MMWR

Recommendations and Reports

July 27, 1990 / 39(RR-11);17-23

Guidelines for Investigating Clusters of Health Events - APPENDIX. Summary of Methods for Statistically Assessing Clusters of Health Events

The following summaries are provided as a resource to investigators who may become involved with the statistical aspects of reported clusters of health events and who are not likely to have a direct effect on the day-to-day management of the clusters. TEMPORAL CLUSTERING Ederer, Myers, and Mantel approach

Ederer, Myers, and Mantel (1) developed a test for temporal clustering using a cell-occupancy approach. They divided the time period into k disjoint subintervals. Under the null hypothesis of no clustering, the n cases are randomly distributed among the subintervals (i.e., are multinomially distributed). The test statistic m is the maximum number of cases occurring in a subinterval. If the health event is rare and of unknown etiology, m is summed over several locations and time periods. The sum is tested by using a single degree of freedom chi-square test. Ederer, Myers, and Mantel (1) and Mantel, Kryscio, and Myers (2) provide tables of the exact null distribution of m for selected values of k and n. Scan Test

Naus proposed a test of temporal clustering that is known as the scan test (3). The test statistic, the maximum number of cases observed in an interval of length t, is found by "scanning" all intervals of length t in the time period (resulting in overlapping intervals). In certain cases, this approach is intuitively more appealing than the disjoint interval approach of Ederer, Myers, and Mantel (1), but more complicated mathematically. However, situations exist for which the disjoint interval approach is the more satisfactory choice. Statistical significance of the scan test is assessed by using tables of p-values calculated by Naus (4) and Wallenstein (5) for selected interval lengths, time lengths, and sample sizes. Unfortunately, the computations necessary to obtain other exact p-values for the scan statistic are complex and often not feasible. However, Knox and Lancashire (6) have derived a set of relatively simple formulas for an approximation to the exact p-value.

Naus compared the power of the scan test with that of the Ederer, Myers, and Mantel test and concluded that if the scanning interval is small and the data are continuous over the interval, the scan test is the more powerful of the two (7). Weinstock proposed a generalization of the scan test that adjusts for changes in the population at risk (8). Bailar, Eisenberg, and Mantel Test of Temporal Clustering

Bailar, Eisenberg, and Mantel suggested a test of temporal clustering based on the number of pairs of cases in a given area that occur within a specified length of time d of each other (9). The numbers of close pairs occurring in q areas are summed. The test statistic is assumed to be approximately normally distributed. Larsen Test

Larsen, Holmes, and Heath developed a rank order procedure for detecting temporal clustering (10). The time period is divided into disjoint subintervals that are numbered sequentially (i.e., ranked). The test statistic K is the sum of absolute differences between the rank of the subinterval in which a case occurred and the median subinterval rank. Small values of K indicate unimodal clustering. Generally, the K statistics for multiple geographic areas are summed. The resulting statistic is asymptotically normal with simple mean and variance. This test is sensitive only to unimodal clustering; it cannot distinguish multiple clustering from randomness. Tango Clustering Index

Tango developed a test of temporal clustering based on the distribution of counts in disjoint equal time intervals (11). The test is useful when the data are grouped. The test statistic (cluster index) is a quadratic form involving the relative frequencies in each interval and a measure of distance between intervals. The clustering index obtains a maximum value of 1 when all cases occur in the same interval. Although the statistic is easy to calculate, the asymptotic distribution using Tango's formula is not. However, Tango will provide upon request an algorithm written in BASIC to obtain the asymptotic distribution.

Whittemore and Keller showed that the distribution of Tango's index is asymptotically normal with simple mean and variance (12). SPATIAL CLUSTERING Geary Contiguity Ratio

Geary developed a test of spatial clustering that assesses whether rates for adjacent areas are more similar than would be expected if they were randomly distributed among the geographic areas (13). The test statistic, the contiguity ratio, is the ratio of the sum of mean squared differences between rates for pairs of adjacent areas to the weighted sum of mean squared differences between rates for all pairs of areas. If the rates are geographically distributed at random, the contiguity ratio is close to one; otherwise, it is less than one. Geary derived an expression for the approximate variance of the ratio. If the number of areas is not too small, the ratio is asymptotically normally distributed. Hechtor and Borhan provide another computational formula for the statistic (14). Ohno, Aoki, and Aoki Test

Ohno, Aoki, and Aoki (15) and Ohno and Aoki (16) developed a simple test for spatial clustering that uses rates for geographic areas (e.g., census tracts, counties, or states) rather than data for individual cases. The test assesses whether the rates in adjacent areas are more similar than would be expected under the null hypothesis of no clustering.

