Pergamon

European Journal of Cancer, Vol. 35, No. 1, pp. 91–96, 1999

Printed in Great Britain
0959-8049/99/$ - see front matter

PII: S0959-8049(98)00345-1

Original Paper

Space–Time Clustering of Acute Lymphoblastic Leukaemia in Parts of the U.K. (1984–1993)

E.A. Gilman,[1] R.J.Q. McNally[2] and R.A. Cartwright[1]

[1]Leukaemia Research Fund Centre for Clinical Epidemiology, University of Leeds, 17 Springfield Mount, Leeds, LS2 9NG; and [2]CRC Paediatric & Familial Cancer Research Group, Royal Manchester Children's Hospital, Stancliffe, Manchester, U.K.

Age-related differences in the incidence and immunological subtypes of acute lymphoblastic leukaemia (ALL) suggest that it may be composed of more than one disease entity, each with different aetiologies. Childhood leukaemia (of which the majority of cases are ALL) has been suspected of having an infectious aetiology, but few studies have systematically examined ALL for clustering by age group. The aim of this study was to examine ALL for evidence of space–time clustering of date and place of diagnosis by age group. Knox space–time analysis was carried out separately for three different age groups: childhood (0–14 years), young adult (15–34 years) and older adults (35–79 years). Data on 968 cases of ALL aged 0–79 years, arising during 1984–1993 in the areas covered by a specialist population based register of leukaemias and lymphomas in parts of the U.K., were used in the analysis. Space–time clustering of diagnoses was limited to children aged 0–14 years. It was more prominent in those diagnosed in the period 1984–1988, than in those diagnosed in 1989–1993. The clustering may indicate an infectious aetiology for childhood ALL, or could be the result of episodic exposures to some environmental hazard. © **1999 Elsevier Science Ltd. All rights reserved.**

Key words: acute lymphoblastic leukaemia, cluster
Eur J Cancer, Vol. 35, No. 1, pp. 91–96, 1999

INTRODUCTION

CASES OF acute lymphoblastic leukaemia (ALL) have a distinctive age distribution. Incidence rates peak in childhood at around 3–4 years of age, followed by a steep decline, the disease being particularly uncommon between 25–60 years of age, but then the incidence slightly rises again in old age [1]. In childhood ALL the majority of cases involve undifferentiated precursors of B lymphocytes and display a characteristic antigen (common-ALL), but in adults immunological subtypes other than common-ALL are seen [2]. These age-related differences could indicate that ALL is composed of more than one disease entity, with each one involving different aetiological factors. Childhood leukaemia (of which the majority of cases are ALL) has long been suspected of having an infectious aetiology [3–5]. Kinlen has suggested that leukaemia rates in children may be raised following unusual levels of population mixing (such as occur during the development of new towns) which may bring a previously isolated population with a high proportion of susceptibles into contact with a hitherto unencountered infectious agent [6]. It has also been suggested that ALL could have a model of transmission similar to that of polio, where late exposure to infection causes serious consequences, whilst early exposure gives rise to trivial disease [7]. Late exposure is associated with high socio-economic status, families with small sibships, perhaps in more immunologically isolated communities, all of which are features of cases of ALL. Whilst there is accumulating evidence that childhood ALL has an infectious aetiology, there is little evidence of this in adults, apart from the association of HTLV-1 infection with increased risk of adult T-cell leukaemia, after a long latency period (over 30 years). Little is known overall about the aetiology of the leukaemias, apart from evidence for a few risk factors: ionising radiation, chemicals such as benzene and chloramphenicol and some immunodeficiency disorders [8, 9].

Correspondence to E.A. Gilman, e-mail: e.a.gilman@leeds.ac.uk
Received 5 Jun. 1998; revised 15 Sep. 1998; accepted 22 Sep. 1998.

If ALL does have an infectious aetiology, it might be expected to occur in clusters, relating to the patterns of exposure to the infectious agent involved. Most of the positive reports of space–time clustering of leukaemia (excesses of cases occurring close together in both space and time) have involved cases in children [10–16]. However, some investigations have not found evidence of space-time clustering [17, 18], whilst others have reported spatial clustering (localised persistent excesses of cases), but the spatial boundaries and age limits used have varied [19–24]. There have been few systematic examinations of whether ALL shows different clustering patterns in different age groups and most studies of ALL have concentrated on children, or used cases of all ages.

