

ABSTRACT

Objectives. This article presents a space-time scan statistic, useful for evaluating space-time cluster alarms, and illustrates the method on a recent brain cancer cluster alarm in Los Alamos, NM.

Methods. The space-time scan statistic accounts for the preselection bias and multiple testing inherent in a cluster alarm. Confounders and time trends can be adjusted for.

Results. The observed excess of brain cancer in Los Alamos was not statistically significant.

Conclusions. The space-time scan statistic is useful as a screening tool for evaluating which cluster alarms merit further investigation and which clusters are probably chance occurrences. (*Am J Public Health.* 1998;88:1377-1380)

Evaluating Cluster Alarms: A Space-Time Scan Statistic and Brain Cancer in Los Alamos, New Mexico

Martin Kulldorff, PhD, William F. Athas, PhD, Eric J. Feuer, PhD, Barry A. Miller, MSPH, and Charles R. Key, MD, PhD

Health officials are often asked to evaluate local disease cluster alarms. After the case definition is established, an early question is whether the cluster has occurred by chance or whether the excess is so great that it is probably due to some common elevated risk factor of limited geographical and/or temporal extension. Only in the latter case would a more thorough investigation be warranted, in an attempt to identify those risk factors.

Simply comparing the standardized disease incidence rate within the cluster area and time frame with what is observed in a larger geographic area and time frame does not lead to a suitable statistical test. In such a comparison, the spatial and temporal boundaries of the cluster are defined from an observed set of cases, leading to preselection bias in the statistical analysis. Moreover, any geographic region always contains some high-rate areas occurring by chance alone. If every cluster of cases found to be "statistically significant" according to such a procedure were to be thoroughly investigated, health officials would investigate mostly random data. Instead, we propose the use of a space-time scan statistic.

A Cancer Cluster in Los Alamos?

Los Alamos, a remote New Mexico community of some 18 000 inhabitants, was established in 1943 as part of the Manhattan Engineer District, a top-secret wartime effort to develop, assemble, and test the first atomic bomb. Most of the workforce is employed at the Los Alamos National Laboratory, a nuclear research and design facility.

In 1991, a community resident expressed concern over what was perceived as a neigh-

Martin Kulldorff is with the Division of Cancer Prevention and Eric J. Feuer and Barry A. Miller are with the Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Md. Dr Kulldorff is also with the Department of Statistics, Uppsala University, Uppsala, Sweden. William F. Athas is with the Division of Epidemiology, Evaluation, and Planning, New Mexico Department of Health, Santa Fe. Charles R. Key is with the New Mexico Tumor Registry, University of New Mexico Cancer Center, Albuquerque.

Requests for reprints should be sent to Martin Kulldorff, PhD, EPN 344, National Cancer Institute, 6130 Executive Blvd, Bethesda, MD 20892-7368.

This paper was accepted February 17, 1998.

TABLE 1—Geographic Analysis of Brain Cancer Incidence Rates, Using the Space–Time Statistic: New Mexico, 1973–1991

	Years	No. of Cases	No. Expected	Standardized Incidence Rate	Log Likelihood Ratio	P ^a
No adjustment for temporal trend						
Most likely cluster: Bernalillo, Cibola, Los Alamos, Sandoval, Santa Fe, Socorro, Torrance, and Valencia counties	1985–1989	223	170.2	1.31	8.87	.074
Secondary cluster: Chaves County	1988–1989	16	6.1	2.64	5.63	.720
Secondary cluster: Curry and Roosevelt counties	1982–1983	13	5.5	2.37	3.73	.999
Secondary cluster: Los Alamos County	1986–1989	10	3.9	2.59	3.39	>.999
Adjustment for temporal trend ^b						
Most likely cluster: Los Alamos and Santa Fe counties	1986–1989	43	23.6	1.82	6.53	.450

Note. Rates adjusted for age, race, and sex. Note that Chaves County had both a higher number of cases and a higher incidence rate than was observed in Los Alamos.

^aSimulated *P* value, calculated with 9999 Monte Carlo replications.

^bIncidence increased by approximately 1.2% per year.

borhood cluster of 12 recent brain tumor deaths. The concerns were quickly reported in the local press and gained national attention with articles in the *New York Times*,¹ *People Weekly*,² and other publications. As a result, additional cases were reported by local residents. A joint laboratory–community working group was formed and public hearings were held. A heated discussion ensued in the letters column of the local newspaper, with some citizens taking the concerns very seriously and others dismissing them as a witch hunt.

