

Key Points: Module 8 - Interpreting Effects

Statistical Significance:

Null hypothesis:

There is no difference between a new treatment and a control treatment or placebo

Null hypothesis is rejected if the difference in treatment arms is statistically significant

P-value:

The probability of observing the results, assuming that the null hypothesis is true

Provides an estimate of the probability that the observed difference between two groups is due to chance (given the assumption that there is in fact no difference between the two groups)

If this probability is high (> 0.05) then we conclude that the observed difference is due to chance

If this probability is low (< 0.05) then we conclude that the observed differences are unlikely due to chance

A p-value of 0.05 implies a 5% (1 in 20) probability that findings are due to chance, yielding a false positive result

Confidence intervals:

Define the range that contains the “true result”

Like $p < 0.05$, the 95% CI is an arbitrary but widely used threshold for determining statistical significance

In general, the CI is much tighter with larger sample sizes

Type I error

Null hypothesis is rejected when it should not have been

Concludes that a treatment has an effect, when in fact no treatment effect (or difference) exists

False positive result (“Convicting an Innocent Man”)

Increases markedly when multiple tests (outcomes) are performed

Type II error

Null hypothesis is not rejected when it should have been

Concludes that a treatment has no effect, when in fact a treatment effect (or difference) does exist

False negative result (“Letting an Guilty Man Go Free”)

Power of a Test is related to type-II error (Probability of rejecting the null hypothesis when the alternative hypothesis is true)

More likely with small sample sizes

Clinical Importance

Requires clinical judgment

Cannot be proven by math or statistics

Statistically significant results may not have clinical importance (e.g. when sample size is large)

Clinically important results may not be statistically significant (e.g. when sample size is small)

Judgment re: clinical importance depends on the outcome selected

Quantifying Treatment Effects

Absolute risk reduction (ARR): Treatment - Control

Relative risk ratio: Treatment / Control

Relative risk reduction: (Treatment - Control) / Control

Number needed to treat (NNT): $1 / \text{ARR}$

NNT (derived from ARR)

More intuitive for clinicians to understand

Magnitude of ARR is less misleading than RRR

Summary of Statistical Significance and Clinical Importance

Large sample + large effect: Statistically significant and clinically important

Large sample + small effect: Statistically significant; may not be clinically important

Small sample + large effect: May be clinically significant, but not powered to show statistical significance (potential for type II error)

Small sample + small effect: Difficult to draw any conclusions (potential for type I and type II error)