Estimating relative risks (RRs) and confidence intervals (CIs) for other comparisons when only given RRs and CIs for comparisons relative to an unexposed group

Author: P N Lee
Date : 10.11.98

The method

Suppose, in a case-control study with subjects divided into n+1 groups, an unexposed group (i=0) and n exposed groups (i=1, ...n), one is provided with estimates, for each exposed group, of the relative risk compared to the unexposed group (R_i) and its upper and lower 95% confidence limits (U_i, L_i). We wish to estimate relative risks for other comparisons, e.g. all exposed combined vs unexposed, one exposure vs another, high exposures vs low exposure + unexposed, etc.

The approach used is to try to reconstruct the underlying table of numbers of cases and controls (or effective numbers in the case of adjusted data) from the evidence provided plus two additional pieces of data:

1) The proportion of unexposed subjects in the controls, P
2) The relative frequency of controls to cases overall, Z.

Let the underlying table to be estimated be as follows:

	Exposure 0	1	2	......	n	Total
Cases	A_0	A_1	A_2		A_n	A
Controls	B_0	B_1	B_2		B_n	B

As a first step, estimate the variance of log R_i by

$$V_i = \left\{\frac{\log_e(U_i/L_i)}{3.92}\right\}^2 \qquad (i = 1, ...n)$$

We wish to estimate the 2(n+1) numbers A_i, B_i (i=0, ..n) from the following 2(n+1) equations

(1)	$P = B_0/B$	1 equation
(2)	$Z = B/A$	1 equation
(3)	$R_i = A_iB_0/A_0B_i$	n equations
(4)	$V_i = 1/A_0+1/B_0+1/A_i+1/B_i$	n equations

Given an estimate of $A_0 = \hat{A}_0$ and of $B_0 = \hat{B}_0$
we can rewrite (3) as $\hat{A}_i = \hat{B}_i R_i \hat{A}_0/\hat{B}_0$
and rewrite (4) as $V_i - 1/\hat{A}_0 - 1/\hat{B}_0 = 1/\hat{A}_i + 1/\hat{B}_i$
$= (1/\hat{B}_i)(1+\hat{B}_0/\hat{A}_0 R_i)$

so that $\hat{B}_i = (1+\hat{B}_0/\hat{A}_0 R_i)/(V_i - 1/\hat{A}_0 - 1/\hat{B}_0)$
and $\hat{A}_i = (1+\hat{A}_0 R_i/\hat{B}_0)/(V_i - 1/\hat{A}_0 - 1/\hat{B}_0)$

Hence we can estimate $\hat{A} = \sum_{i=0}^{n} \hat{A}_i$

and $\hat{B} = \sum_{i=0}^{n} \hat{B}_i$

and thus $\hat{P} = \hat{B}_0/\hat{B}$

and $\hat{Z} = \hat{B}/\hat{A}$

$\hat{A}_0$, $\hat{B}_0$ is a solution if $p = \hat{p}$ and $Z = \hat{Z}$

Hence the estimation can be viewed as finding values of $\hat{A}_0$, $\hat{B}_0$ that minimize

$$\left(\frac{p-\hat{p}}{p}\right)^2 + \left(\frac{Z-\hat{Z}}{Z}\right)^2.$$

This gives a basis for estimating A_0 and B_0 (see below) and hence the whole table.

Having a stable table of estimates one can then readily derive relative risks and confidence intervals for a given contrast. Thus, to compare levels 3+4+5 (exposed) with level 0+2 (base) say

derive the 2x2 table $A_0+A_2 = A^*$ $\quad$ $A_3+A_4+A_5 = A^{**}$

$B_0+B_2 = B^*$ $\quad$ $B_3+B_4+B_5 = B^{**}$

Compute the relative risk $R^* = A^{**}B^*/A^*B^{**}$

Compute the variance of the logarithm of the relative risk

$$V^* = 1/A^*+1/B^*+1/A^{**}+1/B^{**}$$

and hence the upper and lower 95% confidence limits for log R

$$\log U^* = \log R^* + 1.96\sqrt{V^*}$$

$$\log L^* = \log R^* - 1.96\sqrt{V^*}$$

U* and L* are thus derived by taking exponentials.

