INTERACTION OF ALCOHOL AND TOBACCO IN LARYNGEAL CANCER

W. DANA FLANDERS AND KENNETH J ROTHMAN


Both alcohol and tobacco use are accepted risk factors for laryngeal cancer. The authors used case-control data from previous studies to estimate the value of a previously proposed index of interaction between these two risk factors. In addition to the weighting procedure over exposure categories that was previously proposed for estimating a summary index, they applied maximum-likelihood techniques to facilitate the estimation. Overall, they found moderate synergy between alcohol and tobacco in increasing the risk of laryngeal cancer, in that exposure to both factors increased the risk about 50% more than the increase predicted if the effects of tobacco and alcohol were simply additive.

alcohol drinking; laryngeal neoplasms; smoking

The larynx is the site of 1.4 per cent of all new cases of cancer and 2.4 per cent of new cases among US males (1). Known risk factors for laryngeal cancer include sex, age, tobacco smoking, alcohol drinking, and exposure to certain employment-related agents such as asbestos (2). Though the effects of tobacco and alcohol use on laryngeal cancer risk are well documented, the extent to which alcohol and tobacco interact in increasing the risk of laryngeal cancer is not known.

We evaluated the interaction between these risk factors using data from two sources: the interview data from the Third National Cancer Survey (3), and published data from a case-control study by Wynder et al. (4). Our approach was to estimate interaction with a previously proposed parameter (5), which is the ratio of the observed effect of joint exposure to the effect predicted from the sum of the effects of each factor acting separately.

METHODS

The Third National Cancer Survey was a study of all new cases of cancer occurring in seven cities and two states during the period 1969-1971. A 10 per cent probability sample of these cases was interviewed to obtain more detailed medical and epidemiologic information. The data collection and potential problems of the study have been described in detail (3, 6). Our analysis is based on the results of the interview sample and uses a case-control approach. We excluded as controls subjects with cancer in sites that have been strongly associated with alcohol use, tobacco use, or certain occupational exposures, the latter because the same control series was used for an analysis of occupational risk factors which we reported elsewhere (7). Cancer sites excluded were oral
cavity, pharynx, esophagus, liver, small intestine, colon, pancreas, stomach, lung, bronchus, pleura, bladder, and kidney. We excluded from the case group and from the control group all females and any subject for whom information on age, sex, alcohol use, or tobacco use was missing.

After these exclusions, 87 male interviewees with laryngeal cancer remained, constituting the case group, and 956 male interviewees with cancer of other sites remained, constituting the control group. The most common cancer sites among the controls were prostate, rectum, and hematopoietic system. From information ascertained at interview on duration and average daily amount of use of three forms of tobacco (cigarettes, cigars, and pipes), we estimated the lifetime consumption of each type to obtain the total lifetime tobacco consumption using the following equivalencies: 1 tobacco unit = 1 cigarette = 0.2 cigars = 0.4 pipefuls. In an analogous fashion, we estimated the total lifetime consumption of alcohol using the following equivalencies: 1 alcohol unit = \( \frac{1}{2} \) oz liquor = 12 oz beer = 6 oz wine (1 oz = 44 ml).

Wynder et al. (4) conducted a matched case-control study of 258 male and 56 female laryngeal cancer cases. Two controls for each male case were selected from hospitalized patients, excluding those with diagnoses thought to be related to smoking and drinking. As with the Third National Cancer Survey data, we restricted our analysis to the 224 male cases and 414 male controls for whom information on both alcohol and tobacco use was available. Matching factors included year of interview, hospital status, and age at diagnosis. Smoking and alcohol data were not presented in sufficient detail to estimate each subject’s total lifetime use. Instead, we used the data reported on daily consumption of alcohol and tobacco to measure exposure to these factors. The same equivalencies for tobacco and alcohol units as for the Third National Cancer Survey data were used for the data of Wynder et al.

In the context of evaluating the public health costs of exposure to environmental agents, it is generally appropriate that no interaction between two agents correspond to additivity of effects (8). More explicitly, if the effect of exposure is taken as the incidence rate in those exposed minus the incidence rate in those unexposed, this criterion of additivity implies that the effect of joint exposure equals the sum of the effects of exposure to each agent alone. We have used a measure of interaction based on this criterion for independence (5). The measure, \( S \), is the ratio of the observed effect with joint exposure divided by the effect predicted for joint exposure assuming additivity of the effects. No interaction corresponds to \( S = 1 \). In our analyses, we estimated the value of \( S \) for each combination of various categories of alcohol and tobacco use.

