Statistics in Medicine

Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates

JULIE A MORRIS, MARTIN J GARDNER

Gardner and Altman explained the rationale for using estimation and confidence intervals in making inferences from analytical studies and described their calculation for means or proportions and their differences.[1] In this paper we present methods for calculating confidence intervals for other common statistics obtained from medical investigations. The techniques for obtaining confidence intervals for estimates of relative risk are described. These can come either from an incidence study, where, for example, the frequency of a congenital malformation at birth is compared in two defined groups of mothers, or from a case-control study, where a group of patients with the disease of interest (the cases) is compared with another group of people without the disease (the controls).

The methods of obtaining confidence intervals for standardised disease ratios and rates in studies of incidence, prevalence, and mortality are described. Such rates and ratios are commonly calculated to enable appropriate comparisons to be made between study groups after adjustment for confounding factors like age and sex. The most frequently used standardised indices are the standardised incidence ratio (SIR) and the standardised mortality ratio (SMR).

A worked example is included for each method. The calculations have been carried out to full arithmetical precision, as is recommended practice,[2] although intermediate steps are shown as rounded results. Some of the methods given in this paper are large sample approximations and are not reliable for studies with fewer than about 20 cases. Appropriate design principles for these types of study have to be adhered to since confidence intervals convey only the effects of sampling variation on the precision of the estimated statistics and cannot control for other errors such as biases due to the selection of inappropriate controls or in the methods of collecting the data.

Confidence intervals for relative risks (odds ratios)

INCIDENCE STUDY

Suppose that the incidence or frequency of some outcome is assessed in two groups of individuals defined by the presence or absence of some characteristic. The data from such a study can be tabulated as follows:

Group characteristic	Outcome		Total
	Yes	No	
Present	A	C	A+C
Absent	B	D	B+D

The outcome probabilities in exposed and unexposed individuals are estimated from the study groups by A/(A+C) and B/(B+D) respectively. An estimate, R, of the relative risk (or risk ratio) from exposure is given by the ratio of these proportions:

$$R=\frac{A/(A+C)}{B/(B+D)}.$$

Confidence intervals for the population value of R can be constructed through a logarithmic transformation.[3] The standard error of $\log_e R$ is:

$$SE(\log_e R)=\sqrt{\frac{1}{A}-\frac{1}{A+C}+\frac{1}{B}-\frac{1}{B+D}}.$$

A $100(1-\alpha)\%$ confidence interval for R is found by first calculating the two quantities:

$$W=\log_e R-(N_{1-\alpha/2}\times SE(\log_e R))$$

and

$$X=\log_e R+(N_{1-\alpha/2}\times SE(\log_e R)),$$

where $N_{1-\alpha/2}$ is the appropriate value from the standard Normal distribution for the $100(1-\alpha/2)$ percentile. This is widely available in tables.

The confidence interval for the population value of R is then given by exponentiating W and X as:

$$e^W \text{ to } e^X.$$

Worked example

Susceptibility to rubella in antenatal patients screened in three public health laboratories in England and Wales was studied,[4] with the following results:

Group characteristic	Susceptibility to rubella		Total
	Yes	No	
Asians	161	2 475	2 636
Non-Asians	748	34 020	34 768

An estimate of the relative risk of susceptibility to rubella for Asians compared with non-Asians is:

$$R=\frac{161/2636}{748/34768}=2\cdot84.$$

Department of Medical Statistics, Withington Hospital, West Didsbury, Manchester M20 8LR

JULIE A MORRIS, MSC, medical statistician

MRC Environmental Epidemiology Unit (University of Southampton), Southampton General Hospital, Southampton SO9 4XY

MARTIN J GARDNER, PHD, professor of medical statistics

Correspondence to: Mrs Morris.

