

## Brief Communication

A Study of Clustering of Childhood Leukemia by Hospital of Birth<sup>1</sup>

Melville R. Klauber

Department of Preventive Medicine, University of Utah College of Medicine, Salt Lake City, Utah 84112

**Summary.** An attempt was made to obtain epidemiologic evidence that the newborn human may be particularly susceptible to leukemogenic agents, perhaps leukemogenic viruses. The latter has been well established in some animal species. To this end an investigation was undertaken to try to detect clustering by hospital of birth of 234 children who died of leukemia before age 5 years out of approximately 1,000,000 children born in California hospitals in the years 1958-1960. In addition a method of Mantel was used to test for within-hospital clustering of day of birth for leukemia cases. No significant clustering was detected.

**Introduction.** A number of recent studies has been performed to try to detect time-space clustering of childhood leukemia for date of onset (or diagnosis) and residence (1, 3, 8, 13). Consistent positive results have not been obtained for a predetermined form of any one of the approaches referred to above (2, 10). Fraumeni *et al.* (4) found no clustering of year of birth and town of residence (at the time of reporting) for children under age 5 years in New York State (excluding New York City) reported 1943 through 1957. For the 1959 births in California, Klauber and Jackson (7) found a positive, but not a statistically significant, association between leukemia at ages 0 through 4 years and neonatal deaths for other children with the same hospital of birth. This suggests the possibility of clustering by hospital of birth as well as hospital-time of birth clustering.

The decreasing susceptibility with age of newborn mice and rats to leukemia cell-free filtrates (virus) is well established (5). Similar results for the induction of solid tumors by the injection of human adenovirus 12 (14) and SE polyoma virus (11) have been obtained.

In the present study the following specific hypotheses are tested: Children who were born in California during the years 1958 through 1960 and who died of leukemia before 5 years of age have an excess representation within certain unspecified hospitals of birth. A test of temporal-spatial clustering is also performed, i.e., a test of the hypothesis that those cases who were born in the same hospital are clustered as to date of birth compared to random allocation of these dates of birth.

**Methods.** Children under age 5 who died of leukemia and who were born during 1958-1960 in California were identified

through 1958-1965 death certificates for the whole U.S. A search of birth certificates, California Tumor Registry files, and, where possible, hospital records was performed in order to obtain the hospital of birth.

The distribution of leukemia deaths by hospital of birth was determined in the following way: The expected number of leukemia deaths for each hospital, assuming random allocation, was computed for each of the 3 annual cohorts separately. The expected numbers for each hospital were summed over the three years. The numbers of hospitals were cross tabulated by expected number of leukemia deaths (to the nearest integer) and observed numbers. In addition, the method of Haldane (6) was used to test for significant departure from equality of the proportions of leukemia deaths by hospital of birth.

The approach of Mantel (9) was used to test whether there was a tendency for time of birth to cluster for those children who died of leukemia and who were born in the same hospital. This test, unlike the approach mentioned above, involves only the leukemia cases.

Let  $G_{ij} = 0$  if cases  $i$  and  $j$  were born in different hospitals,  
 $= 1$  if cases  $i$  and  $j$  were born in the same hospital,  
 and let  $U_{ij} = 1/(|T_i - T_j| + K)$ ,

where  $K$  is a positive constant and  $T_i$  and  $T_j$  represent the times in days from a common origin for cases  $i$  and  $j$  respectively. Mantel's test statistic is  $A = \sum G_{ij} U_{ij}$ . Significance was tested with two predetermined values of  $K$  and 3 and 30 (days) by comparing the  $A$  obtained from the actual data with the empirical distribution of  $A$ . The choice of  $K$  was made in accordance with Mantel's suggestion that  $K$  be commensurate with anticipated possible or probable distances in time between related cases. The value  $K = 3$  corresponds to an hypothesis of baby-to-baby transmission of an agent, and the value of  $K = 30$  corresponds to an hypothesis of a more lasting insult in the environment or an indirect association between cases, e.g., a member of the hospital staff might be infected with a leukemogenic virus. Other values of  $K$  (1, 5, and 15) were also used in order to try to further ascertain the nature of the clustering should any exist. For each  $K$  a new set of 500 randomizations was performed using a well-tested version of the multiplicative congruential method.

**Results.** There were 247 children under age 5 who died of leukemia and who were born during 1958-1960 in California. The hospital of birth could be determined for 234 cases; four cases were not born in a hospital, the original birth certificates were not obtainable for 2 children who were adopted, and there

<sup>1</sup> This investigation was supported by USPHS Grant CA-05924 from the National Cancer Institute.

Received December 26, 1967; accepted May 9, 1968.

Table 1

Expected number of leukemia deaths <sup>a</sup>	No. of hospitals	Observed number of leukemia deaths							
		0	1	2	3	4	5	6	7
0-0.5	327	273	45	8	0	1	0	0	0
0.5-1.5	116	42	50	18	6	0	0	0	0
1.5-2.5	17	2	5	5	3	0	2	0	0
2.5-3.5	8	1	1	2	2	0	1	0	1
Totals	468 <sup>b</sup>	318	101	33	11	1	3	0	1

Number of California hospitals which were the place of birth in 1958-1960 of various observed numbers of children who died of leukemia under age 5 years, by expected number of leukemic deaths.

<sup>a</sup> Approximately one leukemia death could be expected per 4,500 live births.

