Childhood cancers: space-time distribution in Britain

E A Gilman and E G Knox

Abstract

Study objectives – To examine a national data set of all childhood cancers for evidence of space-time interactions within three distinct sets of dates and places (at birth, at diagnosis, and at death), to show whether the patterns found for these events represent separate phenomena or statistically interdependent processes, and to see whether the childhood leukaemias and the childhood solid cancers have separate distinctive patterns in these respects.

Design – This was a space-time cluster analysis. The large number of cases enabled division of the data into two sets, one for hypothesis generation and the other for hypothesis testing.

Setting – England, Scotland, and Wales.

Subjects – A national collection of 22 360 children aged 0–15 years with fatal cancers and leukaemias in the period 1953 to 1980.

Main results – There was significant clustering among the leukaemias and lymphomas on date and place of birth (particularly among cases born within 1 km and up to 5 months apart), and on date and place of diagnosis (particularly among cases diagnosed from 3 to 5 km apart and up to 9 months apart). There was no clustering among the solid cancers. These findings were confirmed in two separate analyses of two separate sets of data. **Conclusions** – The birth clustering was significant among pairs diagnosed at differing ages, and diagnosis clustering was significant among pairs born at different times, and it was concluded that the two types of clustering must be regarded as separate and statistically independent phenomena. Both the birth and the diagnosis clusters comprised many independent pairs of cases, with no large multiple case clusters. This suggests the involvement of multiple time-space localised exposures to hazards with short and constant latent intervals; probably an infectious agent or an environmental toxin. Given the separate nature of the two types of clustering, exposure to more than one hazard may be involved.

(*J Epidemiol Community Health* 1995;49:158–163)

It has been suggested that childhood leukaemias and cancers sometimes occur in small and transient clusters, and that this is consistent with either an infective cause or a focal-episodic toxic cause for the disease. Many workers have investigated clustering in relation to the date

and place of onset, or of diagnosis, using a variety of methods.¹ The results have been inconsistent – some studies have found evidence of space-time clustering,^{2–6} while others have not.^{7–9} Some investigations have also examined childhood leukaemia data for clustering in relation to the date and place of birth. These studies have likewise provided some evidence, albeit weak evidence, of space-time clustering.^{10–13}

Most published studies have consisted of relatively localised collections of cases (for example, for a county or city). Some have been based on anecdotal data, or on small data sets, and a range of space-time intervals have been examined for clustering. Sometimes, formal analyses have been based on intervals suggested by previous, less formal impressions of excesses within the same data set, and the statistical significance of the clustering is then difficult to assess. These difficulties are exacerbated by the large number of alternative time and distance comparisons available for testing, as well as the multiple opportunities for selecting an area for study because clustering was already suspected.

Only recently has a complete data set from a whole country become available for such examination.^{5,6} Even there, the data were limited to the dates and places of onset and a simultaneous analysis for clustering of dates and places of birth was not possible.

The study reported here gives the results of examining the space-time distributions of births, disease onset, and deaths within a national collection of children with fatal cancers and leukaemias in the period 1953 to 1980. With such a large number of cases (over 22 000), covering a broad geographical area, the data could be divided into two sets: one set to be used for hypothesis generation, including the definition of the distance and time criteria to be used subsequently; and the other set for hypothesis testing.

The objectives were: (i) to test for and to characterise any space-time interactions within 3 distinct sets of dates and places (at birth, at diagnosis, and at death) and to show whether patterns found for these different events represented separate phenomena or statistically interdependent processes; and (ii) to repeat the examination for the childhood leukaemias and solid cancers, both separately and together, to see whether these diagnostic groups show their own distinctive patterns.

Methods

The data are derived from the Oxford survey of childhood cancers (OSCC).^{14–16} The spatial

Institute of Public and Environmental Health,
Department of Public Health and Epidemiology, The Medical School, The University of Birmingham, Edgbaston, Birmingham, B15 2TT
E A Gilman
E G Knox

Correspondence to:
Dr E A Gilman.

Accepted for publication
March 1994

information consisted of the grid references (to the nearest 100 m) of the home addresses at birth and at death, for each of 22 360 children (aged 0–15 years) in England, Scotland, and Wales who had died of cancer or leukaemia between 1953 and 1980 inclusive. The address at diagnosis was not separately recorded and the address at death was used as a proxy for the place of cancer diagnosis. The grid references were obtained by matching the postcodes of the case addresses with the *Central Postcode Directory*, a computer tape of all the postcodes in Britain (approximately 1.3 million) together with their map coordinates.¹⁷ The temporal information available for each case included the dates of birth, diagnosis, and death.

