

---

# Two by Two Tables Containing Counts (TwobyTwo)

Kevin M. Sullivan, PhD, MPH, MHA, Associate Professor, Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

---

## INTRODUCTION

This chapter provides the formulae and examples for calculating crude and adjusted point estimates and confidence intervals for: risk ratios and differences; odds ratios; incidence rate ratios and differences; and etiologic and prevented fractions. The tests for interaction are also presented. First, the estimates from a single 2x2 table (“count” data) are presented followed by estimates adjusted or summarized across stratified data.

## FORMULAE AND EXAMPLE FOR A SINGLE 2X2 TABLE (COUNT DATA)

For a single 2x2 table, the notation is as depicted in table 15-1. The formulae for calculating the risk ratio, risk difference, and odds ratio and their confidence intervals are shown below. For the confidence intervals, the Taylor series approach is provided because it is a reasonably good confidence interval method when the sample size is large. There are other more complicated methods for computing confidence intervals when the data are sparse but they are not shown here, such as maximum likelihood and exact methods. Note that the term “risk” is used assuming a cohort study was performed and the risk of disease was assessed. If a study was based on *prevalent* disease, then substitute the term “prevalence” for “risk,” e.g., prevalence ratio and prevalence difference.

**Table 15-1.** Notation and table setup for a 2x2 table

	Exposed	Nonexposed	
Disease	<i>a</i>	<i>b</i>	<i>m</i> <sub>1</sub>
No Disease	<i>c</i>	<i>d</i>	<i>m</i> <sub>0</sub>
	<i>n</i> <sub>1</sub>	<i>n</i> <sub>0</sub>	<i>n</i>

$$\text{Estimated risk in the exposed} = \hat{R}e = a / n_1$$

$$\text{Estimated risk in the nonexposed} = \hat{R}u = b / n_0$$

$$\text{Estimated risk in the population} = \hat{R} = m_1 / n$$

## Point and Variance Estimates, Confidence intervals

The point and variance estimates and the confidence interval formulae are provided in Table 15.1. For some parameters there will not be a variance formula. The confidence limits for the Etiologic Fraction in the Exposed is based on the calculated upper lower bounds of the confidence limits for the risk ratio ( $RR_{UB}$  and  $RR_{LB}$ , respectively) with risk data. A similar approach is used when the Etiologic Fraction in the Exposed is based on the Odds Ratio.

## Statistical Tests

There are many statistical tests that can be performed on a single 2x2 table. Common tests include the Chi-square test (corrected, uncorrected, and Mantel-Haenszel) and exact tests (Fisher and mid-p exact). In this chapter the uncorrected and Mantel-Haenszel chi-square tests will be presented; however, these test should be used when the number of “expected” observations in each cell are  $\geq 5$ . When the expected number of observations in any cell is  $< 5$ , then one of the exact tests should be used. How to calculate exact p-values is beyond the scope of this text and requires an iterative calculation. The expected number of observations in a cell is calculated by multiplying the row and column total and dividing by the total sample size.

The uncorrected chi square is calculated as

$$uncorr\chi_1^2 = \frac{(n)(ad - bc)^2}{n_1 n_0 m_1 m_0}$$

and the Mantel Haenszel chi square as

$$mh\chi_1^2 = \frac{(n-1)(ad - bc)^2}{n_1 n_0 m_1 m_0}$$

Table 5.1. Estimates and confidence intervals for epidemiologic parameters for a single table