Software

To carry out the analyses a program was written using Corel Quattro Pro spreadsheet facilities. The program required the user to input the values of R_i, U_i and L_i as well as p and Z and a starting estimate of A_0 and B_0. The interactive process used to minimize $\left(\frac{p-\hat{p}}{p}\right)^2+\left(\frac{Z-\hat{Z}}{Z}\right)^2$ used the following options: Search: Newton; Derivatives: Forward; Estimates: Tangent; Precision: 1E-06; Maximum number of iterations: 100.

Example with sensitivity analysis

Fontham (1994, JAMA, 271, 1752-1759) presents in Table 3 the following:

Pack-years of exposure	Cases	Controls	Crude OR (95% CI)	Adjusted OR (95% CI)
All lung carcinomas				
0	267	562	1.00	1.00
≤15.0	146	300	1.02(0.80-1.31)	1.08(0.83-1.39)
15.1-39.9	92	190	1.02(0.76-1.36)	1.04(0.76-1.42)
40.0-79.9	80	126	1.34(0.98-1.83)	1.36(0.97-1.91)
≥80.0	24	27	1.87(1.06-3.31)	1.79(0.99-3.25)
Total	609	1205		

Using the adjusted odds ratios and 95% CIs, an estimate of p = 562/1205 = 0.466390 and of Z = 1205/609 = 1.978654, and starting estimates of $A_0 = 200$ and $B_0 = 400$, the procedure gave an estimate of the adjusted OR for overall exposure of 1.15 (0.94-1.42). This is what we would use in the absence of other information.

To test the procedure further we used the crude ORs. Combining them using our procedure gives 1.12 (0.92-1.36), which is exactly the same as that which could be calculated directly using the 2x2 table

	Cases	Controls
Unexposed	267	562
Exposed	342	642

Next we tried varying our assumed values of p and Z, which, in practice, may not be known exactly, and also the starting values of A_0 and B_0. Results are shown in four attached tables, one for $A_0 = 200$, $B_0 = 400$; one for $A_0 = 250$, $B_0 = 550$; one for $A_0 = 100$, $B_0 = 200$; one for $A_0 = 200$, $B_0 = 200$.

It can be seen from the table that there are a number of cases where the search failed to give an answer (NS). However the specific p, Z combinations varied according to the starting estimates of A_0, B_0 chosen and it was possible to get a solution with some starting estimates for all p, Z combinations except one (p = 0.6, Z = 1.9). Where it converges the answers obtained given p and Z hardly vary at all according to the starting position. However there is a tendency for the estimates to vary slightly by p and Z, though not very much. Certainly given p and Z are reasonably well estimated, the resulting RR and CI estimates are almost constant.

More work needs to be done. Partly this will be to try to avoid non-convergence. Partly it will involve using data stratified into two or three levels to check whether the procedure regenerates the exact stratified RR and CI estimate.

TABLE 1

Results using $A_0 = 200$, $B_0 = 400$

			p			
Z	0.30	0.40	0.466390	0.50	0.55	0.60
1	1.11(0.90-1.38)	1.12(0.92-1.37)	1.13(0.93-1.37)	1.13(0.93-1.37)	1.13(0.94-1.36)	NS
1.2	1.11(0.90-1.38)	1.12(0.91-1.37)	NS	1.13(0.93-1.37)	1.13(0.93-1.36)	NS
1.5	1.11(0.89-1.38)	1.12(0.91-1.37)	NS	1.12(0.93-1.36)	1.13(0.93-1.36)	NS
1.8	NS	1.12(0.91-1.37)	NS	1.12(0.93-1.36)	NS	1.13(0.94-1.35)
1.9	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.13(0.93-1.36)	NS
1.978654	1.11(0.89-1.37)	NS	1.12(0.92-1.36)	NS	1.12(0.93-1.36)	NS
2.1	1.11(0.89-1.37)	NS	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.36)	NS
2.2	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.92-1.36)	NS	NS
2.5	1.11(0.89-1.37)	NS	1.12(0.92-1.36)	NS	1.12(0.93-1.36)	1.13(0.94-1.35)
3	1.10(0.89-1.37)	1.11(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	NS	1.12(0.94-1.35)