The standard error of the $\log_e R$ is:

$$\sqrt{\frac{1}{161}-\frac{1}{2636}+\frac{1}{748}-\frac{1}{35768}}=0{\cdot}0845$$

from which for a 95% confidence interval

$$W=\log_e 2{\cdot}84-(1{\cdot}96\times 0{\cdot}0845)=0{\cdot}8782$$

and

$$X=\log_e 2{\cdot}84+(1{\cdot}96\times 0{\cdot}0845)=1{\cdot}2094.$$

The 95% confidence interval for the population value of R is then given as:

$$e^{0{\cdot}8782} \text{ to } e^{1{\cdot}2094} \text{ that is, from } 2{\cdot}41 \text{ to } 3{\cdot}35.$$

UNMATCHED CASE-CONTROL STUDY

Suppose that groups of cases and controls are studied to assess exposure to a suspected causal factor. The data can be tabulated as follows:

	Exposed		
Study group	Yes	No	Total
Cases	a	b	a+b
Controls	c	d	c+d
Total	a+c	b+d	n

An approximate estimate of the relative risk for the disease associated with exposure to the factor can be obtained from a case-control study through the odds ratio.[5] The odds ratio (OR) is given as:

$$OR=\frac{ad}{bc}.$$

A confidence interval for the population value of OR can be constructed using several methods which vary in their ease and accuracy. The method described here (sometimes called the logit method) was devised by Woolf[6] and is widely recommended as a satisfactory approximation. The exception to this is when any of the numbers a, b, c, or d is small, when a more accurate but complex procedure should be used if suitable computer facilities are available. Further discussion and comparison of methods can be found in sections 4.2 and 4.3 of Breslow and Day.[7]

The logit method uses the Normal approximation to the distribution of the logarithm of the odds ratio ($\log_e OR$) in which the standard error of $\log_e OR$ is:

$$SE(\log_e OR)=\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}.$$

A $100(1-\alpha)\%$ confidence interval for the population value of OR is found by first calculating the two quantities:

$$Y=\log_e OR-(N_{1-\alpha/2}\times SE(\log_e OR))$$

and

$$Z=\log_e OR+(N_{1-\alpha/2}\times SE(\log_e OR)),$$

where $N_{1-\alpha/2}$ is the appropriate value from the standard Normal distribution for the $100(1-\alpha/2)$ percentile.

The confidence interval for OR is then given by exponentiating Y and Z as:

$$e^Y \text{ to } e^Z.$$

Worked example

The ABO secretor state was determined for 114 patients with spondyloarthropathies and 334 controls[8] with the following results:

	ABO secretor state		
Study group	Yes	No	Total
Cases	54	60	114
Controls	89	245	334
Total	143	305	448

The estimated odds ratio is $OR=\frac{54\times 245}{60\times 89}=2{\cdot}48$. The standard error of $\log_e OR$ is:

$$SE(\log_e OR)=\sqrt{\frac{1}{54}+\frac{1}{60}+\frac{1}{89}+\frac{1}{245}}=0{\cdot}2247.$$

For a 95% confidence interval

$$Y=\log_e 2{\cdot}48-(1{\cdot}96\times 0{\cdot}2247)=0{\cdot}4678$$

and

$$Z=\log_e 2{\cdot}48+(1{\cdot}96\times 0{\cdot}2247)=1{\cdot}3487.$$

The 95% confidence interval for the population value of OR is then given as:

$$e^{0{\cdot}4678} \text{ to } e^{1{\cdot}3487} \text{ that is, from } 1{\cdot}59 \text{ to } 3{\cdot}85.$$

MORE THAN TWO LEVELS OF EXPOSURE

If there are more than two levels of exposure one can be chosen as a baseline with which each of the others is compared. Odds ratios and their associated confidence intervals are then calculated for each comparison.

