This method of standardization is sometimes referred to as **direct standardization.** The use of age-standardized risks is somewhat controversial because results may differ depending on which standard is used. However, space limitations often make it impossible to present age-specific results in a paper, and the reader can get a quick summary of the overall results from the age-standardized risks.

The use of standardized risks is a good descriptive tool for controlling for confounding. In the next section, we will discuss how to control for confounding in assessing disease–exposure relationships in a hypothesis-testing framework using the Mantel-Haenszel test. Finally, standardization can be performed based on stratification by factors other than age. For example, standardization by both age and sex is common. Similar methods can be used to obtain age–sex standardized risks, and standardized risk ratios as given in Definition 13.14.

In this section, we have introduced the concept of a confounding variable (C) which is a variable related to both the disease (D) and exposure (E) variables. Furthermore, we classified confounding variables as positive confounders if the directions of association between C versus D and C versus E, respectively, are in the same direction and as negative confounders if the associations between C versus D and C versus E are in opposite directions. We also discussed when it is or is not appropriate to control for a confounder, according to whether C is or is not in the causal pathway between E and D. Finally, because age is often an important confounding variable, it is reasonable to consider descriptive measures of proportions and relative risk that control for age. Age-standardized proportions and risk ratios are such measures.

SECTION 13.5 Methods of Inference for Stratified Categorical Data—The Mantel-Haenszel Test

Example 13.19

Cancer A 1985 study identified a group of 518 cancer cases ages 15–59 and a group of 518 age- and sex-matched controls by mail questionnaire [4]. The main purpose of the study was to look at the effect of passive smoking on cancer risk. In the study, passive smoking was defined as exposure to the cigarette smoke of a spouse who smoked at least one cigarette per day for at least 6 months. One potential confounding variable was smoking by the test subjects themselves (i.e., personal smoking), because personal smoking is related to both cancer risk and spouse smoking. Therefore, it was important to control for personal smoking before looking at the relationship between passive smoking and cancer risk.

To display the data, a 2×2 table relating case–control status to passive smoking can be constructed for both nonsmokers and smokers. The data are given in Table 13.7 for nonsmokers and Table 13.8 for smokers.

The passive-smoking effect can be assessed separately for nonsmokers and smokers. Indeed, we notice from Tables 13.7 and 13.8 that the odds ratio in favor of a case being exposed to cigarette smoke from a spouse who smokes versus a control is $(120 \times 155)/(80 \times 111) = 2.1$ for nonsmokers, whereas the corresponding odds ratio for smokers is $(161 \times 124)/(130 \times 117) = 1.3$. Thus for both subgroups the trend is in the direction of more passive smoking among cases than controls. The

TABLE 13.7 **Relationship of passive smoking to cancer risk among nonsmokers**

	Passive smoker		
Case–control status	Yes	No	Total
Case	120	111	231
Control	80	155	235
Total	200	266	466

Source: Reprinted with permission of the *American Journal of Epidemiology, 121*(1), 37–48, 1985.

TABLE 13.8 **Relationship of passive smoking to cancer risk among smokers**

	Passive smoker		
Case–control status	Yes	No	Total
Case	161	117	278
Control	130	124	254
Total	291	241	532

Source: Reprinted with permission of the *American Journal of Epidemiology, 121*(1), 37–48, 1985.

key question is how to combine the results of the two tables to obtain an overall test of significance for the passive-smoking effect.

In general, the data will be stratified into k subgroups according to one or more confounding variables to make the units within a stratum as homogeneous as possible. The data for each stratum consist of a 2×2 contingency table relating exposure to disease, as shown in Table 13.9 for the ith stratum.

TABLE 13.9 **Relationship of disease to exposure in the ith stratum**

		Exposure		
		Yes	No	Total
Disease	Yes	a_i	b_i	$a_i + b_i$
	No	c_i	d_i	$c_i + d_i$
		$a_i + c_i$	$b_i + d_i$	n_i

Based on our work on Fisher's exact test, the distribution of a_i follows a **hypergeometric distribution.** The test procedure will be based on a comparison of the observed number of units in the (1, 1) cell of each stratum (denoted by $O_i = a_i$) with the expected number of units in that cell (denoted by E_i). The test procedure is the same regardless of the order of the rows and columns; that is, which row (or column) is designated as the first row (or column) is arbitrary. Based on the hypergeometric distribution (Equation 10.9), the expected number of units in the (1, 1) cell of the ith stratum is given by

