STATISTICS IN MEDICINE
Statist. Med. **18**, 3389–3400 (1999)

SOME OF NATHAN MANTEL'S CONTRIBUTIONS TO EPIDEMIOLOGY†

MITCHELL H. GAIL*

National Cancer Institute, Division of Cancer Epidemiology and Genetics, Executive Plaza South, Room 8032, 6120 Executive Blvd., MSC 7244, Bethesda, MD 20892-7244, U.S.A.

SUMMARY

Nathan Mantel's many contributions to epidemiology grew out of his role as a statistical consultant and collaborator. His abilities to understand scientific issues, appreciate their subtleties, and produce simple, compelling analyses were remarkable. Several of the procedures he developed to meet consulting needs remain widely used by statisticians and other researchers. Examples are presented to illustrate his approach and intuitive brilliance, including work on the evaluation of diagnostic tests, the Mantel–Haenszel procedure and its extensions, the use of a prospective logistic risk model to analyse case-control data, a method for evaluating aggregation of cancer disease sites in pairs of diseased relatives, and methods for detecting clustering of disease and temporal-spatial association of diseased cases. Published in 1999 by John Wiley & Sons, Ltd. This article is a US Government work and is in the public domain in the United States.

INTRODUCTION

Shortly after I joined the Clinical and Diagnostic Trials Section at the National Cancer Institute in 1972, I wrote my first publishable statistics paper, 'Mixed quasi-independent models for categorical data'.[1] The idea was to model a contingency table distribution as a sum of component distributions, each corresponding to a table with some structural zero cells and with probabilities otherwise determined by the quasi-independence assumption. Before submitting the paper, I brought it to Nathan, hoping for some acknowledgement of its merit. The next day he returned it to me and pointed out a failure of convergence to maximum likelihood in the motivating example. He also suggested initial parameter values that would lead to a correct iterative solution. With this corrective, the paper was accepted, and I avoided the fate of so many authors in *Biometrics*, a letter from Nathan pointing out an error in logic or analysis.

I was stunned at the speed with which Nathan had understood what I had been doing and gone straight to the error in my numerical calculation. Another time I mentioned to Nathan that I was going to be giving lectures on successive Wednesdays in Pittsburgh, but that unfortunately, the special air fares for round trips were only available for stays that encompassed the weekend. Immediately he suggested that I fly from Washington to Pittsburgh on a round-trip flight originating in Washington and return that same day on another round-trip flight originating in Pittsburgh; the following week I could use the return portion of the Pittsburgh flight to get to

* Correspondence to: Mitchell H. Gail, National Cancer Institute, Division of Cancer Epidemiology and Genetics, Executive Plaza South, Room 8032, 6120 Executive Blvd., MSC 7244, Bethesda, MD 20892-7244, U.S.A.

Published in 1999 by John Wiley & Sons, Ltd.

Pittsburgh and the return portion of the Washington flight to get back to Washington, both on the same day.

Insight and originality in facing specific problems was what Sam Greenhouse[2] had in mind when, in speaking of the art of statistics, he said:

> 'And I might mention that among statisticians the world over, we (at NIH) had probably the greatest artist of all – Nathan Mantel. No one could match him in quickly identifying the information in the data relevant to the question and the swiftness with which he was able to choose an optimum method of analysis. The statistical procedures which bear his name are really nothing compared to his ability to analyze data.'

I shall give some examples that illustrate Nathan's originality and powerful intuition in solving problems in epidemiology. The discussion will be limited to problems in evaluating diagnostic tests, to the Mantel–Haenszel approach to combining information from various sources, to a proof that a prospective logistic model that applies to a population can be used to analyse case-control data, to an occupancy problem in testing for familial aggregation of disease, and to two procedures for detecting clustering. Each of these areas grew out of Nathan's role as a statistical consultant. In fact, Nathan's research was not organized around the broad development of general statistical theories. His strength lay in attacking particular consulting problems and data sets and coming up with solutions marked by simplicity, brilliance, attention to the underlying science, and detailed step-by-step exposition of the data, its analysis and interpretation. Nathan said of his research:[3]

> 'Well, I generally don't generate ideas of my own. Someone has to come to me with a problem. And, apparently, I'm pretty good at coming up with solutions or ideas for solutions. Identifying problems is what is important – solutions just follow.'

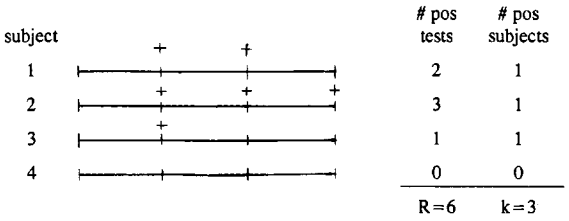
In fact, his solutions were often so useful that they have become eponymous statistical tools, widely used by statisticians and other researchers.

