SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853

TECHNICAL REPORT NO. 841

March 1989

DETECTION OF ASSOCIATIONS BETWEEN
DISEASES IN ANIMAL
CARCINOGENICITY EXPERIMENTS

by

Toby J. Mitchell, Oak Ridge National Laboratory
Bruce W. Turnbull, Cornell University

Support for the second author was provided in part from Grant #GM 28364 from the U.S. National Institutes of Health.

DETECTION OF ASSOCIATIONS BETWEEN DISEASES IN ANIMAL CARCINOGENICITY EXPERIMENTS

Toby J. Mitchell

Oak Ridge National Laboratory, P.O. Box 2009, Oak Ridge, Tennessee 37831-8083, U.S.A.

and

Bruce W. Turnbull

Cornell University, 321 Upson Hall, Ithaca, New York 14853, U.S.A.

SUMMARY

We present a statistical test for the hypothesis of independence of two disease processes in animal carcinogenicity experiments. It is assumed that the diseases are occult (detectable only at death via necropsy), and progressive, in that each passes (potentially) through three states in each animal: absent, nonlethally present, and lethally present. The test is based on the combination of statistics from a series of age-stratified 3×3 tables in which the animals that die in each age interval are cross-classified into one of these three states for each disease. This test utilizes all the data, whether obtained by sacrifice or by "natural death". It avoids the biases due to lethality noted by Breslow et al. (1974, *Journal of the National Cancer Institute*, 52, 233-239) for the test based on age-stratified 2×2 tables, where the nonlethal and lethal occurrences are pooled. In a simulation experiment, we show that such biases (which usually conceal true positive associations and produce spurious negative associations) can be very severe indeed, whereas our test preserves the nominal significance level quite well. The power of the test appears to vary widely, depending on the incidence, time to onset, and rate of progression of each disease. As an example of the application of our method, we assess the association between liver tumors and reticulum cell sarcoma in a radiocarcinogenesis experiment involving approximately one thousand C3H female mice.

Keywords: Carcinogenicity; Independence of diseases; Disease associations; Disease process; Survival/Sacrifice experiments; Nonlethal tumors; Lethal tumors; Contingency tables; Mantel-Haenszel test.

1. Introduction

The detection of associations between different tumor types, i.e., sites, in animal carcinogenesis experiments is important to the understanding of disease mechanisms. If two tumor types tend to occur together in individual animals (positive association), we might suspect a common etiology, a common hormonal, metabolic, or immunologic disturbance, or common genetic susceptibilities. A negative association might indicate that host factors which depress expression of one tumor type promote the occurrence of another.

Detection of these associations also has important implications for regulatory decisions. Weinberg and Storer (1985) have called attention to the fact that some purported carcinogens have shown "ambiguous" effects -- that is that, at least for moderate doses, these agents increase the incidence of some tumors, but at the same time reduce the incidence of others. In such cases the agent might be labelled carcinogenic, whereas the overall effect might be beneficial. Haseman (1983) examined data from 25 of the National Toxicology Program's two-year cancer bioassay feeding studies and found a significant negative association between incidence levels of liver tumors and leukemia/lymphoma in Fischer 344 rats. Using within-animal analyses, this correlation has also been found in other species of rat and in mice -- Breslow et al. (1974), Storer (1982), Wahrendorf (1983), Young and Gries (1984). Clearly if ambiguous effects can exist, difficulties will be created for the interpretation of carcinogenicity tests and in the formulation and implementation of regulations.

We consider diseases that are irreversible and occult, i.e. detectable only at death via necropsy. A crude measure of association would be the cross-product ratio in a 2×2 table with frequencies consisting of the number of animals with and without each of two specified diseases. However Breslow et al. (1974) pointed out that this can be misleading and that the analysis must

include an adjustment for age at death as well as treatment, sex, and other covariates. This is done by stratifying on these variables and combining results by means of the method of Mantel and Haenszel (1959).

Breslow et al. (1974) also note that another source of bias is the lethality of the diseases. When considering two diseases A and B , for example, the lethality of A will reduce the number of cases in which both diseases appear together and correspondingly increase the number of cases in which only A appears. On a gross level, this will cause the association to appear more negative. When the data are stratified by age at death, the effect of lethality is more complicated, but still seems to impart a negative bias to measures of association. Breslow et al. (1974) and Storer (1982) both noted the tendency of lethal diseases to show apparently significant negative associations with other diseases. The model that we describe in Section 2 indicates that both A and B need to be potentially lethal in order for spurious associations to appear in age-stratified 2×2 tables. In a simulation experiment, which we shall describe in Section 4, both diseases of interest were potentially lethal. We found there that the bias was negative in every case considered and was surprisingly severe, even when the diseases progress relatively slowly.

Breslow et al. (1974) did not adjust quantitatively for the bias caused by lethality. They stated that this would need more extensive data and further elaboration of a competing risks model which they proposed. The log-linear prevalence/lethality models of Mitchell and Turnbull (1979) can be used, to a limited extent, for assessment of association between tumor types. A detailed application of these methods was given by Wahrendorf (1983). Unfortunately, although the Mitchell-Turnbull estimates of prevalences adjust for disease lethality, the true prevalences may not accurately reflect independence (or dependence) among diseases, unless one of the diseases is nonlethal (see Mitchell and Turnbull, 1979, page 228). That a given disease is nonlethal can be determined experimentally only with large sample sizes and extensive

sacrificing, although biologists are willing to declare some diseases completely nonlethal *a priori*. Many diseases of interest, however, are nonlethal only up to a certain stage of development, after which they may affect the time of death. In this paper we propose a test for association that relies on recognition of the nonlethal cases of the diseases of interest, rather than on prior assumption or extensive sacrificing.

We shall use as an example an experiment reported by Storer, Mitchell, and Fry (1988), which was carried out at the Biology Division of the Oak Ridge National Laboratory. In this experiment, evaluations were made of the lethalities of the diseases present in each animal at death. The purpose of the experiment was to investigate the effect of radiation on the incidence of tumors at various sites in several strains of mice. Part of this experiment consisted of three groups of C3H female mice. One group consisted of 251 mice exposed to a dose of 0.5 Gy of gamma radiation at an age of 10 weeks; a similar group of 250 mice was exposed to 1.0 Gy. The third group consisted of 503 unexposed controls. For each mouse, the age at death was recorded, along with the presence or absence of a tumor at each of approximately ten sites. If the tumor was present, the pathologist also made a determination as to whether the tumor was lethal (i.e. related to cause of death) or nonlethal. (For some animals more than one tumor was listed as lethal.) It is often pointed out that the determination of this "state of lethality" may be difficult to make, but in this particular experiment the pathologist felt reasonably confident about it.