The data were divided into two sets, one for hypothesis generation, and the other for hypothesis testing. All cases who were born or died in central Britain, between Ordnance Survey grid lines 200.0 north and 400.0 north, were analysed first (8750 cases in total). (Grid line 200.0 north passes from Pembroke in the west to just north of Watford in the east. Grid line 400.0 north passes through Liverpool in the west to just south of Grimsby in the east.) Data from the remaining cases were later used to test specific hypotheses generated by the “central” set.

Cases were excluded from space-time analysis if the date and place of the event of interest were not known with sufficient accuracy. There were 3261 cases born in the period 1953–64 within grid lines 200.0 and 400.0 north, 2712 (83%) of these were used in the analysis of the distribution of date and place of birth. Of the cases excluded, 547 were because the full home address at the time of birth was not obtained; most (80%) of these exclusions were because the case had not been interviewed by the OSCC. Only two cases (both leukaemias and lymphomas) were excluded from analysis because date of birth was not known. There was very little difference between the proportion of leukaemias and lymphomas excluded from analysis (15%) and the proportion of solid cancers excluded (18%).

For analysis of the date and place of cancer diagnosis, address at death was used as a proxy for address at diagnosis. There were 8128 cases who died within central Britain, of these 592 (7%) were excluded from analysis, 590 because the date of diagnosis was not known, and two cases because both date of diagnosis and address were not known. There was no difference between the leukaemias and lymphomas and the solid cancers in the proportion of cases excluded from analysis.

For date and place of death, only 2 of the 8128 cases in central Britain were excluded from analysis, both were patients with solid cancers who had not been interviewed, whose addresses at death could not be postcoded and grid referenced with sufficient accuracy.

In the second data set, 7130 cases were born outside central Britain in the period 1953–64, and 4738 (67%) of these were used in the analyses of the distribution of date and place of birth, a lower proportion than was used in

the analyses of cases born in central Britain, where 83% of cases were used. All of the cases excluded were because the full home address at time of birth was not obtained, most (82%) of these exclusions were because the case had not been interviewed by the OSCC. The rest were interviewed cases for whom the address at time of birth was not given in sufficient detail to allow it to be accurately postcoded and grid referenced. There was very little difference between the proportion of leukaemias and lymphomas excluded from analysis (32%) and the proportion of solid cancers excluded (35%).

There were 14 625 cases who died outside central Britain, of these 1567 (11%) were excluded from analysis of date and place of cancer diagnosis, most because the date of diagnosis was not known. There was practically no difference between the proportion of leukaemias and lymphomas (11%) and the proportion of solid cancers (10%) excluded from analysis.

For date and place of death, only 18 of the 14 625 cases outside central Britain were excluded from analysis. All were cases whose death addresses could not be postcoded and grid referenced with sufficient accuracy.

The distribution of cases in both space and time was examined using the Knox method, comparing the observed number of pairs of cases within short time and space intervals with the number expected if the temporal and geographic distances between pairs were independent of each other. Significance was tested by calculating a statistic, d , which is distributed as the standard normal deviate.³¹⁸ The confusing effects of a possible spatial-temporal heterogeneity in the population at risk, caused by shifts in population concentration over the long time period covered by the data, were avoided by dividing the data set into shorter time periods and using the modified analysis described by Klauber and Mustacchi.¹⁹ This produces a summary statistic, d_s , which is also distributed as the standard normal deviate. The application of the Knox method to childhood cancer data, and the modification suggested by Klauber and Mustacchi are described in more detail in Gilman and Knox.⁵

Analyses were first performed for all cancers together, and then separately for the leukaemias and lymphomas grouped together, and for the solid cancers. Some analyses were repeated for two separate age groups: those under 60 months of age, and those aged 60 months or older.

To avoid possible artefacts because of the incompleteness of those birth cohorts for whom a full 16 years of follow up were not available, the analyses of date and place of birth were carried out in the “completed” cohorts only – that is, cases born between 1953 and 1964 inclusive. This was to avoid the possibility of spatial/temporal heterogeneity in the extent to which onsets and deaths might have been postponed beyond the time limits of case ascertainment. Analyses of date and place of diagnosis and of death were based on deaths during the entire study period, 1953–80.