The following examples by no means exhaust Nathan's contributions to epidemiology. There are many topics, such as his extensive contributions to the analysis of contingency tables,[4–6] survival analysis,[7] analysis of time-dependent covariates,[8] development and applications of regression methods,[9] including logistic regression,[4,10,11] interaction[12,13] and occupancy problems,[14–16] only a few of which are cited here, that are relevant to epidemiology but not developed in this paper.

DIAGNOSTIC TESTS AND RELATED APPLICATIONS OF TRUNCATED SAMPLING

Mantel was part of a research team that investigated the sensitivity (which was called 'efficiency', E) of a test to detect the presence of intestinal protozoa in stool specimens.[17] Sensitivity is the probability that the test is positive in infected subjects. The test never gave false positive results (100 per cent specific), but the sensitivity of the zinc sulphate technique was estimated at only 59 per cent.

Hypothetical data of this type are shown in Figure 1, where $N = 4$ individuals were studied at each of $n = 4$ times. Individuals 1, 2 and 3 are known to be infected, because each has at least one positive test result, but individual 4 may or may not be infected. Assuming that each trial of the



A = # positive tests in people known independently to be positive
B = # tests in people known independently to be positive

$$A/B = \frac{2+3+0+0}{4+4+3+0} = \frac{5}{11} = .4545$$

Figure 1. Estimating sensitivity from $N = 4$ hypothetical subjects, each tested $n = 4$ times

assay gives a result that is conditionally independent of the other trials, given the true infection status of the individual, the probability that individual 4 would have no positive tests is $P(1 - E)^4 + 1 - P$, where P is the disease prevalence. In the notation of Figure 1, it can be shown that the maximum likelihood estimate of E satisfies

$$\hat{E} = R[1 - (1 - (1 - \hat{E})^n/nk = 6[1 - (1 - \hat{E})^4]/12 \quad (1)$$

which is solved by $\hat{E} = 0{\cdot}4563$. Formula (1) was given by Sawitz and Karpinos,[18] but without mention of maximum likelihood.

Mantel[19] devised a simpler, non-iterative way to estimate E. Each time an individual was tested, Mantel would score that test as eligible for counting if some other test on that individual were positive, in which case the person was known 'independently' to be infected. Such a test in an independently known positive subject contributed unity to a denominator B, and it contributed unity to a numerator, A, if, in addition, the test itself was positive. Thus, the contribution from the four tests on individual 1 to A and B were, respectively, $0 + 1 + 1 + 0 = 2$ and $1 + 1 + 1 + 1 = 4$. For individual 3, on the other hand, the contributions to A and B are, respectively, $0 + 0 + 0 + 0 = 0$ and $1 + 0 + 1 + 1 = 3$, because at time 2 there is no 'independent' evidence that individual 3 is infected. The resulting estimate, A/B, obtained by summing contributions from the four subjects is $A/B = (2 + 3 + 0 + 0)/(4 + 4 + 3 + 0) = 0{\cdot}4545$, very close to the maximum likelihood value, 0·4563.

This test can be justified informally by noting that A/B is the ratio of the number of positive independent tests to the total number of independent tests performed among individuals independently known to be infected. The procedure has several hallmarks of Mantel. It is simple to compute. It can be easily extended to allow for varying numbers of tests on study subjects and to accommodate simultaneous evaluation of several types of perfectly specific tests on the same individuals.

Li and Mantel[20] applied the A/B procedure to estimate the segregation ratio from families with several offspring siblings, at least one of whom had an abnormal phenotype ('affected'). The segregation ratio is the probability that two parents with normal phenotypes will give rise to an affected offspring. In this context, A is the number of affected siblings who have at least one other affected sibling, and B is the number of siblings who have at least one other affected sibling. Li and Mantel noted that the sampling corresponded to truncated binomial samples, because a family

was only included if at least one sibling was affected. They rigorously demonstrated that A/B was a consistent estimate of the segregation ratio for truncated binomial sampling.

Li and Mantel[20] developed a variance formula for A/B and showed that A/B had an efficiency above 90 per cent in their applications. Referring to Figure 1, we note that to calculate A/B, we discard all tests from individuals with no positive tests, and we discard any positive test in an individual with only one positive test. Individuals with no positive tests are not very informative, because one does not know whether the outcome results from low test sensitivity or from the fact the individual is uninfected. This may account for the good efficiency of the A/B procedure. Li and Mantel extended the truncation to families with an least k affected sibling, for $k \geqslant 2$ and showed how to modify the A/B estimate accordingly.

