Papers and Originals

Childhood Leukaemia in Greater London: a Search for Evidence of Clustering

MORWENNA M. TILL,* B.A., M.B., D.C.H.; R. M. HARDISTY,* M.D., M.R.C.P.; M. C. PIKE,† PH.D.
RICHARD DOLL,† M.D., D.SC., F.R.C.P., F.R.S.

Brit. med. J., 1967, 3, 755–758

The suggestion that viruses might play a part in the causation of leukaemia has reawakened interest in the epidemiology of the disease, and has led many investigators to search for evidence of its occurrence in clusters. Little attention was paid to Kellett's (1937) original report of an epidemic in the area around Newcastle upon Tyne; but at least 19 similar reports have been made in the past eight years. Heath, Manning, and Zelkowitz (1964) noted 13 besides their own, and other reports have been made by Knox (1964), Meighan and Knox (1965), Mustacchi (1965), Dowsett (1966), and Mainwaring (1966). It was not until 1963, however, that the importance of distinguishing between clustering in space, in time, and in space and time considered together was fully appreciated (Knox, 1963), and not until the following year that a serious attempt was made to assess the probability that an observed cluster in both space and time could occur owing to chance alone (Knox, 1964).

Knox (1964) noted the addresses and dates of onset of all cases of childhood leukaemia in Northumberland, Durham, and part of Yorkshire over the ten-year period 1951–60 and found that there was a statistically significant excess of pairs of affected children in whom the onset of symptoms began before their sixth birthday, who lived within one kilometre of each other, and in whom the dates of onset were within 60 days. The implications of this finding, however, are limited by the fact that the clustering was found only for one clinical subgroup of the cases and for limited space and time intervals, and that those characteristics which defined the positive cluster were thrown up by the investigation itself and not by a preceding hypothesis. Thus Knox's findings should be regarded as constituting a sound basis for formulating a hypothesis to be tested against further observations and not as proof of the existence of real clustering. We have therefore sought to test the hypothesis by examining the distribution of childhood leukaemia (defined for this study as leukaemia with onset of symptoms before 10 years of age) in another area.

From the point of view of theoretical oncology the concept that the *onset* of leukaemia should occur in clusters is not particularly attractive, as factors of aetiological importance are likely to exert their effect a considerable time before the disease becomes clinically apparent. Clustering in space and time of the births of the leukaemic children would seem to be a much more likely event if a virus played any part in the production of the disease. We have therefore also looked into this possibility.

Material

Copies of the death entries were examined of all the children who died of leukaemia under 15 years of age during the ten years 1952–61 and who were resident in Greater London at the time of their death. Only children in whom the primary cause of death was certified as leukaemia were included and no attempt was made to find children who, though born or developing the disease in Greater London, were resident elsewhere when they died. Altogether 618 children were studied. For each of these the hospital case records were sought and the clinical and haematological details reviewed. In addition Dr. Alice Stewart allowed us access to all the data she had collected on these children during the Oxford Survey of Childhood Malignancies.

As a result of this review 135 children were excluded: (*a*) 105 because they were more than 10 years old when symptoms first occurred; (*b*) 11 because they were born and living outside Greater London at the time of onset of symptoms; and (*c*) 19 because the diagnosis of leukaemia was not confirmed—in eight the hospital records suggested that death was due to some other cause, in two the diagnosis rested solely on the post mortem findings, which in our opinion were inconclusive, and in nine the medical records had been destroyed or could not be traced.

Birth certificates were sought for the remaining 483 children and were obtained for all but 18—6 who were known to have been adopted, 9 who were known to have been born abroad, and 3 others.

The following data were used for the analysis:

(1) *Date of onset*, taken as the date of the first symptoms referable to leukaemia in the medical records. This was always recorded as a specific day even though the exact choice of date was known to be to some extent arbitrary. For example, "one month ago" was taken to mean exactly one calendar month before the date on which the history was recorded; symptoms dating from a specified month were recorded as starting on the 15th of the month, and symptoms said to have been present for "several weeks" or "several months" were dated from three weeks or three months previously. The onset in one infant in whom leukaemia was present at birth was taken as three months antenatally.

Table I shows the frequency distribution of the lengths of history recorded according to these conventions. In each case the length of history was taken to be the time between the onset and the date of the first relevant medical record, even though a firm diagnosis of leukaemia was not always made until some time later. In 235 cases the history was of four weeks or less or was referable to an exact date, and in these the date of onset was probably accurate to within a week. In a further 126 cases the duration recorded was "one month," was referable to a month by name, or was not more than two months with symptoms of abrupt onset, and in these the date of onset was probably accurate to within two weeks. Thus three-quarters of the dates of onset are likely to have been accurate to within two weeks; the remainder were probably accurate to within two months and usually to within much less.