In this paper we report a systematic examination of high quality data on ALL for evidence of space–time clustering of date and place of diagnosis in three different age groups: childhood (0–14 years), young adult (15–34 years) and older adults (35–79 years). Our hypothesis was that childhood cases may have an infectious aetiology, but adult cases do not, hence space–time clustering is expected among childhood ALL only.

PATIENTS AND METHODS

The case data came from a specialist registry of leukaemias and lymphomas, the Leukaemia Research Fund Data Collection Survey (DCS). At its beginning in 1984, this population based registry covered approximately 40% of England and Wales [25], but since 1988 the number of areas contributing data has reduced, following rationalisation to make them co-terminous with areas in which detailed epidemiological studies are being carried out. The DCS collects data on incident cases, employing uniform methods of case ascertainment and stringent diagnostic/histological verification and combines high levels of ascertainment with timeliness of registration [26]. Estimates of overall completeness of ascertainment for leukaemias and lymphomas in 1989–1993 showed a decrease with age, declining from 98.5% in children aged 0–14 years to 94.4% in adults aged 80 years or more. In order to ensure analyses were based on a dataset with the highest, continuous levels of case ascertainment, analysis was restricted to cases of ALL aged 0–79 years arising in those areas covered by the DCS throughout the period 1984–1993.

The data were examined for space–time clustering using the Knox method [27, 28]. The null hypothesis is that cases which occur close together in space should be no more likely to occur close together in time than other cases, and vice versa. All possible pairs of cases are assembled and classified according to their distances apart in space and time. The observed number of pairs within short space and short time intervals is compared with the number expected if the space and time intervals between pairs are independent of each other. For each space–time combination, calculation of expected values effectively uses the marginal totals of a 2×2 contingency table. Cases are assumed to be rare, independent events, distributed as a Poisson variable. The significance of the departure of the observed number of close pairs (O) from the expected number (E) is tested using d, where:

$$d = (O - E)/\mathrm{sqrt}(\mathrm{variance}[O])$$

and is distributed as the standard normal deviate, i.e. values greater than 1.96 are significant at the $P = 0.05$ level. This use of the normal approximation was first suggested by Mantel [28] and is generally accepted [29]. We followed Mantel's approach and used permutation to calculate the variance of the number of close pairs, using the algorithm of Pike and Bull [30].

The information needed for an examination of space–time clustering patterns of diagnoses is date of diagnosis, grid reference of address at diagnosis, sex and diagnosis for each case; population denominators are not required. Cases with incomplete grid references or incomplete dates of birth or diagnosis were excluded. Separate analyses were carried out for the age groups 0–14 years (childhood), 15–34 years (young adults) and 35–79 years (older adults). The age groups used were chosen as a crude surrogate for immunophenotype. For example, we know that the majority of cases in the childhood peak are of one subtype (common-ALL). Ideally, we would have liked to have defined subsets based on immunophenotype. However, this information was not available for all cases and its completeness also depended on the time period over which data were collected, more information being provided in recent years. For analyses of space–time clustering it is important that ascertainment of cases is as complete as possible, otherwise space–time clustering of diagnoses will reflect clustering of case reporting rather than case occurrence. We therefore felt it would not be appropriate to use data on immunophenotype to define subsets in our analyses. Table 1 shows, for each diagnostic group examined, the number of cases included in the analyses. There were 976 cases of ALL diagnosed in the counties covered by the DCS in both data collection periods and of these only 8 (0.8%) were excluded from analysis because of incomplete information.

As data were collected over a long period (1984–1993) there was the possibility that there might have been a shift in the temporal and spatial distribution of the population at risk. If there had been such a shift, then any clustering detected might be a reflection of joint space–time heterogeneities of the population rather than of the cases. To reduce the effects of any such population shifts we analysed the data in two separate 5-year periods.

