CONFIDENCE INTERVALS FOR THE ODDS RATIO
IN CASE-CONTROL STUDIES:
THE STATE OF THE ART

Joseph L. Fleiss

April, 1978

PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

DIVISIONS OF BIOSTATISTICS AND
EPIDEMIOLOGY, &
DEPARTMENT OF MATHEMATICAL STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK

Prepared For
— THE U.S. DEPARTMENT OF ENERGY —
UNDER CONTRACT NO. EY-76-S-02-2874

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

CONFIDENCE INTERVALS FOR THE ODDS RATIO
IN CASE-CONTROL STUDIES: THE STATE OF THE ART

by

Joseph L. Fleiss, Ph.D.
Professor and Head,
Division of Biostatistics, Columbia University
School of Public Health,
600 West 168 St., New York, N.Y. 10032

Presented at the Bermuda Symposium on Retrospective Case-Control Studies
April 6-8, 1978

STUDY ON STATISTICS AND ENVIRONMENTAL
FACTORS IN HEALTH

PREPARED UNDER SUPPORT TO SIMS FROM

DEPARTMENT OF ENERGY (DOE)
ENVIRONMENTAL PROTECTION AGENCY (EPA)
NATIONAL SCIENCE FOUNDATION (NSF)
ROCKEFELLER FOUNDATION
SLOAN FOUNDATION

DIVISION OF BIostatISTICS
COLUMBIA UNIVERSITY
SCHOOL OF PUBLIC HEALTH
NEW YORK, NEW YORK

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

2412

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Confidence Intervals for the Odds Ratio
in Case-Control Studies; The State of the Art*

Abstract

Many approximate procedures are available for setting confidence limits around the odds ratio in a four-fold table. A number of them are analyzed according to three criteria: close agreement between the actual and the nominal confidence levels; exact agreement with a hypothesis test for or against significance; and relative simplicity in calculation. The procedure due to Cornfield comes closer to satisfying all three criteria than any of the others examined. Formulas are provided for an iterative solution to Cornfield's equations.

*This work was supported in part by a grant from the SIAM Institute for Mathematics and Society to develop methods for the study of environmental factors in health.

Introduction

Attention will be restricted in this paper to hypothesis tests and confidence intervals for the odds ratio in data resulting from simple case-control studies. A number of approximate procedures for setting confidence limits about the population odds ratio are compared on the basis of three criteria:

(i) close agreement between the actual and the nominal confidence levels;

(ii) identical inferences for or against statistical significance from a hypothesis test and from the confidence interval excluding or including the value 1; and

(iii) relative simplicity in calculation, by which is meant that all calculations can be performed using desk-top or pocket calculators, with the mathematical operations being no more complicated than extracting square roots or taking logarithms.

Suppose that a sample of n_1 cases and a sample of n_2 controls are studied, with results arrayed as in Table 1. Throughout the paper, X will represent the number of cases who were exposed to the risk factor under study, and m will represent the total number of exposed subjects in the two samples combined.

See next page (1-a) for Table 1.

Let P_1 denote the underlying proportion of cases who were exposed to the risk factor and P_2 the underlying proportion of controls who were.

TABLE 1

Generic Four-fold Table from a Retrospective

Case-Control Study

Sample	Risk Factor		Total
	Exposed	Not Exposed	
Cases	X	$n_1 - X$	n_1
Controls	$m - X$	$n_2 - m + X$	n_2
Total	m	$n_1 + n_2 - m$	$n_1 + n_2$

The parameter of interest is the odds ratio,

$$\omega = \frac{P_1(1-P_2)}{P_2(1-P_1)}, \quad (1)$$

or occasionally, for mathematical convenience, its natural logarithm,

$$\lambda = \ln \omega = \ln \frac{P_1}{1-P_1} - \ln \frac{P_2}{1-P_2}. \quad (2)$$

Properties of the odds ratio which make it an important parameter in case-control studies have been described elsewhere [1-3].