* The Hospital for Sick Children, Great Ormond Street, London W.C.1.
† Medical Research Council's Statistical Research Unit, University College Hospital Medical School, London W.C.1.

(2) *Address at time of onset of symptoms,* taken from the medical records. In the absence of specific evidence to the contrary this was taken to be the same as the address at the time of the nearest subsequent medical attendance. Two children were living outside Greater London at the time of onset and are therefore excluded from analyses of clustering at this stage in the course of the disease. For 252 of the remaining children (52%) the address at onset was the same as the place of residence at birth.

TABLE I.—*Number of Children With Different Lengths of History at Presentation (Onset Under 10 Years of Age)*

Length of History (in weeks)	No. of Children	Cumulative Percentage of Total
2 or less	122	25
-4	84	43
-6	113	66
-8	68	80
-13	42	89
-26	39	97
More than 26	15	100
All durations	483	100

(3) *Date of birth,* taken from the birth certificate when available.

(4) *Place of birth and home address at birth,* taken from the birth certificate. Of the 465 children whose birth certificates were traced, 42 were born outside Greater London and 28 had home addresses outside Greater London. The numbers of children for whom it was possible to examine evidence of space-time clustering in relation to date and place of birth and date of birth and home address at birth were therefore reduced to 423 and 437 respectively. The map references of all addresses were found correct to within 100 metres. When the child was born in hospital the middle of the hospital was taken to give the map reference.

(5) *Cytological and clinical type of leukaemia,* determined after a review of the reports of peripheral blood counts, marrow examinations, and clinical and post mortem findings. Three hundred and seventy-four cases were classified as acute lymphoblastic leukaemia ; 52 as acute myeloblastic leukaemia ; 16 as acute monocytic leukaemia ; and 3, in which abnormal erythroblasts were the predominating cell type, were classified as acute erythroleukaemia. Four children aged 5 to 9 years showed blood, marrow, and clinical findings typical of the adult form of chronic granulocytic leukaemia, and the series also included two children previously described as having the juvenile form of the disease (Cases 1 and 2 in the series described by Hardisty, Speed, and Till, 1964) and eight other children aged ? months to 2 years, so closely resembling these in their haematological and clinical features that it seemed justifiable to classify them in the same group—making a total of 14 cases of chronic granulocytic leukaemia. In 15 cases it was not possible to ascertain the cytological type, though there was no doubt about the diagnosis of acute leukaemia, and these were classified as acute unspecified leukaemia. Finally, nine cases in which the evidence of leukaemia was suggestive but incomplete were unclassified.

The distribution of cases by type, sex, and age at onset is shown in Table II.

Results

Clustering in Space

The distribution of cases throughout the Greater London conurbation is shown in Table III. The population at risk has been taken to be the mean of the population under 10 years of age at the censuses of 1951 and 1961. At these dates the Greater London area was divided into 95 London boroughs and municipal boroughs in the home counties, and the populations of most of these areas were too small for individual examination. The areas have therefore been combined to provide units that correspond as nearly as possible to the boroughs of the new Greater London Council. The highest incidence is recorded in Sutton, Epsom, and Ewell (7.7 per 100,000 children per year) and the lowest in Hounslow (1.7 per 100,000 children per year).

TABLE III.—*Incidence of Leukaemia in Different Parts of Greater London (Children With Onset Under 10 Years of Age)*

Borough	Incidence per 100,000 Children per Year	No. of Cases	
		Observed	Expected at Rate for Whole Area
*City of Westminster	3·9	11	11·8
Camden	3·8	11	12·1
Islington	4·9	19	16·5
Hackney	4·2	16	16·2
Tower Hamlets	4·8	16	14·2
*Greenwich	5·2	18	14·7
Lewisham	3·2	14	18·4
Southwark	3·2	16	21·0
*Lambeth	4·2	14	14·2
*Wandsworth	4·8	29	25·6
Hammersmith	3·3	10	12·8
Kensington and Chelsea	3·9	9	9·8
*Waltham Forest	6·0	22	15·5
*Redbridge	3·3	16	20·1
*Barking	4·9	14	12·1
*Newham	2·8	12	17·8
*Bexley	4·9	19	16·3
*Bromley	5·0	18	15·2
Croydon	5·4	25	19·6
*Sutton	7·7	28	15·5
Merton	4·7	12	10·8
*Kingston-on-Thames	4·0	11	11·7
Richmond-on-Thames	3·9	9	9·8
*Hounslow	1·7	6	15·0
Hillingdon	3·8	13	14·6
Ealing	5·0	20	17·0
Brent	4·2	17	17·0
*Harrow	5·0	16	13·4
*Barnes	3·9	19	20·7
Haringey	3·1	11	15·1
Enfield	2·7	10	15·9

* Boundaries do not correspond exactly with the borough of this name in the new Greater London.