For this test, the rate for each area is classified into one of n categories, and each pair of adjacent areas is identified. The test statistic is the number of adjacent concordant pairs--i.e., the number of pairs of areas that are adjacent and have rates in the same category. An overall clustering measure (summed across all categories) can be obtained as well as category-specific clustering measures. The observed number of adjacent concordant pairs is compared with the expected number by using a chi-square test. Ohno, Aoki, and Aoki provide a simple formula for calculating the expected number of pairs (15). Grimson Test

Grimson, Wang, and Johnson proposed a test of spatial clustering for use in detecting clusters of geographic areas designated as high risk (17). The null hypothesis is that high-risk areas are randomly distributed within a larger area and do not cluster.

Given n high-risk areas, the test statistic is the number of pairs of high-risk areas that are adjacent to each other. This statistic is equivalent to the category-specific statistic from Ohno, Aoki, and Aoki (15). Grimson et al. recommended using a simple Monte Carlo simulation to obtain p-values for the test statistic (17). Whittemore Test

Whittemore, Friend, Brown, and Holly developed a test for spatial clustering across geographic areas that adjusts for different distributions of population subgroups across the region (18). Thus, the test requires population data. The test statistic is the mean distance between all pairs of cases, and can be expressed as a generalization of Tango's clustering index--i.e., a quadratic form involving relative frequencies from subgroups and a matrix of distances between pairs of areas. The statistic is asymptotically normal (mean and variance derived), and the test has good power when disease rates for all subgroups are elevated in the same areas. Power is poor when areas with elevated rates vary for subgroups. The test also has poor power when clusters occur in more than one area. The test can be adapted to detect temporal clustering when the distance matrix represents distances between pairs of time intervals. Cuzick and Edwards Test

Cuzick and Edwards proposed a test for spatial clustering that applies to populations with non-uniform population density (19). The test involves drawing a set of controls from the population and combining them with the cases. Cuzick and Edwards propose two nearest-neighbor tests. The statistic for the first test is the number of persons in the case group whose nearest neighbor also is in the case group. The second test statistic is the sum of the number of cases among the K nearest neighbors for each person who is in the case group. This second test will be more powerful when a few large clusters exist, whereas the first test is more powerful when many small clusters are involved. Cuzick and Edwards provide formulas for the mean and variance and establish asymptotic normality for the test statistics. SPATIAL AND TEMPORAL CLUSTERING Pinkel and Nefzger Cell Occupancy Approach

In 1959, Pinkel and Nefzger proposed a cell occupancy approach to test for spatial-temporal clustering (20). Assuming that r cases are randomly allocated to m space-time cells, these investigators developed an exact test for determining the probability of observing k "close" cases (i.e., cases occurring within a specified distance and length of time of each other).

For this test, the study area and time period are divided into space-time cells based on the space and time distances used to define closeness. The test is sensitive not only to space-time clustering but also to spatial clustering or temporal clustering alone, a property that is not desirable (21). Knox 2 x 2 Contingency Table Test

Knox developed a space-time clustering test that involves dichotomizing the spatial and temporal dimensions (22,23). A 2 x 2 contingency table is formed by classifying the n(n-1)/2 pairs of cases as close in space and time, close in space only, close in time only, or close in neither space nor time.

The test statistic X, the observed number of pairs close in both space and time, is assumed to be approximately Poisson (since although pairs are dependent, X is small compared with the total number of pairs).

Barton and David concluded that, although use of the Poisson approximation is appropriate in some situations, in general it could yield misleading results (24). Mantel outlined methodology for obtaining the exact permutational distribution of X (21). Barton and David Points-on-a-Line Approach

Barton, David, and Herrington (25) and David and Barton (26) adapted an earlier test (27) for use in detecting space-time interaction. The test, analogous to analysis of variance, involves the ratio of within-group variance to overall variance. Pairs of cases separated in time by less than a specified length of time are formed into time clusters (i.e., treatment groups).

The test statistic Q is the ratio of the average squared geographic distance between pairs of cases within clusters to the average squared distance between all pairs of cases. Under the null hypothesis of no space-time interaction, one would expect this ratio to be 1. When clustering is present, Q is smaller than

1. To assess significance, David and Barton suggested using a randomization test to determine the exact distribution of Q (26). Since calculation of the exact distribution often is not feasible, Barton and David suggested using a beta approximation when the number of cases is small and a normal approximation when the number of cases is large (28). When the number of clusters is large, Q is approximately normally distributed; otherwise, an F approximation is more appropriate.