EQUATION 13.12

$$E_i = \frac{(a_i + b_i)(a_i + c_i)}{n_i}$$

The observed and expected numbers of units in the (1, 1) cell are then summed over all strata, obtaining $O = \sum_{i=1}^{k} O_i$, $E = \sum_{i=1}^{k} E_i$, and the test is based on $O - E$. Based on the hypergeometric distribution (Equation 10.9), the variance of O_i is given by

EQUATION 13.13

$$V_i = \frac{(a_i + b_i)(c_i + d_i)(a_i + c_i)(b_i + d_i)}{n_i^2(n_i - 1)}$$

Furthermore, the variance of $O = V = \sum_{i=1}^{k} V_i$. The test statistic is given by $X^2_{MH} = \left(|O - E| - .5\right)^2 / V$, which should follow a chi-square distribution with 1 *df* under the null hypothesis of no association between disease and exposure. H_0 is rejected if X^2_{MH} is large. The abbreviation *MH* refers to Mantel-Haenszel; this procedure is known as the Mantel-Haenszel test and is summarized as follows:

EQUATION 13.14

Mantel-Haenszel Test To assess the association between a dichotomous disease and a dichotomous exposure variable after controlling for one or more confounding variables, use the following procedure:

(1) Form k strata, based on the level of the confounding variable(s), and construct a 2×2 table relating disease and exposure within each stratum, as shown in Table 13.9.

(2) Compute the total observed number of units (O) in the (1, 1) cell over all strata, where

$$O = \sum_{i=1}^{k} O_i = \sum_{i=1}^{k} a_i$$

(3) Compute the total expected number of units (E) in the (1, 1) cell over all strata, where

$$E = \sum_{i=1}^{k} E_i = \sum_{i=1}^{k} \frac{(a_i + b_i)(a_i + c_i)}{n_i}$$

(4) Compute the variance (V) of O, where

$$V = \sum_{i=1}^{k} V_i = \sum_{i=1}^{k} \frac{(a_i + b_i)(c_i + d_i)(a_i + c_i)(b_i + d_i)}{n_i^2(n_i - 1)}$$

(5) The test statistic is then given by

$$X^2_{MH} = \frac{\left(|O - E| - .5\right)^2}{V}$$

which under H_0 follows a chi-square distribution with 1 *df*.

(6) For a two-sided test with significance level α, if

$$X^2_{MH} > \chi^2_{1,1-\alpha}$$

then reject H_0. If $X^2_{MH} \leq \chi^2_{1,1-\alpha}$

then accept H_0.

(7) The exact *p*-value for this test is given by

$$p = Pr\left(\chi^2_1 > X^2_{MH}\right)$$

(8) Use this test only if the variance $V \geq 5$.

(9) Which row or column is designated as first is arbitrary. The test statistic X^2_{MH} and the assessment of significance are the same regardless of the order of the rows and columns.

The acceptance and rejection regions for the Mantel-Haenszel test are depicted in Figure 13.1. The computation of the *p*-value for the Mantel-Haenszel test is illustrated in Figure 13.2.

Example 13.20 **Cancer** Assess the relationship between passive smoking and cancer risk using the data stratified by personal smoking status in Tables 13.7 and 13.8.

Solution Denote the nonsmokers as stratum 1 and the smokers as stratum 2.

