

# Sample size

## Sample size

Calculating the number of patients required to perform a satisfactory investigation is an important prerequisite to any study. An incorrect sample size is probably the most frequent reason for research being invalid. Often, surgical trials are marred by the possibility of error caused by the inadequate number of patients investigated. Type I error (false positive). Type II error (false negative). Benefit is missed when it was there to be found (false negative). Calculating the number of patients required in the study can overcome this bias. Unfortunately, it often reveals a larger number of patients is needed for the study than can possibly be obtained from available local resources. This usually means expanding enrolment by running a multicentre study – which has the added benefit of improving the external validity of findings. More patients will need to be randomised than the final sample size to take into account patients who die, drop out or are lost to follow-up; this is known as the attrition rate. A longer time from trial entry to primary outcome assessment will result in an increased attrition rate of participants. The following is an example calculation for a study to recruit patients into two groups. In order to calculate a sample size, it is now common practice to set the level of power for the study at 90% with a 5% significance level. This means that, if there is a difference between study groups, there is a 90% chance of detecting it. Based on previous studies, realistic expectations of differences between groups (i.e. the magnitude of the effect seen from utilising the intervention under study), according to the best available evidence, should be used to calculate the sample size. The formula below uses the results of a reduction in event rate from 30% to 20% (e.g. a new treatment expected to reduce the complication rate such as wound infection from 30% =  $r$  to 20% =  $s$ ). 
$$\frac{[r(100 - r) + s(100 - s)] \times 9 \times 2 \times (r - s)}{[30(100 - 30) + 20(100 - 20)]}$$
 e.g.  $\frac{9 \times 2 \times (30 - 20)}{[30(100 - 30) + 20(100 - 20)]} = 333$  needed in each group

Calculating the number of patients required to perform a satisfactory investigation is an important prerequisite to any study. An incorrect sample size is probably the most frequent reason for research being invalid. Often, surgical trials are marred by the possibility of error caused by the inadequate number of patients investigated. Type I error (false positive). Type II error (false negative). Benefit is missed when it was there to be found (false negative). Calculating the number of patients required in the study can overcome this bias. Unfortunately, it often reveals a larger number of patients is needed for the study than can possibly be obtained from available local resources. This usually means expanding enrolment by running a multicentre study – which has the added benefit of improving the external validity of findings. More patients will need to be randomised than the final sample size to take into account patients who die, drop out or are lost to follow-up; this is known as the attrition rate. A longer time from trial entry to primary outcome assessment will result in an increased attrition rate of participants. The following is an example calculation for a study to recruit patients into two groups. In order to calculate a sample size, it is now common practice to set the level of power for the study at 90% with a 5% significance level. This means that, if there is a difference between study groups, there is a 90% chance of detecting it. Based on previous studies, realistic expectations of differences between groups (i.e. the magnitude of the effect seen from

utilising the intervention under study), according to the best available evidence, should be used to calculate the sample size. The formula below uses the results of a reduction in event rate from 30% to 20% (e.g. a new treatment expected to reduce the complication rate such as wound infection from 30% =  $r$  to 20% =  $s$ ).  $[r(100 - r) + s(100 - s)] \times 9 \times 2 \times (r - s) [30(100 - 30) + 20(100 - 20)]$  e.g.  $9 \times 2 \times (30 - 20) = 333$  needed in each group Sample size

Calculating the number of patients required to perform a satisfactory investigation is an important prerequisite to any study. An incorrect sample size is probably the most frequent reason for research being invalid. Often, surgical trials are marred by the possibility of error caused by the inadequate number of patients investigated. none (false positive). /uni25CF Type II error. Benefit is missed when it was there to be found (false negative). Calculating the number of patients required in the study can overcome this bias. Unfortunately, it often reveals larger number of patients is needed for the study than can possibly be obtained from available local resources. This usually means expanding enrolment by running a multicentre study - which has the added benefit of improving the external validity of findings. More patients will need to be randomised than the final sample size to take into account patients who die, drop out or are lost to follow-up; this is known as the attrition rate. A longer time from trial entry to primary outcome assessment will result in an increased attrition rate of participants. The following is an example calculation for a study to recruit patients into two groups. In order to calculate a sample size, it is now common practice to set the level of power for the study at 90% with a 5% significance level. This means that, if there is a difference between study groups, there is a 90% chance of detecting it. Based on previous studies, realistic expectations of differences between groups (i.e. the magnitude of the effect seen from utilising the intervention under study), according to the best available evidence, should be used to calculate the sample size. The formula below uses the results of a reduction in event rate from 30% to 20% (e.g. a new treatment expected to reduce the complication rate such as wound infection from 30% =  $r$  to 20% =  $s$ ).  $[r(100 - r) + s(100 - s)] \times 9 \times 2 \times (r - s) [30(100 - 30) + 20(100 - 20)]$  e.g.  $9 \times 2 \times (30 - 20) = 333$  needed in each group

---

Revision #1

Created 2025-12-31 15:08:54 UTC by Omar Ayman

Updated 2025-12-31 15:08:54 UTC by Omar Ayman