An advantage of Barton and David's test is that actual distances are used, and the only arbitrariness is in the selection of the critical time point. A disadvantage of the test is that the small distances, which are of most interest, have less influence on the statistic than do the large distances. In fact, the large distances may so dominate the statistic that they mask any clustering. Mantel Generalized Regression Approach

Mantel developed a "generalized regression" approach to the detection of clustering in space and time (21). The test statistic Z is the sum over all pairs of cases of a function of the distance between two cases multiplied by a function of the time between two cases. Knox's test can be derived as a special case of Mantel's test. Mantel recommended using reciprocal transformations of the distances to increase the influence of close distances and decrease the influence of long distances. Mantel (21) and Siemiatycki (29) concluded that the test has low power if no transformation is made.

A constant must be added to the distances before making the reciprocal transformation because of the possibility of very small or zero time and/or space distances. Unfortunately, the constants chosen influence the value of the test statistic and the outcome of the test of significance if the normal approximation is used. Mantel suggested that, for best results, the constants be close to the expected distances between close pairs. Glass, Mantel, Guns, and Spears (30) and Siemiatycki (29) found that as the size of the constants increases, the test statistic tends to decrease.

A test of statistical significance is obtained by obtaining the exact randomization distribution of Z, by using Monte Carlo simulation to obtain an approximation to the distribution of Z, or by assuming that Z is asymptotically normally distributed (Mantel derived expressions for the measured variance) (21). Klauber (31) and Siemiatycki (29) found the distribution of Z to be highly skewed and showed that although the use of the normal approximation is appropriate when Z is highly significant or nonsignificant, its use is inappropriate when Z has borderline significance.

One asset of Mantel's test is that actual space and time distance are used, thus avoiding arbitrary cutpoints and loss of information. Another advantage to this approach is its applicability to two or more samples (31,32). Pike and Smith Extension to Knox Test

Pike and Smith extended Knox's test to diseases with long latent periods by defining a geographic area and period of time of infectivity and susceptibility (33). Pairs of cases are considered close in space if their geographic areas of infectivity and susceptibility overlap, and close in time if their periods of infectivity and susceptibility overlap. The test statistic is the number of pairs close in both space and time. Lloyd and Roberts Test

Lloyd and Roberts outlined a test for either spatial or temporal clustering that Smith and Pike noted in 1974 can be viewed as a special case of Knox's test (34). Lloyd and Roberts suggested using the number of pairs among all possible pairs of cases that are close in time (or in space) as the test statistic. A test of significance is obtained by calculating the mean number of close pairs for sets of randomly selected controls and by assuming a Poisson distribution with this mean. Smith and Pike indicated that the randomization distribution of the test statistic could be obtained, and they suggested that matched controls be used in the procedure (35).

References

1. Ederer F, Myers MH, Mantel N. A statistical problem in space

and time: do leukemia cases come in clusters? Biometrika 1964;20:626-38.

2. Mantel N, Kryscio RJ, Myers MH. Tables and formulas for extended use of the Ederer-Myers-Mantel disease clustering procedure. Am J Epidemiol 1976;104:576-84.

3. Naus JI. The distribution of the size of the maximum cluster of points on a line. Journal of the American Statistical Association 1965;60:532-8.

4. Naus JI. Some probabilities, expectations, and variance for the size of the smallest intervals and largest clusters. Journal of the American Statistical Association 1966;61:1191-9.