O_1 = observed number of nonsmoking cases who are passive smokers = 120

O_2 = observed number of smoking cases who are passive smokers = 161

FIGURE 13.1 **Acceptance and rejection regions for the Mantel-Haenszel test**

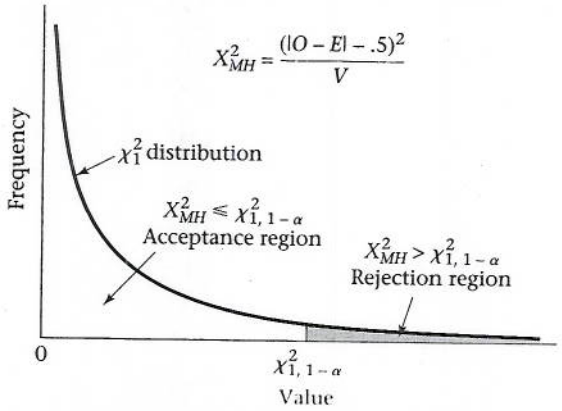
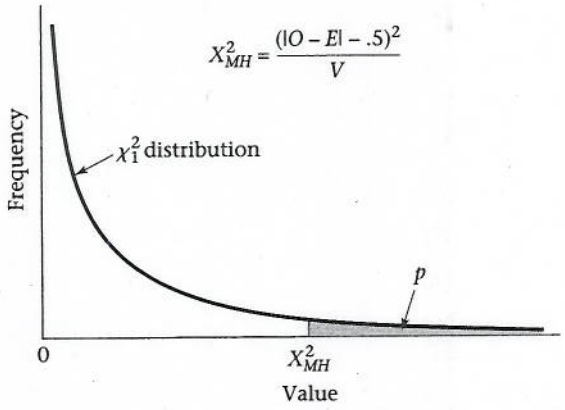


FIGURE 13.2 Computation of the *p*-value for the Mantel-Haenszel test



Furthermore,

$$E_1 = \frac{231 \times 200}{466} = 99.1$$

$$E_2 = \frac{278 \times 291}{532} = 152.1$$

Thus, the total observed and expected numbers of cases who are passive smokers are, respectively,

$$O = O_1 + O_2 = 120 + 161 = 281$$

$$E = E_1 + E_2 = 99.1 + 152.1 = 251.2$$

Therefore, there are more cases who are passive smokers than would be expected based on their personal smoking habits. Now compute the variance to assess if this difference is statistically significant.

$$V_1 = \frac{231 \times 235 \times 200 \times 266}{466^2 \times 465} = 28.60$$

$$V_2 = \frac{278 \times 254 \times 291 \times 241}{532^2 \times 531} = 32.95$$

Therefore, $V = V_1 + V_2 = 28.60 + 32.95 = 61.55$

Thus, the test statistic X^2_{MH} is given by

$$X^2_{MH} = \frac{(|281 - 251.2| - .5)^2}{61.55} = \frac{858.17}{61.55} = 13.94 \sim \chi^2_1 \text{ under } H_0$$

Since $\chi^2_{1,.999} = 10.83 < 13.94 = X^2_{MH}$, it follows that $p < .001$. Thus, there is a highly significant positive association between case–control status and passive-smoking exposure, even after controlling for personal cigarette-smoking habit.

13.5.1 Estimation of the Odds Ratio for Stratified Data

The Mantel-Haenszel test provides a test of significance of the relationship between disease and exposure. However, it does not give a measure of the strength of the association. Ideally, we would like a measure similar to the odds ratio presented for a single 2×2 contingency table in Definition 13.6. Assuming that the underlying odds ratio is the same for each stratum, an estimate of the common underlying odds ratio is provided by the Mantel-Haenszel estimator as follows:

EQUATION 13.15

Mantel-Haenszel Estimator of the Common Odds Ratio for Stratified Data In a collection of k 2×2 contingency tables, where the ith table corresponding to the ith stratum is denoted as in Table 13.9, the Mantel-Haenszel estimator of the common odds ratio is given by

$$\hat{OR}_{MH} = \frac{\sum_{i=1}^{k} a_i d_i / n_i}{\sum_{i=1}^{k} b_i c_i / n_i}$$

Example 13.21

Cancer Estimate the odds ratio in favor of being a passive smoker for cancer cases versus controls after controlling for personal smoking habit.

Solution

From Equation 13.15, Table 13.7, and Table 13.8,

$$\hat{OR}_{MH} = \frac{(120 \times 155/466) + (161 \times 124/532)}{(80 \times 111/466) + (130 \times 117/532)} = \frac{77.44}{47.65} = 1.63$$

Thus, the odds in favor of being a passive smoker for a cancer case is 1.6 times as large as that for a control. Since cancer is a relatively rare disease, we can also interpret these results as indicating that the risk of cancer for a passive smoker is 1.6 times as great as for a nonpassive smoker, even after controlling for personal smoking habit.