Results

The results of the Knox analyses are presented in Tables 1–10 (only tables 1, 2, 3, 5 and 7 are published here, the remainder are available from the authors). The range of space and time intervals used to test for clusters were: (space) <1 km, <2 km, <3 km, <4 km, <5 km, <10 km, <20 km; and (time) 0, <1, <2, <3, <4, <5, <6,

<9, <12, <18, <24, <48 months. For clarity of presentation the tables show details for intervals up to 5 km and up to 12 months only. The numbers within particular intervals are cumulative; for example, pairs of cases born within 3 km and 3 months of each other include those pairs born within 2 km and 3 months, and they in turn include pairs born within 2 km and 2 months, and so on.

Table 1 Knox space-time analyses of date and place of birth: leukaemias and lymphomas in central Britain 1953–64 (1435 cases)

Time (<months)		Space (<km)					All distances
		1	2	3	4	5	
0	Pairs	7	20	36	58	84	7214
	O/E	1.41	1.34	1.26	1.25	1.27†	
	(Cases)	(14)	(40)	(70)	(106)	(156)	(1435)
1	Pairs	18	57	97	154	216	21138
	O/E	1.23	1.30*	1.16	1.13	1.11	
	(Cases)	(36)	(102)	(170)	(242)	(314)	(1435)
2	Pairs	32	89	154	252	359	35046
	O/E	1.32	1.22	1.11	1.11	1.11*	
	(Cases)	(64)	(158)	(252)	(353)	(453)	(1435)
3	Pairs	43	125	218	336	481	48947
	O/E	1.27	1.23*	1.12	1.06	1.07	
	(Cases)	(85)	(212)	(330)	(420)	(527)	(1435)
4	Pairs	51	154	269	422	599	62714
	O/E	1.18	1.18*	1.08	1.04	1.04	
	(Cases)	(101)	(252)	(382)	(478)	(584)	(1435)
5	Pairs	59	179	321	509	731	76479
	O/E	1.12	1.13	1.06	1.03	1.04	
	(Cases)	(114)	(289)	(426)	(537)	(653)	(1435)
6	Pairs	65	205	370	594	851	90158
	O/E	1.04	1.10	1.03	1.02	1.03	
	(Cases)	(122)	(318)	(464)	(591)	(708)	(1435)
9	Pairs	98	300	553	866	1227	130538
	O/E	1.09	1.11	1.07	1.03	1.02	
	(Cases)	(171)	(416)	(586)	(718)	(823)	(1435)
12	Pairs	130	381	722	1144	1617	170259
	O/E	1.11	1.08	1.07	1.04	1.03	
	(Cases)	(212)	(492)	(660)	(795)	(889)	(1435)
All times	Pairs	710	2135	4085	6639	9468	1028895
	(Cases)	(677)	(1008)	(1148)	(1205)	(1255)	(1435)

O/E = ratio of observed number of pairs to expected number.

* Crude O/E significantly different from 1.00, $p < 0.05$.

† O/E remained significant after calculation of d_i (see text). Time and space intervals are cumulative.

Table 2 Knox space-time analyses of date and place of birth: solid cancers in central Britain 1953–64 (1277 cases)

Time (<months)		Space (<km)					All distances
		1	2	3	4	5	
0	Pairs	8	18	26	37	48	5823
	O/E	1.86	1.40	1.06	0.95	0.88	
	(Cases)	(16)	(35)	(49)	(71)	(89)	(1277)
1	Pairs	12	38	73	109	145	16966
	O/E	0.96	1.01	1.02	0.96	0.91	
	(Cases)	(24)	(69)	(122)	(172)	(219)	(1277)
2	Pairs	19	63	129	190	262	28450
	O/E	0.90	1.00	1.07	1.00	0.98	
	(Cases)	(38)	(112)	(199)	(270)	(344)	(1277)
3	Pairs	29	98	190	277	371	39415
	O/E	1.00	1.12	1.14	1.05	1.01	
	(Cases)	(57)	(158)	(265)	(351)	(431)	(1277)
4	Pairs	36	120	237	355	472	50794
	O/E	0.96	1.07	1.10	1.04	0.99	
	(Cases)	(70)	(188)	(315)	(416)	(500)	(1277)
5	Pairs	44	145	285	431	582	62036
	O/E	0.96	1.06	1.09	1.04	1.00	
	(Cases)	(83)	(221)	(361)	(474)	(556)	(1277)
6	Pairs	50	167	338	513	690	73092
	O/E	0.93	1.03	1.09	1.05	1.01	
	(Cases)	(94)	(253)	(408)	(524)	(611)	(1277)
9	Pairs	66	232	459	730	992	105813
	O/E	0.84	0.99	1.03	1.03	1.00	
	(Cases)	(123)	(320)	(493)	(636)	(731)	(1277)
12	Pairs	95	316	606	944	1285	137743
	O/E	0.93	1.04	1.04	1.02	1.00	
	(Cases)	(169)	(396)	(578)	(709)	(798)	(1277)
All Times	Pairs	602	1804	3444	5457	7624	814726
	(Cases)	(607)	(914)	(1012)	(1071)	(1118)	(1277)

O/E = ratio of observed number of pairs to expected number. Time and space intervals are cumulative.