A SERIES OF UNMATCHED CASE-CONTROL STUDIES

A combined estimate is sometimes required when independent estimates of the same odds ratio are available from each of K sets of data—for example, in a stratified analysis to control for confounding variables. A common approach is to use the Mantel-Haenszel pooled estimate of the odds ratio (OR_{M-H}) which is given by:

$$OR_{M-H}=\frac{\Sigma a_i d_i/n_i}{\Sigma b_i c_i/n_i},$$

where a_i, b_i, c_i, and d_i are the frequencies in the ith 2×2 table, $n_i=a_i+b_i+c_i+d_i$, and the summation Σ is over $i=1$ to K for the K tables (see page 140 of Breslow and Day).[7] No method of calculating confidence intervals has been developed for this estimate.

An alternative is to use the logit method to give a pooled estimate of the odds ratio (OR_L) and then derive a confidence interval for the odds ratio in a similar way to that for a single 2×2 table. The logit combined estimate (OR_L) is defined by:

$$\log_e OR_L=\frac{\Sigma w_i \log_e OR_i}{\Sigma w_i},$$

where $OR_i=a_i d_i/b_i c_i$ is the odds ratio in the ith table and

$$w_i=1/\left(\frac{1}{a_i}+\frac{1}{b_i}+\frac{1}{c_i}+\frac{1}{d_i}\right).$$

A $100(1-\alpha)\%$ confidence interval for OR_L is found by calculating:

$$M=\log_e OR_L-(N_{1-\alpha/2}/\sqrt{w})$$

and

$$N=\log_e OR_L+(N_{1-\alpha/2}/\sqrt{w}),$$

where $w=\Sigma w_i$ and $N_{1-\alpha/2}$ is the appropriate value from the standard Normal distribution for the $100(1-\alpha/2)$ percentile.

The confidence interval for the population value of OR_L is given by exponentiating M and N as:

$$e^M \text{ to } e^N.$$

Further discussion of methods and a worked example are given in section 4.4 of Breslow and Day.[7] The logit method is unsuitable if any of the numbers a_i, b_i, c_i, or d_i are small. This will happen, for example, with increasing stratification, and in such cases a more complex exact method is available (see section 4.4 of Breslow and Day).[7]

MATCHED CASE-CONTROL STUDY

If each of n cases of a disease is matched to one control to form n pairs and

each individual's exposure to a suspected causal factor is recorded the data can be tabulated as follows:

Exposure status		
Case	Control	Number of pairs
Yes	Yes	r
Yes	No	s
No	Yes	t
No	No	u

For this type of study an approximate estimate (in fact the Mantel-Haenszel estimate) of the relative risk of the disease associated with exposure is again given by the odds ratio which is now calculated as:

$$OR=\frac{s}{t}.$$

An exact $100(1-\alpha)\%$ confidence interval for the population value of OR is found by first determining a confidence interval for s (the number of case-control pairs with only the case exposed)—see pages 461 and 462 of Armitage and Berry.[5] Conditional on the sum of the numbers of "discordant" pairs (s+t) the number s can be considered as a binomial variable with n=s+t and p=s/(s+t).

The $100(1-\alpha)\%$ confidence interval for p can be obtained from tables based on the binomial distribution.[9] If this confidence interval is denoted by A_L to A_U the $100(1-\alpha)\%$ confidence interval for the population value of OR is then given by:

$$\frac{A_L}{1-A_L} \text{ to } \frac{A_U}{1-A_U}.$$

Worked example

Thirty five patients who died in hospital from asthma were individually matched for sex and age with 35 control subjects who had been discharged alive from the same hospital in the preceding year.[10] The inadequacy of monitoring of all patients while in hospital was independently assessed with the following paired results:

Inadequacy of monitoring		
Deaths	Survivors	Number of pairs
Yes	Yes	10
Yes	No	13
No	Yes	3
No	No	9

The estimated odds ratio of dying in hospital associated with inadequate monitoring is $OR=\frac{13}{3}=4.33$.