Before leaving the topic of diagnostic tests, I would like to remind the reader of a result by Greenhouse and Mantel.[21] A quantitative diagnostic test might be termed positive if it exceeds some cut-off value, c. To compare the sensitivities of two tests on samples of diseased subjects, it is necessary that these tests be operating at the same specificity. For one test, the cut-off value c could be estimated by $\hat{G}(\hat{c}) = 0{\cdot}05$, for example, to assure a 95 per cent specificity, where $\hat{G}$ is he empirical distribution of test outcomes in normal subjects. If $\hat{F}$ is the empirical distribution of that test in diseased subjects, the corresponding estimated sensitivity is $1 - \hat{F}(\hat{c})$. Greenhouse and Mantel developed the variance of $1 - \hat{F}(\hat{c})$ and showed that it had two components, one corresponding to binomial sampling in the diseased population, given $\hat{c}$, and another deriving from variability in $\hat{c}$. Based on this idea and its extension, Greenhouse and Mantel proposed procedures for comparing the sensitivities of two diagnostic tests operating at the same specificity, whether the assays were evaluated on independent subjects or on the same subjects.

THE MANTEL–HAENSZEL PROCEDURE

William Haenszel had had considerable experience in the application and interpretation of case-control studies of the effects of smoking on lung cancer, and he invited Mantel 'to help out' in writing something on the use and analysis of retrospective data.[3] Together, they produced a celebrated paper[22] that is among the 200 most cited papers in the scientific literature from 1945 to 1994.[23] The paper stated: 'A primary goal is to reach the same conclusions in a retrospective study as would have been obtained from a forward study, if one had been done'. The paper contains wise guidance on the proper selection of cases and controls, a comparison of some of the strengths and weaknesses of the case-control approach compared to the prospective cohort design ('forward study'), an appreciation of the tentative nature of some conclusions, if many exposure comparisons are made, and various other points on the design and interpretation of case-control studies. However, what differentiated this paper from other fine expositions on case-control studies (see references in reference 24) was the introduction of simple statistical procedures to: (i) provide a summary estimate of exposure effect (the odds ratio) controlled for confounding; and (ii) provide a powerful summary of chi-square test of the null hypothesis that the odds ratio was unity. Further statistical research and practical experience has served to confirm the utility of these procedures, and, in fact, the rate of citation of this paper, about 250 citations per year, now exceeds its citation rate in the 1970s and 1980s.[23]

A central problem in observational studies is the need to control the comparisons for confounding by factors, such as age, that may be related both to exposure and to disease outcome. Although previous efforts to control for confounding had been based on making comparisons of cases and controls within strata defined by levels of confounders,[25,26] it was not until Mantel and

Haenszel published their procedures that simple, reliable tools were available for controlling confounding. Letting A_i, B_i, C_i and D_i denote the numbers of exposed cases, unexposed cases, exposed controls, and unexposed controls, respectively, in stratum i, and letting $T_i = A_i + B_i + C_i + D_i$, $N_{1i} = A_i + B_i$, $N_{2i} = C_i + D_i$, $M_{1i} = A_i + C_i$ and $M_{2i} = B_i + D_i$, Mantel and Haenszel proposed the summary adjusted odds ratio estimator

$$\mathrm{OR} = (\sum A_i D_i / T_i)(\sum B_i C_i / T_i)^{-1}. \tag{2}$$

This estimator can be regarded as a weighted average of the stratum-specific odds ratios,

$$\mathrm{OR} = \sum w_i (A_i D_i / B_i C_i)$$

where

$$w_i = (B_i C_i / T_i)(\sum B_i C_i / T_i)^{-1}.$$

The estimator (2) has several valuable properties. It is simple to compute. It performs well even if many strata are sparse, unlike methods based on unconditional maximum likelihood for a logistic model.[27] Mantel and Haenszel regarded the OR estimate (2) as a weighted average of odds ratios that could vary among strata, useful even when stratum-specific odds ratios were heterogeneous. They noted that maximum likelihood estimates of a common odds ratio could be obtained under the assumption that odds ratios were homogenous across strata but commented that 'the assumption of a constant relative risk can be discarded as usually untenable'. None the less, subsequent research has shown that the simple, non-iterative estimate (2) is very efficient when the odds ratio is homogenous, both in studies with a few large strata[28] and with many sparse strata.[29] The usefulness of this estimator has been enhanced by procedures for estimating its variance[30] and constructing confidence intervals based on it.[31]

The odds ratio estimator can be used even if observations are dependent within strata, as in studies of familial aggregation of disease,[32] and adapted to regression methods for modelling heterogeneity in odds ratios across strata.[33]