We are also interested in estimating confidence limits for the odds ratio in Equation 13.15. A variance estimate of $\ln(\hat{OR}_{MH})$ has been provided by Robins et al. [5], which is accurate under a wide range of conditions, particularly if there are many strata with small numbers of subjects in each stratum. This variance estimate can be used to obtain confidence limits for $\ln(OR)$. We can then take the antilog of each of the confidence limits for $\ln(OR)$ to obtain confidence limits for OR. This procedure is summarized as follows:

EQUATION 13.16

Interval Estimate for the Common Odds Ratio from a Collection of k 2×2 Contingency Tables A two-sided $100\% \times (1 - \alpha)$ CI for the common odds ratio from a collection of k 2×2 tables is given by

$$\exp\left[\ln \hat{OR}_{MH} \pm z_{1-\alpha/2} \sqrt{Var(\ln \hat{OR}_{MH})}\right]$$

where

$$Var(\ln \hat{OR}_{MH}) = \frac{\sum_{i=1}^{k} P_i R_i}{2\left(\sum_{i=1}^{k} R_i\right)^2} + \frac{\sum_{i=1}^{k}\left(P_i S_i + Q_i R_i\right)}{2\left(\sum_{i=1}^{k} R_i\right)\left(\sum_{i=1}^{k} S_i\right)} + \frac{\sum_{i=1}^{k} Q_i S_i}{2\left(\sum_{i=1}^{k} S_i\right)^2} \equiv A + B + C$$

where A, B, and C correspond to the first, second, and third terms on the right-hand side of $Var(\ln \hat{OR}_{MH})$, and

$$P_i = \frac{a_i + d_i}{n_i},\ Q_i = \frac{b_i + c_i}{n_i},\ R_i = \frac{a_i d_i}{n_i},\ S_i = \frac{b_i c_i}{n_i}$$

Example 13.22 **Cancer** Estimate 95% confidence limits for the common odds ratio using the data in Tables 13.7 and 13.8.

Solution Note from Example 13.21 that the point estimate of the odds ratio = $\hat{OR}_{MH} = 1.63$. To obtain confidence limits, we first compute P_i, Q_i, R_i, and S_i as follows

$$P_1 = \frac{120 + 155}{466} = .590, Q_1 = 1 - P_1 = .410$$

$$R_1 = \frac{120(155)}{466} = 39.91, S_1 = \frac{80(111)}{466} = 19.06$$

$$P_2 = \frac{161 + 124}{532} = .536, Q_2 = 1 - P_2 = .464$$

$$R_2 = \frac{161(124)}{532} = 37.53, S_2 = \frac{130(117)}{532} = 28.59$$

Thus,

$$Var(\ln OR_{MH}) = A + B + C$$

where

$$A = \frac{.590(39.91) + .536(37.53)}{2(39.91 + 37.53)^2} = 0.00364$$

$$B = \frac{.590(19.06) + .410(39.91) + .536(28.59) + .464(37.53)}{2(39.91 + 37.53)(19.06 + 28.59)} = 0.00818$$

$$C = \frac{.410(19.06) + .464(28.59)}{2(19.06 + 28.59)^2} = 0.00464$$

Thus, $Var(\ln \hat{OR}_{MH}) = 0.00364 + 0.00818 + 0.00464 = 0.01646$. The 95% CI for $\ln(OR)$ is

$$\ln(1.63) \pm 1.96\sqrt{0.01646} = (0.234,\ 0.737)$$

The 95% CI for OR is

$$(e^{0.234},\ e^{0.737}) = (1.26,\ 2.09)$$

13.5.2 Effect Modification

One assumption made in the estimation of a common odds ratio in Equation 13.15 is that the strength of association is the same in each stratum. If the underlying odds ratio is different in the various strata, then it makes little sense to estimate a common odds ratio.

DEFINITION 13.15

Suppose we are interested in studying the association between a disease variable D and an exposure variable E, but are concerned about the possible confounding effect of another variable C. We stratify the study population into g strata according to the variable C and compute the odds ratio relating disease to exposure in each stratum. If the underlying (true) odds ratio is different across the g strata, then there is said to be **interaction** or **effect modification** between E and C, and the variable C is referred to as an **effect modifier.**

In other words, if C is an effect modifier, then the relationship between disease and exposure differs for different levels of C.

Example 13.23

Cancer Consider the data in Tables 13.7 and 13.8. We estimated that the odds ratio relating cancer and passive smoking is 2.1 for nonsmokers and 1.3 for smokers. If these were the underlying odds ratios in these strata, then personal smoking would be an effect modifier. Specifically, the relationship between passive smoking and cancer is much stronger for nonsmokers than for smokers. The rationale for this is that the home environment of active smokers already contains cigarette smoke and the extra degradation of the environment by spousal smoking may not be that meaningful.