DATE AND PLACE OF BIRTH

The children with leukaemias and lymphomas born in central Britain clustered on date and place of birth (table 1). They showed excesses of pairs born within a range of short time intervals (up to 5 months) and short distances (up to 5 km) of each other. For example, there were 84 pairs born within 5 km and in the same month, against an expected value of 66.14. The birth clusters included pairs with onsets at widely varying ages. Clustering was present among pairs who were both diagnosed below 60 months of age, and in pairs who were both diagnosed aged 60 months or older; and there was also significant space-time clustering among age cross-pairs, pairs where each onset was in a different age group.

By contrast (table 2), the solid cancers showed no significant space-time interactions on date and place of birth, and had few observed : expected (O/E) ratios large enough to suggest clustering; the largest being for pairs born within 1 km or 2 km and in the same month. Among all cancers combined, the pattern of clustering resembled that of the leukaemias and lymphomas, but with smaller O/E ratios, probably reflecting the dilution of the leukaemia/lymphoma clustering by the non-clustering of the solid cancers.

These initial findings suggest that space-time birth clustering is concentrated in the leukaemias and lymphomas; and the data from the rest of Britain confirmed this. The patterns of birth clustering were repeated for the leukaemias and lymphomas, particularly those born within 1 km and up to 5 months apart (table 3), but were absent for the solid cancers (table 4: available from the authors).

In neither the central nor the wider data set were there significant interactions among the diagnosis cross-pairs (where each member of the pair was from a different diagnostic group), reinforcing the conclusion that the space-time distribution of the leukaemia/lymphoma births is different from that of the solid cancer births.

The consistency of the findings in both parts of the country weakens the possibility that the leukaemia and lymphoma birth clusters could be the result of a statistical or a regionally limited ascertainment artefact.

DATE AND PLACE OF DIAGNOSIS

In addition to the birth clustering, there was evidence of clustering on date and place of diagnosis in the data for central Britain (table 5). Again, this was present among the leukaemias and lymphomas, but not among the

solid cancers (table 6: available from the authors). The patterns differed from those observed on date and place of birth, and involved longer time intervals. There were large, significant concentrations of pairs diagnosed

within 2 km or 3 km of each other and over a range of time intervals, from 1 month apart, up to 18 months apart. Again, there were significant interactions among the age cross-pairs. All cancers combined showed the same dilution effects as for the birth clusters.

The presence of diagnosis clustering among the leukaemias and lymphomas in central Britain was confirmed in the second data set, although with slightly longer space-time intervals, that is, up to 5 km apart and at intervals up to 9 months apart (table 7). The absence of diagnosis clustering among the solid cancers was also confirmed in this second data-set (table 8: available from the authors).

Unlike the birth clusters, there were several significant interactions at short space and time intervals among diagnosis cross-pairs in central Britain. However, this was not confirmed in the rest of Britain. It therefore seems, yet again, that the leukaemias and lymphomas behave differently from the solid cancers in these respects.

DATE AND PLACE OF DEATH

For deaths occurring in central Britain, there was no space-time interaction among the leukaemia and lymphoma pairs: for all ages, or pairs within separate age groups, or among age cross-pairs. This was confirmed by the deaths data from the rest of Britain.

However, deaths from solid cancers occurring in central Britain (table 9: available from the authors) showed evidence of clustering over short times and short distances (from 1 km to 3 km and time intervals up to 2 months apart). There was also an interaction among the age cross-pairs, especially for pairs who died within 1 km and in the same month. In the absence of noticeable space-time interactions of solid cancer births or onsets, this seemed to indicate an "epidemic" distribution of terminal infections. However, these patterns were not repeated among the solid cancer deaths outside central Britain (table 10: available from the authors). If this was indeed an epidemic property of terminal illnesses, rather than a statistical aberration, then it was geographically localised, and we conclude that, exceptions apart, the solid cancers do not generally cluster on date and place of death.