From the appropriate table for the binomial distribution with n=13+3=16, p=13/(13+3)=0·81, and $\alpha=0{\cdot}05$ the 95% confidence interval for p is found to be $A_L=0{\cdot}5435$ to $A_U=0{\cdot}9595$. The 95% confidence interval for the population value of the odds ratio is thus:

$$\frac{0{\cdot}5435}{1-0{\cdot}5435} \text{ to } \frac{0{\cdot}9595}{1-0{\cdot}9595} \text{ that is, from } 1{\cdot}19 \text{ to } 23{\cdot}69.$$

MATCHED CASE-CONTROL STUDY WITH 1:M MATCHING

Sometimes each case is matched with more than one control. The odds ratio is then given by the Mantel-Haenszel estimate as:

$$OR_{M-H}=\frac{\Sigma(M-i+1)\times n_{1,i-1}}{\Sigma i\times n_{0,i}},$$

where M is the number of matched controls for each case, $n_{1,i}$ is the number of matched sets in which the case and i controls are exposed, $n_{0,i}$ is the number of sets in which the case is unexposed and i controls are exposed, and the summation is from 1 to M.

A confidence interval for the population value of OR_{M-H} can be derived by one of the methods in section 5.3 of Breslow and Day.[7] Section 5.4 of that reference explains the calculation of a confidence interval for the odds ratio estimated from a study with a variable number of matched controls for each case.

Confidence intervals for standardised ratios and rates

STANDARDISED RATIOS

If O is the observed number of incident cases (or deaths) in a study group and E the expected number based on a reference population the standardised incidence ratio (SIR) or standardised mortality ratio (SMR) is O/E. This is usually called the indirect method of standardisation. The expected number is calculated as:

$$E=\Sigma n_i R_i,$$

where n_i is the number of individuals in age group i of the study group, R_i is the death rate in age group i of the reference population, and Σ denotes summation over all age groups.

The $100(1-\alpha)\%$ confidence interval for the population value of O/E can be found by first regarding O as a Poisson variable and finding its related confidence interval.[11] This is derived from tables based on the Poisson distribution. Denote this confidence interval by O_L to O_U.

The $100(1-\alpha)\%$ confidence interval for O/E is then given by:

$$\frac{O_L}{E} \text{ to } \frac{O_U}{E}.$$

Worked example

Roman *et al* observed 64 cases of leukaemia in children under the age of 15 years in the West Berkshire Health Authority area during 1972-85.[12] They calculated that 45·6 cases would be expected. Using O=64 and E=45·6 the standardised incidence ratio (SIR) is 64/45·6=1·40. Values of $O_L=49{\cdot}3$ and $O_U=81{\cdot}7$ are found from the appropriate table based on the Poisson distribution when O=64 and $\alpha=0{\cdot}05$.

The 95% confidence interval for the population value of the standardised incidence ratio is:

$$\frac{49{\cdot}3}{45{\cdot}6} \text{ to } \frac{81{\cdot}7}{45{\cdot}6} \text{ that is, from } 1{\cdot}08 \text{ to } 1{\cdot}79.$$

Sometimes the standardised incidence ratio (or standardised mortality ratio) is multiplied by 100 and then the same must be done to the figures describing the confidence interval.

RATIO OF TWO STANDARDISED RATIOS

Let O_1 and O_2 be the observed numbers of cases (deaths) in two study groups and E_1 and E_2 the two expected numbers. It is sometimes appropriate to calculate the ratio of the two standardised incidence ratios (standardised mortality ratios) O_1/E_1 and O_2/E_2 and find a confidence interval for this ratio. Again O_1 and O_2 can be regarded as Poisson variables and a confidence interval for the ratio O_1/O_2 is obtained as described by Ederer and Mantel.[13] The procedure then recognises that conditional on the total of O_1+O_2 the number O_1 can be considered as a binomial variable with $n=O_1+O_2$ and $p=O_1/(O_1+O_2)$. The $100(1-\alpha)\%$ confidence interval for p can be obtained from tables based on the binomial distribution. Denote this confidence interval by A_L to A_U. The $100(1-\alpha)\%$ confidence interval for O_1/O_2 can now be found as:

$$B_L=A_L/(1-A_L) \quad \text{to} \quad B_U=A_U/(1-A_U).$$

The $100(1-\alpha)\%$ confidence interval for the population value of the ratio of the two standardised incidence ratios (standardised mortality ratios) is then given by:

$$B_L\times\frac{E_2}{E_1} \text{ to } B_U\times\frac{E_2}{E_1}.$$

Worked example

Roman *et al* published figures for childhood leukaemia during 1972-85 in Basingstoke and North Hampshire Health Authority which gave O=25 and E=23·7 and a standardised incidence ratio of 1·05.[12] To compare the figures for West Berkshire from the previous example with those from Basingstoke and North Hampshire let $O_1=64$, $E_1=45{\cdot}6$, $O_2=25$, and $E_2=23{\cdot}7$.

The ratio of the two standardised incidence ratios is (64/45·6)/(25/23·7)= 1·40/1·05=1·33. From the appropriate table for the binomial distribution with n=64+25=89, p=64/(64+25)=0·72, and α=0·05, the 95% confidence interval for p is found to be A_L=0·6138 to A_U=0·8093. The 95% confidence interval for O_1/O_2 is thus 0·6138/(1−0·6138) to 0·8093/(1−0·8093) that is, from 1·59 to 4·24.

The 95% confidence interval for the population value of the ratio of the two standardised incidence ratios is then given by:

$$1{\cdot}59\times\frac{23{\cdot}7}{45{\cdot}6} \quad \text{to} \quad 4{\cdot}24\times\frac{23{\cdot}7}{45{\cdot}6} \text{ that is, from 0·83 to 2·21.}$$

STANDARDISED RATES

If a rate rather than a ratio is required the standardised rate (SR) in a study group is given by:

$$SR=\frac{\Sigma(N_i r_i)}{\Sigma N_i},$$

where N_i is the number of individuals in age group i of the reference population, r_i is the disease rate in age group i of the study group, and Σ indicates summation over all age groups. This is usually known as the direct method of standardisation. If n_i is the number of individuals in age group i of the study group the approximate standard error of SR is:

$$SE(SR)=\frac{\sqrt{\Sigma(N_i^2 r_i/n_i)}}{\Sigma N_i},$$

assuming that the rates r_i are small.

The 100(1−α)% confidence interval for the population value of SR is then given by:

$$SR-(N_{1-\alpha/2}\times SE(SR)) \quad \text{to} \quad SR+(N_{1-\alpha/2}\times SE(SR)),$$

where $N_{1-\alpha/2}$ is the appropriate value from the Normal distribution for the 100(1−α/2) percentile.

Worked example

The following observations were made in a study of the radiological prevalence of Paget's disease of bone in British male migrants to Australia[14]:

Age (years)	Study group			Standard population[15]
	Cases	n_i	r_i per 100	
55-64	4	96	4·2	2773
65-74	13	237	5·5	2556
75-84	8	105	7·6	1113
≥85	7	32	21·9	184
Totals	32	470	6·8	6626

The standardised prevalence rate (SR) is 5·7 per 100 with SE(SR)=1·17 per 100. The 95% confidence interval for the population value of SR is then given by:

5·7−(1·96×1·17) to 5·7+(1·96×1·17) that is, from 3·4 to 8·0 per 100.

We thank Dr N E Day and Mr D G Altman for their comments on an earlier version of this article, Dr J R Eason for providing the unpublished data, and Mrs Brigid Howells for her careful typing.