To obtain a summary measure of interaction, we took a weighted average of the estimates obtained for each combination of alcohol and tobacco use, using weights proportional to the inverse of the variance of each estimate (5). In the calculation of these weighted averages, we had to adopt a procedure to avoid division by zero, since some cell frequencies were zero, and these enter the calculations as divisors. Our approach was to add 0.5 to each cell frequency. Changing the observed frequencies in this way may introduce a bias into the estimation procedure, so we also derived an overall estimate for \( S \) using the method of maximum likelihood, which avoids the problems associated with adding an arbitrary quantity to each cell frequency (see Appendix). By iteration we estimated values of the parameters which maximize the likelihood function; these maximum-likelihood estimates are asymptotically normally distributed with variance-covariance matrix equal to the negative of the inverse of the matrix of second partial derivatives.
To simplify the likelihood function and to reduce the number of parameters, we assumed that the interaction index was constant for all combinations of alcohol and tobacco use which did not involve the lowest level of alcohol use or the lowest level of tobacco use. The reasonableness of the assumption of uniformity of $S$ was assessed using a likelihood-ratio test.

**RESULTS**

The frequencies of alcohol and tobacco use among male laryngeal cancer cases and controls in the Third National Cancer Survey (TNCS) (3) and in the study of Wynder et al. (4) are shown in tables 1 and 2, respectively. The data are sparse in those categories representing combinations of heavy alcohol and light tobacco use, both in the TNCS data and the data of Wynder et al. Measures of effect and of interaction involving these categories of heavy alcohol and light tobacco use are consequently unstable.

The estimated index of interaction points in the direction of synergy, or positive interaction, for six of the nine combinations of alcohol and tobacco for the TNCS data (table 3), and for all of the

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distribution of males with laryngeal cancer and controls, based on the interview data from the Third National Cancer Survey (3), according to lifetime alcohol and tobacco consumption</strong></td>
</tr>
<tr>
<td>Lifetime tobacco consumption**</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>0–49 Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>50–549 Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>550–899 Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>900+ Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
</tbody>
</table>

*Lifetime tobacco consumption in tobacco units, divided by 365. 1 tobacco unit = 1 cigarette = 0.2 cigars = 0.4 pipefuls.
†Lifetime alcohol consumption in alcohol units, divided by 365. 1 alcohol unit = 2 oz liquor = 6 oz wine = 12 oz beer (1 oz = 44 ml).

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distribution of males with laryngeal cancer and controls based on the data of Wynder et al. (4), according to daily alcohol and tobacco use</strong></td>
</tr>
<tr>
<td>Daily tobacco use†</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>1–15</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>16–34</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>35+</td>
</tr>
<tr>
<td>Controls</td>
</tr>
</tbody>
</table>

*Note—we have slightly modified the classification of alcohol use from the original publication of Wynder et al., which used the following equivalencies. 1 alcohol unit = 1 oz liquor = 9 oz wine = 8 oz beer, and the categories of daily alcohol use were 0, 1–6 and 6+ (4).
†Average daily tobacco use in tobacco units, where 1 tobacco unit = 1 cigarette = 0.2 cigars = 0.4 pipefuls.
‡Average daily alcohol use in alcohol units, where 1 alcohol unit = 1¾ oz liquor = 6 oz wine = 12 oz beer (1 oz = 44 ml).

<table>
<thead>
<tr>
<th>TABLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category-specific indices of interaction by lifetime tobacco and alcohol use, based on the interview data from the Third National Cancer Survey (3)</strong></td>
</tr>
<tr>
<td>Lifetime tobacco use¶</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>0–49</td>
</tr>
<tr>
<td>50–549</td>
</tr>
<tr>
<td>550–899</td>
</tr>
<tr>
<td>900+</td>
</tr>
</tbody>
</table>

*Entries in this table were derived from the data in table 1 after adding 0.5 to each cell frequency. A value of 1.0 indicates no synergy.
¶Lifetime tobacco consumption in tobacco units, divided by 365. 1 tobacco unit = 1 cigarette = 0.2 cigars = 0.4 pipefuls.
†Lifetime alcohol consumption in alcohol units, divided by 365. 1 alcohol unit = 1¾ oz liquor = 6 oz wine = 12 oz beer (1 oz = 44 ml).
combinations for the data of Wynder et al. (table 4). A weighted average of the nine estimates for the TNCS data is 1.4, using weights inversely proportional to the variance of each estimate. The weighted average of the six estimates for the data of Wynder et al. is 2.5, again using weights inversely proportional to the variance of each estimate.