Discussion

POSSIBLE CAUSES OF ARTEFACTUAL SPACE-TIME CLUSTERING

Firstly, consideration must be given to the possibility of artefactual causes for the space-time clustering within these data. Cases had been excluded from the reported analyses if the date and place of the event of interest were not sufficiently precise. This was done in the interests of accuracy, but might have resulted in clustered exclusions. However, the exclusion rates were similar among the leukaemias/lymphomas and solid cancers, although the patterns of clustering in these diagnostic groups were different. The greatest proportion of exclusions related to the address at birth, so

Table 3 Knox space-time analyses of date and place of birth: leukaemias and lymphomas outside central Britain 1953-64 (2478 cases)

Time (<mo)		Space (<km)					All distances
		1	2	3	4	5	
0	Pairs	15	37	61	87	120	21439
	O/E	1.75†	1.31	1.08	0.95	0.92	
	(Cases)	(30)	(73)	(117)	(160)	(214)	(2478)
1	Pairs	33	88	165	255	368	64364
	O/E	1.28	1.04	0.97	0.93	0.94	
	(Cases)	(66)	(168)	(302)	(428)	(562)	(2478)
2	Pairs	54	146	289	444	629	106959
	O/E	1.26	1.04	1.02	0.97	0.97	
	(Cases)	(108)	(270)	(472)	(638)	(808)	(2478)
3	Pairs	77	203	410	648	896	148948
	O/E	1.29*	1.03	1.04	1.02	0.99	
	(Cases)	(148)	(360)	(624)	(819)	(993)	(2478)
4	Pairs	95	250	512	815	1141	190397
	O/E	1.24*	1.00	1.02	1.00	0.99	
	(Cases)	(181)	(430)	(726)	(929)	(1096)	(2478)
5	Pairs	118	300	622	993	1406	231771
	O/E	1.27†	0.98	1.02	1.00	1.00	
	(Cases)	(224)	(501)	(830)	(1034)	(1201)	(2478)
6	Pairs	132	349	712	1149	1628	272649
	O/E	1.21*	0.97	0.99	0.99	0.98	
	(Cases)	(244)	(564)	(911)	(1111)	(1267)	(2478)
9	Pairs	171	503	1018	1659	2323	393400
	O/E	1.08	0.97	0.98	0.99	0.97	
	(Cases)	(306)	(726)	(1099)	(1301)	(1457)	(2478)
12	Pairs	217	654	1320	2151	3040	512518
	O/E	1.06	0.97	0.98	0.98	0.98	
	(Cases)	(377)	(862)	(1245)	(1455)	(1619)	(2478)
All Times	Pairs	1230	4043	8101	13090	18647	3069003
	(Cases)	(1228)	(1823)	(2034)	(2148)	(2231)	(2478)

O/E=ratio of observed number of pairs to expected number.

* Crude O/E significantly different from 1.00, $p<0.05$.

† O/E remained significant after calculation of d_i (see text). Time and space intervals are cumulative.

Table 5 Knox space-time analyses of date and place of diagnosis: leukemias and lymphomas in central Britain 1953-80 (3862 cases)

Time (<mo)		Space (<km)					All distances
		1	2	3	4	5	
0	Pairs	15	46	79	124	180	24392
	O/E	1.16	1.15	1.02	0.99	1.01	
	(Cases)	(30)	(92)	(158)	(243)	(344)	(3856)
1	Pairs	43	148	264	405	559	72678
	O/E	1.12	1.24*	1.14*	1.09	1.06	
	(Cases)	(86)	(280)	(468)	(677)	(873)	(3860)
2	Pairs	76	248	426	654	932	121105
	O/E	1.19	1.24†	1.11*	1.06	1.06	
	(Cases)	(146)	(426)	(681)	(965)	(1243)	(3860)
3	Pairs	110	336	588	897	1296	169118
	O/E	1.23*	1.21†	1.10*	1.04	1.05	
	(Cases)	(206)	(560)	(876)	(1193)	(1494)	(3860)
4	Pairs	132	419	761	1162	1664	217477
	O/E	1.15	1.17†	1.10*	1.05	1.05	
	(Cases)	(245)	(679)	(1048)	(1400)	(1707)	(3861)
5	Pairs	154	501	926	1420	2003	265553
	O/E	1.10	1.15*	1.10*	1.05	1.04	
	(Cases)	(283)	(788)	(1204)	(1592)	(1870)	(3861)
6	Pairs	186	585	1078	1669	2357	313828
	O/E	1.12	1.13*	1.08*	1.04	1.03	
	(Cases)	(332)	(880)	(1328)	(1713)	(1996)	(3861)
9	Pairs	262	827	1542	2413	3424	458795
	O/E	1.08	1.10*	1.06*	1.03	1.02	
	(Cases)	(435)	(1124)	(1629)	(2030)	(2311)	(3861)
12	Pairs	339	1082	2035	3190	4514	603164
	O/E	1.06	1.09*	1.06*	1.04	1.03	
	(Cases)	(549)	(1333)	(1858)	(2251)	(2515)	(3861)
All Times	Pairs	3944	12272	23665	38096	54289	7955591
	(Cases)	(2650)	(3234)	(3439)	(3570)	(3658)	(3862)