A good estimator of the odds ratio has been found [4] to be

$$\hat{o} = \frac{(X + \frac{1}{2})(n_2 - m + X + \frac{1}{2})}{(m - X + \frac{1}{2})(n_1 - X + \frac{1}{2})}, \quad (3)$$

and a good estimator of the standard error of $\hat{o}$ has been found [5] to be

$$\text{s.e.}(\hat{o}) = \hat{o} \times S(X), \quad (4)$$

where

$$S(X) = \sqrt{\frac{1}{X+\frac{1}{2}} + \frac{1}{m-X+\frac{1}{2}} + \frac{1}{n_1-X+\frac{1}{2}} + \frac{1}{n_2-m+X+\frac{1}{2}}}. \quad (5)$$

Numerical Example

Consider the hypothetical data in Table 2. The value of the standard χ^2 statistic, corrected for continuity, for testing the hypothesis that the population odds ratio is unity is $\chi^2 = 3.86$, which is just barely significant at the 0.05 level. Since the exact significance level, calculated by means of the hypergeometric distribution [6], is

0.0485, the verdict from the chi square test with the continuity correction is seen to be correct.

See next page (3-a) for Table 2.

The next step in the analysis might be to construct confidence limits about the population odds ratio. The particular numerical example of Table 2 was selected to illustrate, among other things, a sensible requirement for the lower limit, say ω_L , of the interval. Because the null hypothesis was rejected, ω_L should be greater than 1.0; because it was just barely rejected, however, ω_L should not be much greater than 1.0.

The distribution theory required to construct exact confidence intervals around the population odds ratio has been known since 1935 [6], but the required calculations are too extensive for routine use. Computer programs [7] and tables [8] have recently become available. For the data of Table 2, the exact 95% confidence interval for the odds ratio is

$$1.005 \leq \omega \leq 7.316. \quad (6)$$

Note in particular that the lower limit, $\omega_L = 1.005$, exceeds unity but not by much.

Because the computer programs for the exact method are not widely available, and because the published tables do not cover all possible cases, one must resort in practice to approximate methods for constructing confidence intervals. A number of the approximate methods which are in use will now be analyzed.

TABLE 2

Hypothetical Four-fold Table

Sample	Risk Factor		Total
	Exposed	Not Exposed	
Cases	20	30	50
Controls	10	40	50
Total	30	70	100

Intervals Based on the Sample Odds Ratio

A simple but very crude approximate procedure is to take

$$o \pm 1.96 \times \text{s.e.}(o) \quad (7)$$

as a 95% confidence interval for ω , where o is defined in (3) and $\text{s.e.}(o)$ in (4). For the data at hand, $o = 2.593$ and $\text{s.e.}(o) = 1.164$, so the approximate interval becomes

$$0.312 \leq \omega \leq 4.874. \quad (8)$$

Among other defects, this interval suffers the obvious error that it includes the value unity, contradicting the evidence from the test of significance that the parameter was different from zero. Another deficiency becomes apparent when the same approach is used in setting confidence limits about ω^{-1} , the ratio of odds against exposure to the risk factor.

The point estimate of ω^{-1} is o^{-1} , the reciprocal of the quantity defined in (3), and its estimated standard error is given by

$$\text{s.e.}(o^{-1}) = \frac{1}{o} S(X), \quad (9)$$

where $S(X)$ is defined in (5). The approximate interval for ω^{-1} is then

$$o^{-1} \pm 1.96 \times \text{s.e.}(o^{-1}) \quad (10)$$

which becomes

$$0.046 \leq \omega^{-1} \leq 0.725. \quad (11)$$

By taking reciprocals, the corresponding interval for ω becomes

$$1.379 \leq \omega \leq 21.729. \quad (12)$$

No comparison with the exact interval given in expression (6) is required to appreciate that something is wrong, but only a comparison with the interval given in (8) which was constructed from identical

principles. The two are so discrepant that the obvious negative conclusion is warranted. This approach should be avoided unless the sample sizes are extremely large, literally in the thousands.

The standard error of the sample odds ratio does have a use, but it is purely a descriptive one. It is a good, rough indicator of the precision with which the underlying parameter has been estimated. It should not be used in making formal inferences, however.

Intervals Based on the Logarithm of the Odds Ratio

One reason for the breakdown of the procedure just described is the asymmetry of the range of possible values for the odds ratio. The infinitely wide interval from 1 on up is available for indexing "positive" association, whereas the finite interval from 0 to 1 is all that is available for indexing "negative" association. When the logarithm of the odds ratio is taken, however, symmetry results. Zero becomes the value indicative of no differential risk, the interval from $-\infty$ to 0 is available for indexing "negative" association, the interval from 0 to ∞ is available for indexing "positive" association, and a negative log odds ratio, say $-\lambda$, indicates exactly the same strength of association as does the positive value, λ .