For the TNCS data, the maximum-likelihood estimate of $S$ was 1.5, assuming uniformity, with 90 per cent confidence limits from 0.6 to 2.3. The $\chi^2$ statistic from the likelihood-ratio test of heterogeneity of $S$ was 8.4 with eight degrees of freedom, corresponding to a $p$ value of about 0.4, showing that the TNCS data are compatible with the assumption of uniformity of $S$. We evaluated possible confounding by age for the TNCS data by including a term in the probability model for each of the two age strata, namely age 35 to 64 years and 65 years or over (see Appendix).

Controlling for age did not substantially change the estimated index of interaction, the point estimate from the TNCS data after control of age being 1.7. Controlling for race in the same manner also had little effect on the index of interaction.

For the data of Wynder et al., the maximum-likelihood estimate of $S$, under the assumption of uniformity, is 2.5, with 90 per cent confidence limits from 1.1 to 3.9. The $\chi^2$ statistic from the likelihood-ratio test for heterogeneity of the index of interaction, with five degrees of freedom is 7.54, corresponding to a $p$ value of 0.18. Thus, the Wynder et al. data are also tolerably compatible with the assumption of uniformity of $S$.

In table 5, we present "smoothed" estimates of the rate ratio for laryngeal cancer by category of alcohol and tobacco use, based on the TNCS data. These estimates are derived from the solution of the likelihood equation assuming uniformity of $S$ (see equation 5 of Appendix). For the combination of heaviest alcohol and tobacco use, the rate ratio estimate is 18.5; for all combinations of use, some effect is apparent in the smoothed estimates.

**DISCUSSION**

Estimates of the interaction measures obtained from the data of the two studies, 1.5 for the Third National Cancer Survey

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Category-specific indices of interaction by daily alcohol and tobacco use, based on the data of Wynder et al. (4)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily tobacco use†</td>
<td>Daily alcohol use‡</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1-14</td>
<td>-</td>
</tr>
<tr>
<td>15-34</td>
<td>-</td>
</tr>
<tr>
<td>35+</td>
<td>-</td>
</tr>
</tbody>
</table>

* The indices were derived from the data in table 2 after adding 0.5 to all frequencies. A value of 1.0 indicates no synergy
† Average daily tobacco use in tobacco units, where 1 tobacco unit = 1 cigarette = 0.2 cigars = 0.4 pipefuls.
‡ Average daily alcohol use in alcohol units, where 1 alcohol unit = 1/4 oz liquor = 6 oz wine = 12 oz beer (1 oz = 44 ml)

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Maximum-likelihood estimates of rate ratios by category of lifetime tobacco and alcohol use, based on the interview data from the Third National Cancer Survey (3)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lifetime tobacco consumption†</td>
<td>Lifetime alcohol consumption‡</td>
</tr>
<tr>
<td>0-49</td>
<td>0-49</td>
</tr>
<tr>
<td>50-549</td>
<td>1.5</td>
</tr>
<tr>
<td>550-899</td>
<td>3.5</td>
</tr>
<tr>
<td>900+</td>
<td>7.9</td>
</tr>
</tbody>
</table>

* Based on the data in table 1 and on the assumption of uniformity of the index of interaction
† Lifetime tobacco consumption in tobacco units, divided by 365. 1 tobacco unit = 1 cigarette = 0.2 cigars = 0.4 pipefuls
‡ Lifetime alcohol consumption in alcohol units, divided by 365. 1 alcohol unit = 1/4 oz liquor = 6 oz wine = 12 oz beer (1 oz = 44 ml)
(3) and 2.5 for the data of Wynder et al. (4), indicate moderate synergy of alcohol and tobacco in increasing laryngeal cancer risk. With effects measured as the difference in incidence rate between exposed and unexposed, the interaction index can be interpreted as a measure of relative increase in the effect, or the number of excess cases, among those exposed to both factors. A value for the index of 1.5 would imply that the effect of alcohol and tobacco acting together is 50 per cent greater than would be predicted by assuming additivity of effects, or that in a steady state the number of cases of laryngeal cancer occurring among those exposed to both alcohol and tobacco and attributable to these exposures is 50 per cent greater than the number expected if effects were simply additive.