The issue remains, how can we detect if another variable C is an effect modifier? We will use a generalization of the Woolf procedure for obtaining confidence limits for a single odds ratio given in Equation 13.11. Specifically, we wish to test the hypothesis H_0: $OR_1 = \cdots = OR_k$ versus H_1: at least two of the OR_i are different from each other. We will base our test on the test statistic $X^2 = \sum_{i=1}^{k} w_i(\ln \widehat{OR}_i - \overline{\ln OR})^2$ where $\ln \widehat{OR}_i$ = the estimated log odds ratio relating disease to exposure in the ith stratum of the potential effect modifier, C, $\overline{\ln OR}$ = the estimated "weighted average" log odds ratio over all strata, and w_i is a weight that is inversely proportional to the variance of $\ln \widehat{OR}_i$. The purpose of the weighting is to weight strata with lower variance (which usually correspond to strata with more subjects) more heavily. If H_0 is true, then X^2 will be small, because each of the stratum-specific log odds ratios will be relatively close to each other and to the "average" log odds ratio. Conversely, if H_1 is true, then X^2 will be large. Under H_0, it can be shown that X^2 follows a chi-square distribution with $k - 1$ *df*. Thus, we will reject H_0 if $X^2 > \chi^2_{k-1,1-\alpha}$ and accept H_0 otherwise. This procedure is summarized as follows:

EQUATION 13.17

Chi-Square Test for Homogeneity of Odds Ratios over Different Strata (Woolf Method) Suppose we have a dichotomous disease variable D and exposure variable E. We stratify our study population into k strata according to a

confounding variable C. Let OR_i = underlying odds ratio in the ith stratum. To test the hypothesis H_0: $OR_1 = \cdots = OR_k$ versus H_1: at least two of the OR_i are different with a significance level α, use the following procedure.

(1) Compute the test statistic $X^2_{\text{HOM}} = \sum_{i=1}^{k} w_i(\ln \widehat{OR}_i - \overline{\ln OR})^2 \sim \chi^2_{k-1}$ under H_0 where $\ln \widehat{OR}_i$ = log odds ratio in the ith stratum = $\ln[a_i d_i/(b_i c_i)]$, and a_i, b_i, c_i, d_i are the cells of the 2×2 table relating disease to exposure in the ith stratum as shown in Table 13.9,

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

$$\overline{\ln OR} = \textstyle\sum_{i=1}^{k} w_i \ln \widehat{OR}_i \Big/ \sum_{i=1}^{k} w_i$$

(1a) An alternative computational form of the test statistic is

$$X^2_{\text{HOM}} = \sum_{i=1}^{k} w_i (\ln \widehat{OR}_i)^2 - \left(\sum_{i=1}^{k} w_i \ln \widehat{OR}_i\right)^2 \Big/ \sum_{i=1}^{k} w_i$$

(2) If $X^2_{\text{HOM}} > \chi^2_{k-1,1-\alpha}$, then reject H_0.
If $X^2_{\text{HOM}} \le \chi^2_{k-1,1-\alpha}$, then accept H_0.

(3) The exact p-value = $Pr(\chi^2_{k-1} > X^2)$

Example 13.24 **Cancer** Assess whether the odds ratios relating passive smoking to cancer is different for smokers versus nonsmokers using the data in Tables 13.7 and 13.8.

Solution Let stratum 1 refer to nonsmokers and stratum 2 to smokers. Referring to Tables 13.7 and 13.8, we see that

$$\ln \widehat{OR}_1 = \ln\left(\frac{120 \times 155}{80 \times 111}\right) = \ln(2.095) = 0.739$$

$$w_1 = \left(\frac{1}{120} + \frac{1}{111} + \frac{1}{80} + \frac{1}{155}\right)^{-1} = (0.036)^{-1} = 27.55$$

$$\ln \widehat{OR}_2 = \ln\left(\frac{161 \times 124}{130 \times 117}\right) = \ln(1.313) = 0.272$$

$$w_2 = \left(\frac{1}{161} + \frac{1}{117} + \frac{1}{130} + \frac{1}{124}\right)^{-1} = (0.031)^{-1} = 32.77$$

Thus, based on step 1a in Equation 13.17, the test statistic is given by

$$\begin{aligned} X^2_{\text{HOM}} &= 27.55(0.739)^2 + 32.77(0.272)^2 - \left[27.55(0.739) + 32.77(0.272)\right]^2 \Big/ (27.55 + 32.77) \\ &= 17.486 - (29.284)^2 / 60.32 \\ &= 17.486 - 14.216 = 3.27 \sim \chi^2_1 \text{ under } H_0 \end{aligned}$$

Referring to Table 6 in the Appendix, we note that $\chi^2_{1,.90} = 2.71$, $\chi^2_{1,.95} = 3.84$. Since $2.71 < 3.27 < 3.84$, it follows that $.05 < p < .10$. Thus, there is no significant effect modification; i.e., the odds ratios in the two strata are not significantly different.