O/E=ratio of observed number of pairs to expected number.

* Crude O/E significantly different from 1.00, $p<0.05$.

† O/E remained significant after calculation of d_i (see text). Time and space intervals are cumulative.

Table 7 Knox space-time analyses of date and place of diagnosis: leukaemias and lymphomas outside central Britain 1953–80 (6511 cases)

Time (<months)		Space (<km)					All distances
		1	2	3	4	5	
0	Pairs	29	82	182	286	421	69943
	O/E (Cases)	1.20 (58)	1.06 (162)	1.19† (343)	1.15* (510)	1.17† (717)	(6503)
1	Pairs	88	247	509	791	1154	210500
	O/E (Cases)	1.21 (172)	1.06 (465)	1.10* (872)	1.05 (1240)	1.06* (1643)	(6509)
2	Pairs	136	391	815	1297	1881	350705
	O/E (Cases)	1.13 (263)	1.01 (706)	1.06 (1290)	1.04 (1804)	1.04 (2257)	(6509)
3	Pairs	186	560	1147	1838	2660	491106
	O/E (Cases)	1.10 (354)	1.03 (964)	1.07* (1653)	1.05* (2257)	1.05* (2733)	(6509)
4	Pairs	236	708	1489	2370	3423	631922
	O/E (Cases)	1.08 (451)	1.01 (1183)	1.08* (2012)	1.05* (2648)	1.05* (3130)	(6509)
5	Pairs	299	915	1848	2937	4226	772064
	O/E (Cases)	1.12* (561)	1.07* (1449)	1.09† (2313)	1.07† (2951)	1.06† (3433)	(6509)
6	Pairs	355	1061	2152	3448	4982	911909
	O/E (Cases)	1.13* (649)	1.05 (1619)	1.08† (2558)	1.06* (3198)	1.06† (3680)	(6509)
9	Pairs	488	1538	3115	4983	7224	1332351
	O/E (Cases)	1.06 (864)	1.04 (2114)	1.07† (3097)	1.05* (3752)	1.05* (4197)	(6510)
12	Pairs	633	1997	4012	6457	9360	1750555
	O/E (Cases)	1.05 (1075)	1.03 (2470)	1.05* (3469)	1.03* (4087)	1.04* (4470)	(6510)
All Times	Pairs	7303	23463	46398	75553	109331	21193305
	(Cases)	(4792)	(5655)	(5937)	(6105)	(6220)	(6511)

O/E = ratio of observed number of pairs to expected number.

* Crude O/E significantly different from 1.00, $p < 0.05$.

† O/E remained significant after calculation of d_i (see text). Time and space intervals are cumulative.

the impact of the birth-place exclusions was assessed by repeating the analyses using place of death as proxy if the place of birth was not available. The results were similar, although at reduced levels of significance.

The diagnosis clusters could have been the result of clustered failures to recognise the diagnosis, or to clustered recognition and reporting of cases, where the discovery of an initial case had led to increased local vigilance. However, the second contingency would have led to corresponding deficits of pairs of cases separated by somewhat longer time intervals; and no such deficits were found. In addition, clustering on date and place of diagnosis was recently reported in a data-set exhibiting very high levels of ascertainment. This consisted of childhood leukaemias and non-Hodgkin's lymphomas registered in Britain between 1966 and 1983.⁵ Space-time clustering was again present among pairs of all ages, within separate age bands, and among age band, cross-pairs. Ascertainment among this set of cases (which included both fatal and non-fatal cases) was around 99%,²⁰ and effectively excludes the proposition that localised variations in the reporting and diagnosis of cases could explain the clusters.

This conclusion is strongly supported by the consistent difference in clustering behaviour between the leukaemias and the lymphomas on the one hand, and the solid cancers on the other; and by the consistency of both findings in different parts of the country.