Table 1 summarizes the experience of the three groups with regard to two diseases of particular interest, namely reticulum cell sarcoma (a disease related to lymphoma) and liver tumors. The frequencies are cross-classified into one of three states for each disease -- absent at death, present nonlethally at death, present lethally at death. A more detailed description of the data for the control group appears in Table 2, where the mice are classified according to age at death into 50 day intervals. We omit the analogous tables for the other two groups. Since the exact age at death is available, other interval widths could be used. The choice of constant 50 day

intervals is a convenient one. For this data set and for the simulation experiments of Section 4, alternate choices of interval widths yielded similar results.

[Table 1 here.]

[Table 2 here.]

In the next section we define notation and elaborate further on a model for disease progression and lethality suggested in the paper by Breslow et al. (1974, pp. 238-239), but not used in their analysis. (See also Gart et al. (1986, page 166).) In Section 3 we show how the model can be used to construct a test of a null hypothesis of independence based on a series of stratified 3×3 tables. The test statistic is the same as that used by Breslow et al. (1974) in their analysis of stratified 2×2 tables, but now the null distribution is modified to take into account information on the state of lethality of each disease. We apply this method to the data summarized in Table 1 and compare the results that would be obtained using the 2×2 classification. We find in this case that the significance of the negative association between reticulum cell sarcoma and liver tumors is reduced when state of lethality is taken into account. In a simulation experiment, described in Section 4, we then show that the effect of utilizing state of lethality information can be quite dramatic.

2. The Model

We propose the following approximating model for the progression of various diseases of interest in individuals in an animal population. It is intended to utilize data on whether each disease of interest was observed to have contributed to the death of the animal or was observed

incidentally as the result of death by other causes. The classification of occurrences of a disease as lethal or nonlethal, as well as the accounting for treatment related toxicity, was shown to be essential by Hoel and Walburg (1972) and by Peto (1974) in assessing differences in disease incidence among groups. (See also Peto et al. (1980), McKnight and Crowley (1984).) We shall use it here to develop a test of association that avoids the kind of spurious negative correlations noted by Breslow, et al. (1974).

Our model can be constructed by superimposing mortality on the "potential history" of an animal, where the history includes, as a function of time, t , descriptions of all relevant disease processes or other characteristics of the animal. A particular history is "potential" in that it is assumed to exist over all t ; its full realization is prevented by death, which is treated as a censoring variable. The use of "potential" or "latent" event-times is a common one for time-to-tumor models (Gart et al., 1986, Section 6.3), and was the formulation suggested by Breslow et al. (1974, p. 238) for the association problem. For a discussion about this formulation and its relation to the cause-specific hazards model, see Kalbfleisch, Krewski and Van Ryzin (1983, Section 3.3). We denote by H_t the state of an animal at time t ; $\{H_s : 0 \leq s \leq t\}$ is the potential history of an animal up to time t . Let A and B be two irreversible diseases whose independence we wish to test, and let C be a composite of all other diseases or characteristics that can affect mortality. We denote by A_t, B_t, C_t the states of A, B , and C , at time t , and we denote the time of death by W .

We first assume that survival to t depends only on $\{H_s : 0 \leq s \leq t\}$, through $\{A_s, B_s, C_s : 0 \leq s \leq t\}$:

$$Pr[W > t \mid \{H_s : s \geq 0\}] = Pr[W > t \mid \{H_s : 0 \leq s \leq t\}] = Pr[W > t \mid \{A_s, B_s, C_s : 0 \leq s \leq t\}] \quad (A1)$$

The first equality in (A1) stipulates that death is a "nonprognostic" censoring variable; the second stipulates the sufficiency of $\{A_s, B_s, C_s : 0 \leq s \leq t\}$ for determining the probability of survival to t .

We assume that A and B progress from absent (State 1) to present but nonlethal (State 2) to present and lethal (State 3), where States 2 and 3 may each have several substates. (A disease can go directly from State 1 to State 3, but we would consider this unusual.) The distinction between States 2 and 3 of A is made formal by defining the substate a_t to be nonlethal if and only if:

$$Pr[W > t \mid A_s = a_s, B_s = b_s, C_s = c_s; 0 \leq s \leq t] = Pr[W > t \mid A_t = 0, B_s = b_s, C_s = c_s; 0 \leq s \leq t] \quad (A2)$$

where " $A_t = 0$ " means that A is absent at t (and hence absent at any time prior to t). That is, the history of A up to t can be replaced with the statement that A is absent up to t without changing the probability of survival to t , no matter what the histories (up to t) of the other diseases that may affect death. A similar requirement must be met for a substate b_t of B to be considered nonlethal. Under this definition of nonlethality, the lethal state of a disease is generally broader than the determination of that disease as a "cause of death". (A related type of classification is the basis for the analyses proposed by Peto (1974) and Peto et al. (1980) for testing differences in disease incidence among different groups of animals. Breslow et al. (1974, Table 5) also used three categories, indirectly related to lethality, for classifying the presence of diseases. There, lymphomas were categorized as absent, localized, or generalized, and hepatomas and lung adenomas as absent, small, or large.) Let U_A be the (potential) time to onset of Disease A and let V_A be the (potential) time of first entry of A into a lethal state, where $U_A \leq V_A$. Either of these potential times can be infinite. For Disease B , U_B and V_B are defined similarly.

In a homogeneous group of animals, e.g., animals of the same species, sex, and treatment, we assume that the pairs $(\{H_t : t \geq 0\}, W)$ are independent across animals. (Note that this

assumption is violated if, at any time during the experiment, a fixed number of animals is sacrificed. If only one animal is scheduled to be sacrificed at a given time, for example, the choice of one animal for sacrifice removes sacrifice as a potential risk for any other animal at that time. This results in a dependency among the death times. For the time being, therefore, we shall assume that at any sacrifice time, the decision to sacrifice or not is made independently for each animal, and does not depend on anything that has occurred in the rest of the population. Note that this assumption does accommodate the killing of moribund animals, which is a common practice. In the next section, we shall discuss more general sacrifice protocols.)

The observed data for each animal in a particular group consist of its age at death (W , say) and, for each disease of interest, indicators of whether that disease was absent ($W < U$), present nonlethally ($U \leq W < V$), or present lethally ($V \leq W$). As in Turnbull and Mitchell (1978) and Mitchell and Turnbull (1979), it simplifies matters considerably if we use a discrete time approximation to the model. The time scale is divided up into M intervals $(t_1, t_2], (t_2, t_3], \dots, (t_M, t_{M+1}]$. For each of the M intervals we have a 3×3 contingency table with frequencies n_{ijm} ($1 \leq i, j \leq 3$; $1 \leq m \leq M$) as shown in Table 3.