In general, it is important to test for homogeneity of the stratum-specific odds ratios. If the true odds ratios are different, then it makes no sense to obtain a pooled-odds ratio estimate such as given by the Mantel-Haenszel estimator in Equation 13.15. Instead, separate odds ratios should be reported.

13.5.3 Estimation of the Odds Ratio in Matched-Pair Studies

There is a close connection between McNemar's test for matched-pair data in Equation 10.12 and the Mantel-Haenszel test procedure for stratified categorical data in Equation 13.14. Matched pairs are a special case of stratification where each matched pair corresponds to a separate stratum of size 2. It can be shown that McNemar's test is a special case of the Mantel-Haenszel test for strata of size 2. Furthermore, the Mantel-Haenszel odds-ratio estimator in Equation 13.15 reduces to $\widehat{OR} = \frac{n_A}{n_B}$ for matched-pair data where n_A = number of discordant pairs of type A and n_B = number of discordant pairs of type B. Also, it can be shown that the variance of $\ln(\widehat{OR})$ for a matched-pair study is given by $Var[\ln(\widehat{OR})] = \frac{1}{n\hat{p}\hat{q}}$, where n = total number of discordant pairs = $n_A + n_B$, $\hat{p}$ = proportion of discordant pairs of type A = $n_A/(n_A + n_B)$, $\hat{q} = 1 - \hat{p}$. This leads to the following technique for estimating the disease–exposure odds ratio in matched-pair studies.

EQUATION 13.18

Estimation of the Odds Ratio in Matched-Pair Studies Suppose we wish to study the relationship between a dichotomous disease and exposure variable, in a case–control design. We control for confounding by forming matched pairs of subjects with disease (cases) and subjects without disease (controls), where the 2 subjects in a matched pair are the same or similar on one or more confounding variables.

(1) The odds ratio relating disease to exposure is estimated by

$$\widehat{OR} = n_A/n_B$$

where

n_A = number of matched pairs where the case is exposed and the control is not exposed

n_B = number of matched pairs where the case is not exposed and the control is exposed

(2) A two-sided $100\% \times (1 - \alpha)$ CI for OR is given by (e^{c_1}, e^{c_2}), where

$$c_1 = \ln(\widehat{OR}) - z_{1-\alpha/2}\sqrt{\frac{1}{n\hat{p}\hat{q}}}$$

$$c_2 = \ln(\widehat{OR}) + z_{1-\alpha/2}\sqrt{\frac{1}{n\hat{p}\hat{q}}}$$

$$n = n_A + n_B$$

$$\hat{p} = \frac{n_A}{n_A + n_B}, \quad \hat{q} = 1 - \hat{p}$$

(3) The same methodology can be used for prospective or cross-sectional studies where exposed and unexposed individuals are matched on one or more confounding variables and disease outcomes are compared between exposed and unexposed individuals. In this setting,

n_A = number of matched pairs where the exposed subject has disease and the unexposed subject does not

n_B = number of matched pairs where the exposed subject does not have disease and the unexposed subject does and steps 1 and 2 are as just indicated

(4) This method should only be used if n = number of discordant pairs ≥ 20.

Example 13.25 **Cancer** Estimate the odds ratio relating type of treatment to 5-year mortality using the matched-pair data in Table 10.14.

Solution We have from Table 10.14 that

n_A = number of matched pairs where treatment A patient dies within 5 years and treatment B patient survives for 5 years = 5

n_B = number of matched pairs where treatment B patient dies within 5 years and treatment A patient survives for 5 years = 16

Thus, $\widehat{OR} = 5/16 = 0.31$. To obtain 95% confidence limits we see that $n = 21$, $\hat{p} = 5/21 = .238$, $\hat{q} = .762$, and $n\hat{p}\hat{q} = 3.81$. Thus, $\ln(\widehat{OR}) = -1.163$, $Var[\ln(\widehat{OR})] = 1/3.81 = 0.263$ and a 95% CI for $\ln(OR)$ is $(-1.163 - 1.96\sqrt{0.263}, -1.163 + 1.96\sqrt{0.263}) = (-2.167, -0.159)$. The corresponding 95% CI for OR is $(e^{-2.167}, e^{-0.159}) = (0.11, 0.85)$.