Artefactual birth clustering can result from the presence of concordant twins; and a genetically determined similarity of their latent

periods could lead to clustering of their dates and places of diagnosis. The data set used for the analyses of birth clustering (cases who were born between 1953 and 1964 inclusive) included one pair of twins concordant for malignant oligodendroglioma (born in central Britain). For the analyses of diagnosis clustering the data included this pair (diagnosed in central Britain) and two pairs concordant for leukaemia (diagnosed outside central Britain). The oligodendroglioma diagnoses were separated by 48 months. One leukaemia pair was diagnosed in the same month, and the other pair separated by 22 months. The removal of these twin pairs from the data did not change the pattern or significance of the clustering already described.

EPIDEMIOLOGICAL INTERPRETATION OF FINDINGS

One of our objectives was to see whether different diagnostic groups of childhood cancer exhibited different and characteristic clustering patterns. The analyses showed that the leukaemias and lymphomas cluster on date and place of birth, and on date and place of diagnosis, but that the solid cancers do not. These findings were confirmed in two separate analyses of two separate sets of data.

Another objective was to see whether the clustering patterns of different events at different ages were statistically inter-related, or whether they represented separate phenomena – the diagnosis clustering “primary” and solely related to events around the time of diagnosis; or is it “secondary” to birth clustering, reflecting the delayed effects of prenatal or perinatal events? This question has been raised before, but without simultaneous information on both births and onsets, the authors were unable to resolve it fully.^{5,6}

Clustering of dates and places of diagnosis could be secondary to birth clustering only if there was little or no migration between the times of birth and diagnosis, and if the latent period between cancer initiation and promotion were nearly constant. Secondary diagnosis clusters would then be concentrated among pairs diagnosed at similar ages. However, primary diagnosis clustering, depending upon contemporary events rather than upon clustered birth events, would not necessarily be age concordant.

All the evidence presented here denies the possibility that the later (diagnosis) clustering is secondary to the earlier (birth) clustering. Firstly, as was discussed earlier, the birth clustering found in the leukaemias and lymphomas was significant among pairs with different ages at diagnosis, indicating that the period between cancer initiation and recognition is variable. Secondly, diagnosis clustering itself does not seem to be age group specific: it was present in the cross-pairs between age groups (and hence between birth cohorts), and the space-time interactions among the age cross-pairs were similar to the pattern shown by the analysis of pairs of all ages. It should be noted that the age heterogeneity of the diagnosis clusters

is also incompatible with the hypothesis that the birth clusters could be statistically secondary to the diagnosis clusters.

The results of the analyses reported here therefore support the assertion that, among the leukaemias and lymphomas, clustering on date and place of diagnosis and clustering on date and place of birth are both "primary". The birth clustering was significant among pairs diagnosed at differing ages, and diagnosis clustering was significant among pairs born at different times. The two types of clustering must be regarded as separate and statistically independent phenomena.

POSSIBLE BIOLOGICAL MECHANISMS

The leukaemia and lymphoma clusters comprised many independent pairs of cases, with little sharing of cases between different pairs, and there were no large multiple case clusters. This suggests the involvement of multiple time-space localised exposures to hazards with short and constant latent intervals; probably an infectious agent or an environmental toxin. Given the separate nature of the two types of clustering, exposure to more than one hazard may be involved.

Such exposures, either during the pregnancy or around the time of birth, could induce space-time patterns of the kinds observed for date and place of birth. If a toxic source were involved, the widespread geographical scatter of the leukaemia and lymphoma clusters would necessitate its common presence in residential areas, but fluctuating in the severity of emission or the closeness of contact. If an infectious agent were involved, it must also be quite widespread. Upper respiratory diseases during pregnancy have been associated with an increased risk of leukaemia or lymphoma for the unborn child,^{21 22} and these, together with the treatments they invoke, must stand as candidates.

The diagnosis clusters could represent "promotion" events which either accelerate the disease process, or prompt haematological examinations which lead to the recognition of the leukaemia. Local epidemics of infections might interact with pre-leukaemic children, whose immune systems are already under stress, to produce clustered onsets of symptoms. However, there is recent evidence that acute space-time clustering may be superimposed on a longer term tendency towards geographical clustering, which would be more difficult to explain in these terms.^{6 23} A common process may be involved in both the space-time and the geographical concentrations, perhaps a repeated exposure to some locally fluctuating

oncogenic agent. Concentrations of children may have been "initialized" by earlier micro-epidemics of infection, possibly before birth, and the onsets triggered by reinfection, or by a toxic suppression of immunity.

