



QIMR Berghofer
Medical Research Institute

Sample size calculation

Overview

- When do I need to do a sample size calculation?
- What a sample size number actually means
- What pieces of information do I need to do the calculation?

When do I need to do a sample size calculation?

- The aim of a sample size calculation is to determine the number of participants or experimental units needed to detect a clinically significant treatment effect
- When do I need to do one?
 - If you are testing a hypothesis or require confidence interval precision
 - Are worth consideration if you are addressing objectives
- Why?
 - Underpowered studies do not have a large enough sample to answer the research question
 - Overpowered studies are a waste of time, resources and money
 - Requirement for grants and ethics committee approval

What a sample size number actually means

- The number calculated by a sample size calculation is the number of patients or experimental units required for the study
- This is the minimal number required to detect the effect you have hypothesised as being clinically significant
- It does not account for any loss of patients or experimental units

Definitions

- **Type 1 error** (α) is the probability of incorrectly rejecting the null hypothesis, ie making a false positive conclusion
 - Common to fix at 5%
- **Power** ($1-\beta$) is the probability of correctly rejecting a null hypothesis, ie avoiding a false negative conclusion
 - Common to fix at 80%
- **Type 2 error** (β) is the probability of incorrectly retaining the null hypothesis, ie making a false negative conclusion
 - Common to fix at 20%
- **Size of the effect** is the smallest clinically significant difference between the groups

Hypothesis testing basics

H_0 = null hypothesis, eg treatment has no effect

H_A = alternative hypothesis, eg treatment has an effect

- A hypothesis test involves the calculation of a test statistic to determine whether to reject the null hypothesis, assuming the null hypothesis is true
- The test statistic belongs to a sampling distribution under the null hypothesis and allows a p-value to be calculated

What pieces of information do I need to do the calculation?

- Understand your research question and study design
- Sample size calculation requires:
 - Identifying the **primary** outcome variable
 - **Minimum difference** worth detecting
 - An **estimate of SD** for mean or **preliminary proportion**
 - Probability levels for **significance and power**

Do not have all the information required?

- Use previous studies as a guide
- Think about what effect size would be compelling (minimal difference that is clinically meaningful)
- Do a small pilot study

Summary

- Sample size calculations should be completed during the research proposal stage
- The sample size number that is calculated is the minimum number required
- Understand the research question and be able to identify the primary outcome variable
- Use the literature to obtain information on effect size and variability
- Sample size calculations require a lot of thought and are not as easy as they look, seek assistance from a statistician
- Choose values wisely for the sample size calculation

How to contact the Statistics Unit

- **Email:**

Statistical.Services@qimrberghofer.edu.au

- **Location:**

Level 12, Bancroft Building, QIMR Berghofer, 300 Herston Road, Herston

- **Website:**

<https://www.qimrberghofer.edu.au/our-research/scientific-services/qimr-berghofer-statistics-unit/>