[Table 3 here.]

In addition, we have s_m animals known to survive the m th interval (i.e. past t_{m+1}); we shall denote by s_0 the sample size entering the first interval at time t_1 .

Consider the (1,1) cell in Table 3. This consists of all animals dying in interval $(t_m, t_{m+1}]$ that had neither Disease A nor B. The probability g_{11m} that an animal appears in this cell is:

$$\begin{aligned} g_{11m} &= Pr[t_m < W \leq t_{m+1}, U_A > W, U_B > W] \\ &= \int_{t_m}^{t_{m+1}} Pr[U_A > t, U_B > t] r_{11}(t) dt \end{aligned}$$

where $r_{11}(t)$ is the conditional density of W given $U_A > t$ and $U_B > t$, evaluated at t . Denoting the interval width by $\delta_m = t_{m+1} - t_m$ and the midpoint by $t'_m = (t_m + t_{m+1})/2$, we approximate this integral by:

$$\delta_m r_{11}(t'_m) p_{11}(t'_m),$$

where

$$p_{11}(t'_m) = Pr[U_A > t'_m, U_B > t'_m].$$

Similar arguments can be used to obtain analogous approximations to the probability of an animal appearing in any of the other cells in the table for the m th interval. That is:

$$g_{ijm} \approx \delta_m r_{ijm} p_{ijm} \quad (1 \leq i \leq 3; 1 \leq j \leq 3)$$

where p_{ijm} is the probability that Disease A is potentially in state i and Disease B is potentially in state j at time t'_m in the absence of other causes of death, and $r_{ijm} = r_{ij}(t'_m)$ is the corresponding conditional density of W , defined analogously to $r_{11}(t'_m)$.

We define A and B to be independent if $\{A_t : t \geq 0\}$ and $\{B_t : t \geq 0\}$ are independent random processes. Independence of A and B implies

$$H: p_{ijm} = \alpha_{im} \beta_{jm}$$

for some α_{im}, β_{jm} , where $1 \leq i \leq 3$, $1 \leq j \leq 3$, and $1 \leq m \leq M$. In Section 3, we shall regard H as the null hypothesis, keeping in mind, however, that independence of A and B implies H but not vice-versa. Thus our data may be able to reject the hypothesis of independence but never confirm it. This is a familiar result (Clifford 1977).

In general, relationships among the p_{ijm} 's for given m will be concealed by differences among the r_{ijm} 's. However, it is possible to test H if we are willing to assume that the mortality factors (r_{ijm} 's) are unaffected by the nonlethal presence of a disease, specifically:

$$r_{11m} = r_{12m} = r_{21m} = r_{22m} = q_m, \text{ say,}$$

$$r_{31m} = r_{32m} = q_m^A, \text{ say,}$$

$$r_{13m} = r_{23m} = q_m^B, \text{ say,}$$

$$r_{33m} = q_m^{AB}, \text{ say.}$$

Even these relationships, which appear at first to follow directly from the definition of nonlethality, will not hold in general without the following requirement of conditional independence between each disease of interest and the states of the remaining diseases:

$$\begin{aligned} & Pr[A_s = a_s, B_s = b_s, C_s = c_s; 0 \leq s \leq t \mid V_A > t] \\ &= Pr[A_s = a_s; 0 \leq s \leq t \mid V_A > t] Pr[B_s = b_s, C_s = c_s; 0 \leq s \leq t \mid V_A > t] \end{aligned} \quad (A3)$$

and similarly with A and B reversed. That is, given that A (or B) is not yet in a lethal state, it is independent of the other diseases. It is somewhat disappointing that a test for the independence of A and B should require both to be independent of other diseases, at least to the extent stated by (A3). However, it is unlikely that A and B would be independent if they were both associated with another disease, so often (A3) would be implicit in the null hypothesis.

Given a total of $n_m = s_{m-1} - s_m$ animals dying in $(t_m, t_{m+1}]$, the cell frequencies n_{ijm} are, under (A1), (A2), (A3), and H , multinomially distributed with probabilities proportional to the entries in Table 4.

[Table 4 here.]

It follows from the structure of Table 4 that, if the nonlethal and lethal categories of each disease are pooled, the resulting 2×2 table will show independence if $\alpha_{1m} \alpha_{3m} \beta_{1m} \beta_{3m} = 0$. In particular, this will occur if either A or B cannot be lethal (i.e., $\alpha_{3m} = 0$ or $\beta_{3m} = 0$). If both diseases are potentially lethal, then the resulting 2×2 table will show a spurious negative (positive) association if $q_m q_m^{AB} - q_m^A q_m^B$ is negative (positive).

Remark: The structure shown in Table 4 can also be derived using a compartment model, which does not involve potential event times, similar to that of Turnbull and Mitchell (1978). In this model, transitions among illness states occur instantaneously between intervals and transitions from alive to dead occur during intervals. Let A_t be the random variable that denotes the state of Disease A at time t . In Turnbull and Mitchell (1978), there were only two such states, absent and present, but there may be more in general. Let C be a composite of all diseases that affect mortality, including B , and let I_{A_t} be the property that A_t and C_t are conditionally independent, given that A is not present lethally at t . It can be shown that if I_{A_t} holds in the population of animals alive at t , where t is just prior to a death transition, it also holds for t just after the transition, for the population of animals dying during the transition and also for the population of survivors. Moreover, I_{A_t} is preserved by illness state transitions if the transitions involving the nonlethal states of A (including absence of A) are independent of those involving C . Under this model, therefore, the ratio of the expected number of animals in which Disease A is present nonlethally (expectation of row 2 of Table 3) to the expected number of animals in which Disease A is absent (expectation of row 1 of Table 3) is the same for all three states of B in the table. A similar argument for Disease B leads to the same kind of result for the ratio of the expectation of the first two columns of the table. With an appropriate reparametrization, the cell probabilities then have the structure shown in Table 4.

3. A Conditional Test

In this section we describe a method for testing H , which is a necessary condition for the independence of A and B . We first concentrate our attention on obtaining an exact test for independence in a single table. For notational convenience, we omit the interval subscript m on n_{ijm} , n_m , etc.