Space–time clustering was examined using five space intervals (≤ 1 km, ≤ 2 km, ≤ 3 km, ≤ 4 km, ≤ 5 km) and 10 time intervals (same month, ≤ 1, ≤ 2, ≤ 3, ≤ 4, ≤ 5, ≤ 6, ≤ 9, ≤ 12, ≤ 18 months), i.e. 50 tests were carried out on each

Table 1. *Details of acute lymphoblastic leukaemia cases used in space–time clustering analyses*

	Total cases	Cases from counties in both data collection periods*	Cases used in analysis†		
			1984–1988	1989–1993	Total
Age group (years)					
0–14	842	549	255	290	545
15–34	267	179	97	79	176
35–79	350	248	125	122	247
0–79	1459	976	477	491	968

*Counties contributing cases to both data collection periods were: Avon, Cornwall, Cumbria, Devon, Dorset, Gloucestershire, Humberside, Lancashire, North Yorkshire, Somerset, West Yorkshire.
†Reasons for exclusion were one or all of: incomplete grid reference, date of birth, or date of diagnosis.

subset of cases examined. However, the problem of multiple significance testing is not as large as it may seem, since the 50 tests are not independent of one another, e.g. pairs of cases diagnosed within 2 km and 3 months of each other are also included in pairs of cases diagnosed within 3 km and 4 months of each other and so on. This correlation between the results of tests means that the true significance level of an individual test is not greatly affected [18, 29]. In this situation, adjusting results for multiple significance testing using the Bonferroni correction (dividing the critical significance level by the number of tests performed, giving $P = 0.05/50 = 0.001$ as the critical P value for 50 independent tests at the 0.05 level of significance [31]) would be too conservative. Instead, following an approach used by Rodrigues and associates [32], we adjusted the critical P value for the most significant individual test by allowing a factor of 2 for multiple testing within the same subset and then dividing by the number of separate subsets of data examined. Our prior hypothesis was that clustering would be found in childhood cases only, the data were examined in two periods and so the critical P value was taken as 0.0125 (0.05/4). Unadjusted P values are given and interpreted as usual in the results, but for space–time excesses to be interpreted as evidence of clustering, the most extreme P value for a significance test in a given age group must be below the adjusted critical value of 0.0125.

Table 2. Cases of acute lymphoblastic leukaemia, aged 0–14 years, diagnosed during 1984–1988

Time (≤ mth)		Space (≤ km) 1	2	3	4	5	All distances
Same month	Pairs	1	3	4	6	6	538
	O/E	3.01	2.96	1.99	1.92	1.41	
	P*		0.0436				
1	Pairs	3	5	6	10	13	1616
	O/E	3.01	1.64	0.99	1.07	1.01	
	P	0.0356					
2	Pairs	4	9	10	16	20	2603
	O/E	2.49	1.84	1.03	1.06	0.97	
	P	0.0444					
3	Pairs	5	10	13	21	29	3557
	O/E	2.28	1.49	0.98	1.02	1.03	
	P	0.0406					
4	Pairs	6	11	17	27	37	4549
	O/E	2.14	1.28	1.00	1.02	1.02	
	P	0.0358					
5	Pairs	6	12	20	33	45	5541
	O/E	1.75	1.15	0.97	1.03	1.02	
6	Pairs	9	18	28	43	57	6455
	O/E	2.26	1.48	1.16	1.15	1.11	
	P	0.0042					
9	Pairs	9	21	36	54	72	9180
	O/E	1.59	1.21	1.05	1.01	0.99	
12	Pairs	11	26	44	71	91	11 644
	O/E	1.53	1.19	1.01	1.05	0.98	
18	Pairs	13	35	57	96	127	16 538
	O/E	1.27	1.12	0.92	1.00	0.97	
All times	Pairs (Cases)	20 (39)	61 (99)	121 (126)	188 (148)	257 (162)	32 385 (255)

*P value shown only if $P < 0.05$ for test of O/E significantly different from 1.00. Time and space intervals are cumulative. O/E, ratio of observed number of pairs to expected number.