A second major methodologic contribution of the paper by Mantel and Haenszel was the development of a one degree-of-freedom chi-square test of the null hypothesis that all stratum-specific ratios were unity, with good power against an alternative in which all odds ratios were greater (or less) than unity. The proposed statistic was

$$\chi_1^2 = \{|\sum A_i - \sum E(A_i)| - 0{\cdot}5\}^2 / \sum V_i \tag{3}$$

where $E(A_i) = N_{1i} M_{1i} / T_i$ and $V_i = N_{1i} N_{2i} M_{1i} M_{2i} / T_i^2 (T_i - 1)$. This test is easy to compute, valid with sparse data and efficient against the alternative of a common odds ratio different from unity. Indeed, Cochran[34] had shown that an asymptotically equivalent statistic was efficient for testing against this alternative, when combining data from a few large strata. Birch[35] and Day and Byar[36] showed that equation (3) is a score test against this alternative based on a conditional likelihood with all margins N_{1i}, N_{2i}, M_{1i} and M_{2i} fixed. This conditional score test had been advocated as yielding the uniformly most powerful unbiased test.[37]

The paper by Mantel and Haenszel included two other methodological elements of considerable practical importance. They treated the analysis of pair-matched data as a special case, with each pair defining a separate stratum. They showed that equation (2) reduced to the ratio of the number discordant pairs with the case exposed to the number of discordant pairs with the control exposed. Moreover, equation (3) yielded a continuity-corrected version of the McNemar test for paired proportions.[38] The paper also included extensions to allow for multiple unordered

exposure categories, including a summary χ^2_{r-1} test analogous to equation (3) for testing equivalence of odds ratios in r exposure categories (see also reference 35).

Four years later, Mantel[39] published a paper extending the range of applications of the procedures in Mantel and Haenszel.[22] He pointed out that these procedures were general techniques for combining data from various sources (strata), whether the data arose in retrospective studies, in prospective studies, or in laboratory experiments (see also reference 40). These procedures could be applied, for example, to provide age-adjusted comparisons of death rates across age strata in cohort studies, or to test for the effectiveness of a factor in bioassay while adjusting for the effect of another exposure. This paper also mentions that the chi-square test (3) can be used to compare two actuarial survival curves, with each interval at risk regarded as a separate stratum. Mantel later developed this theme to produce the first version of the logrank test for comparing censored time-to-response distributions.[7]

The paper on extensions also introduced a one degree-of-freedom chi-square test for detecting a trend in proportions, while combining information across strata. Suppose a score Y_j is attached to the jth exposure category, to which M_j subjects had been exposed, and of those exposed, X_i had the response 1 (for example, diseased) and $M_j - X_j$ the response 0 (for example, free of disease). Cochran[34] and Armitage[41] had previously used weighted least squares regression to test for a linear trend in the proportions with disease, X_j/M_j, against the score Y_j. Mantel[39] extended this analysis to allow for stratification by considering the distribution of the sum over strata of the stratum-specific centred quantities

$$\sum Y_j X_j - E(\sum Y_j X_j) = \sum Y_j X_j - (\sum M_j Y_j)(\sum X_j)/(\sum M_j). \tag{4}$$

This quantity is proportional to the slope parameters in the regression of X_j/M_j on Y_j in each stratum.[34,41] Mantel, however, obtained the mean and variance of this quantity within each stratum by considering the permutational distribution of $\sum X_j Y_j$ conditional on $\{M_j\}$ and $\sum X_j$. Under the null hypothesis of no association, one can imagine an urn containing M_j of the scores Y_j for categories $j = 1, 2, \ldots, J$, and the total score allocated to the diseased group, $\sum X_j Y_j$, can be regarded as distributed as the sum of a random sample of $\sum X_j$ objects, drawn without replacement from the urn. Thus the mean of the stratum-specific centred quantity above is zero and its variance is obtained from the theory of sampling from a finite population. The summary 1 degree-of-freedom chi-square is taken as the sum of these centred quantities, all squared, divided by the sum of the corresponding finite sample variances.

A nice application of this procedure was the demonstration by Stark and Mantel[42] that birth order, scored as 1, 2, 3, 4 or 5 (for a fifth or later birth), had no statistically significant association with the risk of Down's syndrome, once one adjusted for maternal age and study period ($\chi^2_1 = 2{\cdot}49$). In contrast, the adjusted association with maternal age was impressive ($\chi^2_1 = 1840$).

This paper extending the Mantel–Haenszel procedure to test for trends in proportions[39] and the adaptation to test for differences in survival curves[7] is currently cited about as frequently as the original paper.[22] In referring to the original paper and its extensions, Mantel recently remarked:[3] 'It turned out that the procedures in the paper could be extended so that they met perhaps 90 to 95 per cent of the kinds of problems that people were encountering'.

LOGISTIC REGRESSION IN CASE-CONTROL STUDIES

Mantel was asked by Tavia Gordon[3] at the National Heart and Lung Institute to assist in the analysis of data from a cohort of about 4000 subjects, of whom 165 developed disease.