Woolf [9] was one of the first to study explicitly the logarithm of the odds ratio. Define

$$L = \ln(o) = \ln \frac{(X + \frac{1}{2})(n_2 - m + X + \frac{1}{2})}{(m - X + \frac{1}{2})(n_1 - X + \frac{1}{2})} . \quad (13)$$

A good estimate of the standard error of L is given by

$$s.e.(L) = S(X), \quad (14)$$

where $S(X)$ is defined in equation (5), and an approximate 95% confidence interval for λ , the logarithm of the population odds ratio, is

$$L \pm 1.96 \times s.e.(L). \quad (15)$$

For the data of Table 2, $L=0.953$ and $s.e.(L) = 0.449$, so the resulting approximate 95% confidence interval becomes

$$0.073 \leq \lambda \leq 1.833. \quad (16)$$

An interval for the odds ratio itself is obtained by taking antilogarithms of the limits of this interval:

$$1.076 \leq \omega \leq 6.253. \quad (17)$$

An important property of this approach is that identically the same interval for the population odds ratio results if one analyzes the data with respect to odds against rather than odds for exposure.

Nevertheless, the approach errs in yielding an interval which is appreciably narrower than the exact interval given in (6). This phenomenon is true for the data of Table 2, and has been found to be true in other cases by Gart and Thomas [10]. The consequence is that the actual confidence level associated with the interval given by (15) is less than the nominal level of 95%, and thus that the interval is subject to greater uncertainty than is apparent from the nominal confidence level. The problem is only partially solved by redefining the estimated standard error of L by deleting the quantities $\frac{1}{2}$ in formula (5), thus increasing the estimate somewhat. The resulting interval is

$$1.060 \leq \omega \leq 6.342, \quad (18)$$

which is still too narrow.

Miettinen [11] recently proposed a simple modification of the approach represented by expression (15). Let M be any measure of association which assumes the value 0 under the null hypothesis, and let $V(M)$ denote the square of its estimated standard error. Then, when the sample sizes are large, $M^2/V(M)$ will tend to assume the same value no matter what the measure M , and this common value will be that of the standard χ^2 statistic. In particular,

$$\frac{L^2}{V(L)} = \chi^2, \quad (19)$$

so that

$$V(L) = \frac{L^2}{\chi^2} \quad (20)$$

and

$$\text{s.e.}(L) = \frac{L}{\chi}. \quad (21)$$

Miettinen suggested that

$$L\left(1 \pm \frac{1.96}{\chi}\right) \quad (22)$$

be taken as an approximate 95% confidence interval for λ , and that

$$\text{antilog}\left(L\left(1 - \frac{1.96}{\chi}\right)\right) \leq \omega \leq \text{antilog}\left(L\left(1 + \frac{1.96}{\chi}\right)\right) \quad (23)$$

be taken as an approximate 95% confidence interval for the population odds ratio.

For the data in Table 2, $\chi = \sqrt{3.86} = 1.964$, and Miettinen's interval becomes

(24)

$$1.002 \leq \omega \leq 6.708.$$

The lower limit of this interval agrees closer with that of the exact interval (6) than do the lower limits of the other intervals examined, but its upper limit is still too low.

In an exchange of correspondence with Halperin [12] on the validity of his approach, Miettinen [13] also pointed out that his intervals tended to be too narrow. They are easily constructed, however, and have the important property of being consistent with the result of the chi square test. If the verdict by the chi square test is "significance", then Miettinen's interval for the odds ratio will necessarily not include the value 1; if the verdict is "nonsignificance," then Miettinen's interval will necessarily include the value 1.

Yet other approximate procedures have been suggested for setting confidence limits around the odds ratio [14-16], but none has been found [10] to be uniformly better than one proposed by Cornfield [1]. Cornfield's procedure requires a series of iterations for solution; the results of Miettinen's procedure (23) are useful as a first approximation. Cornfield's approximate method will be described below following the statement of some theoretical statistical results.

Some Statistical Theory

Consider the data of Table 1, and suppose that the population odds ratio is equal to w . If the four marginal frequencies in Table 1 are considered fixed, then X is approximately normally distributed with

$$\text{mean } (X) = x \quad (25)$$

and

$$\begin{aligned}\text{variance } (X) &= \frac{1}{\frac{1}{x} + \frac{1}{m-x} + \frac{1}{n_1-x} + \frac{1}{n_2-m+x}} \\ &= \frac{1}{W(x)},\end{aligned}\tag{26}$$

say, where x is the unique positive solution to the quadratic equation

$$\frac{x(n_2-m+x)}{(m-x)(n_1-x)} = \omega.\tag{27}$$

The explicit equation for x as a function of ω is given in the appendix.

