.5)^2(5.0^2+5.0^2)=768.\]