

RESEARCH METHODS & REPORTING

How to obtain the confidence interval from a P value

Douglas G Altman *professor of statistics in medicine*¹, J Martin Bland *professor of health statistics*²

¹Centre for Statistics in Medicine, University of Oxford, Oxford OX2 6UD; ²Department of Health Sciences, University of York, Heslington, York YO10 5DD

Confidence intervals (CIs) are widely used in reporting statistical analyses of research data, and are usually considered to be more informative than P values from significance tests.^{1,2} Some published articles, however, report estimated effects and P values, but do not give CIs (a practice *BMJ* now strongly discourages). Here we show how to obtain the confidence interval when only the observed effect and the P value were reported.

The method is outlined in the box below in which we have distinguished two cases.

(a) Calculating the confidence interval for a difference

We consider first the analysis comparing two proportions or two means, such as in a randomised trial with a binary outcome or a measurement such as blood pressure.

For example, the abstract of a report of a randomised trial included the statement that “more patients in the zinc group than in the control group recovered by two days (49% v 32%, P=0.032).”⁵ The difference in proportions was $Est = 17$ percentage points, but what is the 95% confidence interval (CI)?

Following the steps in the box we calculate the CI as follows:

$$z = -0.862 + \sqrt{[0.743 - 2.404 \times \log(0.032)]} = 2.141;$$

$$SE = 17/2.141 = 7.940, \text{ so that } 1.96 \times SE = 15.56 \text{ percentage points};$$

$$95\% \text{ CI is } 17.0 - 15.56 \text{ to } 17.0 + 15.56, \text{ or } 1.4 \text{ to } 32.6 \text{ percentage points.}$$

(b) Calculating the confidence interval for a ratio (log transformation needed)

The calculation is trickier for ratio measures, such as risk ratio, odds ratio, and hazard ratio. We need to log transform the estimate and then reverse the procedure, as described in a previous Statistics Note.⁶

For example, the abstract of a report of a cohort study includes the statement that “In those with a [diastolic blood pressure] reading of 95-99 mm Hg the relative risk was 0.30 (P=0.034).”⁷ What is the confidence interval around 0.30?

Following the steps in the box we calculate the CI as follows:

$$z = -0.862 + \sqrt{[0.743 - 2.404 \times \log(0.034)]} = 2.117;$$

$$Est = \log(0.30) = -1.204;$$

$$SE = -1.204/2.117 = -0.569 \text{ but we ignore the minus sign, so } SE = 0.569, \text{ and } 1.96 \times SE = 1.115;$$

$$95\% \text{ CI on log scale} = -1.204 - 1.115 \text{ to } -1.204 + 1.115 = -2.319 \text{ to } -0.089;$$

$$95\% \text{ CI on natural scale} = \exp(-2.319) = 0.10 \text{ to } \exp(-0.089) = 0.91.$$

Hence the relative risk is estimated to be 0.30 with 95% CI 0.10 to 0.91.

Limitations of the method

The methods described can be applied in a wide range of settings, including the results from meta-analysis and regression analyses. The main context where they are not correct is in small samples where the outcome is continuous and the analysis has been done by a *t* test or analysis of variance, or the outcome is dichotomous and an exact method has been used for the confidence interval. However, even here the methods will be approximately correct in larger studies with, say, 60 patients or more.

P values presented as inequalities

Sometimes P values are very small and so are presented as $P < 0.0001$ or something similar. The above method can be applied for small P values, setting P equal to the value it is less than, but the z statistic will be too small, hence the standard error will be too large and the resulting CI will be too wide. This is not a problem so long as we remember that the estimate is better than the interval suggests.

When we are told that $P > 0.05$ or the difference is not significant, things are more difficult. If we apply the method described here, using $P = 0.05$, the confidence interval will be too narrow. We must remember that the estimate is even poorer than the confidence interval calculated would suggest.

Steps to obtain the confidence interval (CI) for an estimate of effect from the P value and the estimate (*Est*)**(a) CI for a difference**

- 1 calculate the test statistic for a normal distribution test, z , from P^2 : $z = -0.862 + \sqrt{[0.743 - 2.404 \times \log(P)]}$
- 2 calculate the standard error: $SE = Est/z$ (ignoring minus signs)
- 3 calculate the 95% CI: $Est - 1.96 \times SE$ to $Est + 1.96 \times SE$.

(b) CI for a ratio

For a ratio measure, such as a risk ratio, the above formulas should be used with the estimate *Est* on the log scale (eg, the log risk ratio). Step 3 gives a CI on the log scale; to derive the CI on the natural scale we need to exponentiate (antilog) *Est* and its CI.⁴

Notes

All P values are two sided.

All logarithms are natural (ie, to base e).⁴

For a 90% CI, we replace 1.96 by 1.65; for a 99% CI we use 2.57.

- 1 Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *BMJ* 1986;292:746-50.
- 2 Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010. Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869.
- 3 Lin J-T. Approximating the normal tail probability and its inverse for use on a pocket calculator. *Appl Stat* 1989;38:69-70.
- 4 Bland JM, Altman DG. Statistics Notes. Logarithms. *BMJ* 1996;312:700.
- 5 Roy SK, Hossain MJ, Khatun W, Chakraborty B, Chowdhury S, Begum A, et al. Zinc supplementation in children with cholera in Bangladesh: randomised controlled trial. *BMJ* 2008;336:266-8.
- 6 Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;326:219.
- 7 Lindblad U, Råstam L, Rydén L, Ranstam J, Isacson S-O, Berglund G. Control of blood pressure and risk of first acute myocardial infarction: Skaraborg hypertension project. *BMJ* 1994;308:681.

Cite this as: *BMJ* 2011;343:d2090

© BMJ Publishing Group Ltd 2011