The data of Wynder et al. (4) lead to an estimate of synergy slightly greater than that based on the TNCS data (3). Wynder et al. found no cases who were nonsmokers and moderate or heavy drinkers; the estimated relative risks for these categories of smoking-drinking, compared with the nonsmoking-nondrinking category, is 0, clearly a severe underestimate. Underestimation of the effect of alcohol drinking among nonsmokers results in overestimation of $S$, perhaps accounting for part of the discrepancy between the results based on the TNCS data and those from the study of Wynder et al. Aside from chance variation, other factors which may contribute to the difference in the index estimated from the two data sets include differences in the measure of alcohol and tobacco use (total lifetime use for the TNCS data and daily use for the study of Wynder et al.), and geographical or temporal differences in types of alcohol and tobacco use.

This evidence for interaction should be interpreted cautiously because uncontrolled confounding between alcohol and tobacco within categories may lead to elevation of $S$ above the null value. Since the uncontrolled confounding, if any, should be considerably less in categories which are not open-ended, we calculated the maximum-likelihood estimate of $S$ based only on these categories. The result was 2.5 for the TNCS data, and 1.4 for the Wynder et al. data, providing some assurance that confounding between alcohol and tobacco does not account for the entirety of the synergistic relation between the two factors.

An interaction (excess over additivity) between alcohol and tobacco in the etiology of laryngeal cancer would have both biologic and public health implications. Biologically, one possible mechanism to account for the interaction (10) would be a physical contact between alcohol and the tobacco carcinogens, perhaps with the alcohol acting as a solvent and thereby facilitating entry of tobacco carcinogens into epithelial cells (11). On the other hand, since alcohol is not believed to come into direct contact with the larynx during ingestion, a more likely possibility might be that alcohol vapor in expired air interacts with tobacco carcinogens. Additional possibilities are that alcohol affects metabolism of laryngeal epithelium, or that it affects metabolic conversion of carcinogens in the liver, thus indirectly interacting with the carcinogenic components of tobacco smoke (11). The public health implication of a greater-than-additive effect of joint exposure is the improved prospect for prevention by interventions aimed at either of the interacting risk factors. For example, an antismoking campaign would prevent more laryngeal cancer if alcohol drinkers were persuaded to stop smoking than if the same number of non-drinkers were persuaded to stop smoking.

**References**

APPENDIX

To express the likelihood function, we use the following notation:

\[ a_{ij} \] = number of cases at level \( i \) of alcohol and level \( j \) of tobacco, \( 1 \leq i \leq I, 1 \leq j \leq J; \]
\[ b_{ij} \] = number of controls at level \( i \) of alcohol and level \( j \) of tobacco, \( 1 \leq i \leq I, 1 \leq j \leq J; \]
\[ n_{ij} = a_{ij} + b_{ij}; \]
\[ n_{ca} = \sum_{i=1}^{I} \sum_{j=1}^{J} a_{ij}; \]
\[ n_{ce} = \sum_{i=1}^{I} \sum_{j=1}^{J} b_{ij}; \]
\[ R_{ij} = \] the rate ratio comparing the rate at level \( i \) of alcohol and level \( j \) of tobacco with the rate at level 1 of both alcohol and tobacco;
\[ S_{ij} = \] the index of interaction at level \( i \) of alcohol and level \( j \) of tobacco;
\[ I = \] number of levels of alcohol use; and
\[ J = \] number of levels of tobacco use.

The index of interaction (5) for joint-exposure category \( i,j \) may be written as:

\[ S_{ij} = \frac{(R_{ij} - 1)}{(R_{11} + R_{1j} - 2)}. \] (1)

Note that if \( i = 1 \) or \( j = 1 \), the \( S_{ij} = 1 \) since \( R_{11} = 1 \) by definition.

If the number of cases at level \( i \) of alcohol and level \( j \) of tobacco is distributed as a binomial random variable with parameters \( P_{ij} \) and \( n_{ij} \), similar to the assumption of Hogan et al. (9) in their likelihood model for a test of interaction, then the likelihood function may be written as:

\[ L = \prod_{i=1}^{I} \prod_{j=1}^{J} \left( \frac{n_{ij}}{a_{ij}} \right) P_{ij}^{a_{ij}} (1 - P_{ij})^{n_{ij} - a_{ij}}. \] (2)

The rate ratio for category \( i,j \) may be expressed as a function of \( P_{ij} \) and \( P_{11} \) using the familiar relation

\[ R_{ij} = \left[ P_{ij}(1 - P_{ij}) \right] \left[ (1 - P_{1i})/P_{11} \right]. \] (3)