The risk of disease ($Y = 1$) was assumed to conform to a logistic risk model

$$P(Y = 1|X) = \frac{\exp(\mu + \beta X)}{1 + \exp(\mu + \beta X)}$$

where $Y = 1$ or 0. At the time, computer facilities were limited, and computational capabilities for fitting the logistic model were insufficient for a cohort of this size.[10] Mantel suggested that all cases ($Y = 1$) be analysed but only a fraction of the subjects with $Y = 0$, just as in a case-control study. More generally, assuming that a fraction π_1 of cases ($Y = 1$) and π_0 of non-cases ($Y = 0$) were sampled at random for analysis, Mantel calculated that

$$\begin{aligned} P(Y = 1|X, \text{sampled}) &= \frac{\pi_1 P(Y = 1|X)}{\pi_1 P(Y = 1|X) + \pi_0 P(Y = 0|X)} \\ &= \exp(\mu^* + \beta x)/\{1 + \exp(\mu^* + \beta x)\} \end{aligned}$$

where $\mu^* = \mu + \log(\pi_1/\pi_0)$. Although motivated by a need to find a way to analyse a large cohort study with limited computing resources, this conceptualization applies perfectly to the idealized case-control study, namely a random sample of those with ($Y = 1$) and without disease ($Y = 0$). This is the first clear exposition of the fact that if a logistic risk model applies in the general population, it also applies, except for a change of intercept, in a case-control sample from that population. Siegel and Greenhouse reached similar conclusions provided a normal discriminant model held for cases and controls,[43] but this is a more restrictive starting point than the assumption of a logistic model for the general population. In a pathbreaking paper, Anderson[44] noted that the result held provided the logistic model described the general population, and he also showed that fitting prospective logistic model to case-control data yielded the maximum likelihood estimate of β for the retrospective likelihood based on $P(X|Y)$.

FAMILIAL AGGREGATION OF CANCER

Mantel was asked to consult on a project to study the distribution of cancer sites among pairs of relatives, both of whom had cancer.[45] If, for example, stomach cancer was common in pairs with no recent common ancestors, such as husbands and wives, a common environmental exposure in the family, such as diet, might be implicated, whereas, if such an association were found in genetically related pairs, such as siblings, it might be due either to common environmental or hereditary factors.

Hypothetical data of this type are shown for 23 husband–wife pairs in Figure 2. To test for aggregation of sites, Chen *et al.*[45] compared the sum of the diagonal elements in such a table, $\sum m_{ii}$, with its expectation under the null hypothesis of no association. Conditioning on the margins of this table, Chen *et al.* obtained a one degree-of-freedom chi-square

$$\chi_1^2 = \{|12 - (9 \times 9 + 5 \times 7 + 9 \times 7)/23| - 0{\cdot}5\}^2/5{\cdot}1859 = 2{\cdot}66. \quad (5)$$

The mean 7·7826 and variance 5·1859 were obtained from standard theory for the multivariate hypergeometric distribution. Mantel derived this theory by considering two urns. In this example, one urn for the husbands would contain 9 stomach, 5 lung and 9 skin cancer labels, while an independent urn for the wives would contain 9, 7 and 7 corresponding labels. Pairs of labels are drawn by selecting one pairmate each from each urn, without replacement. Needed means and variances follow from this model.

	Wife			
Husband	Stomach	Lung	Skin	
Stomach	6	2	1	9
Lung	1	2	2	5
Skin	2	3	4	9
	9	7	7	23 pairs (N)

Figure 2. Hypothetical distribution of sites of cancers in husband–wife pairs

	Sister			
Sister	Stomach	Lung	Skin	
Stomach	6	3	3	12
Lung		2	5	7
Skin			4	4
				23 pairs (N)

Figure 3. Hypothetical distribution of sites of cancers in sister pairs

Suppose, however, that the data in Figure 2 represented results from pairs of siblings. Sisters, for example, are genetically exchangeable. There is usually no reason to distinguish one sibling from the other. Does this change our expected value for $\sum m_{ii}$ or its variance? Chen *et al.* depicted data for siblings as in Figure 3. These data are obtained from Figure 2 by combining the symmetrically disposed off-diagonal cells from Figure 2. Thus, there were $3 + 2 = 5$ sister pairs involving lung and skin cancer, without regard to which sister had lung cancer and which had skin cancer. This insightful representation of the data corresponds to an entirely different urn model in which a single urn now contains $2 \times 6 + 3 + 3 = 18$ stomach cancer labels, $2 \times 2 + 3 + 5 = 12$ lung cancer labels, and $2 \times 4 + 3 + 5 = 16$ skin cancer labels. These labels are drawn from the single urn in pairs without replacement. Even the expectation of m_{ii} is different from the case with distinguishable relatives. For example, the expected value of $\sum m_{ii}$ is $(18 \times 17 + 12 \times 11 + 16 \times 15)/2\,(46 - 1) = 7{\cdot}5333$, instead of previous value 7·7826 under generalized hypergeometric sampling. The corresponding chi square is $\chi_1^2 = \{|12 - 7{\cdot}5333| - 0{\cdot}5\}^2/5{\cdot}0489 = 3{\cdot}12$, which is somewhat larger than the previous value, 2·66.