RESULTS

Children

Among children aged 0–14 years, diagnosed during 1984–1988, there were significant excesses of pairs diagnosed within 1 km and at intervals from 1 month up to 6 months apart (O/E ratios ranged from 3.01 to 2.26); also within 2 km and in the same month (O/E = 2.96); together with non-significant excesses of pairs diagnosed at distances up to 3 or 4 km and within the same month (Table 2). These data suggest that childhood leukaemias cluster on date and place of diagnosis, with excesses of pairs diagnosed between 1 and 4 km apart and at intervals up to 6 months apart. The pattern of clustering seen among children diagnosed in 1989–1993 (Table 3) was weaker, but it also showed prominent (but non-significant) excesses of pairs diagnosed within 1, 2 or 3 km and the same month and a significant excess of pairs diagnosed within 4 km and the same month (O/E = 2.01). The most extreme P value in either period was for cases diagnosed within 1 km and 6 months in 1984–1988 ($P = 0.0042$, Table 2) and this was below the critical value of $P = 0.0125$ set for the most significant individual test.

Dividing the cases in both periods into finer age groups showed that clustering was concentrated in children aged 0–9 years, with similar patterns among cases aged 0–4 years and aged 5–9 years (data not shown), and among the cross-pairs between these age groups, where one member of the pair was 0–4 years old and the other was 5–9 years old (Table 4, results are shown for cases diagnosed in 1984–1988 only). Among cases aged 10-14 years there were no significant

Table 3. Cases of acute lymphoblastic leukaemia, aged 0–14 years, diagnosed during 1989–1993

Time (≤ mth)		Space (≤ km) 1	2	3	4	5	All distances
Same month	Pairs	1	3	4	9	9	714
	O/E	1.73	1.98	1.35	2.01	1.49	
	P*				0.0294		
1	Pairs	1	3	8	15	18	2177
	O/E	0.57	0.65	0.89	1.10	0.98	
2	Pairs	1	6	13	26	30	3455
	O/E	0.36	0.82	0.91	1.20	1.03	
3	Pairs	1	8	17	33	40	4753
	O/E	0.26	0.79	0.86	1.11	1.00	
4	Pairs	3	12	23	41	51	6088
	O/E	0.61	0.93	0.91	1.07	0.99	
5	Pairs	4	13	25	47	57	7372
	O/E	0.67	0.83	0.82	1.02	0.92	
6	Pairs	4	17	31	55	67	8571
	O/E	0.58	0.93	0.87	1.02	0.93	
9	Pairs	5	25	46	80	102	12 376
	O/E	0.50	0.95	0.90	1.03	0.98	
12	Pairs	10	35	65	109	136	15 905
	O/E	0.77	1.04	0.98	1.09	1.01	
18	Pairs	14	48	94	152	200	22 610
	O/E	0.76	1.00	1.00	1.07	1.05	
All times	Pairs (cases)	34 (53)	89 (121)	174 (168)	263 (197)	354 (208)	41 905 (290)

*P value shown only if $P = < 0.05$ for test of O/E significantly different from 1.00. Time and space intervals are cumulative. O/E, ratio of observed number of pairs to expected number.

Table 4. *Cases of acute lymphoblastic leukaemia, cross-pairs aged* (0–4) (5–9) *years, diagnosed during* 1984–1988

Time (≤ mth)		Space (≤ km) 1	2	3	4	5	All distances
Same month	Pairs	0	1	1	3	3	176
	O/E	0.00	3.05	1.52	2.89	2.14	
1	Pairs	0	1	1	3	5	530
	O/E	0.00	1.01	0.51	0.96	1.18	
2	Pairs	0	3	3	5	7	809
	O/E	0.00	1.99	0.99	1.05	1.09	
3	Pairs	1	4	5	9	11	1057
	O/E	1.31	2.03	1.27	1.44	1.31	
4	Pairs	2	5	7	11	14	1349
	O/E	2.05	1.99	1.39	1.38	1.30	
5	Pairs	2	6	8	14	18	1637
	O/E	1.69	1.97	1.31	1.45	1.38	
6	Pairs	3	7	11	18	23	1897
	O/E	2.18	1.98	1.56	1.61	1.52	
	*P**				0.0424	0.0430	
9	Pairs	3	8	14	23	29	2703
	O/E	1.53	1.59	1.39	1.44	1.35	
12	Pairs	5	11	19	31	37	3387
	O/E	2.04	1.74	1.50	1.55	1.37	
	P				0.0139		
18	Pairs	5	12	21	36	46	4810
	O/E	1.43	1.34	1.17	1.27	1.20	
All times	Pairs	7	18	36	57	77	9656

**P* value shown only if $P = < 0.05$ for test of O/E significantly different from 1.00. Time and space intervals are cumulative. O/E, ratio of observed number of pairs to expected number.

