ELSEVIER

European Journal of Cancer 40 (2004) 585–589

European Journal of Cancer

www.ejconline.com

Little or no space–time clustering found amongst cases of childhood lymphoma in North West England

R.J.Q. McNally[a,*], F.E. Alexander[b], O.B. Eden[c], J.M. Birch[a]

[a]*Cancer Research UK Paediatric & Familial Cancer Research Group, Central Manchester and Manchester Children's University Hospitals NHS Trust, Manchester M27 4HA, UK*

[b]*Department of Public Health Sciences, The University of Edinburgh Medical School, Teviot Place, Edinburgh EH8 9AG, UK*

[c]*Academic Unit of Paediatric Oncology, Central Manchester and Manchester Children's University Hospitals NHS Trust and Christie Hospital Trust, Wilmslow Road, Manchester M20 4BX, UK*

Received 22 May 2003; received in revised form 15 September 2003; accepted 3 October 2003

Abstract

We have examined space–time clustering amongst cases of lymphoma in children, aged 0–14 years, using population-based data from the North West of England for the period 1954–2001. There was little or no evidence for space–time clustering amongst all the lymphomas or amongst those sub-groups identified in advance.

Keywords: Aetiology; Children; Environment; Epidemiology; Infection; Lymphoma; Space–time clustering

1. Introduction

We have previously reported significant evidence of seasonal variation in the onset of childhood Hodgkin's disease, diagnosed during 1954–1996, and based on time of first symptom [1]. The finding of seasonal variation is consistent with an infectious aetiology. There is evidence for an involvement of infections in certain lymphomas. Specifically, Epstein–Barr virus (EBV) has been associated with Hodgkin's disease (HD) [2] and Burkitt's lymphoma [3]; the percentage of cases with viral DNA in the malignant cells is highest in children and older adults [4]. There have been few studies of space–time clustering of lymphomas generally amongst children. However, HD has been studied more specifically. One study [5] found no evidence for space–time clustering of HD amongst children, but did find space–time clustering amongst young adults (aged 15–34 years). Analyses of case clustering [6] have usually involved cases in young people, but not children, and have shown contradictory findings. Some studies have found space–time clustering between cases [7,8], but other studies did not [9,10]. Spatial clustering has been identified predominantly in younger adult cases (<35 years) [11–13]. We have now analysed for space–time clustering, using cases of lymphoma included in the Manchester Children's Tumour Registry (MCTR), during 1954–2001. We have used places of birth and diagnosis, and times of birth, first symptom and diagnosis, so that a direct comparison may be made with our previous study [1].

2. Patients and methods

All cases of children aged 0–14 years, diagnosed with lymphoma between 1 January 1954 and 31 December 2001, and registered by the MCTR were analysed. Reporting practice to the registry has been consistently good during this period. The registry retains histopathological material and diagnostic re-review is undertaken periodically. Ordnance Survey (OS) eight-digit grid references were allocated to each case with respect to addresses at time of birth and diagnosis, locating each address to within 0.1 km. The following diagnostic groups were specified *a priori* for analysis: (i) Hodgkin's disease (HD); (ii) non-Hodgkin's lymphoma (NHL); (iii) Hodgkin's disease and non-Hodgkin's lymphoma; (iv) all lymphomas.

* Corresponding author. Tel. +44-161-727-2507/2522; fax: +44-161-727-2508.

E-mail address: richard.mcnally@man.ac.uk (R.J.Q. McNally).

0959-8049/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2003.10.004

There are six possible space–time interactions between: (i) times and places of diagnosis; (ii) times and places of birth; (iii) time of diagnosis and place of birth; (iv) time of birth and place of diagnosis; (v) time of first symptom and place of birth; and (vi) time of first symptom and place of diagnosis. The interpretation of these interactions will depend on the extent of migration between birth and diagnosis among cases [14]. If the cases have not migrated then these reduce to two possibilities: an interaction between proximity of 'place of residence' and time of diagnosis or time of first symptom or time of birth. However, since a large proportion of the children moved between birth and time of first symptom or diagnosis, we believe that migration may have some effect, and there may be differences in the space–time clustering effects that are dependent on spatial definitions. For example, an interaction between times and places of birth only would indicate an aetiological exposure around the place of birth, prenatally or shortly after the time of birth. It would also indicate that the disease has a variable latent period. Furthermore, an interaction between times and places of diagnosis would indicate a later effect which may be explained by an aetiological exposure around the place of diagnosis and close to the time of diagnosis. It would also indicate that the disease has a short latent period following the exposure. Additional details are given in Birch and colleagues in Ref. [14].