Using these techniques, Chen *et al.* demonstrated aggregation of cancer sites not only for sister pairs and brother pairs, but for husband–wife pairs, indicative of a possible role for environmental as well as hereditary factors.

This problem nicely illustrates Mantel's ability to identify the essential features of a problem, including easily overlooked subtleties, and to come up with a simple, compelling solution.

DISEASE CLUSTERING

Mantel was asked by Fred Ederer[3] to consult on a study to determine whether there was a tendency for temporal clustering of leukaemia cases within units defined by town and five-year periods.[46] Suppose, for example that two cases occurred in the first year of such a unit, and none occurred in each of the other 4 years. Under the null hypothesis that these two cases were equally

likely to have been allocated to each of the 5 years, the probability of the reverse order statistics $r_1 = 2$, $r_2 = r_3 = r_4 = r_5 = 0$ is

$$\left(\frac{2!}{2!0!0!0!0!}\right)(0{\cdot}2)^2\left(\frac{5!}{4!0!1!}\right) = 0{\cdot}20. \tag{6}$$

The first two factors are simply the multinomial probability of the observed outcome for a given ordering of the years, and the last factor is the number of ways that four of the years would have no counts, none of the years would have one count, and one of the years would have two counts. A general expression was given for arbitrary reverse order statistics $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_s$, and the expected value of r_1 and its variance were computed for each fixed $r = \sum r_i$ and s. A summary chi-square statistic was proposed as

$$\{|\sum r_1 - \sum E(r_1)| - 0{\cdot}5\}^2 / \sum \mathrm{var}(r_1)$$

where the summation is over strata, in this example composed of a cross-classification on towns and 5-year periods. Ederer *et al.*[46] found strong evidence of clustering for polio and hepatitis using these methods, but not for childhood leukaemia.

Although temporal or spatial clustering of cases could be an indication of a contagious element, a high correlation of spatial and temporal proximity among cases can also provide a clue to contagion. As Mantel[47] pointed out, one can look for a temporal-spatial association by examining cases alone, without having to estimate the sizes of the underlying populations or the underlying disease rates. Letting X_{ij}, measure the closeness in space between cases i and j, and Y_{ij} the closeness in time, Mantel proposed to study the statistic $Z = \sum X_{ij} Y_{ij}$, where the summation is over all pairs of cases. In the special case where $X_{ij} = 1$ or 0 according to whether there is proximity in space or not and $Y_{ij} = 1$ or 0 according as to whether there is proximity in time, this statistic reduces to the numbers of pairs of cases that are close both in time and space, namely the Knox statistic.[48] Mantel developed the permutational mean and variance of Z by considering, the permutational distribution of $\{Y_{ij}\}$ for fixed $\{X_{ij}\}$. The simplest way to generate this distribution is to hold each case's position fixed but to permute the cases' times of disease onset to obtain $n!$ rearrangements of times. These rearrangements induce $n!$ new versions of $\{Y_{ij}\}$. As Mantel[47] noted, Monte Carlo methods can be used to estimate the distribution of Z by sampling from the permutation set. Mantel[47] also showed how to calculate the mean and variance of Z under these permutations and appealed to the 'assumption that Z is approximately normally distributed' for testing $E(Z) \neq 0$. Under mild conditions on the functions X_{ij} and Y_{ij}, asymptotic normality does follow because Z is a U statistic.[49] The ideas in this paper were modified to develop a permutational test of clustering in space[50] and to develop a permutational test of whether individuals who were close with respect to a set of p variables were also close with respect to a set of q other variables.[51] Despite the elegance of the approach in reference 47, subsequent research has shown that the power of this test can be limited for detecting contagion in diseases with long and variable latency.[52]

DISCUSSION

Several of these examples illustrate Nathan Mantel's brilliance and penetrating understanding of the problems he confronted. A deeper appreciation can be obtained by reading his papers.

The consulting problem itself is usually central, and the data are presented forcefully and clearly. Nathan wrote:[53]

> I prefer to have experimental results recorded so there is a minimum obstruction between the reader and the data. It is so arranged as to permit a rather full impact of the source of the data. For this purpose, the usual statistical significance tests can sometime be just so much window dressing, and are frequently not so helpful as simple descriptive statistics derived from the data.

The tools he developed flowed from the problem, and unnecessary complexity was avoided.