As usual when obtaining an exact test for independence in contingency tables, we eliminate the effect of nuisance marginal probability parameters by conditioning on the marginal row and column totals, $n_{1.}, n_{2.}, n_{3.}, n_{.1}, n_{.2}, n_{.3}$. Further, to eliminate the effect of the nuisance lethality parameters $q_m, q_m^A, q_m^B, q_m^{AB}$, we also condition on $n' = n_{11} + n_{12} + n_{21} + n_{22}$, the number of deaths in the interval for which neither A nor B is in a lethal state. (When included with the other conditions, this is sufficient also to fix $n_{31} + n_{32}, n_{13} + n_{23}$, and n_{33} .) We note that we have imposed 6 independent restrictions on the 9 cell frequencies. Once n_{11}, n_{13} , and n_{31} are given, all remaining entries in the table are determined. Therefore the conditional probability of observing any particular set of table frequencies, given n' and the marginal totals, is determined by the joint conditional distribution of (n_{11}, n_{13}, n_{31}) . We now proceed to derive this under the null hypothesis H as stated in Section 2.

Let

$$h(x; n, N, K) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}} \quad (3.1)$$

for $\max(n-N+K, 0) \leq x \leq \min(n, K)$ denote the hypergeometric distribution probabilities with parameters n, N, K . If random variable X has this distribution then

$$E[X] = \frac{nK}{N} \quad (3.2)$$

$$E[X(X-1)] = \frac{K(K-1)n(n-1)}{N(N-1)} \quad (3.3)$$

Under H , n_{13} has a hypergeometric distribution with parameters $n_{13} + n_{23}, n_{1.} + n_{2.}$, and $n_{1.}$. This can be seen by merging the first two columns of the original 3×3 table and deleting the third row. The resulting table is a standard 2×2 table with fixed marginals; the distribution of n_{13} is then seen to be hypergeometric as claimed. Given n_{13} (and the original conditions), n_{31} has a

hypergeometric distribution, independent of n_{13} , with parameters $n_{31}+n_{32}$, $n_{.1}+n_{.2}$, and $n_{.1}$. This can be seen by noting that the third column of the original 3×3 table is now fixed. Delete this column and merge the first two rows of the resulting 3×2 table. The result is a standard 2×2 table with fixed marginals, and the claimed hypergeometric distribution of n_{31} follows. Since all the parameters of this distribution are fixed by the original conditions (and do not depend on n_{13}), n_{31} and n_{13} are independent. Finally, given n_{13} and n_{31} , as well as the original conditions, the third row and the third column of the original 3×3 table are fixed. Delete these; the remaining table is a standard 2×2 table with fixed marginals. It follows that n_{11} has a hypergeometric distribution with parameters $n_{1.}-n_{13}$, n' , $n_{.1}-n_{31}$. Thus, the exact conditional probability of observing any particular table under the null hypothesis H can be computed.

As an example consider the fourteenth interval (951-1000 days) in Table 2. There are 16 possible tables with the given marginal totals $n_{1.} = 20$, $n_{2.} = 3$, $n_{3.} = 1$, $n_{.1} = 15$, $n_{.2} = 2$, $n_{.3} = 7$, and $n' = 16$. These tables are listed in Table 5a in increasing order of associated null probabilities. To compute the probability of the table that was observed, for example, where $n_{13} = 7$, $n_{31} = 1$, and $n_{11} = 11$, (line 15 of Table 5a), use (3.1) to obtain $P(n_{13}) = 0.31621$, $P(n_{31}) = 0.88235$, and $P(n_{11} | n_{13}, n_{31}) = 0.65$; then take the product to yield 0.18135.

We choose as our test statistic n_{11} , which is equivalent to the statistic used by Breslow, et al. (1974), and also to the ratio of the odds for presence of A when B is present to the odds for presence of A when B is absent. To see this, we note that

$$n_{22}+n_{23}+n_{32}+n_{33} = n_{11}+n-n_{1.}-n_{.1},$$

so the odds ratio is

$$\frac{n_{11}(n_{11}+n-n_{1.}-n_{.1})}{(n_{1.}-n_{11})(n_{.1}-n_{11})},$$

which is easily seen to increase with n_{11} ; all other quantities are fixed by the conditioning.

The marginal distribution of n_{11} under H can be obtained from the joint distribution of n_{11} , n_{13} , and n_{31} . In our example, it can be computed from Table 5a and is shown in Table 5b. The observed value of 11 for n_{11} is clearly not significant.

[Table 5 here.]

The mean of n_{11} can be derived by applying the standard formula (3.2) for the mean of a hypergeometric distribution, first by taking the expectation of n_{11} given n_{13} and n_{31} and then by taking the expectation of the result. The variance of n_{11} is obtained by first finding, in a similar fashion, the expectation of $n_{11}(n_{11}-1)$ using (3.3). Letting $a = n_{1.} + n_{2.}$ and $b = n_{.1} + n_{.2}$, we find that

$$E[n_{11}] = \frac{n' n_{1.} n_{.1}}{ab} \quad (3.4)$$

$$E[n_{11}(n_{11}-1)] = \frac{n'(n'-1)n_{1.}(n_{1.}-1)n_{.1}(n_{.1}-1)}{a(a-1)b(b-1)} \quad (3.5)$$

$$Var[n_{11}] = E[n_{11}] \left[\frac{(n'-1)(n_{1.}-1)(n_{.1}-1)}{(a-1)(b-1)} + 1 - E[n_{11}] \right]. \quad (3.6)$$

Note that if n' or $n_{1.}$ or $n_{.1}$ are zero then $E[n_{11}] = Var[n_{11}] = 0$; if this is not the case but a or b equal one then $Var[n_{11}] = E[n_{11}] \{1 - E[n_{11}]\}$, i.e. the first term in the variance is ignored.

In our example we have $n' = 16$, $n_{1.} = 20$, $n_{.1} = 15$, $a = 23$, $b = 17$ and so $E[n_{11}] = 12.276$ and $Var[n_{11}] = 0.724$. This can be checked by directly calculating the moments of the distribution in Table 5b.