Knox [15] tests were applied to the data with thresholds fixed, *a priori*, as: close in space, less than 5 km, and close in time, less than 1 year apart. One-sided tests were used to detect a significant interaction. The strength of interactions (S) was indicated by calculating $[(O-E)/E]\times 100$ counts of pairs which are close in space and close in time (O, observed number of close pairs, and E, expected number of close pairs). To adjust for the effects of different population densities, the tests were repeated replacing geographical distance thresholds by distance to the Nth nearest neighbour, using all locations of all the cases in the data-set except addresses for the same child at a different time. N was chosen such that the mean distance was 5 km ($N=11$ for both birth and diagnosis locations).

Two problems are apparent with the Knox test: boundary problems and the arbitrariness of the thresholds chosen, and these are discussed elsewhere in Ref. [16]. To overcome these, a second order procedure based on K-functions [17] is used, employing both geographical distance and nearest neighbour (NN) approaches as described above.

Data were analysed in two age-groups (0–9 years and 10–14 years) and for cross-clustering between the 'younger' and 'older' cases. This division corresponds to the age distribution of the subtypes of HD. The mixed cellularity subtype predominates in the younger group, whilst the nodular sclerosing subtype predominates in the older group. Data were also analysed by examining clustering pairs which contained at least 1 male case ('male: any') and clustering pairs which contained at least 1 female case ('female: any'). The 50% of addresses that were closer to their 11th NN than the median distance were classified as being located in a more densely populated area, whilst the 50% of addresses that were further from their 11th NN than the median distance were classified as being located in a less densely populated area [16]. Analysis by population density was undertaken by considering clustering pairs which contained at least 1 case from a more densely populated area ('more densely populated: any') and clustering pairs which contained at least 1 case from a less densely populated area ('less densely populated: any'). The methods are given in more detail in Birch and colleagues in Ref. [14] and McNally and colleagues in Ref. [16].

Statistical significance was indicated if $P<0.05$ for the cases in the specified diagnostic group, using at least 2 of the 4 methods (the geographical or NN versions of the Knox test, or the geographical or NN versions of the K-function method) and including a NN threshold version.

3. Results

The study included 414 cases of lymphoma, comprising 188 cases of HD, 211 cases of NHL (including 8 cases of Burkitt's lymphoma) and 15 cases of other lymphoreticular neoplasms (comprising 5 malignant lymphomas, not otherwise specified (NOS); 8 malignant histiocytosis and 2 malignant lymphoproliferative disease). These latter 15 cases tended to be from the early years of the registry and the material available for review was limited and/or of poor quality such that more precise diagnoses could not be allocated. It was considered safer to exclude these cases from the separate analyses. However, they were included in the analyses of all lymphomas. There was little or no evidence for space–time clustering in any group based on place of birth/time of birth, place of birth/time of first symptom, nor on place of birth/time of diagnosis (Table 1). Furthermore, there was no evidence for space–time clustering in any group based on place of diagnosis/time of birth, place of diagnosis/time of first symptom, nor on place of diagnosis/time of diagnosis (Table 2). There was some evidence for cross-clustering between younger and older cases of HD, based on place of birth/time of birth ($O=14$, $E=6.3$, $S=122.6\%$, $P=0.005$, using the geographical distance version of the Knox test; $O=9$, $E=4.2$, $S=112.5\%$, $P=0.03$, using the NN threshold version of the Knox test; $I=26.17$, $P=0.02$, using the geographical distance version of the K-function method; and $I=28.95$, $P=0.04$, using the NN threshold version of the K-function method). There was little or no evidence of space–time clustering for any of the other additional analyses (data not shown).

Table 1
Space–time clustering tests for lymphomas, in children aged 0–14 years from North West England, based on place of birth, and diagnosed during the period 1954–2001