5. Wallenstein S. A test for detection of clustering over time. Am J Epidemiol 1980;111:367-72.

6. Knox EG, Lancashire R. Detection of minimal epidemics. Stat Med 1982;1:183-9.

7. Naus JI. A power comparison of two tests of non-random clustering. Technometrics 1966;8:493-517.

8. Weinstock MA. A generalized scan statistic test for the detection of clusters. Int J Epidemiol 1981;10:289-93.

9. Bailar III JC, Eisenberg H, Mantel N. Time between pairs of leukemia cases. Cancer 1970;25:1301-3. 10. Larsen RJ, Holmes CL, Heath CW. A statistical test for measuring unimodal clustering: a description of the test and of its application to cases of acute leukemia in metropolitan Atlanta, Georgia. Biometrics 1973;29:301-9. 11. Tango T. The detection of disease clustering in time. Biometrics 1984;40:15-26. 12. Whittemore A, Keller JB. A letter to the editor. On Tango's index of disease clustering in time. Biometrics 1986;42:218. 13. Geary RC. The contiguity ratio and statistical mapping. Incorp Statist 1954;5:115-45. 14. Hechtor HH, Borhan NO. Mortality and geographic distribution of atherosclerotic heart disease. Public Health Rep 1965;80:11-24. 15. Ohno Y, Aoki K, Aoki N. A test of significance for geographic clusters of disease. Int J Epidemiol 1979;8:273-81. 16. Ohno Y, Aoki K. Cancer deaths by city and county in Japan: a test of significance for geographic clustering of disease. Soc Sci Med 1981;15D:251-8. 17. Grimson RC, Wang KC, Johnson PWC. Searching for hierarchical clusters of disease: spatial patterns of sudden infant death syndrome. Soc Sci Med 1981;15D:287-93. 18. Whittemore AS, Friend N, Brown BW, Holly EA. A test to detect clusters of disease. Biometrika 1987;74:631-5. 19. Cuzick JC, Edwards R. Spatial clustering for inhomogeneous populations. Proceedings of the Joint Statistical Meetings, Washington, DC, August 7, 1989. 20. Pinkel D, Nefzger D. Some epidemiological features of childhood leukemia in the Buffalo, NY area. Cancer 1959;12:351-7. 21. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Research 1967;27:209-20. 22. Knox G. Detection of space-time interactions. Applied Statistics 1964;13:25-30. 23. Knox G. Epidemiology of childhood leukemia in Northumberland and Durham. Br J Prev Soc Med 1964;18:17-24. 24. Barton DE, David FN. The random intersection of two graphs. In: David FN, ed. Research papers in statistics. New York: John Wiley & Sons Inc, 1966:45-59. 25. Barton DE, David FN, Herrington M. A criterion for testing contagion in time and space. Ann Hum Genet 1965;29:97-103. 26. David FN, Barton DE. Two space-time interaction tests for epidemicity. Br J Soc Med 1966;20:44-8. 27. Barton DE, David FN. The analyses of chromosome patterns in the normal cell. Ann Hum Genet 1962;25:323-32. 28. Barton DE, David FN. Randomization basis for multivariate tests. I. The bivariate case: randomness of n points in a plane. Bull Int Stat Inst 1962;i39 II:455-67. 29. Siemiatycki J. Mantel's space-time clustering statistic: computing higher moments and a comparison of various data transforms. Journal of Statistical Computation and Simulation 1978;7:13-31. 30. Glass AG, Mantel N, Guns FW, Spears GFS. Time-space clustering of childhood leukemia in New Zealand. J Natl Cancer Inst 1971;47:329-36. 31. Klauber MR. Two sample randomization test for space-time clustering. Biometrics 1971;27:129-42. 32. Klauber MR. Space-time clustering tests for more than two samples. Biometrics 1975;31:719-26. 33. Pike MC, Smith PG. Disease clustering: a generalization of Knox's approach to the detection of space-time interactions. Biometrics 1968;24:541-6. 34. Lloyd S, Roberts CJ. A test for space clustering and its applications to congenital limb defects in Cardiff. Br J Prev Soc Med 1973;27:186-91. 35. Smith PG, Pike MC. A note on a "close pairs" test for space clustering. Br J Prev Soc Med 1974;28:63-4. *Representatives of the Association of State and Territorial Health Officials, the Agency for Toxic Substances and Disease Registry, and CDC, consisting of Carl W. Armstrong, M.D., David H. Culver, Ph.D., Richard L. Ehrenberg, M.D., Patricia A. Honchar, Ph.D., Dedun Ingram, Ph.D., Jeffrey A. Lybarger, M.D., Stanley I. Music, M.D., Richard B. Rothenberg, M.D., Karen K. Steinberg, Ph.D., Stephen B. Thacker, M.D., and G. David Williamson, Ph.D. **Timothy E. Aldrich, Ph.D., Alan P. Bender, D.V.M., Valerie Beral, M.D., Glyn G. Caldwell, M.D., Beth Fiore, Roger C. Grimson, Ph.D., Clark W. Heath, Jr., M.D., Dennis M. Perotta, h.D., and Lowell E. Sever, Ph.D.

**Questions or messages regarding errors in formatting should be addressed to mmwrq@cdc.gov.

Page converted: 08/05/98

Morbidity and Mortality Weekly Report
Centers for Disease Control and Prevention
1600 Clifton Rd, MailStop E-90, Atlanta, GA 30333, U.S.A

This page last reviewed 5/2/01