Equations 1–3 may be combined to express \( P_{ij} \) as a function of \( S_{ij}, P_{11} \) and \( P_{1j} \). If we
define \( S_{11} = 1.0 \), the likelihood function may then be rewritten as:

\[
L = \prod_{i=1}^{l} \prod_{j=1}^{J} \left( \frac{n_{ij}}{a_{ij}} \right) \left[ S_{ij} \left( \frac{P_{ij}}{1 - P_{ij}} + \frac{P_{11}}{1 - P_{11}} - \frac{2P_{11}}{1 - P_{11}} \right) + \frac{P_{11}}{1 - P_{11}} \right]^{a_{ij}}
\]

\[
\left[ 1 + S_{ij} \left( \frac{P_{ij}}{1 - P_{ij}} + \frac{P_{11}}{1 - P_{11}} - \frac{2P_{11}}{1 - P_{11}} \right) + \frac{P_{11}}{1 - P_{11}} \right]^{n_{ij}}
\]

Equation 4 may be derived by substituting equation 3 for \( R_w, R_{ii} \) and \( R_u \) into equation 1 to express \( S_{ij} \) in terms of \( P_{ij}, P_{11}, P_{ij} \) and \( P_{11} \), then using the result to solve for \( P_u \) in terms of \( S_{ij}, P_{11}, P_{ij} \) and \( P_{11} \), and finally substituting for \( P_u \) in equation 2.

The number of parameters in this likelihood function can be reduced by assuming that \( S_{ij} \) is constant for \( i \neq 1 \) or \( j \neq 1 \). With this simplifying assumption the likelihood becomes:

\[
L = \prod_{i=2}^{l} \prod_{j=2}^{J} \left( \frac{n_{ij}}{a_{ij}} \right) \left[ S \left( \frac{P_{ij}}{1 - P_{ij}} + \frac{P_{11}}{1 - P_{11}} - \frac{2P_{11}}{1 - P_{11}} \right) + \frac{P_{11}}{1 - P_{11}} \right]^{a_{ij}}
\]

\[
\left[ 1 + S \left( \frac{P_{ij}}{1 - P_{ij}} + \frac{P_{11}}{1 - P_{11}} - \frac{2P_{11}}{1 - P_{11}} \right) + \frac{P_{11}}{1 - P_{11}} \right]^{n_{ij}}
\]

\[
\times \prod_{i=1}^{J} \left( \frac{n_{1j}}{a_{1j}} \right)^{a_{1j}} \left[ 1 + \frac{P_{1j}}{1 - P_{1j}} \right]^{n_{1j}}
\]

\[
\times \prod_{j=2}^{J} \left( \frac{n_{ij}}{a_{ij}} \right)^{a_{ij}} \left[ 1 + \frac{P_{ij}}{1 - P_{ij}} \right]^{n_{ij}}
\]

The sampling procedure used in most case-control studies may not be adequately reflected in the above likelihood function, based on a binomial distribution of the number of cases in each exposure category, since, for example, in case-control studies the investigator usually determines the number of cases and/or controls. Theoretically, for most case-control studies it may be more appropriate to base the likelihood function on the assumption that cases and controls have independent multinomial sampling distributions over category of exposure. Therefore, we also derived a likelihood function based on independent multinomial distributions of exposure for cases and controls.

If the distribution of category of alcohol and tobacco use is multinomial for cases with parameters \( n_{ca} \) and \( \lambda_{ij} \) for \( i = 1, \ldots, I \) and \( j = 1, \ldots, J \) and is multinomial for controls with parameters \( n_{co} \) and \( l_{ij} \) for \( i = 1, \ldots, I \) and \( j = 1, \ldots, J \), then the likelihood function is:

\[
L = \prod_{j=1}^{J} \prod_{i=1}^{l} \left( \frac{n_{ca}}{a_{ij}}, \ldots, a_{ij} \right)^{x_{ij}} \left( \frac{n_{co}}{b_{11}}, \ldots, b_{ij} \right)^{t_{ij}}
\]

where

\[
\sum_{j=1}^{J} \sum_{i=1}^{l} \lambda_{ij} = 1 \quad \text{and} \quad \sum_{j=1}^{J} \sum_{i=1}^{l} l_{ij} = 1.
\]