	Knox test (observed space–time pairs,[a] expected space–time pairs, strength,[b] P value[c])		K-function analysis,[f] (observed integral,[g] P value[h])	
	Geographical distance[d]	NN threshold[e]	Geographical distance[i]	NN threshold[j]
a. Analyses by time of birth				
Hodgkin's disease	$O=20$; $E=13.9$	$O=11$; $E=8.9$	$I=11.37$	$I=3.75$
	$S=44.1\%$ ($P=0.07$)	$S=24.2\%$ ($P=0.28$)	($P=0.16$)	($P=0.38$)
Non-Hodgkin's lymphoma	$O=33$; $E=37.3$	$O=11$; $E=13.5$	$I=-4.42$	$I=-0.30$
	$S=-11.4\%$ ($P=0.72$)	$S=-18.4\%$ ($P=0.69$)	($P=0.58$)	($P=0.46$)
Hodgkin's disease and	$O=89$; $E=90.2$	$O=37$; $E=41.9$	$I=-8.55$	$I=-11.55$
Non-Hodgkin's lymphoma	$S=-1.3\%$ ($P=0.52$)	$S=-11.6\%$ ($P=0.75$)	($P=0.72$)	($P=0.76$)
All lymphomas	$O=93$; $E=93.8$	$O=39$; $E=44.5$	$I=-6.31$	$I=-10.19$
	$S=-0.8\%$ ($P=0.50$)	$S=-12.3\%$ ($P=0.77$)	($P=0.66$)	($P=0.72$)
b. Analyses by time of first symptom				
Hodgkin's disease	$O=8$; $E=10.5$	$O=13$; $E=14.0$	$I=-5.75$	$I=-9.91$
	$S=-23.4\%$ ($P=0.72$)	$S=-7.4\%$ ($P=0.54$)	($P=0.65$)	($P=0.71$)
Non-Hodgkin's lymphoma	$O=27$; $E=27.2$	$O=20$; $E=13.2$	$I=4.10$	$I=13.75$
	$S=-0.9\%$ ($P=0.47$)	$S=51.3\%$ ($P=0.05$)	($P=0.39$)	($P=0.20$)
Hodgkin's disease and	$O=75$; $E=72.5$	$O=74$; $E=59.8$	$I=8.34$	$I=18.31$
Non-Hodgkin's lymphoma	$S=3.4\%$ ($P=0.40$)	$S=23.8\%$ ($P=0.04$)	($P=0.26$)	($P=0.14$)
All lymphomas	$O=81$; $E=76.0$	$O=80$; $E=62.9$	$I=9.63$	$I=18.33$
	$S=6.6\%$ ($P=0.30$)	$S=27.2\%$ ($P=0.02$)	($P=0.22$)	($P=0.14$)
c. Analyses by time of diagnosis				
Hodgkin's disease	$O=9$; $E=12.9$	$O=5$; $E=8.3$	$I=-15.21$	$I=-15.14$
	$S=-30.3\%$ ($P=0.83$)	$S=-39.3\%$ ($P=0.83$)	($P=0.90$)	($P=0.84$)
Non-Hodgkin's lymphoma	$O=37$; $E=40.3$	$O=16$; $E=14.6$	$I=-3.67$	$I=4.50$
	$S=-8.3\%$ ($P=0.66$)	$S=9.6\%$ ($P=0.39$)	($P=0.62$)	($P=0.37$)
Hodgkin's disease and	$O=94$; $E=96.1$	$O=50$; $E=44.6$	$I=-3.02$	$I=11.72$
Non-Hodgkin's lymphoma	$S=-2.2\%$ ($P=0.56$)	$S=12.0\%$ ($P=0.23$)	($P=0.59$)	($P=0.22$)
All lymphomas	$O=99$; $E=100.1$	$O=53$; $E=47.5$	$I=-1.28$	$I=8.95$
	$S=-1.1\%$ ($P=0.52$)	$S=11.7\%$ ($P=0.23$)	($P=0.51$)	($P=0.28$)

[a] Cases are close in time if dates of diagnosis/birth/first symptom differ by less than 1 year.
[b] Strength (S) = (Observed − Expected)/Expected × 100 counts of pairs which are close in time and space.
[c] One-sided P value derived from the Poisson distribution.
[d] When using geographical distance cases are close in space if their locations are <5 km apart.
[e] When using nearest neighbour (NN) thresholds, cases are close in space if the locations of one (or both) was nearer than the other's 11th NN in the total data-set, for both diagnosis and birth.
[f] Cases are close in time if dates differ by $<t$ where t is in the range 1–15 months.
[g] $I=\int R(s,t)\,\mathrm{d}s\,\mathrm{d}t$, where $R(s,t)=[K(s,t)-K1(s)K2(t)]/\sqrt{[K1(s)K2(t)]}$. $K(s,t)$ = proportion of pairs whose distance apart is $\leqslant t$ in time and $\leqslant s$ in space, $K1(s)$ = proportion of pairs whose distance apart is $\leqslant s$, and $K2(t)$ = proportion of pairs whose distance apart is $\leqslant t$.
[h] P value obtained by simulation (999 runs) with dates randomly re-allocated to the cases in the analysis.
[i] Cases are close in space if distances between their locations differ by $<s$ where s is in the range 0.5–7.5 km.
[j] Cases are close in space if either is within the distance to the Nth nearest neighbour of the other (in the total data-set), where N is in the range 4–18 for both diagnosis and birth.