Remark: So far, we have assumed that deaths by sacrifice, if any, occur independently for each animal. However, it is more common to fix, at a given sacrifice time, the number of animals

that will be sacrificed. (This number may be predetermined or it may depend on the state of the surviving population at that time.) In this case, the sacrificed animals at a given time t'_m can be classified into a separate 3×3 table S_m . The cell probabilities in S_m have a form similar to that of Table 4 except that the q 's are defined differently. For example, $q_m = Pr[W > t'_m | U_A > t'_m, U_B > t'_m]$. The sacrifice tables could be considered separately to obtain the mean and variance of the number of sacrificed animals with neither disease at each sacrifice time; they would then be treated as though they represented additional "intervals" and combined with the information from the tables of natural deaths in the calculation of the test statistic Z as indicated below. A more efficient approach would be to define the intervals so that sacrifices occurred at the midpoints. In the m th interval, both the table of natural deaths D_m and the table of sacrifices S_m provide information about the state of the disease processes at t'_m , i.e., about the parameters p_{ijm} . The other nuisance parameters (the q 's) are not shared by D_m and S_m . It can be shown that, under the appropriate conditioning to eliminate the effect of all nuisance parameters, the mean and variance of n_{11m} (the total number of animals in D_m and S_m that died in the m th interval with neither disease) can be obtained by pooling both tables and applying (3.4) and (3.6). Our proof is much longer than it is interesting, so we have omitted it here. In the sequel, we shall assume that the sacrificed animals in each interval have been pooled with the natural deaths.

We are now ready to construct a single degree of freedom test of H based on combining information from all M intervals. We reintroduce the interval indicator m as a subscript on the cell frequencies. Since we have conditioned on the numbers dying in each interval the quantities n_{11m} are independently distributed. Proceeding as in Mantel and Haenszel (1959), our test statistic is:

$$Z = \frac{O - E}{\sqrt{V}}$$

where

$$O = \sum_{m=1}^M n_{11m}, \quad E = \sum_{m=1}^M E[n_{11m}], \quad V = \sum_{m=1}^M \text{Var}[n_{11m}]$$

which we compare with the percentage points of the standard normal distribution. It should be noted that there is a possibility that $O - E = V = 0$, in which case we say the experiment is "not informative." Of course, the technique of summing n_{11m} over all intervals to gain power will be self-defeating if there is a cancellation effect resulting from a change in direction of association over time. When there is sufficient data, a plot of $n_{11m} - E[n_{11m}]$ against m may expose this situation.

Referring to the data for the control group in Table 2, we find that we observe $O = 359$ deaths with neither disease present, we expect $E = 362.83$ under H , with a variance $V = 2.868$. Thus $Z = -2.26$ for a P -value of .012. This would lead us to doubt the hypothesis H of independence of reticulum cell sarcoma and liver tumors in this group.

Suppose that state of lethality were ignored (or not available), and that the lethal and non-lethal categories were combined for each disease. One would then have a sequence of M 2×2 tables reflecting frequencies for findings of presence or absence of Disease A and Disease B , respectively. These could then be analysed using the method of Breslow et al. (1974). Information from the 2×2 tables are combined using the method of Mantel and Haenszel (1959). The

$O = \sum_{m=1}^M n_{11m}$ statistic is the same as before, however now

$$E = \sum_{m=1}^M n_{1.m} n_{.1m} / n_m,$$

and

$$V = \sum_{m=1}^M \frac{n_{1.m}(n_{2.m} + n_{3.m})n_{.1m}(n_{.2m} + n_{.3m})}{n_m^2(n_m - 1)}.$$

The Z -statistic is -2.89 and evidence against the hypothesis of independence in favor of a negative association in this group would appear to be much stronger ($P=0.002$). The results of similar calculations for all three treatment groups and for different interval widths are given in Table 6. The effect of lethality here is apparently to impart a negative bias; this is consistent with the remarks of Breslow et al. (1974).

We may wish to combine information from several groups of animals, representing different treatment by sex combinations, say. In this case, we will have several series of 3×3 tables. It is clear that the same technique can be used to combine all series of tables -- the O , E , and V values are summed over the groups and an overall Z statistic computed. The results of these calculations for our experiment are summarized in Table 6. For these data, the pooled test is not recommended, however, because of the apparent differences among dose groups with respect to association of the two diseases.

Table 6 also shows the same calculations when the ages at death are grouped into 25- or 15-day intervals instead of 50 days. As can be seen, the results appear fairly insensitive to the choice of interval width. This insensitivity was also observed in the simulation study to be described in the next section.

[Table 6 here.]

4. Simulation Experiments

Simulation experiments were carried out to gain an impression of the kind of results that one might expect from applying the 3×3 test proposed in Section 3 to various situations.

The following model was used in the simulations. For Disease A , the potential age at onset was generated as a function of two random variables: a binary variate I_A which indicates (with probability p_A) whether U_A is finite, and a Weibull variate X_A which is assigned to the value of

U_A given U_A is finite. The shape parameter of X_A was fixed at 2, so the survivor function of U_A is $p_A \exp [(-(u/a_A)^2)]$, where a_A is the scale parameter for X_A . We call p_A the "incidence" parameter for Disease A. The potential time from onset until entry into the "lethal" state is $V_A - U_A = Y_A$, say, which was generated as a Weibull variate with survivor function $\exp [(-(y/b_A)^2)]$ independent of U_A . The potential time from entry into the lethal state V_A until A causes death of the animal was taken to be a constant, κ days. Disease B was modelled similarly, but with possibly different values of the parameters p_B , a_B , b_B , and with random variables X_B , Y_B independent of X_A , Y_A .

Dependence between A and B was introduced only through the incidence variates I_A and I_B . We define p_{11} to be the probability that both U_A and U_B are finite, i.e., the animal is "susceptible" to both diseases. Also, p_{12} is the probability that U_A is finite and U_B is infinite (animal "susceptible" to disease A only). The parameters p_{21} and p_{22} are defined analogously. Obviously, $p_{11} + p_{12} = p_A$ and $p_{21} + p_{22} = p_B$. Let $r = p_{11}p_{22} / p_{12}p_{21}$ be the odds ratio. A simple kind of positive (negative) association can be imposed by taking $r > 1$ ($r < 1$). Of course, if $r = 1$, A and B are independent, so the null hypothesis holds.

The potential time to death from "natural" causes was generated from the Weibull distribution with survivor function $\exp [(-(w/c)^\gamma)]$, where $c = 1005$ days and $\gamma = 8$ in all experiments. This yields 5th, 50th, and 95th percentiles of 700, 960 and 1150 days, respectively. Although these lifetimes are somewhat larger than those usually seen in mouse controls, it provides a "standard" life of about 1000 days as a convenient reference point for the disease progression distributions. The variance of this distribution is similar to that seen in C3H female mouse controls.

