

Collated responses from R-help on confidence intervals for risk ratios

Michael E Dewey*

November, 2006[†]

Introduction

This document arose out of a problem assessing a confidence interval for the risk ratio quoted by authors of a paper I was reviewing. As an example which is close to theirs but made slightly simpler we suppose we have two groups each of size 500 and in one group we observed 8 events and in the other group zero events giving us an observed risk ratio of ∞ .

Notation

We assume we have r_i events in group i out of n_i trials ($i = 1, 2$). We have a proportion $p_i = \frac{r_i}{n_i}$ in each group.

Methods proposed

Wald method

We use the usual formulae for the log risk ratio and its variance but add a small constant to the numerator r_i .

*Section of Epidemiology, Division of Psychological Medicine, PO 60, Institute of Psychiatry, De Crespigny Park, London, SE5 8AF, UK. <mailto:m.dewey@iop.kcl.ac.uk>

[†]Printed November 22, 2006 at 10:52

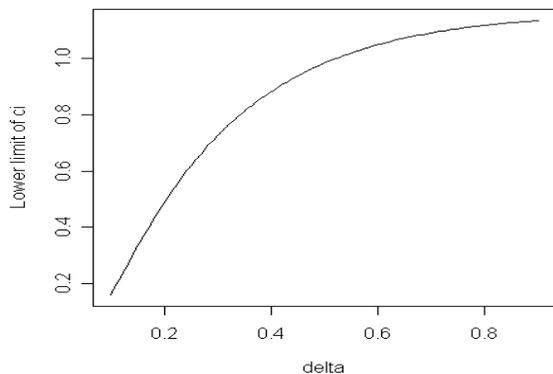


Figure 1: Lower limit of confidence interval for the risk ratio for a range of values of delta

```
# functions for Wald method
#
rr <- function(x1,n1,x2,n2,conflev, delta = 0.5) {
  x1.d <- x1 + delta
  x2.d <- x2 + delta
  lrr <- log((x1.d/n1) / (x2.d/n2))
  se.lrr <- sqrt(1/x1.d - 1/n1 + 1/x2.d - 1/n2)
  mult <- abs(qnorm((1-conflev)/2))
  res <- list(lrr = lrr, se.lrr = se.lrr,
             low.lim = lrr - mult * se.lrr, high.lim = lrr + mult * se.lrr)
  res
}
wrapper <- function(x) {exp(rr(8, 500, 0, 500, 0.95, x)$low.lim)}
plot(wrapper, 0.1, 0.9, xlab = "delta", ylab = "Lower limit of ci")
```

Example 1: R code for Wald method

$$RR = \log \frac{\left(\frac{r_1 + \delta}{n_1}\right)}{\left(\frac{r_2 + \delta}{n_2}\right)} \quad (1)$$

The usual choice for δ is 0.5 although there does not seem to be any theory behind this.

$$\text{var}(RR) = \frac{1}{r_1 + \delta} - \frac{1}{n_1} + \frac{1}{r_2 + \delta} - \frac{1}{n_2} \quad (2)$$

In fact the outcome is quite sensitive to the choice of δ as can be seen from

```

#
# function for transforming odds ratio
#
rr.from.or <- function(p1, or) {
  or/((1-p1) + p1 * or)
}
rr.from.or(8/500,
  fisher.test(matrix(c(8, 492, 0, 500), ncol = 2, byrow = TRUE))$conf.int[1])

```

Example 2: R code for the OR to RR transformation

Figure 1. Example 1 shows R code for calculating the confidence interval and for producing the plot in Figure 1.

Score test based intervals

Code has been published for generating confidence intervals by inverting a score test. It is available from http://web.stat.ufl.edu/~aa/cda/R/two_sample/R2/

Using the interval for the odds ratio

We know (Zhang and Yu, 1998) that

$$RR = \frac{OR}{(1 - p_1) + p_1 OR} \quad (3)$$

where p_1 is the rate in the unexposed group (called p_0 in Zhang and Yu (1998)). So having estimated the lower limit of the interval for the odds ratio (using `fisher.test` or one of a variety of other R packages) we can transform it into a risk ratio. Example 2 shows code for doing this for the example. McNutt et al. (2003) suggest that this method is not in general a good thing and may be biased.

Using the profile likelihood

We can fit a model to the table using Poisson regression and then use profile likelihood methods to find the value of the relative risk at which the log likelihood differs from that for the full model by 3.84 (for 95% confidence interval). Figure 2 shows a plot of the deviance against values for the log

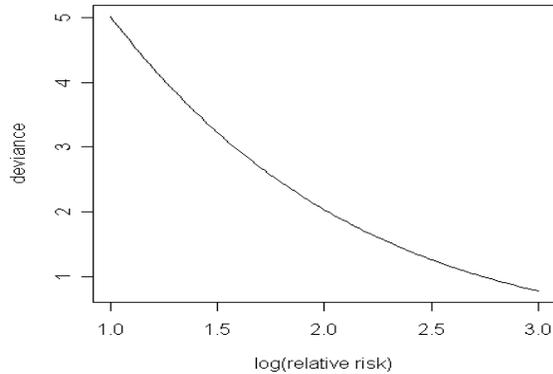


Figure 2: Deviance and log relative risk

```

#
# functions for likelihood based method
#
prof <- function(delta) { # vectorised to enable plotting
  n.delta <- length(delta)
  dev <- numeric(n.delta)
  for (i in 1:n.delta) {
    dev[i] <- glm(r ~ offset(treat * delta[i] + log(n)),
      family = poisson, data = adverse)$deviance
  }
  dev
}
prof.uniroot <- function(delta) {
  prof(delta) - qchisq(0.95, 1)
}
adverse <- data.frame(r = c(8,0), n = c(500,500), treat = c(1,0))
uniroot(prof.uniroot, interval = c(-2,2))

```

Example 3: R code for the likelihood method

relative risk. The critical value is obtained around 1.3 Example 3 shows code for evaluating the likelihood and for plotting Figure 2. It also contains code for finding the point at which the deviance differs by the prescribed amount from the deviance for the saturated model (in this case zero).

Conclusions

Table 1 summarises the values obtained using the various methods.

Method	Estimate	95%ci	
Wald	17	0.98	293.7
Score test	∞	2.09	∞
Transform odds ratio	∞	1.70	∞
Profile likelihood	∞	3.69	∞

Table 1: Comparison of confidence intervals using the proposed methods

- It seems strange to use the Wald method as it is sensitive to the choice of δ and also fails to include the sample estimate of relative risk ∞ .
- General problems with the transformation of the odds ratio suggest another method might be better.

The choice seems therefore to lie between the score test based method and that based on likelihood.

To do

- Work out why the values are so different (incompatible assumptions about the 'other' end of the ci?)
- Work out how to provide starting values so `confint` from MASS works for the likelihood method.

Acknowledgments

Thanks to David Duffy, Spencer Graves, Jukka Jokinen, Terry Therneau, and Wolfgang Viechtbauer. The views and calculations presented here are my own and any errors are my responsibility not theirs.

References

- L-A McNutt, C Wu, X Xue, and J P Hafner. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *American Journal of Epidemiology*, 157:940–943, 2003.
- J Zhang and K F Yu. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *Journal of the American Medical Association*, 280:1690–1691, 1998.