Parameter	Point Estimate	Variance Estimate	Confidence Interval
<i>Parameters based on risks (from randomized trials and cohort studies) or prevalences (cross-sectional studies)</i>			
Risk Ratio	$\hat{RR} = \frac{\hat{R}_e}{\hat{R}_u}$	$\hat{Var}(\ln \hat{RR}) \approx \frac{1 - \hat{R}_e}{n_1 \hat{R}_e} + \frac{1 - \hat{R}_u}{n_0 \hat{R}_u}$	$\hat{RR} \exp \left[ \pm Z_{1-\alpha/2} \sqrt{\hat{Var}(\ln \hat{RR})} \right]$
Risk Difference	$\hat{RD} = \hat{R}_e - \hat{R}_u$	$\hat{Var}(\hat{RD}) \approx \frac{\hat{R}_e(1 - \hat{R}_e)}{n_1} + \frac{\hat{R}_u(1 - \hat{R}_u)}{n_0}$	$\hat{RD} \pm Z_{1-\alpha/2} \sqrt{\hat{Var}(\hat{RD})}$
Etiologic Fraction in the Population	$E\hat{F}_p = \frac{\hat{R} - \hat{R}_u}{\hat{R}}$		
Etiologic Fraction in the Exposed	$E\hat{F}_e = \frac{\hat{R}_e - \hat{R}_u}{\hat{R}_e}$	Based on variance estimate for the RR	$LB = \frac{\hat{RR}_{LB} - 1}{\hat{RR}_{LB}}; UB = \frac{\hat{RR}_{UB} - 1}{\hat{RR}_{UB}}$
Prevented Fraction in the Population	$P\hat{F}_p = \frac{\hat{R}_u - \hat{R}}{\hat{R}_u}$		
Prevented Fraction in the Exposed	$P\hat{F}_e = \frac{\hat{R}_u - \hat{R}_e}{\hat{R}_u}$	Based on variance estimate for the RR	$LB = 1 - \hat{RR}_{UB}; UB = 1 - \hat{RR}_{LB}$
<i>Parameters based on the odds and odds ratio (from randomized trials, cohort studies, case-control, or cross-sectional studies)</i>			
Odds Ratio	$\hat{OR} = \frac{ad}{bc}$	$\hat{Var}(\ln \hat{OR}) \approx \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$	$\hat{OR} \exp \left[ \pm Z_{1-\alpha/2} \sqrt{\hat{Var}(\ln \hat{OR})} \right]$
Etiologic Fraction in the Population	$E\hat{F}_{pOR} = \frac{p'(OR - 1)}{p'(OR - 1) + 1}$		
Etiologic Fraction in the Exposed	$E\hat{F}_{eOR} = \frac{\hat{OR} - 1}{\hat{OR}}$	Based on variance estimate for the OR	$LB = \frac{\hat{OR}_{LB} - 1}{\hat{OR}_{LB}}; UB = \frac{\hat{OR}_{UB} - 1}{\hat{OR}_{UB}}$
Prevented Fraction in the Population	$P\hat{F}_{pOR} = p'(1 - \hat{OR})$		
Prevented Fraction in the Exposed	$P\hat{F}_{eOR} = 1 - \hat{OR}$	Based on variance estimate for the OR	$LB = 1 - \hat{OR}_{UB}; UB = 1 - \hat{OR}_{LB}$

LB=lower bound; UB=upper bound

P'=...

To work through an example of the calculations, a study was performed in children 12-23.9 months of age. In this study, the prevalence of anemia was estimated. The results are shown in Table 15-2.

**Table 15-2.** Example data; prevalence of anemia in children 12-23.9 months of age by sex

	Male	Female	
Anemic	205	129	334
Not Anemic	89	86	175
	294	215	509

The **prevalence estimates** are:

Prevalence in males = 205/294 = 0.697 or 69.7%

Prevalence in females = 129/215 = 0.600 or 60.0%

The **Prevalence Ratio estimate** is as follows (using the formulae for the risk ratio):

Prevalence ratio = .697/.600 = 1.16

Variance of the prevalence ratio =

$$\frac{1-.697}{294*.697} + \frac{1-.6}{215*.6} = .004579$$

95% confidence interval; replace the Z value in the formula to 1.96 for the calculation of a two-sided 95% confidence interval (for a 90 confidence interval, the Z value is 1.645, and for a 99% confidence, 2.576):

$$1.16\exp[\pm 1.96\sqrt{.004579}] = 1.16\exp[\pm .132629]$$

(1.02, 1.32)

The interpretation would be that males in this study were 1.16 times more likely to have anemia than females; the 95% confidence interval around this estimate is 1.02, 1.32.