The parameters p , a , and b for each disease were each chosen from two levels. The incidence parameter was either chosen to be "low" ($p = 0.1$) or "high" ($p = 0.5$). The disease onset distribution was either "early" ($a = 360$ days, median onset age = 300 days) or "late" ($a = 840$ days,

median = 700 days). The time Y from onset to entry to the lethal state, here termed the "course" of the disease, was either "fast" ($b = 12$ days, median time = 10 days) or "slow" ($b = 240$ days, median = 200 days). These parameter values were chosen after discussion with biologists involved in the Oak Ridge C3H mice study. Finally the duration of the lethal state until death was fixed at $\kappa = 3$ days for both diseases. (This choice was rather arbitrary, and a more extensive simulation study would allow consideration of longer and more variable values for time spent in the lethal state.) Taking all possible configurations of the disease process parameters yields eight possible disease types; there are 36 distinct pairs of these for assignment to (A, B) , allowing for the symmetry of A and B . Of these, we chose a subset of 10 that was well balanced with respect to p , a , and b . These are the cases shown in Tables 7 and 8. Although we made an attempt to consider a variety of situations, these cases should be considered as a collection of examples rather than a comprehensive "study."

The left hand side of Table 7 shows the results we obtained under the null hypothesis ($r=1$) in each of the 10 cases, where the sample size was taken to be 500 and the interval width was 50 days. There were 1000 simulated experiments per case; the number of informative experiments is shown together with the rejection rates. In each simulation experiment, we also calculated the test statistic for the test used by Breslow et al. (1974), based on age-stratified 2×2 tables, in which nonlethal and lethal occurrences are pooled. These results are given on the right hand side of Table 7.

In calculating the rejection rates in Table 7, we treated noninformative experiments the same as informative experiments in which the test statistic was not significant. The rejection rate per informative experiment can be recovered from the information in the table if desired.

[Table 7 here.]

Perhaps the most striking aspect of Table 7 is the extreme negative bias shown by the stratified 2×2 test in all the situations considered. Whereas it might be expected from the remarks of Breslow et al. (1974) that this test would have a high rate of reporting negative associations where none exist, the extent of the bias shown by our simulations was surprising to us.

The test based on 3×3 tables maintained the nominal 5% rejection rate quite well. Pooling all experiments, the rejection rate was 4.27% of the total number of experiments, and 5.17% of the number of informative experiments. For reasons that we do not understand, the number of positive rejections was about twice the number of negative rejections.

For each simulation experiment, we also calculated the test statistic using 100-day intervals instead of 50-day intervals. The effect on the rejection rate was surprisingly small. Pooling all experiments, the rejection rate was 4.54% of the total number of experiments, and 5.32% of the total number of informative experiments. These results suggest that the approximation to the integrals in Section 2 does not adversely affect the size of the test, even when transitions occur rapidly relative to the size of the interval.

Even with 500 animals in the group, there can be situations in which the probability is small that the experiment will be informative about disease associations. Note, in particular, the third and ninth cases in Table 7.

Table 8 gives the rejection rates for the 3×3 test when there is a true negative association ($r=0.1$) and a true positive association ($r=10$), again for groups of 500 animals and for 50-day intervals. The main impression given by this table is that one shouldn't be too hopeful about

detecting interactions, even with 500 animals. The chance of detecting interactions is clearly better if both diseases are slowly lethal, and quite bad if both diseases are rapidly lethal.

The chance of detecting positive interactions appears to be greater than the chance of detecting negative interactions. Although this impression is consistent with the tendency toward positive rejections noted in the null case, we should view it cautiously, since we do not really have a good measure of the "magnitude" of the interaction between two disease processes. (In the absence of such a measure, we cannot necessarily view the two cases $r=0.1$ and $r=10$ as being equal in magnitude but opposite in direction.)

[Table 8 here.]

We obtained information similar to that presented in Tables 7 and 8 for 100-day intervals, for sample sizes of 200 and 1000, and for odds ratios (r) of 0.0625 and 16.0. We also did an earlier simulation, under the null hypothesis, for each of the 36 disease type combinations based on groupings of 10-, 25-, 50-, and 100-day intervals and for group sizes of $N = 200$ and $N = 1000$ animals, and repeated it for a larger value of κ (10 days). The results of these additional simulations were qualitatively similar to those we have discussed here.

5. Concluding Remarks

We do not think it possible to test for independence between two diseases using a series of time-stratified 2×2 tables unless one of the diseases is nonlethal. This is because independence relationships that might exist at the start of a given interval cannot otherwise be expected to hold either among the animals that die in that interval or among the survivors of that interval. However, if one is willing to record and use the state (nonlethal or lethal) of each disease found in each animal, then the test we have proposed here can be used. This test avoids biases due to lethality and uses all the data, whether obtained by sacrifice or by natural death.

We recognize that the ability of pathologists to determine accurately and consistently whether or not the presence of a disease is lethal or nonlethal should not be taken for granted. However, the task is in some ways less demanding than the designation of cause of death. What is critical is that diseases that are declared to be in a non-lethal state in a given animal really had no effect on the death of that animal; this is the essence of Assumption A2. When in doubt, it would be prudent to classify a particular occurrence as lethal rather than nonlethal. Falsely declaring that a disease is in a lethal state when it is really in a nonlethal state can reduce the power of the test, but the null distribution should not be affected, since this kind of misclassification does not violate our assumptions.

Unfortunately, the power of the 3×3 test that we propose can be low, depending on the incidence, onset time, and course of the diseases. It is clear for example that testing for associations in groups of size fifty, say, as used in the National Toxicology Program carcinogen bioassay series, would be of little value. However this situation is somewhat alleviated by the fact that the same diseases of interest often occur in many different experiments, and in many different groups of animals within the same experiment. If the nature of the association between two diseases tends to be the same in all of these groups, then the Z-statistics can be combined in the manner indicated in Section 3. We expect also that the situation can be improved further in many cases by means of well chosen sacrifice schedules.

ACKNOWLEDGEMENTS

We are grateful to Dr. J. B. Storer, Dr. R. L. Ullrich, and Dr. R. J. M. Fry of the Oak Ridge National Laboratory for making their data available to us. We would also like to acknowledge the role of Dr. D. A. Wolf, also of the Oak Ridge National Laboratory, who wrote the computer program that originally implemented the method.