Inspection of the map of London shows that there is no tendency for high and low areas to lie close to one another, and statistical analysis reveals that the amount of variation between the boroughs is no more than can easily be attributed to chance ($\chi^2 = 33.4$, n = 30, P = 0.3).

Clustering in Time

Table IV shows the distribution of cases by year of onset, separately for the two principal cytological types, and for cases with ages of onset under and over 6 years. As the data are derived from information about deaths in the years 1952–61 a relatively small number of cases is to be expected in 1960–1. Between 1952 and 1959 the annual number shows little variation.

Table V shows the distribution of cases by month of onset. No seasonal peaks are evident, neither for all cases taken together, nor for lymphoblastic cases alone, nor for children with onset under 6 years of age ; in all groups the numbers arising in summer and winter—defined by Knox (1964) as May to October and November to April—are practically identical.

TABLE II.—*Number of Children with Different Types of Leukaemia by Sex and by Age at Onset (Onset under 10 Years of age)*

Type	Age (in years) at Onset										Sex		Total
	0	1	2	3	4	5	6	7	8	9	M	F	
Acute lymphoblastic	21	33	64	71	53	31	29	24	25	23	213	161	374
Acute myeloblastic	8	7	2	4	5	7	8	6	3	2	31	21	52
Acute monocytic	0	0	2	1	1	0	3	0	5	4	11	5	16
Other acute and unclassified	7	1	3	5	2	2	2	4	1	0	18	9	27
Chronic granulocytic	3	5	2	0	0	1	0	1	0	2	7	7	14
Total	39	46	73	81	61	41	42	35	34	31	280	203	483

TABLE IV.—*Number of Children Developing Leukaemia in Different Years, by Cytological Type and by Age at Onset (Onset Under 10 Years of Age in Greater London)*

Category	Year											All Years
	Before 1952	52	53	54	55	56	57	58	59	60	61	
Acute lymphoblastic	12	35	34	39	46	41	41	46	37	27	14	372
Acute myeloblastic	3	4	9	7	8	4	4	2	5	5	1	52
Others	3	7	9	3	5	3	7	5	3	8	4	57
Onset before 6th birthday	13	35	34	31	40	33	31	41	34	36	11	339
Onset after 6th birthday	5	11	18	18	19	15	21	12	11	4	8	142
All categories	18	46	52	49	59	48	52	53	45	40	19	481

TABLE V.—*Number of Children Developing Leukaemia in Different Months, by Cytological Type and by Age at Onset (Onset under 10 Years of Age in Greater London)*

Category	Month												All Months
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	
Acute lymphoblastic	25	26	31	36	27	36	31	29	22	40	31	38	372
Acute myeloblastic	5	5	4	2	4	7	6	3	3	3	6	4	52
Others	8	6	6	1	8	4	7	4	3	2	3	5	57
Onset before 6th birthday	29	26	31	31	25	36	32	28	19	29	21	32	339
Onset after 6th birthday	9	11	10	8	14	11	12	8	9	16	19	15	142
All categories	38	37	41	39	39	47	44	36	28	45	40	47	481

Clustering in Space and Time

The data were examined for evidence of clustering in both space and time together by Knox's (1964) method—that is, by seeing whether pairs of cases that occurred close together in space also occurred close together in time. All cases were examined as one set, and separately in various subsets consisting of lymphoblastic cases only, cases with onset under 6 years of age, and cases with deaths under 6 years of age. Clustering was looked for in regard to place and date of onset, place and date of birth, and place of mother's residence at the child's birth and the child's date of birth. Subsets of data relating to date of birth were also examined for children born in 1952–5, to ensure that all members of a cohort who died under 6 years of age were included.

Since the data were derived from information about deaths in 1952–61 the number of cases known to have had their *onset* in years before 1952 and in 1961 (the last year of the period studied) must be deficient; examination of the interaction between date and place of onset was therefore limited to the period 1952–60.

The various sets of data used and the types of interaction examined are shown in Table VI. In each case evidence of clustering was looked for over all combinations of 12 critical times apart (15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, and 180 days) and 12 critical distances apart (0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 3.5, and 4 kilometres). In only three instances was any evidence of positive clustering found —that is, significantly more pairs of cases within the critical times and distances of one another than would be expected from the whole subset of data under examination. No statistically significant clustering was found in the subset originally specified by Knox—that is, clustering of onsets among children in whom the disease began before their sixth birthday—nor was any significant clustering found when the examination of this subset of children was limited to those with lymphoblastic leukaemia.