The **Prevalence Difference estimate** is as follows (using the formulae for the risk difference):

$$\text{Prevalence difference} = .697 - .600 = .097 \text{ or } 9.7\%$$

Variance of the prevalence difference =

$$\frac{.697(1-.697)}{294} + \frac{.6(1-.6)}{215} = .0018346$$

95% confidence interval:

$$.097 \pm 1.96\sqrt{.0018346} = .097 \pm 1.96(.04283)$$

(.013, .181) or (1.3%, 18.1%)

The interpretation would be that the prevalence of anemia is 9.7% higher in males compared to females (in terms of an absolute difference), with a 95% confidence interval from 1.3% to 18.1%.

The **odds ratio estimate**, or in this example the *prevalence odds ratio estimates*, is as follows:

$$\text{Odds ratio} = (205*86)/(129*89) = 1.54$$

Variance of the odds ratio =

$$\frac{1}{205} + \frac{1}{129} + \frac{1}{89} + \frac{1}{86} = .0354938$$

95% confidence interval =

$$1.54\exp[\pm 1.96\sqrt{.0354938}] = 1.54\exp[\pm .369258]$$

(1.06, 2.23)

The interpretation would be that the odds of anemia in males is 1.54 times the odds in females with a 95% confidence interval of 1.06 to 2.23. Note that the odds ratio is larger than the risk ratio because the prevalence of anemia is high ( $334/509 = 66\%$ ).

The uncorrected chi-square tests would be calculated as:

$$uncorr\chi_1^2 = \frac{(509)(205 \times 86 - 129 \times 89)^2}{294 \times 215 \times 334 \times 175} = 5.20903$$

which would have a p-value = .022. The Mantel Haenszel chi square would be calculated as:

$$mh\chi_1^2 = \frac{(509-1)(205 \times 86 - 129 \times 89)^2}{294 \times 215 \times 334 \times 175} = 5.19879$$

which would have a p-value of .023. The conclusion would be that there was a statistically significant association between the sex of the child and the prevalence of anemia. Note that the statistical test for a 2x2 table can be used with the risk ratio, risk difference, or odds ratio. Also, it is calculated the same whether the data are from an unmatched case-control study, a cohort study, or a clinical trial.

## FORMULAE AND EXAMPLE FOR STRATIFIED DATA (COUNT DATA)

For stratified analyses, the same calculations for the crude table can be used for stratum-specific estimates. For adjusted or summary estimates, a slightly different notation is used as shown in Table 15-2. In this table, the subscript *i* to denote estimates from a specific stratum. The general approach for *adjusted* point estimates is to weight each of the stratum-specific estimates by a weighting method and then sum the results.

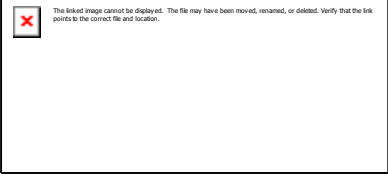

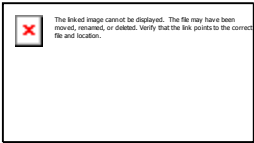

**Table 15-2.** Notation and table setup for stratified 2x2 tables

	Exposed	Nonexposed	
Disease	$a_i$	$b_i$	$m_{1i}$
No Disease	$c_i$	$d_i$	$m_{0i}$
	$n_{1i}$	$n_{0i}$	$n_i$

For the risk ratio and the odds ratio, two different approaches are given for estimating the adjusted point estimate and confidence interval, one referred to as the *directly* adjusted ratio and the other

referred to as the *Mantel-Haenszel* adjusted ratio. The directly adjusted approach requires “large” numbers in each stratum. The weights for directly adjusted values are the inverse of the variance; this approach provides a greater weight to strata with the least amount of variance and less weight to strata with a large variance. The

Mantel-Haenszel method works better when data are sparse.