13.5.4 Testing for Trend in the Presence of Confounding—Mantel-Extension Test

Example 13.26 **Sleep Disorders** Sleep-disordered breathing is very common among adults. To estimate the prevalence of this disorder, 3513 employees 30–60 years of age who worked for three large state agencies in Wisconsin were sent a mail questionnaire concerning their sleep habits [6]. Subjects were classified as habitual snorers if they reported either (1) snoring, snorting, or breathing pauses every night or almost every night or (2) extremely loud snoring. The results are given by age and sex group in Table 13.10.

TABLE 13.10 **Prevalence of habitual snoring by age and sex group**

	Women			Men		
Age	Yes	No	Total	Yes	No	Total
30–39	196	603	799	188	348	536
40–49	223	486	709	313	383	696
50–60	103	232	335	232	206	438
Total	522	1321	1843	733	937	1670

We would like to assess whether the prevalence of habitual snoring increases with age.

In this study, we would like to assess whether there is a trend in the prevalence rates with age after controlling for sex. To address this issue, we need to generalize the chi-square test for trend given in Equation 10.24 to allow for stratification of our study sample by relevant confounding variables. We can also describe this problem as a generalization of the Mantel-Haenszel test given in Equation 13.14 to the case where we are combining results from several $2 \times k$ tables (rather than just 2×2 tables). Suppose we have s strata and k ordered categories for the exposure variable. Consider the $2 \times k$ table relating the dichotomous disease variable D to the ordered categorical exposure variable E for subjects in the ith stratum (see Table 13.11). We assume that there is a score for the jth exposure category denoted by x_j, $j = 1, \ldots, k$.

TABLE 13.11 **Relationship of disease to exposure in the ith stratum, $i = 1, \ldots, s$**

		Exposure				
		1	2	...	k	
Disease	+	n_{i1}	n_{i2}	...	n_{ik}	n_i
	−	m_{i1}	m_{i2}	...	m_{ik}	m_i
		t_{i1}	t_{i2}	...	t_{ik}	N_i
Score		x_1	x_2	...	x_k	

The total observed score among subjects with disease in the ith stratum = O_i $\sum_{j=1}^{k} n_{ij}x_j$. The expected score among diseased subjects in the ith stratum under the null hypothesis that the average score for subjects with and without disease in a stratum is the same = $E_i = \left(\sum_{j=1}^{k} t_{ij}x_j\right)\frac{n_i}{N_i}$. If diseased subjects tend to have higher exposure scores on average than nondiseased subjects, then O_i will be greater than E_i for most strata. If diseased subjects tend to have lower exposure scores than nondiseased subjects, then O_i will be less than E_i for most strata. Therefore, we will base our test on $O - E$ where $O = \sum_{i=1}^{s} O_i$, $E = \sum_{i=1}^{s} E_i$. The test procedure is given as follows:

EQUATION 13.19

Chi-Square Test for Trend-Multiple Strata (Mantel-Extension Test)

(1) Suppose we have s strata. In each stratum, we have a $2 \times k$ table relating disease (2 categories) to exposure (k ordered categories) with score for the jth category = x_j as shown in Table 13.11.

(2) To test the hypothesis H_0: $\beta = 0$ versus H_1: $\beta \neq 0$, where

$$p_{ij} = \alpha_i + \beta x_j$$

p_{ij} = proportion of subjects with disease among subjects in the ith stratum and jth exposure category

We compute the test statistic

$$X_{TR}^2 = \left(|O - E| - 0.5\right)^2 / V \sim \chi_1^2 \text{ under } H_0$$

where

$$O = \sum_{i=1}^{s} O_i = \sum_{i=1}^{s} \sum_{j=1}^{k} n_{ij} x_j$$

$$E = \sum_{i=1}^{s} E_i = \sum_{i=1}^{s} \left[\left(\sum_{j=1}^{k} t_{ij} x_j \right) \frac{n_i}{N_i} \right]$$

$$V = \sum_{i=1}^{s} V_i = \sum_{i=1}^{s} \frac{n_i m_i \left(N_i s_{2i} - s_{1i}^2 \right)}{N_i^2 \left(N_i - 1 \right)}$$

$$s_{1i} = \sum_{j=1}^{k} t_{ij} x_j, \; i = 1, \ldots, s$$

$$s_{2i} = \sum_{j=1}^{k} t_{ij} x_j^2, \; i = 1, \ldots, s$$

(3) If $X_{TR}^2 > \chi_{1,1-\alpha}^2$, we reject H_0.
If $X_{TR}^2 \leq \chi_{1,1-\alpha}^2$, we accept H_0.