Nathan thrived at NIH because of opportunities for consultation and collaboration. People from various institutes knocked on his door looking for advice and insight, and their problems stimulated his interest and creative imagination. Nathan was able to impart some of the art of data analysis to two outstanding young statisticians, Charles C. Brown and David P. Byar, with whom he wrote several important publications. Dave Byar, too, loved to explore a data set until he understood it completely and would sometimes say, 'We tortured the data until it confessed'. Dave used only a few tools in his work. On his blackboard were some Taylor series expansions that never changed, and dominating the adjacent space was a big black-and-white photograph of Nathan Mantel. I imagine that when Dave faced a particularly challenging problem in data analysis and Nathan was not available for consultation, Dave looked to this picture for inspiration. That might not be a bad idea for the rest of us.

REFERENCES

1. Gail, M. H. 'Mixed quasi-independent models for categorical data', *Biometrics*, **28**, 703–712 (1972).
2. Greenhouse, S. W. 'Some reflections on the beginnings and development of statistics in "Your Father's NIH"', *Statistical Science*, **12**, 82–87 (1997).
3. Gail, M. H. 'A conversation with Nathan Mantel', *Statistical Science*, **12**, 88–97 (1997).
4. Mantel, N. 'Models for complex contingency tables and polychotomous dosage response curves', *Biometrics*, **22**, 83–95 (1966).
5. Gail, M. H. and Mantel, N. 'Counting the number of $r \times c$ tables with fixed margins', *Journal of the American Statistical Association*, **72**, 859–862 (1977).
6. Mantel, N. and Byar, D. P. 'Marginal homogeneity, symmetry and independence', *Communications in Statistics – Theory and Methods*, **7**, 953–976 (1978).
7. Mantel, N. 'Evaluation of survival data and two new rank order statistics arising in its consideration', *Cancer Chemotherapy Reports*, **50**, 163–170 (1966).
8. Mantel, N. and Byar, D. P. 'Evaluation of response-time data involving transient states: An illustration using heart-transplant data', *Journal of the American Statistical Association*, **69**, 81–86 (1974).
9. Mantel, N. 'Why step down methods in variable selection', *Technometrics*, **12**, 621–625 (1970).
10. Mantel, N. 'Synthetic retrospective studies and related topics', *Biometrics*, **29**, 479–486 (1973).
11. Mantel, N. and Brown, C. 'A logistic analysis of Sowden's data on respiratory symptoms in British Coal Miners', *Biometrics*, **29**, 649–665 (1973).
12. Mantel, N., Brown, C. and Byar, D. P. 'Tests for homogeneity of effect in an epidemiologic investigation', *American Journal of Epidemiology*, **106**, 125–129 (1977).
13. Halperin, M., Ware, J. H., Byar, D. P., Mantel, N., Brown, C. C., Koziol, J., Gail, M. H. and Green, S. B. 'Testing for an interaction in an $I \times J \times K$ contingency table', *Biometrika*, **64**, 271–275 (1977).
14. Mantel, N. and Pasternack, B. S. 'A class of occupancy problems', *American Statistician*, **22**, 23–24 (1968).
15. Mantel, N. and Bailar, J., III. 'A class of permutational and multinomial tests arising in epidemiological research', *Biometrics*, **26**, 687–700 (1970).