Parameter	Point Estimate	Confidence Interval
Risk Ratio – Directly Adjusted	 <p>where</p>  	$\hat{RR}_{Direct} \exp \left( \pm \frac{Z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^s w_i}} \right)$
Risk Ratio – Mantel-Haenszel Adjusted	$\hat{RR}_{MH} = \frac{\sum_{i=1}^s \frac{a_i n_{oi}}{n_i}}{\sum_{i=1}^s \frac{b_i n_{li}}{n_i}}$	 <p>where</p> $SE[\ln \hat{RR}_{MH}] = \sqrt{\frac{\sum_{i=1}^s (m_{1i} n_{1i} n_{0i} - a_i b_i n_i) / n_i^2}{\left[ \sum_{i=1}^s \frac{a_i n_{oi}}{n_i} \right] \left[ \sum_{i=1}^s \frac{b_i n_{li}}{n_i} \right]}}$
Risk Difference – Directly Adjusted	$\hat{RD}_{Direct} = \frac{\sum_{i=1}^s w_i \hat{RD}_i}{\sum_{i=1}^s w_i}$ <p>where</p> $w_i = \frac{1}{\frac{a_i c_i}{n_{li}^3} + \frac{b_i d_i}{n_{oi}^3}}, \hat{RD}_i = \frac{a_i}{n_{li}} - \frac{b_i}{n_{oi}}$	$\hat{RD}_{Direct} \pm \frac{Z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^s w_i}}$
Odds Ratio – Directly Adjusted	$\hat{OR}_{Direct} = \exp \left( \frac{\sum_{i=1}^s w_i \ln(\hat{OR}_i)}{\sum_{i=1}^s w_i} \right)$ <p>where</p> $\hat{OR}_i = \frac{a_i d_i}{b_i c_i}, w_i = \frac{1}{\frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}}$	$\hat{OR}_{Direct} \exp \left( \pm \frac{Z_{1-\alpha/2}}{\sqrt{\sum_{i=1}^s w_i}} \right)$
Odds Ratio – Mantel-Haenszel Adjusted	$\hat{OR}_{MH} = \frac{\sum_{i=1}^s \frac{a_i d_i}{n_i}}{\sum_{i=1}^s \frac{b_i c_i}{n_i}}$	$\hat{OR}_{MH} \exp(\pm Z_{1-\alpha/2} SE[\ln \hat{OR}_{MH}])$ <p>where</p>

		$\text{SE}[\ln \hat{O}R_{\text{MH}}] =$ $\sqrt{\frac{\sum_{i=1}^s P_i R_i}{2 \left[ \sum_{i=1}^s R_i \right]^2} + \frac{\sum_{i=1}^s (P_i S_i + Q_i R_i)}{2 \sum_{i=1}^s R_i \sum_{i=1}^s S_i} + \frac{\sum_{i=1}^s Q_i S_i}{2 \left[ \sum_{i=1}^s S_i \right]^2}}$ $P_i = (a_i + d_i) / n_i$ $Q_i = (b_i + c_i) / n_i$ $R_i = a_i d_i / n_i$ $S_i = b_i c_i / n_i$

### Tests for Interaction for the Risk Ratio, Risk Difference, and the Odds Ratio

The tests for interaction presented here are generally referred to as the “Breslow-Day test of homogeneity” and are based on a chi square test.

The test for interaction for the risk ratio is:

$$\chi^2_{s-1} = \sum_{i=1}^S \frac{[\ln(\hat{RR}_i) - \ln(\hat{RR}_{Direct})]^2}{\hat{Var}[\ln(\hat{RR}_i)]}$$

where the  $\text{Var}[\ln(\hat{RR}_i)] = 1/w_i$  from the direct RR point estimate calculation.

The test for interaction for the risk difference is

$$\chi^2_{s-1} = \sum_{i=1}^S \frac{[\hat{RD}_i - \hat{RD}_{Direct}]^2}{\hat{Var}(\hat{RD}_i)}$$

where the  $\text{Var}(\hat{RD}_i) = 1/w_i$  from the direct RD point estimate calculation.