Table 5. *Cases of acute lymphoblastic leukaemia, aged* 15–34 *years, diagnosed during* 1984–1988

Time (≤ mth)		Space (≤ km) 1	2	3	4	5	All distances
Same month	Pairs	0	0	0	0	0	101
	O/E	0.00	0.00	0.00	0.00	0.00	
1	Pairs	0	0	0	0	0	280
	O/E	0.00	0.00	0.00	0.00	0.00	
2	Pairs	0	0	0	1	1	407
	O/E	0.00	0.00	0.00	0.33	0.27	
3	Pairs	1	1	1	2	2	570
	O/E	1.36	0.48	0.37	0.47	0.39	
4	Pairs	1	3	4	5	5	729
	O/E	1.06	1.13	1.16	0.91	0.76	
5	Pairs	1	3	4	5	6	875
	O/E	0.89	0.94	0.97	0.76	0.76	
6	Pairs	1	3	4	6	7	1033
	O/E	0.75	0.80	0.82	0.77	0.75	
9	Pairs	1	3	5	9	10	1490
	O/E	0.52	0.55	0.71	0.80	0.74	
12	Pairs	1	5	8	13	18	1991
	O/E	0.39	0.69	0.85	0.87	1.00	
18	Pairs	2	8	12	18	24	2780
	O/E	0.56	0.79	0.91	0.86	0.96	
All times	Pairs (Cases)	6 (11)	17 (30)	22 (36)	35 (47)	42 (49)	4656 (97)

No *P* value was < 0.05 for test of O/E significantly different from 1.00. Time and space intervals are cumulative. O/E, ratio of observed number of pairs to expected number.

clusters in either time period (data not shown; O/E ratios ranged from 0 to 1.77).

Adults

Cases in the age group 15–34 years showed no significant excesses in either calendar period (Table 5, results are shown for cases diagnosed in 1984–1988 only). Among cases aged 35–79 years there was one significant excess of pairs diagnosed within 3 km and 1 month (O/E = 2.35) in 1984–1988 (Table 6) and three significant excesses in 1989–1993, involving pairs diagnosed within 1 km and 1, 9 and 18 months of each other (Table 7).

None of the *P* values for the space–time excesses seen in the adults was below the critical value of $P = 0.0125$.

It was concluded that clustering on date and place of diagnosis was confined to childhood ALL and was more prominent in cases diagnosed in 1984–1988 than in 1989–1993.

DISCUSSION

Before discussing the main findings, attention must be paid to possible sources of artefactual clustering in this dataset. This could be caused by selective exclusion of cases, but as was shown in Table 1, only a very small number of cases (0.8%) were excluded from analysis because of incomplete information. More cases were excluded due to analysis being restricted to those counties covered by the complete period of data collection 1984–1993. However, this restriction removed all cases from the relevant counties and did not affect those from the remaining counties at all, so it is difficult to see how this could have caused space–time clustering. The dataset has a very high level of case ascertainment: a cross-check of 1986 leukaemia and lymphoma registrations indicated completeness of around 92% [33], whilst registrations of ALL in 1989–1993 are estimated to be 98% complete [26], making it unlikely that clustering could be due to localised variations in the reporting of cases. Finally, clustered recognition of cases, perhaps due to increased vigilance following the detection of earlier cases, could cause artefactual space–time clustering. However, this would also produce corresponding deficits of pairs of cases at the same distances as the clustered pairs, but separated by longer time intervals. Such a pattern was not seen here.