References

1 Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br Med J* 1986;**292**:746-50.
2 Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. *Br Med J* 1983;**286**:1489-93.
3 Katz D, Baptista J, Azen SP, Pike MC. Obtaining confidence intervals for the risk ratio in cohort studies. *Biometrics* 1978;**34**:469-74.
4 Miller E, Nicoll A, Rousseau SA, *et al.* Congenital rubella in babies of south Asian women in England and Wales: an excess and its causes. *Br Med J* 1987;**294**:737-9.
5 Armitage P, Berry G. *Statistical methods in medical research.* 2nd ed. Oxford: Blackwell, 1987: 456-9.
6 Woolf B. On estimating the relationship between blood group and disease. *Ann Hum Genet* 1955;**19**:251-3.
7 Breslow NE, Day NE. *Statistical methods in cancer research: volume 1—the analysis of case-control studies.* Lyons: International Agency for Research on Cancer, 1980.
8 Shinebaum R, Blackwell CC, Forster PJG, Hurst NP, Weir DM, Nuki G. Non-secretion of ABO blood group antigens as a host susceptibility factor in the spondyloarthropathies. *Br Med J* 1987;**294**:208-10.
9 Lentner C, ed. *Geigy scientific tables. Vol 2.* 8th ed. Basle: Ciba-Geigy, 1982.
10 Eason J, Markowe HLJ. Controlled investigation of deaths from asthma in hospitals in the North East Thames region. *Br Med J* 1987;**294**:1255-8.
11 Breslow NE, Day NE. *Statistical methods in cancer research. Vol 2. The analysis of cohort studies.* Oxford: University Press, 1988.
12 Roman E, Beral V, Carpenter L, *et al.* Childhood leukaemia in the West Berkshire and the Basingstoke and North Hampshire District Health Authorities in relation to nuclear establishments in the vicinity. *Br Med J* 1987;**294**:597-602.
13 Ederer F, Mantel N. Confidence limits on the ratio of two Poisson variables. *Am J Epidemiol* 1974;**100**:165-7.
14 Gardner MJ, Guyer PB, Barker DJP. Radiological prevalence of Paget's disease of bone in British migrants to Australia. *Br Med J* 1978;i:1655-7.
15 Barker DJP, Clough PWL, Guyer PB, Gardner MJ. Paget's disease of bone in 14 British towns. *Br Med J* 1977;i:1181-3.

(Accepted 4 February 1988)

To what extent is the menstrual cycle likely to affect the academic performance of an 18 year old girl?

In 1968 Dalton reported lower average marks among girls at boarding school taking examinations during their period or within the four days before it.[1] The study included 128 girls, but lack of statistical data makes it difficult to interpret satisfactorily.[2] Examinations altered the cycle in 42% of girls, which meant that more girls than expected menstruated during their examination weeks,[1] and the relation between low marks and menstruation could have been due to the common factor of stress. A more recent study of 244 medical students showed no effect of the cycle on examination performance, even among those who might have been vulnerable because of severe menstrual symptoms or personality predisposition.[3] A study of 13 students doing teacher training showed no significant variation over different phases of the cycle, though for difficult tasks there was a small association between symptoms and diminished performance.[4] There is some evidence that the endocrine response to examination stress is less pronounced in women than men,[5] and the overall consensus among recent studies of examination performance is that severe changes do not occur in most women although some may be adversely affected.[2] The premenstrual syndrome can be treated by dietary measures, relaxation, and exercise as well as by hormonal measures such as the contraceptive pill,[6] but there is no information about the effect of treatment on examination performance.—

JAMES OWEN DRIFE, senior lecturer in obstetrics and gynaecology, Leicester.

1 Dalton K. Menstruation and examinations. *Lancet* 1968;ii:1386-8.
2 Clare AW. Hormones, behaviour and the menstrual cycle. *J Psychosom Res* 1985;**29**:225-33.
3 Walsh RN, Budtz-Olsen I, Leader C, Cummins RA. The menstrual cycle, personality and academic performance. *Arch Gen Psychiatry* 1981;**38**:219-21.
4 Slade P, Jenner FA. Performance tests in different phases of the menstrual cycle. *J Psychosom Res* 1980;**24**:5-8.
5 Johansson G, Laakso ML, Karonen SL, Peder M. Examination stress affects plasma levels of TSH and thyroid hormones differently in females and males. *Psychosom Med* 1987;**49**:390-6.
6 Massil HY, O'Brien PMS. Approach to the management of premenstrual syndrome. *Clin Obstet Gynecol* 1987;**30**:443-52.