REFERENCES

- Breslow, N.E., Day, N.E., Tomatis, L. and Turusov, V.S. (1974). Associations between tumor types in a large-scale carcinogenesis study of CF-1 mice. *Journal of the National Cancer Institute* 52, 233-239.
- Clifford, P. (1977). Nonidentifiability in stochastic models of illness and death. *Proceedings of the National Academy of Science U.S.A.* 74, 579-594.
- Gart, J.J., Krewski, D., Lee, P.N., Tarone, R.E. and Wahrendorf, J. (1986). *The Design and Analysis of Long-Term Animal Experiments*. International Agency for Research on Cancer, Lyon.
- Haseman, J.K. (1983). Patterns of tumor incidence in two-year cancer bioassay feeding studies in Fischer rats.1 *Fundamental and Applied Toxicology* 3,11-9.
- Hoel, D.G. and Walburg, H.E. (1972). Statistical analysis of survival experiments. *Journal of the National Cancer Institute* 49, 361-372.
- Kalbfleisch, J.D., Krewski, D.R. and Van Ryzin, J. (1983). Dose-response models for time-to-response toxicity data (With discussion). *Canadian Journal of Statistics* 11, 25-49.
- Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 22, 719-748.
- McKnight, B. and Crowley, J. (1984). Tests for differences in tumor incidence based on animal carcinogenesis experiments. *Journal of the American Statistical Association* 79, 639-648.
- Mitchell, T.J. and Turnbull, B.W. (1979). Log-linear models in the analysis of disease prevalence data from survival/sacrifice experiments. *Biometrics* 35, 221-234.
- Peto, R. (1974). Guidelines on the analysis of tumor rates and death rates in experimental animals. *British Journal of Cancer* 29,101-105.
- Peto, R., Pike, M., Day, N.E., Gray, R.G., Lee, P.N., Parish, S., Peto, J., Richards, S. and Wahrendorf, J. (1980). Guidelines for simple sensitive significance tests for carcinogenic effects in long-term animal experiments. In *Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal*. IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans. Supplement 2, 311-385. International Agency for Research on Cancer, Lyon.
- Storer, J.B. (1982). Associations between tumor types in irradiated BALB/c female mice. *Radiation Research* 92, 396-404.
- Storer, J.B., Mitchell, T.J., and Fry, R.J.M. (1988). Extrapolation of the relative risk of radiogenic neoplasms across mouse strains and to man. *Radiation Research* 114, 331-353.

- Turnbull, B.W. and Mitchell, T.J. (1978). Exploratory analysis of disease prevalence data from survival/sacrifice experiments. *Biometrics* 34, 555-570.
- Wahrendorf, J. (1983). Simultaneous analysis of different tumor types in a long-term carcinogenicity study with scheduled sacrifices. *Journal of the National Cancer Institute* 70, 915-921.
- Weinberg, A.M. and Storer, J.B. (1985). Ambiguous carcinogens and their regulation (with discussion). *Risk Analysis* 5, 151-167.
- Young, S.S. and Gries, C.L. (1984). Exploration of the negative correlation between proliferative hepatocellular lesions and lymphoma in rats and mice -- establishment and implications. *Fundamental and Applied Toxicology* 4, 632-640.

Table 1
*Total incidences of reticulum cell sarcoma (RCS) and liver tumors
in C3H female mice (Storer, Mitchell, and Fry, 1988)*

Control Group ($n = 503$)				
Liver Tumors				
		Absent	Nonlethal	Lethal
RCS	Absent	359	31	74
	Nonlethal	5	0	0
	Lethal	32	0	2
0.5 Gy Group ($n = 251$)				
Liver Tumors				
		Absent	Nonlethal	Lethal
RCS	Absent	166	28	42
	Nonlethal	2	1	2
	Lethal	10	0	0
1.0 Gy Group ($n = 250$)				
Liver Tumors				
		Absent	Nonlethal	Lethal
RCS	Absent	162	18	54
	Nonlethal	1	0	1
	Lethal	11	2	1

Table 2

*Joint incidences of liver tumors and reticulum cell sarcoma (RCS)
by state for 503 C3H female mice (Control Group)*

RCS* Liver Tumor*		1	1	1	2	2	2	3	3	3	Surviving
		1	2	3	1	2	3	1	2	3	
m	Interval (days)	n_{11m}	n_{12m}	n_{13m}	n_{21m}	n_{22m}	n_{23m}	n_{31m}	n_{32m}	n_{33m}	s_m
1	201-250	1	0	0	0	0	0	0	0	0	502
2	351-400	3	0	0	0	0	0	0	0	0	499
3	401-450	4	0	0	0	0	0	0	0	0	495
4	451-500	9	0	0	0	0	0	0	0	0	486
5	501-550	10	1	0	0	0	0	0	0	0	475
6	551-600	25	1	0	0	0	0	1	0	0	447
7	601-650	16	3	3	0	0	0	1	0	0	424
8	651-700	32	4	5	0	0	0	2	0	1	380
9	701-750	38	4	13	0	0	0	4	0	0	321
10	751-800	61	6	11	0	0	0	5	0	1	238
11	801-850	77	4	11	1	0	0	6	0	0	140
12	851-900	36	3	9	1	0	0	7	0	0	83
13	901-950	25	1	14	0	0	0	4	0	0	39
14	951-1000	11	2	7	3	0	0	1	0	0	15
15	1001-1050	11	2	1	0	0	0	1	0	0	0
TOTAL		359	31	74	5	0	0	32	0	2	503

* Key to States: 1=Absent, 2=Nonlethal, 3=Lethal

Table 3.

3×3 Classification of frequencies for interval m

		DISEASE B		
		Absent	Nonlethal	Lethal
DISEASE A	Absent	n_{11m}	n_{12m}	n_{13m}
	Nonlethal	n_{21m}	n_{22m}	n_{23m}
	Lethal	n_{31m}	n_{32m}	n_{33m}

Table 4.

Form of the cell probabilities in the 3×3 table for interval m , under Assumptions A1, A2, A3, and H. The probabilities are proportional to the entries in the table.

		DISEASE B		
		Absent	Nonlethal	Lethal
DISEASE A	Absent	$q_m \alpha_{1m} \beta_{1m}$	$q_m \alpha_{1m} \beta_{2m}$	$q_m^B \alpha_{1m} \beta_{3m}$
	Nonlethal	$q_m \alpha_{2m} \beta_{1m}$	$q_m \alpha_{2m} \beta_{2m}$	$q_m^B \alpha_{2m} \beta_{3m}$
	Lethal	$q_m^A \alpha_{3m} \beta_{1m}$	$q_m^A \alpha_{3m} \beta_{2m}$	$q_m^{AB} \alpha_{3m} \beta_{3m}$

Table 5a

*Enumeration and exact probabilities for interval $m = 14$ (951-1000 days)
Control Group*

