CAUTIONS ABOUT THE ESTIMATION OF SUMMARY RELATIVE RISK FROM VETERINARY EPIDEMIOLOGIC STUDIES

Alan P. Bender, D.V.M., Ph.D.*
George E. Policello, II, Ph.D.+
C. Richard Dorn, M.P.H., D.V.M.++
Robert A. Robinson, D.V.M., Ph.D.#

Epidemiologic studies of animal populations have often been used to evaluate hypotheses about the natural history of animal diseases. The case-control epidemiologic design has been particularly popular in veterinary research. This design compares cases of a given disease to suitable controls for their "exposure" to putative risk factors. An association between the disease and the factor is said to exist when a statistically significant larger proportion of the cases have a positive history for the factor. One method for analysis of the case-control study arrays the data into strata of 2 x 2 tables. Each table represents the counts of cases and controls that were "exposed" and "not exposed" to the factor under study. The strata consist of categories of other variables such as age, breed, or sex that are associated with both the disease state and the exposure status. The odds ratio is often used as the estimate of relative risk within each of the 2 x 2 tables (Cornfield and Haenszel 1960) and a summary relative risk for all the strata is usually calculated. The combination of evidence from multiple 2 x 2 tables is appealing to veterinary researchers who generally have many stratifying variables for which a single summary relative risk estimate would be convenient.

Of the many methods for combining estimates from stratified 2 x 2 tables the Mantel-Haenszel (1959) procedure (M-H) is the most popular and it has several statistical advantages over the other methods (Birch 1974, McKinlay 1978). Using the M-H procedure as an example it is the intent of this article to highlight known limitations on

*Chief, Chronic Disease Epidemiology Section, Minnesota Department of Health, 717 S.E. Delaware Street, Minneapolis, Minnesota 55440
+Research Statistician, Bell Laboratories, Room 2B-200, 6200 East Broad Street, Columbus, Ohio 43231
++Chairman, Department of Veterinary Preventive Medicine, The Ohio State University, Columbus, Ohio 43220
#Professor, Department of Large Animal Clinical Sciences, University of Minnesota, St. Paul, Minnesota 55108

the calculation of summary relative risks and to indicate how the use of log-linear models can help identify the situations when the summary statistic may be misleading.

THE MANTEL-HAENSZEL ESTIMATOR

The odds ratio for a 2 x 2 table is usually taken as an approximation to the relative risk of the disease under study (Cornfield and Haenszel 1960). For a 2 x 2 x k table (k strata), the M-H estimator is often interpreted as an overall approximate relative risk summarizing the odds ratios in the k subtables. The M-H procedure has traditionally been used to control for confounding variables (Dayal 1978). Rothman (1975) has stated that control of confounding variables is central to the methodologic considerations of analytical epidemiology. He defined confounding as a distortion in the estimate of association (odds ratio) between a suspected factor and the disease state produced by relationships between the factor and disease with an extraneous variable-the confounding variable. A similar definition of confounding was given by Miettinen (1970). The following example illustrates the dramatic effect that confounding can have upon the estimation of association and also illustrates that the M-H procedure can prevent this confounding from occurring.

For each level of the stratifying variable in table 1 there is a significant ($p < 0.05$) odds ratio of 5.21. When the tables are summed (collapsed) over the stratifying variable the odds ratio becomes 1.17 and confounding has occurred. The application of the M-H procedure to the stratified subtables yields a result of 5.21; a significant ($p < 0.005$) summary estimate of the approximate relative risk. If the M-H procedure had not been used, then a distortion-confounding-of the single estimate of the relative risk in the collapsed table would have occurred. In this example the two sample tables, before collapsing, have identical odds ratios. This implies there is a common population odds ratio at all levels of the stratifying variable. It is the preservation of this common value that is a major goal of the M-H procedure.

Table 1. Hypothetical data representing a case-control study where confounding would occur if stratification was not maintained.

	STRATUM			
	1		2	
	Factor Present	Factor Absent	Factor Present	Factor Absent
Case	10	190	198	2
Control	2	198	190	10

Thus, the M-H procedure is well suited to control for confounding. However, the concept of "control" has apparently been subject to very diverse interpretations. There has appeared in the veterinary literature many case-control studies that have claimed to have used M-H or similar procedures to either "adjust", "standardize", "control", or "account" for the stratifying variable. These terms could carry the misleading impression that the application of the procedure made the analyses and conclusions independent of the stratifying variables. A notion of standardization similar to that used in the direct standardization of mortality rates (where the differences in the age structure are accounted for in the estimates) is clearly not valid for these summary relative risk procedures.