(4) The exact p-value = $Pr(\chi_1^2 > X_{TR}^2)$.

(5) This test should only be used if $V \geq 5$.

Example 13.27

Use the data in Table 13.10 to assess whether the prevalence of habitual snoring increases with age, after controlling for sex.

Solution

In this example, we have two strata, corresponding to women ($i = 1$) and men ($i = 2$), respectively. We will use scores of 1, 2, 3 for the three age groups. We have

$$O_1 = 196(1) + 223(2) + 103(3) = 951$$

$$O_2 = 188(1) + 313(2) + 232(3) = 1510$$

$$O = 951 + 1510 = 2461$$

$$E_1 = \left[799(1) + 709(2) + 335(3)\right] 522/1843 = 912.6$$

$$E_2 = \left[536(1) + 696(2) + 438(3)\right] 733/1670 = 1423.0$$

$$E = 912.6 + 1423.0 = 2335.6$$

$$s_{11} = 799(1) + 709(2) + 335(3) = 3222$$

$$s_{21} = 799(1^2) + 709(2^2) + 335(3^2) = 6650$$

$$s_{12} = 536(1) + 696(2) + 438(3) = 3242$$

$$s_{22} = 536(1^2) + 696(2^2) + 438(3^2) = 7262$$

$$V_1 = \frac{522(1321)\left[1843(6650) - 3222^2\right]}{1843^2(1842)} = 206.61$$

$$V_2 = \frac{733(937)\left[1670(7262) - 3242^2\right]}{1670^2(1669)} = 238.59$$

$$V = 206.61 + 238.59 = 445.21$$

Thus the test statistic is given by

$$X^2_{TR} = \frac{\left(|2461 - 2335.6| - .5\right)^2}{445.21} = \frac{124.9^2}{445.21} = 35.06 \sim \chi^2_1$$

Because $\chi^2_{1,.999} = 10.83$ and $X^2_{TR} = 35.06 > 10.83$, it follows that $p < .001$. Therefore, there is a significant trend relating the prevalence of habitual snoring with age, with older subjects snoring more frequently. This analysis was performed while controlling for the possible confounding effects of sex.

In this section, we have learned about analytic techniques for confounding in epidemiologic studies. If we have a dichotomous disease variable (D), a dichotomous exposure variable (E), and a categorical confounder (C), then we can use the Mantel-Haenszel test to assess the association between D and E while controlling for C. Referring to the master flowchart in the back of the book (p. 780), at box 6 we would answer yes to (1) 2×2 contingency table? and at A arrive at the box entitled "Use two-sample test for binomial proportions, or 2×2 contingency-table methods if no confounding is present, or the Mantel-Haenszel test if confounding is present."

If E is categorical but has more than two categories, then we can use the Mantel Extension test for this purpose. Referring to the master flowchart again, we would answer no to (1) 2×2 contingency table? yes to (2) $2 \times k$ contingency table? and yes to (3) interested in trend over k proportions? This would lead us to the box entitled "Use chi-square test for trend if no confounding is present, or the Mantel Extension test if confounding is present."

SECTION 13.6 Power and Sample-Size Estimation for Stratified Categorical Data

Example 13.28

Cancer A study was performed [7] based on a sample of 106,330 women enrolled in the Nurses' Health Study relating ever use of oral contraceptives (OC) at baseline (in 1976) to breast-cancer incidence from 1976 to 1980. Since both OC use and breast cancer are related to age, the data were stratified by 5-year age groups and the Mantel-Haenszel test was employed to test for this association. The results supported the null hypothesis. The estimated odds ratio ($\widehat{OR}_{MH}$) was 1.0 with 95% confidence interval = (0.8, 1.3). The issue is, What power did the study have to detect a significant difference if the underlying $OR = 1.3$?