These results, originally stated without proof in 1951 by Stevens [17], have subsequently been proven by Cornfield [1] and by Hannan and Harkness [18]. Their most important use is in testing hypotheses about the value of the population odds ratio. If ω_0 is the hypothesized value, and if x_0 is the corresponding solution of equation (27), then

$$\chi^2 = \left(|X - x_0| - \frac{1}{2} \right)^2 W(x_0)\tag{28}$$

has, approximately, a chi square distribution with one degree of freedom when the hypothesis is true. Because of the intimate connection pointed out below between hypothesis tests and confidence intervals, equation (28) will also play a role in constructing confidence intervals for the population odds ratio.

As a simple example of the use of these results, consider the problem of testing the hypothesis that $\omega = 1$. For this hypothesized value only, the solution of equation (27) becomes especially simple,

$$x = \frac{mn_1}{n_1 + n_2},\tag{29}$$

as does the expression for the quantity $W(x)$ defined in equation (26),

$$W(x) = \frac{(n_1 + n_2)^3}{n_1 n_2 m (n_1 + n_2 - m)} . \quad (30)$$

The value for χ^2 given by equation (28) then becomes

$$\chi^2 = \frac{\left(\left| x - \frac{mn_1}{n_1 + n_2} \right| - \frac{1}{2} \right)^2 (n_1 + n_2)^3}{n_1 n_2 m (n_1 + n_2 - m)} , \quad (31)$$

which is algebraically identical to any of the usual formulas for chi square with the continuity correction [3]. In this one important case at least, therefore, the above approach succeeds in recapturing exactly the classic one.

This result also has a role to play in constructing a confidence interval for ω . Consider an arbitrary random variable Z whose distribution depends on an unknown parameter θ . A particular parameter value θ^* is said to be consistent with the data at the α level of significance if a test of the hypothesis that $\theta = \theta^*$ fails to reject. A $(1-\alpha) \cdot 100$ percent confidence interval for θ contains those, and only those values of the parameter which are consistent with the data at the α level of significance [19].

For the problem being considered, a 95% confidence interval for the odds ratio consists of all values of ω for which the value of χ^2 given by formula (28) is less than or equal to 3.84, the value required for significance at the 0.05 level. This is the essence of the procedure described by Cornfield [1].

Cornfield's Approximate Confidence Interval

Explicit expressions for the lower and upper confidence bounds resulting from Cornfield's procedure are not possible, because they call for the solution of fourth degree equations. The upper 95% confidence limit, say ω_U , is such that, when equation (27) is solved for x_U as a function of ω_U , and this value is inserted into expression (26) to yield $W(x_U)$, then

$$(X - x_U + \frac{1}{2})^2 W(x_U) = 3.84. \quad (32)$$

Similarly, the lower 95% confidence limit, say ω_L , is such that, when x_L is the corresponding solution of equation (27), then

$$(X - x_L - \frac{1}{2})^2 W(x_L) = 3.84. \quad (33)$$

The appendix gives the mathematical results necessary for convergence to ω_L and ω_U given any initial estimates; the better the initial estimates, the quicker the convergence. Miettinen's bounds (23) are proposed as the first approximations because they are fairly accurate and are easily calculated.

See next page (11-a) for Table 3.

Table 3 presents the series of calculations required to converge to $\omega_U = 7.208$, the upper limit of Cornfield's 95% confidence interval; only three steps were required. Equation (32) is easily seen to be satisfied. For $\omega_U = 7.208$, the corresponding value of x_U , the solution to equation (27), is $x_U = 24.230$, and the value of $W(x_U)$, defined by equation (26), is $W(x_U) = 0.2760$. Then, recalling that $X = 20$ in Table 2,

TABLE 3

Convergence to Cornfield's Upper 95% Confidence Limit
on the Odds Ratio, for the Data of

Table 2

Step	ω	x	$W(x)$	$F(x)$	$f'(\omega)$	$W'(x)$	$F'(x)$
1	6.708	23.966	0.2686	-2.283	0.555	0.0267	4.636
2	7.200	24.226	0.2759	-0.035	0.503	0.0293	4.492
3	7.208	24.230	0.2760	0.000	-	-	-

$$(20 - 24.230 + \frac{1}{2})^2 \times 0.2760 = 3.84, \quad (34)$$

No iteration is required to find ω_L , the lower 95% confidence limit. For Miettinen's lower limit of 1.002 (24), the corresponding value of x_L is 15.009, and that of $W(x_L)$ is 0.1905. Therefore,

$$(20 - 15.009 - \frac{1}{2})^2 \times 0.1905 = 3.84, \quad (35)$$

so that equation (33) is satisfied.