Details of the amount and character of the clustering in the three instances in which statistically significant results were obtained are shown in Table VII. Each subset includes lymphoblastic cases only. In two instances there is weak evidence of clustering of times and places of onset; the excess numbers of pairs are small (3 to 5), but it is notable that the distances and times within which the cases occurred (¾ km. and 60–75 days) are similar to those recorded by Knox (1964). In the third subset the clustering relates to the date of birth and the mother's place of residence when the child was born. The probability of as many or more clusters occurring due to chance alone is lower and the excess numbers of pairs are somewhat greater (9 to 24), but the clusters are less striking in that the distances in both time and space separating individual members of each pair are much greater.

Four sets of data were also examined for clustering at the onset of the disease by an alternative method suggested by David and Barton (1966). Each set consisted of cases with onset in the period 1952-60, the groups being composed of (*a*) all children, (*b*) children with lymphoblastic leukaemia, (*c*) all children with onset before their sixth birthday, and (*d*) children with lymphoblastic leukaemia with onset before their sixth birthday. In none of the groups was any evidence of clustering obtained. Comparison of the results obtained by these two methods showed that David and Barton's (1966) method was less sensitive to the presence of small clustering in this type of material and the method was not used to examine the other sets.

Discussion

Our data have failed to provide clear evidence in support of Knox's hypothesis that leukaemia in children under 6 years

TABLE VI.—*Examination Made for Evidence of Space-Time Clustering*

Subset of Leukaemic Children	Type of Interaction	No. of Children
All with onset 1952–60	Date and place of onset	444
All with onset 1952–60 before 6th birthday	Date and place of onset	315
All with onset 1952–60 and death before 6th birthday	Date and place of birth	282
All with onset 1952–60 and death before 6th birthday	Date of birth, place of residence at birth	292
All with birth 1952–55 and death before 6th birthday	Date and place of birth	116
All with birth 1952–55 and death before 6th birthday	Date of birth, place of residence at birth	118
*Lymphoblastic with onset 1952–60	Date and place of onset	346
Lymphoblastic with onset 1952–60 before 6th birthday	Date and place of onset	252
*Lymphoblastic with onset 1952–60 and death before 6th birthday	Date and place of onset	230
Lymphoblastic with onset 1952–60 and death before 6th birthday	Date and place of birth	224
*Lymphoblastic with onset 1952–60 and death before 6th birthday	Date of birth, place of residence at birth	232
Lymphoblastic birth 1952–55 and death before 6th birthday	Date and place of birth	89
Lymphoblastic birth 1952–55 and death before 6th birthday	Date of birth, place of residence at birth	91

* Evidence of statistically significant clustering, see Table VII.

TABLE VII.—*Categories of Leukaemia with Significant Evidence of Clustering in Space and Time*

Subset of Patients	Interaction	No. of Children	Critical Time (Days)	Critical Distance (Kilometres)	No. of Pairs	Expected No. of Pairs	Standard Deviation	Probability*
Lymphoblastic leukaemia onset 1952–60	Date and place of onset	346	60	0·75	8	4·00	1·96	0·051
			75	0·75	10	5·06	2·20	0·034
Lymphoblastic leukaemia onset 1952–60, death before 6th birthday	Date and place of onset	230	75	0·75	6	2·67	1·60	0·054
Lymphoblastic leukaemia onset 1952–60, death before 6th birthday	Date of birth, place of parents' residence at child's birth	232	120	2·00	31	21·56	4·52	0·033
			135	2·00	35	24·53	4·81	0·025
			165	1·75	32	22·81	4·61	0·040
			165	2·00	41	30·01	5·30	0·028
			165	3·50	102	83·68	8·84	0·023
			180	3·00	88	69·23	8·02	0·014
			180	3·50	116	91·83	9·24	0·007
			180	4·00	139	115·70	10·36	0·017

* Probability of an equal or greater excess number of pairs being observed by chance from the Poisson distribution.

of age tends to occur in clusters, children who live within one kilometre of each other tending to develop symptoms of the disease within 60 days of each other. It should be noted, however, that 9 of the 10 children who formed the clusters in Knox's original data had lymphoblastic leukaemia and all developed their disease under 4 years of age; and our data provide some some weak evidence of clustering at the onset of the disease within similar critical limits ($\frac{3}{4}$ km. and 75 days) in children who *died* of lymphoblastic leukaemia under 6 years of age (in fact the ages of the children at onset were under 4 years in 10 out of 12 cases). In view of the fact that Knox was studying an area with a population scattered over large and small towns and villages, whereas we were studying a single conurbation, this degree of similarity between the results is notable and provides some weak support for the general hypothesis that cases of leukaemia do tend to occur in clusters more often than would be expected by chance alone.