To test for interaction for the odds ratio (OR), the chi square test is calculated as:

$$\chi^2_{s-1} = \sum_{i=1}^S \frac{[\ln(\hat{OR}_i) - \ln(\hat{OR}_{Direct})]^2}{\hat{Var}[\ln(\hat{OR}_i)]}$$

where the  $\text{Var}[\ln(\hat{OR}_i)] = 1/w_i$  from the direct OR point estimate calculation.

### Summary Statistical Test

A statistical test to assess whether there is a statistically significant association between the exposure and outcome variable controlling for the third variable is the Mantel-Haenszel uncorrected chi-square test. This statistic would be used only if it was decided that there was no statistically significant interaction.

$$\chi^2_1 = \frac{\left( \sum_{i=1}^S \frac{a_i d_i - b_i c_i}{n_i} \right)^2}{\sum_{i=1}^S \frac{n_{1i} n_{0i} m_{1i} m_{0i}}{(n_i - 1) n_i^2}}$$

An example of the calculations for stratified data are provided next. Continuing on with the example in table 15-3 on the association between sex and anemia in children,

the data are stratified on mothers education level. Again, because the data were based on prevalent cases, the term “prevalent” will be used rather than “risk.”

**Table 15-3.** Example data; prevalence of anemia in children 12-23.9 months of age by sex stratified on mothers education level.

Mother has <b>low</b> level of education				
	Male	Female		
Anemic	66	36	102	
Not Anemic	28	32	60	
	94	68	162	

Mother has <b>high</b> level of education				
	Male	Female		
Anemic	139	93	232	
Not Anemic	61	54	115	
	200	147	347	

Calculation of the **directly adjusted prevalence ratio** and its 95% confidence interval is shown in Table 15-4.

**Table 15-4.** Calculations for computing directly adjusted prevalence (risk) ratio

Stratum	PR <sub>i</sub>	ln(PR <sub>i</sub> )	w <sub>i</sub>	w <sub>i</sub> ln(PR <sub>i</sub> )
1	1.326	.2821669	56.86628	16.04578
2	1.099	.0944001	162.75481	15.36407
Sum			219.62109	31.40985

The calculated point estimate is:

$$\hat{PR}_{Direct} = \exp\left[\frac{31.40985}{219.62109}\right] = 1.154$$

The 95% confidence interval is:

$$1.154 \exp\left(\pm \frac{1.96}{\sqrt{219.62109}}\right) = 1.154 \exp(\pm .132257)$$

(1.011, 1.317)

The interpretation would be that males were 1.154 times more likely to be anemic than females *controlling* or *adjusting* for the mother’s education level. In addition, we are 95% confident that the true prevalence ratio is captured between 1.011 and 1.317. However, we must still calculate the test for interaction to see if the mother’s education level modifies the sex-anemia relationship. To calculate the test for interaction, the directly adjusted risk ratio needs to be calculated beforehand. Also, note that



$$\text{var}(\ln \hat{P}R_i) = \frac{1}{w_i}$$

Therefore, the **test for interaction for the prevalence/risk ratio** would be:

$$\chi^2_{s-1} = \frac{[.2821669 - .143234]^2}{.017585} + \frac{[.0944001 - .143234]^2}{.0061442}$$

$$\chi^2_{s-1} = 1.097660 + .388130 = 1.485790$$

The p-value for the chi square would be calculated for a chi square value of 1.486 with one degree of freedom (the degrees of freedom is determined from the number of strata minus 1). The p-value from this example is .223. Therefore, we would state that the mother's education level does not significantly modify the sex-anemia relationship. Therefore, the next question is whether the mother's education level confounds the relationship. The crude prevalence ratio was 1.16 and the directly adjusted value was 1.15, which is less than a 1% difference, therefore the conclusion would be that mother's education does not modify nor confound the sex-anemia relationship.

The calculation of the **directly adjusted Mantel-Haenszel prevalence ratio** and its 95% confidence interval is shown in Table 15-5.