<sup>b</sup> There was one additional hospital with an observed number of 8 leukemia deaths for which 8.1 were expected.

was no matching birth certificate nor could the hospital or place of birth be obtained by other means for the 7 remaining cases.

Table 1 shows the observed numbers of hospitals having 0, 1, 2, . . . leukemia deaths among children born in the same hospital by the expected number of leukemia deaths.<sup>2</sup> There are no extraordinary deviations from randomness (Poisson variation) indicated by the table; the most extreme case, one hospital with 4 deaths observed with 0.46 expected is not improbable considering the number of hospitals. The value for Haldane's test was very close to its expected value under the assumption of randomness;  $P = 0.49$ .<sup>3</sup>

The results using Mantel's statistic to test for hospital-time of birth clustering were also not significant; the empirical  $P$ -values were all in the range 0.37 to 0.43 for all five values of  $K$ .

**Discussion.** The hypothesis of within-hospital clustering is similar to that of "cancer houses" discussed by Pearson (12) over 50 years ago; in the present study we ask whether there was an excess of multiple cases developing within cohorts born in the same hospital instead of persons living within the same house.

The present study was an attempt to detect clustering of leukemia cases by actual time and place of birth, only one of a vast number of possibilities if the disease were caused by a virus. The relatively weak or negative space-time clustering results for leukemia compared to those for diseases of known viral etiology certainly is evidence against childhood leukemia being of viral origin and horizontal transmission with many

susceptible persons in the population. On the other hand, the type of clustering that has been obtained by others is compatible with a disease of viral origin, vertical transmission, with susceptibility under the influence of genetic factors, as is suggested by the murine and avian models. The clustering that has been detected may be for some insult or insults which may trigger or activate a virus.

#### Acknowledgments

The author gratefully acknowledges the following: Dr. Edwin W. Jackson for references on the susceptibility of newborn animals to leukemogenic viruses; Nathan Mantel of the National Cancer Institute for a number of very helpful communications and for allowing the author to see a prepublication copy of his paper (9); Dr. Robert Grove, Chief, Division of Vital Statistics, National Center for Health Statistics, for identifying 18 members of the case series who died in states other than California; the California State Department of Public Health for the live births by hospital data used in this paper; and Dr. Louis Schmittroth, Director, University of Utah Computer Center for the use of a UNIVAC 1108 Computer.

#### References

1. Barton, D. E., David, F. N., and Merrington, M. A Criterion for Testing Contagion in Time and Space. *Ann. Human Genet.* London, 29: 97-102, 1965.
2. Ederer, F., Myers, M. H., Eisenberg, H., and Campbell, P. C. Temporal-Spatial Distribution of Leukemia and Lymphoma in Connecticut. *J. Natl. Cancer Inst.*, 35: 625-629, 1965.
3. 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.
4. Fraumeni, J. F., Jr., Ederer, F., and Handy, V. H. Temporal-Spatial Distribution of Childhood Leukemia in New York State. *Cancer*, 19: 996-1000, 1966.
5. Gross, L. *Oncogenic Viruses*, pp. 179-197, pp. 308-309. New York: Pergamon Press Inc., 1961.
6. Haldane, J. B. S. The Mean and Variance of  $X^2$ , When Used as a Test of Homogeneity, When Expectations are Small. *Biometrika*, 31: 346-355, 1940.

<sup>2</sup> The following tabulations will be made available on request: (a) the annual number of live births for each of the 469 hospitals and the corresponding leukemia deaths and (b) a table showing the leukemia cell type by course of disease (acute or chronic) for the 234 leukemia cases.

<sup>3</sup> For the 436 hospitals with 353,895 live births in 1959, Haldane's test indicated a highly statistically significant discrepancy from equality of the proportion of neonatal deaths,  $P < 0.001$ . However, when the denominators were tripled and 234 births were distributed proportional to the hospital neonatal death rates in order to make the situation more comparable to that of leukemia, Haldane's test was no longer significant.

7. Klauber, M. R., and Jackson, E. W. Leukemia and Neonatal Death-Rates. *Lancet*, 1: 1002-1003, 1967.
8. Knox, G. Epidemiology of Childhood Leukemia in Northumberland and Durham. *Brit. J. Prevent. Social Med.*, 18: 17-24, 1964.
9. Mantel, N. The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Res.*, 27: 209-220, 1967.
10. Meighan, S. S., and Knox, G. Leukemia in Childhood: Epidemiology in Oregon. *Cancer*, 18: 811-814, 1965.
11. Mirand, E. A., Grace, J. T., Moore, C. T., and Mount, D. Relationship of Viruses to Malignant Disease. *A.M.A. Arch. Internal Med.*, 105: 469-481, 1960.
12. Pearson, K. On "Cancer Houses" from the Data of the Late Th. Law Webb. *Biometrika*, 8: 430-435, 1913.
13. Pinkel, D., Dowd, J. E., and Bross, I. D. J. Some Epidemiological Features of Malignant Solid Tumors of Children in the Buffalo, N. Y. area. *Cancer*, 16: 28-33, 1963.
14. Yaba, Y., Trenton, J. J., and Taylor, G. Cancer Induction in Hamsters by Human Type 12 Adenovirus. Effect of Age and Virus Dose. *Proc. Soc. Exptl. Biol. Med.*, 108: 343-344, 1962.