In our data the strongest evidence of clustering occurs with date and residence at birth, and a leukaemogenic agent is presumably more likely to be effective at or around this period than at any other in the child's life. The dimensions of the cluster are, however, too large to suggest a single focus of infection (4 km. and 180 days), and, as with Knox's (1964) original data, the statistical tests employed do not have the significance normally attached to them. A large number of possible critical times and distances were examined, and, logically, the analysis does no more than allow us to pick out a new hypothesis, which can be tested only on further data.

With the methods now available we have been able to look for clustering only in relation to a few specified periods in the natural history of the disease, and these would not necessarily be the most relevant if the disease was truly contagious. Under these conditions it would be more appropriate to look for interaction between, say, the date and place of residence of one child during his period of infectivity and the place of residence of another child during a period of susceptibility—which might be several years before the disease was clinically apparent. Without knowledge of the periods when infectivity and susceptibility occur, this is a problem of some complexity. A general method of searching for interaction between different hypothetical periods of infectivity and susceptibility has been developed by one of us (Pike, 1967), but it will be some time before it can be tested adequately.

Summary

The hypothesis that cases of childhood leukaemia occur grouped together in space and in time has been tested by examining the distribution of the disease in Greater London.

Altogether 618 children were certified as having died of leukaemia under 15 years of age in Greater London during the period 1952–61. For each of these children the hospital records were sought and the clinical and haematological details reviewed. Four hundred and eighty-three children were accepted as having developed leukaemia under 10 years of age while resident in Greater London, and for these children information was obtained about the date and place of birth, the parents' place of residence at the child's birth, and the date and place of residence of the child at the onset of the disease.

The incidence of leukaemia varied between the 31 boroughs of the new Greater London Council, but the amount of variation (from 1.7 to 7.7 per 100,000 children per year) could easily be accounted for by chance variation of small numbers.

No evidence was found to suggest that the disease had been unusually common in any one year, or to suggest that it occurred more commonly at any one season, nor was there any evidence of temporal variation in the incidence of any of the principal subgroups of the disease (lymphoblastic and myeloblastic leukaemia and leukaemia occurring before or after the sixth birthday).

Evidence of clustering in space and time taken together was sought by examining whether pairs of cases that occurred close together in space also occurred close together in time. Thirteen subsets of the data were examined, and weak evidence of clustering was found for all cases of lymphoblastic leukaemia in regard to date and place of onset, and for cases of lymphoblastic leukaemia leading to death under 6 years of age in regard both to date and place of onset, and to date of birth and the place of the parents' residence at the child's birth.

The results obtained in relation to the onset of the disease are similar to but not identical with those reported previously by others, and provide some weak support for the general hypothesis under examination.

We are grateful to the Leukaemia Research Fund and the Medical Research Council for grants to one of us (R. M. H.) for the purpose of this study; to Mrs. Ranjana Ash and Mrs. Jean Bridger for locating the map references of the addresses and tracing the case records; to Mrs. Anna Frackiewicz (M.R.C. Clinical Effects of Radiation Research Unit) for extracting copies of the death entries of children in Greater London who died of leukaemia; to Mrs. Serina Bell for assistance in the computer programming; to Dr. Alice Stewart for allowing us access to her records; and to our colleagues in many London hospitals for their willing co-operation in making their case records available.

REFERENCES

David, F. N., and Barton, D. E. (1966). *Brit. J. prev. soc. Med.*, **20**, 44.
Dowsett, E. G. (1966). *Brit. J. Cancer*, **20**, 16.
Hardisty, R. M., Speed, D. E., and Till, M. (1964). *Brit. J. Haemat.*, **10**, 551.
Heath, C. W., jun., Manning, M. D., and Zelkowitz, L. (1964). *Lancet*, **2**, 136.
Kellett, C. E. (1937). *Arch. Dis. Childh.*, **12**, 239.
Knox, G. (1963). *Brit. J. prev. soc. Med.*, **17**, 121.
—— (1964). Ibid., **18**, 17.
Mainwaring, D. (1966). Ibid., **20**, 189.
Meighan, S. S., and Knox, G. (1965). *Cancer*, **18**, 811.
Mustacchi, P. (1965). Ibid., **18**, 362.
Pike, M. C. (1967). *Biometrics* (in press).