**Table 15-5.** Calculations for computing the Mantel-Haenszel prevalence (risk) ratio

Stratum	$a_i n_{0i} / n_i$	$b_i n_{1i} / n_i$	$(m_{1i} n_{1i} n_{0i} - a_i b_i n_i) / n_i^2$
1	27.7037	20.8889	10.17650
2	58.8847	53.6023	19.3933
Sum	86.5884	74.4912	29.5698

The point estimate is

$$\hat{P}R_{MH} = \frac{86.5884}{74.4912} = 1.162$$

To calculate the 95% confidence interval we will first calculate the standard error of the estimate:

$$SE[\ln \hat{P}R_{MH}] = \sqrt{\frac{29.5698}{86.5884 * 74.4912}} = .067708$$

The 95% confidence interval is calculated as:

$$1.162 \exp(\pm 1.96 [.067708]) = 1.162 \exp(\pm .132708)$$

$$(1.018, 1.327)$$

Previously we found that mother's education did not modify the sex-anemia relationship, therefore the interpretation would be that, controlling for mother's education, males were 1.162

times more likely to be anemic than females. However, because there is little confounding (the crude value is 1.15), there is no need to control for mother's education level.

Calculation of the **directly adjusted prevalence difference** and its 95% confidence interval is shown in Table 15-6.

**Table 15-6.** Calculations for computing the direct adjusted prevalence (risk) difference

Stratum	PD <sub>i</sub>	w <sub>i</sub>	w <sub>i</sub> PD <sub>i</sub>
1	0.1727	169.8171	29.3274
2	0.0623	378.6661	23.5909
Sum		548.4832	52.9183

The point estimate is:

$$\hat{PD}_{Direct} = \frac{52.9183}{548.4832} = .0965$$

and the 95% confidence interval is:

$$.0965 \pm \frac{1.96}{\sqrt{548.4832}} = .0965 \pm .0837$$

(.0128, .1802)

Depending on the frequency of disease, it may be useful to describe the difference in term of per 100 individuals (or percent), per 1,000, or some other unit. In this example, the males had a prevalence of anemia 9.7% higher (in absolute terms) than females controlling for maternal education, and we are 95% confident that the truth is captured between 1.3% and 18.1%. However, before the decision is made as to whether or not to present the adjusted difference, the test for interaction should be calculated. Again, note that:

$$\text{var}(\ln \hat{PD}_i) = \frac{1}{w_i}$$

Therefore, the **test for interaction for prevalence/risk differences** would be:

$$\chi^2_{s-1} = \frac{[0.1727 - .0965]^2}{.00589} + \frac{[0.0623 - .0965]^2}{.00264} =$$

$$.98581 + .44305 = 1.42886$$

The chi square value of 1.42886 with one degree of freedom would have a p-value of .232, which would not be statistically significant. The next step would be to determine whether mother's education confounds the sex-anemia relationship. The crude prevalence difference was .097, the same as the adjusted difference, which would lead to the conclusion that there is no important confounding in this analysis.

Calculation of the **directly adjusted (prevalence) odds ratio** and its 95% confidence interval is shown in Table 15-7.

**Table 15-7.** Calculations for computing the direct adjusted (prevalence) odds ratio

Stratum	OR <sub>i</sub>	ln(OR <sub>i</sub> )	w <sub>i</sub>	w <sub>i</sub> ln(OR <sub>i</sub> )
1	2.095	.73955	9.09971	6.72969
2	1.323	.27990	18.91829	5.29523
Sum			28.01800	12.02492

The calculated point estimate is:

$$\hat{OR}_{Direct} = \exp\left[\frac{12.02492}{28.018}\right] = 1.536$$

The 95% confidence interval is:

$$1.536 \exp\left(\pm \frac{1.96}{\sqrt{28.018}}\right) = 1.536 \exp(\pm .370286)$$

(1.061, 2.224)