COMBINATION OF EVIDENCE FROM STRATIFIED 2 x 2 TABLES

In order to summarize data from stratified 2 x 2 tables the individual odds ratios should not be qualitatively different (McKinlay 1978). Certainly, some tables should not have odds ratios less than one and others greater than one if the combined estimate is to be meaningful.

Consider the data in table 2. The data has been arranged so that each stratum has identical weight as used in the M-H procedure. Stratum 1 has an odds ratio of 2.10 and stratum 2 has an odds ratio of 0.48, the reciprocal of 2.10. That is, the estimates of approximate relative risk are reversed for the two strata. This is a situation in which the odds ratios are not consistent over the strata and any summary measure of the overall risk is inappropriate. However, if the epidemiologist was not alert to the lack of homogeneity and applied the M-H procedure, a result of 1.29 ($p < 0.05$) would be obtained for the summary relative risk. This conclusion not only disguises the basic differences in these two subtables but it is also counter-intuitive in that any summary of the data "should" not have been statistically different from 1.0. The two equal sized subtables had exactly opposite odds ratio structures. The false impression of an overall significantly increased risk was caused by inappropriate application of a summary statistic to subtables where different data structures-odd ratios-were present.

Valuable epidemiologic relationships may also be lost through the inappropriate use of summary procedures. The data in table 3 is condensed for illustration from an epidemiologic study of canine multiple primary neoplasia* (MPN) (Bender et al. 1982). There were thirteen breeds in the stratifying variable (B). The other variables were the two female sexes (S) - intact female and spayed female- and the tumor status (T) - multiple primary case and nonneoplastic

*MPN is the development of distinctly different neoplasms in the same dog.

Table 2. Hypothetical data representing a case-control study where the requirement for homogeneity of the odds ratio is not satisfied.

	STRATUM 1			STRATUM 2		
	Factor Present	Factor Absent	Total	Factor Present	Factor Absent	Total
Case	210	100	310	25	100	125
Control	200	200	400	200	385	585
Total	410	300	710	225	485	710

control. Table 4 contains a log-linear model screening summary of these data (Brown 1976, Dixon and Brown 1979). Log-linear models provide a systematic method for evaluating the effect of secondary relationships within the data on the resultant odds ratios. For this example, the odds ratio of interest, the approximate relative risk of MPN in intact females compared to spayed females stratified by breed, is represented as the ST effect (interaction). Table 4 clearly demonstrates that the ST effect (odds ratio) is not consistent over the breed stratification since the STB term is significantly different from zero (Bishop et al. 1975, Heilbron 1981). That is, the individual odds ratios are highly dependent upon the breed variable.

If the epidemiologist was not aware of the structure of these data the application of the M-H procedure to obtain a summary approximate relative risk would yield a result of 1.04. The associated approximate 95% confidence interval is (.94, 1.14) and the incorrect conclusion of no association between sex and tumor status might be made. This conclusion would result in the loss of potentially important epidemiologic information. In particular, table 5 gives the significant breed-specific approximate relative risks that would have been incorrectly summarized. The range of breed-specific odds ratios from 0.33 to more than 2.0 indicates that no summary statistic is adequate for these data. In the words of Mantel and Haenszel (1959) "The overall relative risk estimates are averages and as averages may conceal substantial variation in the magnitudes of the relative risk among subgroups". Also, it should be clear that the summary procedure has not adjusted, controlled standardized, nor accounted for the breed variable in any meaningful way.

Table 3. Cross-Tabulation of tumor status (multiple or none) versus sex (intact female or spayed female) stratified by breed.