In response to rapidly mounting public concern, the New Mexico Department of Health and the New Mexico Tumor Registry conducted a comprehensive review of brain cancer incidence rates for 1970 through 1990.³ No particularly unusual or alarming rate was observed in data for the entire time period; however, a temporal analysis showed an 80% rate elevation (10 cases) during the years 1986 to 1990. The community was informed that such a finding could easily have resulted from random fluctuation in the incidence of a rare disease within a small population. Reconciliation with citizens' claims of a much larger brain-tumor cluster revealed that many of the alleged cases identified by residents involved metastatic brain disease or tumors diagnosed among former county residents. Follow-up monitoring showed a decline in brain cancer incidence among Los Alamos residents during the early 1990s.

A Space–Time Scan Statistic

The 1-dimensional scan statistic has long been used to study disease clusters in time,^{4–7} and a 2-dimensional scan statistic has been pro-

posed for studying purely spatial disease clusters.^{8–10} The space–time scan statistic is defined by a cylindrical window with a circular geographic base and with height corresponding to time. The base is centered around one of several possible centroids located throughout the study region, with the radius varying continuously in size. The height reflects any possible time interval of less than or equal to half the total study period, as well as the study period as a whole. The window is then moved in space and time so that for each possible geographic location and size, it also visits each possible time interval.

In effect, we obtain an infinite number of overlapping cylinders of different size and shape, jointly covering the entire study region. Each cylinder reflects a possible cluster, and the preselection of the cluster alarm is accounted for in terms of both cluster location and cluster size.

Cases are assumed to be Poisson distributed with constant risk over space and time under the null hypothesis, and with different risk inside and outside at least one of the cylinders under the alternative hypothesis. For each cylinder the numbers of disease cases inside and outside the cylinder are noted, together with the expected number of cases reflecting the population at risk and relevant covariates. On the basis of these numbers, the likelihood is calculated for each cylinder.⁸ The cylinder with the maximum likelihood, and with more than its expected number of cases, is denoted the *most likely cluster*.

Although the number of cylinders is infinite, epidemiological data contain a finite set of individuals, so that many of the cylinders will contain exactly the same people. This situation leads to a finite set of cylinders for which

the likelihood actually has to be calculated.

Since disease rates and socioeconomic variables are typically aggregated for census areas, these data are assigned to a centroid for each area. As the scanning window moves across each centroid, all individuals from that area are considered to be within the cylinder.

Significance is evaluated with Monte Carlo simulation, where the null hypothesis of no clusters is rejected at an α level of .05 exactly if the simulated *P* is less than or equal to .05 for the most likely cluster.¹¹ In this sense, simulated *P* values function in the same way as mathematically calculated *P* values. The reported *P* values for secondary clusters are conservative.⁸

If the window size is allowed to expand until it covers most of the geographic region and time period, the likelihood no longer reflects a cluster of increased disease risk inside the cylinder, but rather a decreased risk outside. For this reason, we recommend that the geographic size of the window be limited to half the expected number of cases and that the time size be limited to half the total time period. In addition, we suggest including cylinders covering the whole time period to allow for purely spatial clusters. The upper limit on size should never be chosen so as to be more in accordance with the size of the cluster alarm under evaluation, as that defeats the purpose of the scan statistic, which is to eliminate the preselection bias.

Confounding variables can be adjusted for by calculating the expected number of cases in each area and time period through indirect standardization, conditioning on the total number of cases observed. If there is a temporal trend, the trend can be adjusted for by multiplying the confounder-adjusted expected num-

ber of cases for each census area and time period by the overall or national rate during that particular year, or by any proportion of that rate.

Application to Los Alamos

Incidence and population data came from the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute, collected by the New Mexico Tumor Registry. From 1973 through 1991 there were 1175 cases of malignant neoplasm of the brain and nervous system (*International Classification of Diseases, Ninth Revision [ICD-9] codes 191.0–192.9*). For each case, we know county of residence, year of diagnosis, age in 5-year intervals, race (White, Black, other), and sex. For each year, we can cross tabulate the population in each county by age group, race, and sex. For each county, the centroid is defined as the location of the county seat. (The raw data are available at <http://dcp.nci.nih.gov/bb/datasets.html>.)

Calculations were performed with SaTScan v.1.0.¹² With 9999 Monte Carlo replications, the program took 11 minutes to run on a 100-MHz Pentium PC.