We found no evidence of space–time clustering of diagnoses of ALL except among children aged 0–14 years where it was more marked in 1984–1988 than in 1989–1993. There were significant excesses of pairs diagnosed at distances up to 4 km apart and at intervals up to 6 months, with the most significant individual excess involving 9 pairs of children diagnosed within 1 km and 6 months of each other in 1984–1988. The clustering was concentrated in cases under 10 years of age, but was present in pairs where children had been diagnosed at different ages and hence were born at different times. This suggests that the clustering on date and place of diagnosis in these children could not easily be a reflection, after some fairly regular latent period, of clustering on date and place of birth, but seems to be a primary feature of onsets

Table 6. *Cases of acute lymphoblastic leukaemia, aged* 35–79 *years, diagnosed during* 1984–1988

Time (≤ mth)		Space (≤ km) 1	2	3	4	5	All distances
Same month	Pairs	0	0	1	1	1	111
	O/E	0.00	0.00	1.49	0.98	0.73	
1	Pairs	1	1	5	5	6	351
	O/E	2.45	1.05	2.35	1.55	1.39	
	*P**			0.0382			
2	Pairs	1	2	7	8	9	609
	O/E	1.41	1.21	1.90	1.43	1.21	
3	Pairs	1	2	7	9	10	835
	O/E	1.03	0.88	1.38	1.18	0.98	
4	Pairs	2	3	9	13	16	1087
	O/E	1.58	1.02	1.37	1.31	1.20	
5	Pairs	2	4	10	14	19	1318
	O/E	1.31	1.12	1.25	1.16	1.18	
6	Pairs	3	5	11	17	22	1572
	O/E	1.64	1.17	1.15	1.18	1.14	
9	Pairs	4	7	16	22	31	2230
	O/E	1.54	1.16	1.18	1.08	1.13	
12	Pairs	4	12	20	29	39	2893
	O/E	1.19	1.40	1.14	1.09	1.10	
18	Pairs	4	11	26	38	48	4053
	O/E	0.85	1.00	1.06	1.02	0.97	
All times	Pairs (Cases)	9 (15)	21 (30)	47 (43)	71 (56)	95 (67)	7750 (125)

**P* value shown only if $P = < 0.05$ for test of O/E significantly different from 1.00. Time and space intervals are cumulative. O/E, ratio of observed number of pairs to expected number.

Table 7. *Cases of acute lymphoblastic leukaemia, aged* 35–79 *years, diagnosed during* 1989–1993

Time (≤ mth)		Space (≤ km) 1	2	3	4	5	All distances
Same month	Pairs	0	0	0	0	0	124
	O/E	0.00	0.00	0.00	0.00	0.00	
1	Pairs	1	1	1	1	2	352
	O/E	5.24	3.49	1.40	0.68	0.86	
	P	0.0384	0.1453	0.7205	0.6782	0.8165	
2	Pairs	1	1	1	4	5	581
	O/E	3.18	2.12	0.85	1.64	1.30	
3	Pairs	1	1	2	5	6	783
	O/E	2.36	1.57	1.26	1.52	1.15	
4	Pairs	1	1	2	7	8	1041
	O/E	1.77	1.18	0.95	1.60	1.16	
5	Pairs	1	1	2	7	8	1274
	O/E	1.45	0.97	0.77	1.31	0.95	
6	Pairs	1	1	2	7	8	1490
	O/E	1.24	0.83	0.66	1.12	0.81	
9	Pairs	3	3	6	13	16	2136
	O/E	2.59	1.73	1.38	1.45	1.13	
	P	0.0265					
12	Pairs	3	3	7	15	19	2695
	O/E	2.05	1.37	1.28	1.33	1.06	
18	Pairs	4	4	10	20	26	3812
	O/E	1.94	1.29	1.29	1.25	1.03	
	P	0.0342					
All times	Pairs (Cases)	4 (8)	6 (12)	15 (25)	31 (44)	49 (56)	7381 (122)

**P* value shown only if $P = < 0.05$ for test of O/E significantly different from 1.00. Time and space intervals are cumulative. O/E, ratio of observed number of pairs to expected number.

of childhood ALL. Similar findings have been reported from other studies of childhood leukaemia in Britain, with clustering at intervals of 1 km and 3 months and 3 km and 6 months [12] and at intervals up to 5 km and 9 months apart [14] and since there is no overlap between those data and this dataset, the findings reported here are another, independent demonstration of clustering of onsets of childhood ALL. A recent European collaborative study of childhood leukaemia has also reported significant space–time overlap of onsets of cases in small areas of high childhood leukaemia incidence, suggesting that postnatal exposures may be more important than *in utero* exposures [16].