Cornfield's approximate 95% confidence interval for ω is therefore

$$1.002 \leq \omega \leq 7.208, \quad (36)$$

which comes closer than any of the others examined to the exact interval given in (6). Its actual confidence level is 94.86%, only trivially less than the nominal level of 95%.

Discussion

For simplicity only, the above development considered 95% confidence intervals exclusively. While 95% seems to be the most widely used confidence level, others are occasionally employed. All one need do is replace 1.96 with the corresponding critical normal curve value (e.g., with 2.58 for 99% confidence), and 3.84 with the corresponding critical chi square value (e.g., with 6.63 for 99% confidence).

While the setting for this paper was the retrospective case-control study, the results apply equally to comparative cohort studies of independent samples, and to cross-sectional studies in which none of the marginal frequencies is specified beforehand.

Gart [21] has extended Cornfield's procedure to the problem of setting confidence limits about the assumed common odds ratio in a number of independent fourfold tables. Work is in progress applying it to the related problem of testing the hypothesis that a series of independent odds ratios are equal.

In the preface to [3], the statement appears that "The reader who finds a need for confidence intervals ...will find...that the proper interval is almost always more complicated than simply the point estimate plus or minus a multiple of its standard error." The above analysis, together with analyses by Gart [10, 15, 20], reinforces that warning. The exact method for setting confidence limits on the odds ratio satisfies the first two criteria given in the introduction, but cannot generally be applied without access to a computer or without the expenditure of an inordinate amount of time performing tedious calculations. Approximate methods such as Miettinen's [11] satisfy the

second and third criteria, but provide intervals which are too narrow and thus have actual confidence levels which are lower than the nominal ones.

Cornfield's method [1] satisfies all three criteria, provided one takes the third to mean perhaps spending half an hour or so performing the calculations, instead of the five or ten minutes required with the simpler but inaccurate methods. When one considers that Cornfield's method is extremely accurate, and that it is perfectly consistent with respect to the classical chi square test for significance, the slight increase in arithmetic becomes trivial.

Appendix

An Iterative Solution for Cornfield's Approximate Confidence

Interval for the Odds Ratio

The upper 95% confidence limit for ω , say ω_U , is such that, when x_U is the unique positive solution to the quadratic equation

$$\frac{x_U(n_2 - m + x_U)}{(m - x_U)(n_1 - x_U)} = \omega_U, \quad (A1)$$

and when

$$W(x_U) = \frac{1}{x_U} + \frac{1}{m-x_U} + \frac{1}{n_1-x_U} + \frac{1}{n_2-m+x_U}, \quad (A2)$$

then

$$(X - x_U + \frac{1}{2})^2 W(x_U) = 3.84. \quad (A3)$$

For any value of ω , the positive root of equation (A1) is equal to

$$x = f(\omega) = \frac{1}{2(\omega-1)} \left(A(\omega) - B(\omega) \right), \quad (A4)$$

where

$$A(\omega) = (n_1 + m)\omega + (n_2 - m) \quad (A5)$$

and

$$B(\omega) = \sqrt{[A(\omega)]^2 - 4\omega(\omega-1)mn_1}. \quad (A6)$$

Equation (A3) may be re-expressed as follows to yield an equation which assumes the value zero when the desired upper limit is found:

$$F(x) = (X - x + \frac{1}{2})^2 - \frac{3.84}{W(x)}. \quad (A7)$$

If $\omega_U^{(1)}$ is a first approximation to ω_U , and if $x_U^{(1)}$ is the corresponding solution to equation (A4), then a second, better approximation is, by Newton's method [22],

$$\omega_U^{(2)} = \omega_U^{(1)} - \frac{F(x_U^{(1)})}{F'(x_U^{(1)})}, \quad (A8)$$

where

$$F'(x_U^{(1)}) = f'(\omega_U^{(1)}) \left\{ \frac{3.84 W'(x_U^{(1)})}{W(x_U^{(1)})^2} - 2 \left(x - x_U^{(1)} + \frac{1}{2} \right) \right\}. \quad (A9)$$