With adjustment for age, sex, and race, the most likely cluster is in the Albuquerque–Santa Fe area during 1985 through 1989 (Table 1 and Figure 1). This area contains 48% of the state population. With a *P* value of .074, the cluster is not statistically significant.

Los Alamos is included within the most likely cluster. By itself, Los Alamos had excess incidence during 1986 through 1989, as we would expect based on the cluster alarm, with 10 cases when 3.9 were expected. This excess is not statistically significant when we take the preselection bias and multiple testing into account by means of the space–time scan statistic. It would be significant (*P* = .007) if we simply, and inappropriately, tested for a difference in the standardized incidence rates within and outside this particular area and time period.

When adjustments are also made for the temporal rise in incidence,¹³ the most likely cluster consists of Santa Fe and Los Alamos counties during 1986 through 1989. The *P* value is .45. Hence, the temporal trend may partly explain the most likely cluster in the first analysis.

The power of the test depends on the number of cases in the data set (1175), on the size of the cluster area measured in terms of its expected number of cases under the null hypothesis, and on the true relative risk within, as opposed to outside, the cluster. To achieve an 80% chance of rejecting the null hypothesis due to a hypothetical cluster, the relative risk must be 4.4 for a cluster of size 10 (i.e., 10 cases expected under the null hypothesis, 44

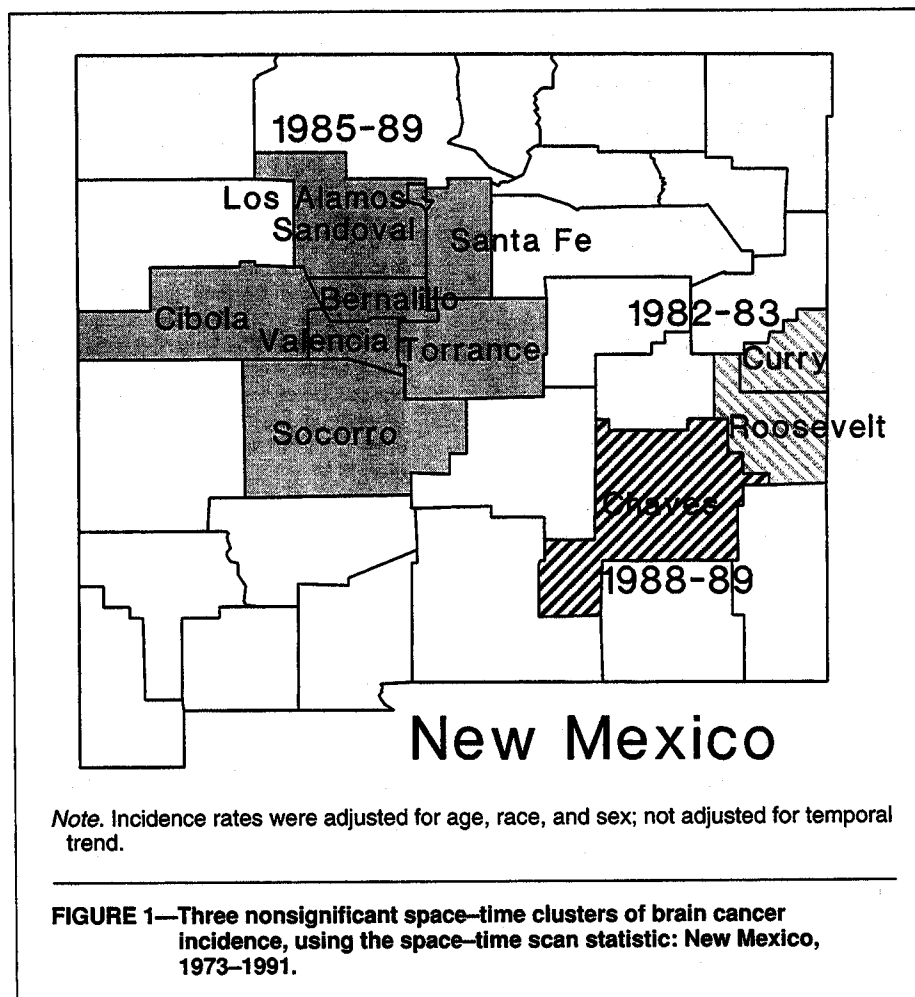


FIGURE 1—Three nonsignificant space–time clusters of brain cancer incidence, using the space–time scan statistic: New Mexico, 1973–1991.

expected under the alternative hypothesis), 3.2 for size 20, 2.4 for size 50, 1.9 for size 100, and 1.7 for size 200.