The sensitivity of any method of detecting clustering will be reduced if the disease has a long latent period. This could result in migration of cases (before disease has been diagnosed) out of the area covered by data collection, so leading to 'missed' cases. Similarly, a variable latent period would make it more difficult to detect clustering of the cause of a disease, using date and place of diagnosis, because cases who were initiated together would not necessarily manifest their disease at the same time. Both these problems may have affected our ability to detect clustering in adult ALL.

The pattern of clustering reported here could be interpreted as evidence that an infectious process is involved in the aetiology of childhood ALL. The findings could also result from a promoting factor which prompts clustered onsets or causes symptoms which lead to clinical examinations during which the malignant disease is discovered. This factor may itself be a recent infection with some unknown agent [34]. The difference in the prominence of clustering in children between the two time periods is intriguing. It may be related to different levels of exposure to infection in the period 1984–1988 compared with 1989–1993, either due to different prevalences of infectious agents, or to a change in patterns of behaviour which resulted in different opportunities for exposure to infection. Published information on notifications of selected infectious diseases and laboratory identifications of viruses over the period 1984–1993 [35] showed that there were markedly fewer identifications in 1989–1993 compared with 1984–1988 of Echovirus (1984–1988, 7880; 1989–1993, 4129), Herpes simplex (1984–1988, 81275; 1989–1993, 53786), Varicella zoster (1984–1988, 2736; 1989–1993, 1343), Influenza B (1984–1988, 5100; 1989–1993, 3316), Polio virus (1984–1988, 2683; 1989–1993, 531) and Rubella (1984–1988, 7728; 1989–1993, 3545). There was also a change in the reported rates of influenza and influenza-like illness, decreasing from a level of around 3360 newly diagnosed episodes per 100 000 in 1984–1988 to around 2272 per 100 000 in the period 1989–1993. Whilst by no means conclusive, these patterns are suggestive that there were differences between the periods in levels of exposure to certain infectious agents.

It is also possible that the clustering is due to episodic exposures to some environmental hazard, or that exposure to

such a hazard could render the immune system more susceptible to subsequent infection and ultimately lead to clustered diagnosis of cases. To take the investigation further will require detailed investigations of the locations, medical and occupational histories of the cases and a suitable set of controls.