We would like to acknowledge the help received from Dr A M Stewart and Mr G W Kneale in using the data held by the Oxford Survey of Childhood Cancers. This study was carried out as part of a research programme funded by the Medical Research Council (project grant number G8605397) and subsequently by the Three Mile Island Public Health Fund (USA).

- 1 Smith PG. Spatial and temporal clustering. In: Schottenfeld D, Fraumeni JF eds. *Cancer epidemiology and prevention*. Philadelphia: Saunders, 1982.
- 2 Pinkel D, Dowd JE, Bross IDJ. Some epidemiological features of malignant solid tumours of children in the Buffalo, NY, area. *Cancer* 1963;16:28-33.
- 3 Knox G. Epidemiology of childhood leukaemia in Northumberland and Durham. *Br J Prev Soc Med* 1964;18:17-24.
- 4 Glass AG, Mantel N, Gunz FW, Spears GFS. Time-space clustering of childhood leukaemia in New Zealand. *J Natl Cancer Inst* 1971;47:329-36.
- 5 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, 1991. Studies on Medical and Population Subjects No 53.
- 6 Knox EG, Gilman EA. Leukaemia clusters in Great Britain. 1. Space-time interactions. *J Epidemiol Community Health* 1992;46:566-72.
- 7 Ederer F, Myers MH, Mantel M. Do leukemia cases come in clusters? *Biometrics* 1964;20:626-38.
- 8 Lock SP, Merrington M. Leukaemia in Lewisham (1957-63). *BMJ* 1967;3:759-90.
- 9 Glass AG, Mantel N. Lack of space-time clustering of childhood leukemia in Los Angeles County, 1960-64. *Cancer Res* 1969;29:1995-2001.
- 10 Fraumeni JF, Ederer F, Handy VH. Temporal-spatial distribution of childhood leukemia in New York state: Special reference to case clustering by year of birth. *Cancer* 1966;19:996-1000.
- 11 Stark CR, Mantel N. Temporal-spatial distribution of birth dates for Michigan children with leukemia. *Cancer Res* 1967;27:1749-75.
- 12 Klauber MR. A study of clustering of childhood leukemia by hospital of birth. *Cancer Res* 1968;28:1790-2.
- 13 Smith PG, Pike MC, Till MM, Hardisty RM. Epidemiology of childhood leukaemia in Greater London: A search for evidence of transmission assuming a possibly long latent period. *Br J Cancer* 1976;33:1-8.
- 14 Knox EG, Stewart AM, Kneale GW, Gilman EA. Prenatal irradiation and childhood cancer. *J Radiol Prot* 1987;7:177-89.
- 15 Bithell JF, Stewart AM. Pre-natal irradiation and childhood malignancy: a review of British data from the Oxford Survey. *Br J Cancer* 1975;31:271-87.
- 16 Stewart AM, Webb J, Hewitt D. A survey of childhood malignancies. *BMJ* 1958;i:1495-508.
- 17 Wilson PR, Elliot DJ. An evaluation of the postcode address file as a sampling frame and its use within OPCS. *J R Statist Soc A* 1987;150:230-40.
- 18 Mantel N. The detection of disease clustering and a generalised regression approach. *Cancer Res* 1967;27:209-20.
- 19 Klauber MR, Mustacchi P. Space-time clustering of childhood leukemia in San Francisco. *Cancer Res* 1970;30:1969-73.
- 20 Stiller CA, O'Connor CM, Vincent TJ, Draper GJ. The National Registry of Childhood Tumours and the leukaemia/lymphoma data for 1966-1983. In: Draper GJ ed. *The geographical epidemiology of childhood leukaemia and non-Hodgkin lymphomas in Great Britain, 1966-83*. London: HMSO, 1991. Studies on Medical and Population Subjects No 53.
- 21 Fedrick J, Alberman ED. Reported influenza in pregnancy and subsequent cancer in the child. *BMJ* 1972;ii:485-88.
- 22 Gilman EA, Kinnier Wilson LM, Kneale GW, Waterhouse JAH. Childhood cancers and their association with pregnancy drugs and illnesses. *Paediatric and Perinatal Epidemiology* 1989;3:66-94.
- 23 Knox EG, Gilman EA. Leukaemia clusters in Great Britain. 2. Geographical concentrations. *J Epidemiol Community Health* 1992;46:573-76.