The interpretation would be that odds of anemia in males was 1.536 times the odds of anemia in females *controlling* or *adjusting* for the mother's education level. In addition, we are 95% confident that the true prevalence odds ratio is captured between 1.061 and 2.224. However, we must still calculate the test for interaction to see if the mother's education level modifies the sex-anemia relationship. To calculate the test for interaction, the directly adjusted odds ratio needs to be calculated beforehand. Also, note that

$$\text{var}(\ln \hat{OR}_i) = \frac{1}{w_i}$$

Therefore, the **test for interaction for the (prevalence) odds ratio** would be:

$$\chi^2_{s-1} = \frac{[.73955 - .42918]^2}{.10989} + \frac{[.27990 - .42918]^2}{.05286}$$

$$\chi^2_{s-1} = .87660 + .42158 = 1.2918$$

The p-value for the chi square would be calculated for a chi square value of 1.2918 with one degree of freedom (the degrees of freedom is determined from the number of strata minus 1). The p-value from this example is .256. Therefore, we would state that the mother's education level does not significantly modify the sex-anemia relationship. Therefore, the next question is whether the mother's education level confounds the relationship. The crude prevalence odds ratio was 1.536 and the directly adjusted value was the same, the conclusion would be that, based on the odds ratio, mother's education does not modify nor confound the sex-anemia relationship.

Calculation of the **Mantel-Haenszel adjusted (prevalence) odds ratio** and its 95% confidence interval is as follows. The values that need to be calculated are shown in Table 15-8.

To calculate the point estimate and the confidence interval, eight values in Table 15-8 need to be calculated.

The calculated point estimate is:

$$\hat{OR}_{MH} = \frac{34.66816}{22.57092} = 1.536$$

The standard error of the natural log of the point estimate is calculated as:

**Table 15-8.** Calculations for computing the Mantel-Haenszel adjusted (prevalence) odds ratio

Stratum	P <sub>i</sub>	Q <sub>i</sub>	R <sub>i</sub>	S <sub>i</sub>
1	.60494	.39506	13.03704	6.22222
2	.55620	.44380	21.63112	16.34870
Sum	1.16114	.83886	34.66816	22.57092

Stratum	P <sub>i</sub> R <sub>i</sub>	P <sub>i</sub> S <sub>i</sub>	Q <sub>i</sub> R <sub>i</sub>	Q <sub>i</sub> S <sub>i</sub>
1	7.88663	3.76407	5.15041	2.45815
2	12.03123	9.09315	9.59999	7.25555
Sum	19.91786	12.85722	14.75040	9.71370

$$SE[\ln \hat{OR}_{MH}] =$$

$$\sqrt{\frac{19.91786}{2[34.66816]^2} + \frac{12.85722 + 14.75040}{2 \times 34.66816 \times 22.57092} + \frac{9.71370}{2[22.57092]^2}}$$

$$= \sqrt{.00829 + .01764 + .00953} = .18831$$

The confidence interval based on the Robins, Greenland, Breslow method is:

$$1.536 \exp(\pm 1.96[.18831])$$

$$1.536 \exp(\pm .36908)$$

The 95% confidence interval is

$$(1.062, 2.222)$$

Previously we found that mother's education did not modify the sex-anemia relationship, therefore the interpretation would be that, controlling for mother's education, the odds of males having anemia were 1.536 times more likely to be anemic than the odds in females. However, because there is little or no confounding (the crude value is 1.536), there is no need to control for mother's education level.

The overall Mantel-Haenszel uncorrected chi-square test would be calculated as with the intermediate calculations shown in Table 15-9.

**Table 15-9.** Calculations for computing the Mantel-Haenszel uncorrected chi-square test

Stratum	$(a_id_i - b_ic_i)/n_i$	$(n_{1i}n_{0i}m_{1i}m_{0i})/[(n_i - 1)n_i^2]$
1	6.81481	9.25832
2	5.28242	18.82774
Sum	12.09723	28.08606

Therefore

$$\chi_1^2 = \frac{(12.09723)^2}{28.08606} = 5.2105$$

which would have a p-value of .022.