The derivation of the form of this likelihood equation under the assumption of uniformity of \( S \) is algebraically tedious and is similar to the derivation of the likelihood
based on independent binomial distribution (equation 5), so we present only the result here. In terms of the multinomial parameters, the likelihood function is:

\[
L = \left[ \prod_{j=1}^{I} \prod_{i=1}^{l} \binom{n_{co}}{b_{1i}, \ldots, b_{ij}} \right] \times \left[ \prod_{j=2}^{J} \prod_{i=2}^{l} \binom{n_{ca}}{a_{1i}, \ldots, a_{ij}} \right] \times \left\{ l_u \left( \frac{\lambda_{11}}{l_{11}} + \frac{\lambda_{1j}}{l_{1j}} - 2 \frac{\lambda_{1i}}{l_{1i}} + \frac{\lambda_{11}}{l_{11}} \right) \right\}^{a_{ij}} \times \left[ \prod_{i=1}^{l} \lambda_{ii}^{a_{ii}} \times \prod_{j=2}^{J} \lambda_{1j}^{a_{1j}} \right],
\]

where

\[
\lambda_{ij} = \frac{1 - \sum_{i=1}^{I-1} \lambda_{i1} - \sum_{j=2}^{J} \lambda_{1j} - \sum_{j=2}^{J} \sum_{i=2}^{l} l_u \left( \frac{\lambda_{11}}{l_{11}} + S \left( \frac{\lambda_{1j}}{l_{1j}} - 2 \frac{\lambda_{1i}}{l_{1i}} \right) \right) - \sum_{j=2}^{J} \sum_{i=2}^{l} l_u S \frac{\lambda_{11}}{l_{11}}}{1 + \sum_{j=2}^{J} \sum_{i=2}^{l} l_u S \frac{l_{1j}}{l_{11}}}. \tag{7}
\]

We calculated maximum-likelihood estimates of the interaction index for the data of Wynder et al. (4) and for the Third National Cancer Survey data (3) using likelihood equation 5 and also using likelihood equation 7. The estimates based on each likelihood function were nearly identical, since each depends in a similar way on the index of interaction and rate ratios.

In controlling for possible confounding by age, we chose to use likelihood equation 5 because, compared with equation 7, it was simpler, it involved far fewer parameters (8 vs. 22 for the Third National Cancer Survey example), and its use yielded estimates nearly identical with estimates obtained using equation 7. To control for confounding by age, the likelihood function may be written for each level \( k \) of the covariate, \( 1 \leq k \leq K \):

\[
L_k = \prod_{i=1}^{I} \prod_{j=1}^{J} \binom{n_{uk}}{a_{uk}} \left[ S_{uk} \left( \frac{P_{1jk}}{1 - P_{1jk}} + \frac{P_{1jk}}{1 - P_{1jk}} - 2P_{1jk} \right) + \frac{P_{1jk}}{1 - P_{1jk}} \right]^{a_{uk}} \left[ 1 + S_{uk} \left( \frac{P_{1jk}}{1 - P_{1jk}} + \frac{P_{1jk}}{1 - P_{1jk}} - 2P_{1jk} \right) + \frac{P_{1jk}}{1 - P_{1jk}} \right]^{n_{uk}} \tag{8}
\]

The notation is the same as above, with the additional subscript, \( k \), referring to the level of age. The likelihood function, \( L \), is then given by

\[
L = \prod_{k=1}^{K} L_k, \tag{9}
\]

where \( K = 2 \) for the Third National Cancer Survey example.

To reduce the number of parameters, we assumed that the odds for disease at the younger level of age and each combination of alcohol and tobacco use,

\[
\frac{P_{uj}}{1 - P_{uj}},
\]
differed by a constant from the corresponding odds at the older level of age,

\[
\frac{P_{ij2}}{1 - P_{ij2}} = \frac{P_{ij1}}{1 - P_{ij1}} + C,
\]

That is,

\[
\frac{P_{ij1}}{1 - P_{ij1}} = \frac{P_{ij2}}{1 - P_{ij2}} + C,
\]

for \(1 \leq i \leq I\), and \(1 \leq j \leq J\).

As before, we assumed that the index of interaction was constant for all combinations of \(i, j\) and \(k\) for which neither \(i\) nor \(j\) is equal to one. With these assumptions, for the Third National Cancer Survey data, the maximum likelihood estimate of \(S\) was 1.7. The likelihood-ratio test of our simplifying assumptions gives a \(p\) value of 0.15. Basing the probability model on other assumptions about the effect of age or including two parameters for age in the model did not lead to substantial change in the point estimate of interaction, providing additional assurance that adequate control of confounding by age was obtained with the simple model we adopted.

The published data of Wynder et al. (4) were not sufficiently detailed to control for age.