	Probability	n_{11}	n_{12}	n_{13}	n_{21}	n_{22}	n_{23}	n_{31}	n_{32}	n_{33}
1	.00139	15	0	5	0	1	2	0	1	0
2	.00232	15	1	4	0	0	3	0	1	0
3	.00349	14	0	6	0	2	1	1	0	0
4	.00697	13	0	7	1	2	0	1	0	0
5	.00697	13	0	7	2	1	0	0	1	0
6	.00697	14	0	6	1	1	1	0	1	0
7	.01744	14	2	4	0	0	3	1	0	0
8	.02093	14	1	5	1	0	2	0	1	0
9	.02093	14	1	5	0	1	2	1	0	0
10	.03023	12	1	7	3	0	0	0	1	0
11	.04883	13	1	6	2	0	1	0	1	0
12	.09068	12	1	7	2	1	0	1	0	0
13	.09765	13	1	6	1	1	1	1	0	0
14	.14648	13	2	5	1	0	2	1	0	0
15	.18135	11	2	7	3	0	0	1	0	0
16	.31737	12	2	6	2	0	1	1	0	0

1.00000

Table 5b

Exact Distribution of n_{11} for interval $m = 14$ (951-1000 days)

x	11	12	13	14	15
$Pr [n_{11} = x]$	0.18135	0.43828	0.30690	0.06976	0.00371

Table 6

*Summary of Results of Two Methods of Analysis
Association of RCS and liver tumors for data of Table 1*

50-day intervals								
Group	# Animals	Obs.(X)	3×3 Analysis			2×2 Analysis		
			Exp.(E)	Var.	Z	Exp.(E)	Var.	Z
0.0 Gy	503	359	362.83	2.868	-2.26	366.26	6.290	-2.89
0.5 Gy	251	166	165.68	2.030	+0.22	167.01	2.680	-0.62
1.0 Gy	250	162	160.42	1.100	+1.51	162.72	2.822	-0.43
Total	1004	687	688.93	5.998	-0.79	695.99	11.792	-2.62

25-day intervals								
Group	# Animals	Obs.(X)	3×3 Analysis			2×2 Analysis		
			Exp.(E)	Var.	Z	Exp.(E)	Var.	Z
0.0 Gy	503	359	362.77	2.735	-2.28	365.99	6.133	-2.82
0.5 Gy	251	166	165.64	1.942	+0.26	166.93	2.496	-0.59
1.0 Gy	250	162	160.20	0.938	+1.86	162.18	2.387	-0.12
Total	1004	687	688.61	5.615	-0.68	695.10	11.016	-2.44

15-day intervals								
Group	# Animals	Obs.(X)	3×3 Analysis			2×2 Analysis		
			Exp.(E)	Var.	Z	Exp.(E)	Var.	Z
0.0 Gy	503	359	362.63	2.442	-2.33	366.66	5.986	-3.13
0.5 Gy	251	166	165.70	1.933	+0.21	166.51	2.173	-0.34
1.0 Gy	250	162	160.32	0.929	+1.74	162.10	2.335	-0.07
Total	1004	687	688.65	5.304	-0.72	695.27	10.494	-2.55

Table 7

*Selected Simulation Results: Estimated rejection rates under H for
5% two-sided test of Section 3 based on 3×3 tables and for
5% two-sided test based on 2×2 tables (ignoring state of lethality).*

1000 Simulations

Sample Size n=500

50-Day Intervals

		3×3 test			2×2 test		
Disease Types		Inf. Expts	Rejection Rate (%)		Inf. Expts	Rejection Rate (%)	
Dis. A	Dis. B		Neg. Assn.	Pos. Assn.		Neg. Assn.	Pos. Assn.
low, E,S	low, E,S	983	0.7	3.4	1000	100.0	0.0
low, E,S	high,L,F	991	1.9	2.5	1000	100.0	0.0
low, E,F	low, L,F	102	0.1	0.3	1000	100.0	0.0
low, E,F	high,E,S	960	1.6	2.3	1000	100.0	0.0
low, L,S	low, L,S	1000	0.1	5.5	1000	61.8	0.0
low, L,S	high,E,F	881	1.5	3.7	1000	98.2	0.0
low, L,F	high,L,S	1000	0.8	3.9	1000	100.0	0.0
high,E,S	high,L,S	1000	2.9	2.7	1000	100.0	0.0
high,E,F	high,E,F	382	2.9	0.3	1000	100.0	0.0
high,L,F	high,L,F	965	1.4	4.2	1000	100.0	0.0

Key: low = low incidence (p=.1), high = high incidence (p=.5)
E = early onset (a=360), L = late onset (a=840)
F = fast course (b= 12), S = slow course (b=240)
Inf. Expts = Number of informative experiments

Time from lethal state entry until death $\kappa=3$ days.

The rejection rate is the proportion of rejections of the null hypothesis in the 1000 simulations, expressed as a percent.

Table 8

*Selected Simulation Results: Estimated rejection rates for
5% two-sided test of Section 3 based on 3×3 tables
when there is negative or positive association between diseases.*

1000 Simulations

Sample Size n=500

50-Day Intervals

Disease Types Dis. A Dis. B		True Negative Assn.			True Positive Assn.		
		Inf. Expts	Rejection Rate (%)		Inf. Expts	Rejection Rate (%)	
			Neg. Assn.	Pos. Assn.		Neg. Assn.	Pos. Assn.
low, E,S	low, E,S	982	7.8	0.1	1000	0.0	61.3
low, E,S	high,L,F	969	13.9	0.0	995	0.2	13.0
low, E,F	low, L,F	64	0.1	0.1	130	0.0	1.2
low, E,F	high,E,S	968	15.5	0.1	930	0.1	23.6
low, L,S	low, L,S	1000	0.1	0.0	1000	0.0	69.6
low, L,S	high,E,F	904	5.9	0.7	868	0.2	11.5
low, L,F	high,L,S	1000	13.9	0.0	1000	0.1	32.0
high,E,S	high,L,S	1000	92.6	0.0	1000	0.0	92.4
high,E,F	high,E,F	279	6.3	0.1	491	0.6	2.0
high,L,F	high,L,F	970	4.4	0.3	969	0.8	16.3

Key: low = low incidence (p=.1), high = high incidence (p=.5)
 E = early onset (a=360), L = late onset (a=840)
 F = fast course (b= 12), S = slow course (b=240)
 Inf. Expts. = Number of Informative experiments

Time from lethal state entry until death $\kappa=3$ days.

The "true negative association" case is based on setting the odds ratio for the incidence parameters to 0.1;
the "true positive association" case is based on setting this odds ratio to 10.0.

The rejection rate is the proportion of rejections of the null hypothesis in the 1000 simulations,
expressed as a percent.