Discussion

The space–time scan statistic should prove useful as a tool for identifying cluster alarms that are not likely to be of public health importance. A complementary approach is to look at the time it takes for new cases to occur in the cluster area.¹⁴ There are also other statistical tests for space–time clustering,^{15–20} including one that uses the scan statistic in the time dimension.²¹ These are not suitable for evaluation of disease cluster alarms because they do not make inferences about the specific location of clusters. Rather, they are designed to test whether clustering is present globally throughout space and time, such as when a disease is infectious.²²

The excess of brain cancer in Los Alamos falls well within the realm of chance. This finding coincides with the final conclusions of the New Mexico Health Department, which used a statistically less formal approach.³ In this article, we have shown how the space–time

scan statistic may be used to complement descriptive statistical methods in order to provide statistical inferences through a clearly defined hypothesis test. We think the space–time scan statistic is an important addition to the public health official's toolbox. □

Acknowledgments

M. Kulldorff was supported in part by the Swedish Council for Research in the Humanities and Social Sciences. C. R. Key acknowledges support from the Department of Energy (grant DE-FG04-91AL75237) and the National Cancer Institute (grant N01-CN-05228).

We thank Todd Gibbson and Mike Livingston of IMS Inc for computer support.

References

- Schneider K. Cancer cases to be examined for radiation link. *New York Times*. July 23, 1991.
- Grogan D, Haederle M. Hunting a silent killer. *People Weekly*. October 28, 1991:125–126.
- Athas WF, Key CR. *Los Alamos Cancer Rate Study: Phase I, Final Report*. Santa Fe: New Mexico Department of Health; 1993.
- Naus J. The distribution of the size of maximum cluster of points on the line. *J Am Stat Assoc*. 1965;60:532–538.

5. Wallenstein S. A test for detection of clustering over time. *Am J Epidemiol.* 1980;111:367-372.
6. Wallenstein S, Weinberg CR, Gould M. Testing for a pulse in seasonal event data. *Biometrics.* 1989;45:817-830.
7. Weinstock MA. A generalized scan statistic test for the detection of clusters. *Int J Epidemiol.* 1981;10:289-293.
8. Kulldorff M. A spatial scan statistic. *Commun Stat Theory Methods.* 1997;26:1481-1496.
9. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. *Stat Med.* 1995;14:799-810.
10. Turnbull B, Iwano EJ, Burnett WS, et al. Monitoring for clusters of disease: application to leukemia incidence in upstate New York. *Am J Epidemiol.* 1990;132:S136-S143.
11. Dwass M. Modified randomization tests for non-parametric hypotheses. *Ann Mathematical Statistics.* 1957;28:181-187.
12. Kulldorff M, Rand K, Williams G. *SaTScan v 1.0, Software for the Spatial and Space-Time Scan Statistics.* Bethesda, Md: National Cancer Institute; 1997. [Available at: <http://dcp.nci.nih.gov/bb/SaTScan.html>]
13. Gloeckler Ries LA, Miller BA, Hankey BF, et al, eds. *SEER Cancer Statistics Review, 1973-1991.* Bethesda, Md: National Institutes of Health; 1994. Publication No. 94-2789.
14. Chen R, Connelly RR, Mantel N. Analyzing post alarm data in a monitoring system, in order to accept or reject the alarm. *Stat Med.* 1993;12:1807-1812.
15. Ederer F, Myers MH, Mantel N. A statistical problem in space and time: do leukemia cases come in clusters? *Biometrics.* 1964;20:626-638.
16. Jacquez GM. A k nearest neighbour test for space-time interaction. *Stat Med.* 1996;15:1935-1949.
17. Knox G. Detection of space-time interactions. *Appl Stat.* 1964;13:25-29.
18. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 1967;27:209-220.
19. Ranta J, Pitkaniemi J, Karvonen M, et al. Detection of overall space-time clustering in a non-uniformly distributed population. *Stat Med.* 1996;15:2561-2572.
20. Raubertas RF. Spatial and temporal analysis of disease occurrence for detection of clustering. *Biometrics.* 1988;44:1121-1129.
21. Wallenstein S, Gould M, Kleinman M. Use of the scan statistic to detect time-space clustering. *Am J Epidemiol.* 1980;111:367-372.
22. Kulldorff M. Statistical methods for spatial epidemiology: tests for randomness. In: Löytönen M, Gatrell A, eds. *GIS and Health.* London, England: Taylor & Francis; 1998.