Breed	Tumor Status	Spayed Female	Intact Female
Beagle	Multiple	35	32
	None	62	36
Boxer	Multiple	79	50
	None	40	45
Chihauhau	Multiple	9	28
	None	24	75
Dachshund	Multiple	70	98
	None	160	110
Doberman	Multiple	20	33
	None	57	119
German Shepherd	Multiple	48	58
	None	242	211
Mongrel	Multiple	446	382
	None	1161	885
Other Pedigree	Multiple	183	200
	None	455	767
Poodle	Multiple	137	296
	None	278	364
Retriever	Multiple	32	38
	None	95	119
Setter	Multiple	20	17
	None	35	89
Spaniel	Multiple	108	57
	None	130	119
Terrier	Multiple	66	88
	None	114	151

This commentary should not be construed as requiring the individual odds ratios to be essentially constant (homogeneous) over the stratifying variable before application of a summary procedure. The definition of homogeneity of odds ratios and even its requirement are still debated (Zelen 1971, Bishop et al. 1975, Halperin et al. 1977, Mantel et al. 1977). However, the need to identify extreme situations where subtables represent conflicting relationships is obvious. This can be done by inspection or through systematic use of multivariate techniques such as log-linear models.

Table 4. Log-linear screening results (partial associations) for case-control study of two canine female sexes (S), tumor status (T), and stratifying variable, breed (B).[a]

Effect	d.f.	X^{2b}	P^c
ST	1	0.56	0.4551
SB	12	259.65	0.0
TB	12	264.08	0.0
STB	12	67.01	0.0

[a]Thirteen breeds and 8573 dogs. Since this is only an example we have ignored other epidemiologically important variables, such as age, for simplicity

[b]Likelihood ratio chi-square statistic

[c]Observed significance level

Table 5. Significant breed-specific approximate relative risks that would be missed by the combination of risk estimates from the thirteen subtables.

Breed	Relative Risk	Significance
Setter	0.33	0.005
Boxer	0.56	0.05
Spaniel	0.57	0.01
Other Pedigree	0.65	0.0005
Poodle	1.65	0.0005
Dachshund	2.04	0.0005

The log-linear model is easily applied to the stratified 2 x 2 tables of a case-control study. In addition to identification of situations in which the data structure may not be conducive to the application of summary relative risk procedures, the log-linear model allows for systematic analysis of other higher order interactions that may be of biologic importance. In order to avoid erroneous conclusions, the examples given here indicate the need for detailed examination of epidemiologic data before interpreting a summary relative risk.

REFERENCES

1. Bender, A.P., G.P. Bender, C.R. Dorn and R. Schneider. 1982. Associations between canine benign and malignant neoplasms. Prev. Vet. Med. (in press)

2. Birch, M.W. 1974. The detection of partial association, I: The 2 x 2 case. J. Roy. Statist. Soc. Ser. B 27: 313-324.

3. Bishop, Y.M.M., S.E. Fienberg and P.W. Holland. 1975. Discrete Multivariate Analysis. Theory and Practice. MIT Press, Cambridge, Mass.

4. Brown, M.B. 1976. Screening effects in multidimensional contingency tables. Appl. Stat. 25: 37-46.

5. Cornfield, J. and W. Haenszel. 1960. Some aspects of retrospective studies. J. Chron. Dis. 11: 523-534.

6. Dayal, H.H. 1978. On the desirability of the Mantel-Haenszel summary measure in case-control studies of multifactor etiology of disease. Am. J. Epidemiol. 108: 506-511.

7. Dixon, W.J. and M.B. Brown ed. 1979. BMDP Biomedical Computer Programs P-Series. University of California Press. Los Angeles, Calif. 297 p.

8. Halperin, M., J.H. Ware and D.P. Byar et al. 1977. Testing for an interaction in an I x J x K contingency table. Biometrika 64: 271-275.

9. Heilbron, D.C. 1981. The analysis of ratios of odds ratios in stratified contingency tables. Biometrics 37: 55-66.

10. Mantel, N. and W. Haenszel. 1959. Statistical aspects of the analysis of data for retrospective studies of disease. J. Natl. Cancer Inst. 22: 719-748.

11. Mantel, N., C. Brown and D.P. Byar. 1977. Tests for homogeneity of effect in an epidemiologic investigation. Am. J. Epidemiol. 106: 125-129.

12. McKinlay, S.J. 1978. The effect of non zero second-order interaction of combined estimates of the odds ratio. Biometrika 65: 191-202.

13. Miettinen, O.S. 1970. Matching and design efficiency in retrospective studies. Am. J. Epidemiol. 91: 111-118.

14. Rothman, K.J. 1975. A pictorial representation of confounding in epidemiologic studies. J. Chron. Dis. 28: 101-108.

15. Zelen, M. 1971. The analysis of several 2 x 2 contingency tables. Biometrika 58: 129-137.