

Does “statistically significant” mean “important”?

How to interpret p-values, effect sizes and the real-world meaning of results.

*More questions.
Healthier choices.
Tomorrow.*



Statistical significance means that a result has crossed a prespecified threshold in a statistical test. It does not tell us how large the effect is, whether it is practically important, or whether it is clinically meaningful.

1 Is there a statistical signal?

p-value

- A p-value indicates how incompatible the observed data are with a specified statistical model; it is not the probability that the result is due to chance.
- A threshold such as $p < 0.05$ must be chosen before interpreting the test.
- It does not show effect size or practical importance.

2 How large is the difference?

Effect size

- Effect size describes the magnitude of a difference or association (e.g. mean difference, risk ratio or odds ratio).
- A statistically significant result can still have a very small effect with little practical importance.

3 How precise is the estimate?

Confidence interval (CI)

- A confidence interval shows a range of values compatible with the data and model at the stated confidence level.
- A narrower interval indicates greater precision.
- For common null values, an interval including 0 (difference) or 1 (ratio) is not statistically significant at the matching level.

Example: a very small difference can be statistically significant

Large study 12,000 per group

Group A		$240/12,000 = 2.0\%$
Group B		$300/12,000 = 2.5\%$

Relative increase: 25%
Difference: 0.5 percentage points
Two-sided two-proportion z-test: $p \approx 0.009$

With 12,000 participants per group, the 0.5-percentage-point difference crosses the chosen significance threshold. Its practical or clinical importance still depends on context.

The same rates in a smaller study

Smaller study 200 per group

Group A:	$4/200 = 2.0\%$
Group B:	$5/200 = 2.5\%$

Difference: 0.5 percentage points
Two-sided two-proportion z-test: $p \approx 0.74$

The same rates can cross the threshold in a large study and not in a small one.



Does the result matter?

This depends on:

- ✓ The size of the effect
- ✓ The outcome (e.g. mortality, quality of life, symptoms)
- ✓ The population (e.g. general or high-risk)
- ✓ The context and comparisons (e.g. with other interventions)
- ✓ The values and preferences of the people who will be affected



Key takeaway: “Statistically significant” does not mean “important”.

Do not ask only “Is it significant?”

Ask instead:

- ✓ How large is the effect?
- ✓ How confident are we in the estimate?
- ✓ Does it matter in this context?



*Critical questions today.
A healthier tomorrow.*