1. Linet MS. *The leukemias: epidemiologic aspects. Monographs in Epidemiology and Biostatistics*. Volume 6. New York, Oxford University Press, 1985.
2. Greaves MF, Pegram SM, Chan LC. Collaborative group study of the epidemiology of acute lymphoblastic leukaemia subtypes: background and first report. *Leukaemia Res* 1985, **9**, 715–733.
3. Kellett CE. Acute leukaemia in one of identical twins. *Arch Dis Child* 1937, **12**, 239–252.
4. Heath CW, Hasterlik RJ. Leukemia among children in a suburban community. *Am J Med* 1963, **34**, 796–812.
5. Alexander FE. Viruses, clusters and clustering of childhood leukaemia: a new perspective?. *Eur J Cancer* 1993, **29**, 1424–1443.
6. Kinlen LJ, Clarke K, Hudson C. Evidence from population mixing in British New Towns 1946–85 of an infective basis for childhood leukaemia. *Lancet* 1990, **336**, 577–582.
7. Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukaemia. *Leukaemia* 1988, **2**, 120–125.
8. Greaves MF. Aetiology of acute leukaemia. *Lancet* 1997, **349**, 344–349.
9. Cartwright, RA. Epidemiology. In Whittaker JA, ed. *Leukaemia*. Oxford, Blackwell Scientific Publications, 1992, 3–33.
10. Knox G. Epidemiology of childhood leukaemia in Northumberland and Durham. *Br J Prev Soc Med* 1964, **18**, 17–24.
11. Glass AG, Mantel N, Gunz FW, Spears GFS. Time–space clustering of childhood leukaemia in New Zealand. *J Natl Cancer Inst* 1971, **47**, 329–336.
12. Gilman EA, Knox EG. Temporal-spatial distribution of childhood leukaemias and non-Hodgkin lymphomas in Great Britain. In Draper GJ, ed. *The Geographical Epidemiology of Childhood Leukaemia and Non-Hodgkin Lymphomas in Great Britain, 1966–1983*. London, HMSO, Studies on Medical and Population Subjects No. 53, 1991, 77–99.
13. Knox EG, Gilman EA. Leukaemia clusters in Great Britain. 1. Space–time interactions. *J Epid Comm Health* 1992, **46**, 566–572.
14. Gilman EA, Knox EG. Childhood cancers: space–time distribution in Britain. *J Epid Comm Health* 1995, **49**, 158–163.
15. Petridou E, Revinthi K, Alexander FE, *et al*. Space–time clustering of childhood leukaemia in Greece: evidence supporting a viral etiology. *Br J Cancer* 1996, **73**, 1278–1283.
16. Alexander FE, Boyle P, Carli P-M, *et al*. Spatial temporal patterns in childhood leukaemia: further evidence for an infectious origin. *Br J Cancer* 1998, **77**, 812–817.
17. Ederer F, Myers MH, Mantel M. Do leukemia cases come in clusters? *Biometrics* 1964, **20**, 626–638.
18. Glass AG, Mantel N. Lack of space–time clustering of childhood leukemia in Los Angeles County 1960–64. *Cancer Res* 1969, **29**, 1995–2001.
19. Heath CW, Manning MD, Zelkowitz L. Case clusters in the occurrence of leukemia and congenital malformations. *Lancet* 1964, **ii**, 136–137.
20. Black D. *Investigations of Possible Increased Incidence of Cancer in West Cumbria*. Report of the Independent Advisory Group. London, HMSO, 1984.
21. Craft AW, Openshaw S, Birch JM. Childhood cancer in the Northern Region 1968–1982: incidence in small geographical areas. *J Epidemiol Comm Health* 1985, **39**, 53–57.
22. Alexander FE. Investigation of localised spatial clustering and extra-Poisson variation. In Draper GJ, ed. *The Geographical Epidemiology of Childhood Leukaemia and Non-Hodgkin Lymphomas in Great Britain, 1966–1983*. London, HMSO, Studies on Medical and Population Subjects No. 53, 1991, 69–76.
23. Draper GJ, ed. *The Geographical Epidemiology of Childhood Leukaemia and Non-Hodgkin Lymphomas in Great Britain, 1966–1983*. London, HMSO, Studies on Medical and Population Subjects No. 53, 1991.
24. Alexander FE, Boyle P, Carli P-M, *et al*. Spatial clustering of childhood leukaemia: summary results from the EUROCLUS project. *Br J Cancer* 1998, **77**, 818–824.
25. Cartwright, RA, Alexander FE, McKinney PA, Ricketts TL *Leukaemia and Lymphoma: an Atlas of Distribution within Areas of England and Wales 1984–1988*. London, Leukaemia Research Fund, 1990.
26. Cartwright RA, McNally RJQ, Rowland DJ, Thomas J. *The Descriptive Epidemiology of Leukaemia and Related Conditions in Parts of the United Kingdom. 1984–1993*. London, Leukaemia Research Fund, 1997.
27. Knox EG. The detection of space–time interactions. *Appl Statistics* 1964, **13**, 25–29.
28. Mantel N. The detection of disease clustering and a generalised regression approach. *Cancer Res* 1967, **27**, 209–220.
29. Chen R, Mantel N, Klingberg MA. A study of 3 techniques for time–space clustering in Hodgkin's disease. *Statistics in Medicine* 1984, **3**, 173–184.
30. Pike MC, Bull D. Knox test for space–time clustering in epidemiology. *Appl Statist* 1974, **23**, 92–95.
31. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *Br Med J* 1995, **310**, 170.
32. Rodrigues LC, Marshall T, Murphy M, Osmond C. Space–time clustering of births in SIDS: do perinatal infections play a role? *Int J Epidemiology* 1992, **21**, 714–719.
33. Alexander FE, Ricketts TJ, McKinney PA, Cartwright RA. Cancer registration of leukaemias and lymphomas: results of a comparison with a specialist registry. *Comm Med* 1989, **11**, 81–89.
34. Alexander FE. Space–time clustering of childhood acute lymphoblastic leukaemia: indirect evidence for a transmissible agent. *Br J Cancer* 1992, **65**, 589–592.
35. OPCS. *Communicable Disease Statistics*. Series MB2, Nos. 13, 15, 16, 19, 21, London, HMSO.

Acknowledgement—Source of funding: Leukaemia Research Fund.