16. Gail, M. H., Weiss, G. H., Mantel, N. and O'Brien, S. J. 'A solution to the generalized birthday problem with application to allozyme screening for cell culture contamination', *Journal of Applied Probability*, **16**, 246–251 (1979).
17. Tobie, J. E., Reardon, L. V., Bozicevich, J., Shih, B.-C., Mantel, N. and Thomas, E. H. 'The efficiency of the zinc sulfate technic in the detection of intenstinal protozoa by successive stool examinations', *American Journal of Tropical Medicine*, **31**, 552–560 (1951).
18. Sawitz, W. and Karpinos, B. D. 'Statistical problems involved in the application of the N.I.H. swab for diagnosis of oxyuriasis', *American Journal of Hygiene*, **35**, 15–26 (1942).
19. Mantel, N. 'Evaluation of a class of diagnostic tests', *Biometrics*, **7**, 240–246 (1951).
20. Li, C. C. and Mantel, N. 'A simple method of estimating the segregation ratio under complete ascertainment', *American Journal of Human Genetics*, **20**, 61–81 (1968).
21. Greenhouse, S. W. and Mantel, N. 'The evaluation of diagnostic tests', *Biometrics*, **6**, 399–412 (1950).
22. Mantel, N. and Haenszel, W. 'Statistical aspects of the analysis of data from retrospective studies of disease', *Journal of the National Cancer Institute*, **22**, 719–748 (1959).
23. Breslow, N. E. 'Statistics in epidemiology: the case-control study', *Journal of the American Statistical Association*, **91,** 14–28 (1996).
24. Gail, M. H. 'Statistics in action', *Journal of the American Statistical Association*, **91**, 1–13 (1996).
25. Wynder, E. L., Cornfield, J., Schroff, P. D. and Doraiswami, K. R. 'A study of environmental factors in carcinoma of the cervix', *American Journal of Obstetrics and Gynecology*, **68**, 1016–1052 (1954).
26. Haenszel, W., Shimkin, M. B. and Mantel, N. 'A retrospective study of lung cancer in women', *Journal of the National Cancer Institute*, **21**, 825–842 (1958).
27. Breslow, N. E. and Day, N. E. *Statistical Methods in Cancer research*, *Volume* 1: *The Analysis of Case-control Studies*, International Agency for Research on Cancer, Lyon, 1980, Chapter VII.
28. Donner, A. and Hauck, W. W. 'The large-sample relative efficiency of the Mantel-Haenszel estimator in the fixed-strata case', *Biometrics*, **42**, 537–545 (1986).
29. Hauck, W. W. and Donner, A. 'The asymptotic relative efficiency of the Mantel-Haenszel estimator in the increasing number-of-strata case', *Biometrics*, **44**, 379–384 (1988).
30. Robins, J., Breslow, N. E. and Greenland, S. 'Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models', *Biometrics*, **42**, 311–323 (1986).
31. Sato, T. 'Confidence intervals for the common odds ratio based on the asymptotic distribution of the Mantel-Haenszel estimator', *Biometrics*, **46**, 71–80 (1990).
32. Liang, K. Y. 'Odds ratio inference with dependent data', *Biometrika*, **72**, 678–682 (1985).
33. Liang, K. Y. 'Extended Mantel-Haenszel estimating procedure for multivariate logistic regression models', *Biometrics*, **43**, 289–299 (1987).
34. Cochran, W. G. 'Some methods of strengthening the common χ^2 test', *Biometrics*, **10**, 417–451 (1954).
35. Birch, M. W. 'The detection of partial association, II. The general case', *Journal of the Royal Statistical Society*, *Series B*, **27**, 111–124 (1965).
36. Day, N. E. and Byar, D. P. 'Testing hypotheses in case-control studies-equivalence of Mantel-Haenszel statistics and logit score tests', *Biometrics*, **35**, 623–630 (1979).
37. Birch, M. W. 'The detection of partial association, I. The 2×2 cases', *Journal of the Royal Statistical Society*, *Series B*, **26**, 313–324 (1964).
38. McNemar, Q. 'Note on the sampling error of the difference between correlated proportions or percentages', *Psychometrika*, **12**, 153–157 (1947).
39. Mantel, N. 'Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure', *Journal of the American Statistical Association*, **58**, 690–700 (1963).
40. Mantel, N. 'A personal perspective on statistical techniques for quasi-experiments', in Owen, D. B. (ed.), *On the History and Probability*, New York, Marcel Dekker, 1976, pp. 103–129.
41. Armitage, P. 'Tests for linear trends in proportions and frequencies', *Biometrics*, **11**, 375–386 (1955).
42. Stark, C. R. and Mantel, N. 'Effects of maternal age and birth order on the risk of mongolism and leukemia', *Journal of the National Cancer Institute*, **37**, 687–698 (1966).
43. Siegel, D. G. and Greenhouse, S. W. 'Multiple relative risk functions in case-control studies', *American Journal of Epidemiology*, **97**, 324–331 (1973).
44. Anderson, J. A. 'Separate sample logistic discremination', *Biometrica*, **59**, 19–35 (1972).
45. Chen, W. Y., Crittenden, L. B., Mantel, N. and Cameron, W. R. 'Site distribution of cancer deaths in husband-wife and sibling pairs', *Journal of the National Cancer Institute*, **27**, 875–892. (1961).

46. Ederer, F., Myers, M. H. and Mantel, N. 'A statistical problem in space and time. Do leukemia cases come in clusters?', *Biometrics*, **20**, 626–638 (1964).
47. Mantel, N. 'The detection of disease clustering and a generalized regression approach', *Cancer Research*, **27**, 209–220 (1967).
48. Knox, G. 'The detection of space-time interactions', *Applied Statistics*, **13**, 25–29 (1964).
49. Hoeffding, W. 'A class of statistics with asymptotically normal distribution', *Annals of Mathematical Statistics*, **19**, 293–325 (1948).
50 Mantel, N. and Bailar, J. C., III. 'Letter: Model I or Model II in competing risk analysis?', *Biometrics*, **26**, 861–863 (1970).
51. Mantel, N. and Valand, R. S. 'A techniques of nonparametric multivariate analysis', *Biometrics* **26**, 547–558 (1970).
52. Chen, R., Mantel, N. and Klingberg, M. A. 'A study of three techniques for time-space clustering in hodgkin's disease', *Statistics in Medicine*, **3**, 173–184 (1984).
53. Mantel, N. 'Some statistical viewpoints in the study of carcinogenesis', *Progress in Experimental Tumor Research*, **11**, 431–443 (1969).