TABLE 2

Results using $A_0 = 250$, $B_0 = 550$

			p			
Z	0.30	0.40	0.466390	0.50	0.55	0.60
1	1.11(0.90-1.38)	1.12(0.92-1.37)	1.13(0.93-1.37)	1.13(0.93-1.37)	1.13(0.94-1.36)	1.13(0.94-1.36)
1.2	1.11(0.90-1.38)	1.12(0.91-1.37)	1.12(0.92-1.37)	NS	1.13(0.94-1.36)	1.13(0.94-1.36)
1.5	1.11(0.89-1.38)	1.12(0.91-1.37)	1.12(0.92-1.37)	NS	NS	1.13(0.94-1.36)
1.8	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.13(0.93-1.36)	NS
1.9	1.11(0.89-1.37)	1.12(0.91-1.37)	NS	1.12(0.93-1.36)	1.13(0.93-1.36)	NS
1.978654	1.11(0.89-1.37)	NS	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.36)	1.13(0.94-1.35)
2.1	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.36)	NS
2.2	1.11(0.89-1.37)	NS	1.12(0.92-1.36)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.13(0.94-1.35)
2.5	1.11(0.89-1.37)	1.11(0.91-1.37)	1.12(0.92-1.36)	1.12(0.92-1.36)	1.12(0.93-1.36)	NS
3	1.10(0.89-1.37)	1.11(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.35)	1.12(0.94-1.35)

TABLE 3

Results using $A_0 = 100$, $B_0 = 200$

	p					
Z	0.30	0.40	0.466390	0.50	0.55	0.60
1	1.11(0.90-1.38)	NS	1.13(0.93-1.37)	1.13(0.93-1.37)	1.13(0.94-1.36)	NS
1.2	1.11(0.90-1.38)	1.12(0.91-1.37)	NS	1.13(0.93-1.37)	1.13(0.94-1.36)	1.13(0.94-1.36)
1.5	1.11(0.89-1.38)	1.12(0.91-1.37)	1.12(0.92-1.37)	1.12(0.93-1.36)	1.13(0.93-1.36)	NS
1.8	1.11(0.89-1.37)	1.12(0.91-1.37)	NS	1.12(0.93-1.36)	1.13(0.93-1.36)	1.13(0.94-1.35)
1.9	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	NS	1.13(0.93-1.36)	NS
1.978654	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.36)	1.13(0.94-1.35)
2.1	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.36)	1.13(0.94-1.35)
2.2	NS	1.12(0.91-1.37)	1.12(0.92-1.36)	NS	1.12(0.93-1.36)	1.13(0.94-1.35)
2.5	1.11(0.89-1.37)	1.11(0.91-1.37)	1.12(092-1.36)	1.12(0.93-1.36)	NS	1.13(0.94-1.35)
3	1.11(0.89-1.37)	1.11(0.91-1.37)	1.12(0.92-1.36)	1.12(0.92-1.36)	1.12(0.93-1.35)	NS

TABLE 4

Results using $A_0 = 200$, $B_0 = 200$

	p					
Z	0.30	0.40	0.466390	0.50	0.55	0.60
1	1.11(0.90-1.38)	NS	NS	NS	1.13(0.94-1.36)	NS
1.2	1.11(0.90-1.38)	1.12(0.91-1.37)	1.12(0.92-1.37)	1.12(0.93-1.37)	1.13(0.93-1.36)	1.13(0.94-1.36)
1.5	NS	1.12(0.91-1.37)	1.12(0.92-1.37)	NS	NS	1.13(0.94-1.36)
1.8	NS	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.13(0.93-1.36)	NS
1.9	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.13(0.93-1.36)	NS
1.978654	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.36)	NS
2.1	1.11(0.89-1.37)	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.36)	1.13(0.94-1.35)
2.2	NS	1.12(0.91-1.37)	1.12(0.92-1.36)	1.12(0.92-1.36)	NS	1.13(0.94-1.35)
2.5	1.11(0.89-1.37)	1.11(0.91-1.37)	1.12(092-1.36)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.13(0.94-1.35)
3	1.10(0.89-1.37)	1.11(0.91-1.37)	1.12(0.92-1.36)	1.12(0.93-1.36)	1.12(0.93-1.35)	NS