In equation (A9),

$$f'(\omega) = \frac{1}{2(\omega-1)} \left\{ (n_1+m) - \frac{(n_1+m)A(\omega) - 2mn_1(2\omega-1)}{B(\omega)} \right\} - \frac{f(\omega)}{\omega-1}, \quad (A10)$$

where $f(\omega)$, $A(\omega)$ and $B(\omega)$ are defined explicitly in equations (A4) - (A6),

and

$$W'(x) = \frac{1}{(m-x)^2} + \frac{1}{(n_1-x)^2} - \frac{1}{x^2} - \frac{1}{(n_2-m+x)^2}. \quad (A11)$$

If the value of x associated with the second approximation, say $x_U^{(2)}$, is such that $F(x_U^{(2)})$ is still not zero (say, if its absolute value exceeds 0.01), the process would have to be repeated.

Convergence to ω_L , the lower 95% confidence limit, proceeds by exactly the same process, except that the continuity correction is taken as $-\frac{1}{2}$ in equations (A7) and (A9).

The above formulas, while hardly simple, are explicit and are easily programmed for analysis using any programmable desk-top or pocket calculator. In practice, three or four iterations should suffice, even when the initial approximations are poor.

REFERENCES

1. Cornfield J: A statistical problem arising from retrospective studies.
In J Neyman (Ed): Proc. 3rd Berkeley Symp. on Mathem. Statist. and Prob., Vol. IV, 135-148, Berkeley: Univ.Calif.Pr., 1956.
2. Mantel N, Haenszel W: Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 22: 719-748, 1959.
3. Fleiss J L: Statistical Methods for Rates and Proportions. New York: Wiley, 1973.
4. Anscombe F J: On estimating binomial response relations. Biometrika 43:461-464, 1956.
5. Gart J J, Zweifel J R: On the bias of various estimates of the logit and its variance, with application to quantal assay. Biometrika 54:181-187, 1967.
6. Fisher R A: The logic of inductive inference. J Roy Statist Soc Ser A 98:39-54, 1935.
7. Thomas D G: Exact limits for the odds ratio in a 2x2 table. Appl Statist 20:105-110, 1970.
8. Thomas D G, Gart J J: A table of exact confidence limits for differences and ratios of two proportions and their odds ratios. J Amer Statist Assoc 72: 73-76, 1977.
9. Woolf B: On estimating the relation between blood group and disease. Ann Hum Genet 19:251-253, 1955
10. Gart J J, Thomas D G: Numerical results on approximate confidence limits for the odds ratio. J Roy Statist Soc Ser B. 34:441-447, 1972.

11. Miettinen O: Estimability and estimation in case-referent studies
Amer J Epidemiol 103:226-235, 1976.
12. Halperin M: Letter to the Editor, Re: "Estimability and estimation
in case-referent studies." Amer J Epidemiol 105: 496-498, 1977.
13. Miettinen O: Author's reply to letter by M. Halperin. Amer J
Epidemiol 105:498-502, 1977.
14. Cox D R: The regression analysis of binary sequences. J Roy
Statist Soc Ser B 20: 215-242, 1958.
15. Gart J J: Approximate confidence limits for the relative risk.
J Roy Statist Soc Ser B 24: 454-463, 1962.
16. Goodman L A: Simultaneous confidence limits for cross-product
ratios in contingency tables. J Roy Statist Soc Ser B 26:86-102,
1964.
17. Stevens W L: Mean and variance of an entry in a contingency table.
Biometrika 38:468-470, 1951.
18. Hannan J, Harkness W L: Normal approximation to the distribution of
two independent binomials, conditional on a fixed sum. Ann Mathem
Statist 34:1593-1595, 1963.
19. Lehmann E L: Testing Statistical Hypotheses. New York: Wiley, 1959.
20. Gart J J: The comparison of proportions: a review of significance
tests, confidence intervals and adjustments for stratification.
Rev Internat Statist Inst 39: 148-169, 1971.
21. Gart J J: Point and interval estimation of the common odds ratio in
the combination of 2x2 tables with fixed marginals. Biometrika
57:471-475, 1970.
22. Franklin P: Differential and Integral Calculus. N Y:McGraw-Hill 1953.