4. Discussion

The analyses presented here have been performed using rigorous statistical methods on high quality incidence data. However, there will be some opportunity for dilution of the population-base of birth addresses, due to migration out of the study region. Such migration is likely to exhibit a random, rather than a systematic pattern and is highly unlikely to introduce artefactual clustering. Thus, the results based on place of birth are likely to be somewhat conservative. Previous studies from the MCTR have identified space–time clustering amongst cases of leukaemia, central nervous system tumours, soft tissue sarcomas and Wilms' tumours [14,16,18,19]. There have been two previous studies of space–time clustering in HD from the Manchester area over limited periods of time only [8,9]. However, these studies considered all ages and did not apply the K-function methodology [17]. Therefore, the previously reported conclusions concerning HD should be regarded with caution. In the present study, there was little or no evidence of space–time clustering for the main analyses, and little or no evidence of space–time clustering for the additional analyses. The finding of cross-clustering between 'younger' and 'older' cases of HD, in the absence of clustering amongst pairs of

Table 2
Space–time clustering tests for lymphomas, in children aged 0–14 years from North West England, based on place of diagnosis, and diagnosed during the period 1954–2001

	Knox test (observed space–time pairs[a], expected space–time pairs, strength[b], P value[c]		K-function analysis[f], (observed integral[g], P value[h])	
	Geographical distance[d]	NN Threshold[e]	Geographical distance[i]	NN Threshold[j]
a. Analyses by time of birth				
Hodgkin's disease	$O=17$; $E=14.2$	$O=9$; $E=9.3$	$I=2.44$	$I=11.00$
	$S=19.5\%$ ($P=0.26$)	$S=-2.7\%$ ($P=0.45$)	($P=0.38$)	($P=0.22$)
Non-Hodgkin's lymphoma	$O=28$; $E=30.7$	$O=11$; $E=13.3$	$I=-11.61$	$I=-5.66$
	$S=-8.8\%$ ($P=0.64$)	$S=-17.1\%$ ($P=0.67$)	($P=0.81$)	($P=0.61$)
Hodgkin's disease and	$O=82$; $E=82.4$	$O=35$; $E=42.7$	$I=-15.31$	$I=-14.83$
Non-Hodgkin's lymphoma	$S=-0.5\%$ ($P=0.49$)	$S=-18.0\%$ ($P=0.86$)	($P=0.87$)	($P=0.82$)
All lymphomas	$O=87$; $E=86.4$	$O=39$; $E=45.6$	$I=-9.92$	$I=-8.70$
	$S=0.7\%$ ($P=0.49$)	$S=-14.5\%$ ($P=0.82$)	($P=0.75$)	($P=0.72$)
b. Analyses by time of first symptom				
Hodgkin's disease	$O=17$; $E=20.6$	$O=9$; $E=13.0$	$I=-4.61$	$I=-10.33$
	$S=-17.3\%$ ($P=0.74$)	$S=-30.7\%$ ($P=0.83$)	($P=0.63$)	($P=0.73$)
Non-Hodgkin's lymphoma	$O=41$; $E=40.5$	$O=23$; $E=16.5$	$I=5.20$	$I=21.22$
	$S=1.2\%$ ($P=0.49$)	$S=39.1\%$ ($P=0.08$)	($P=0.35$)	($P=0.11$)
Hodgkin's disease and	$O=121$; $E=114.2$	$O=66$; $E=58.0$	$I=11.87$	$I=16.94$
Non-Hodgkin's lymphoma	$S=6.0\%$ ($P=0.27$)	$S=13.8\%$ ($P=0.16$)	($P=0.18$)	($P=0.14$)
All lymphomas	$O=128$; $E=120.4$	$O=72$; $E=62.4$	$I=10.92$	$I=15.17$
	$S=6.3\%$ ($P=0.26$)	$S=15.4\%$ ($P=0.13$)	($P=0.19$)	($P=0.15$)
c. Analyses by time of diagnosis				
Hodgkin's disease	$O=21$; $E=25.2$	$O=14$; $E=15.5$	$I=-6.89$	$I=-7.39$
	$S=-16.7\%$ ($P=0.76$)	$S=-9.4\%$ ($P=0.58$)	($P=0.69$)	($P=0.67$)
Non-Hodgkin's lymphoma	$O=47$; $E=50.7$	$O=25$; $E=21.4$	$I=4.74$	$I=18.74$
	$S=-7.2\%$ ($P=0.66$)	$S=16.9\%$ ($P=0.24$)	($P=0.35$)	($P=0.13$)
Hodgkin's disease and	$O=139$; $E=142.5$	$O=80$; $E=72.8$	$I=7.86$	$I=21.11$
Non-Hodgkin's lymphoma	$S=-2.4\%$ ($P=0.59$)	$S=9.9\%$ ($P=0.21$)	($P=0.29$)	($P=0.12$)
All lymphomas	$O=149$; $E=151.4$	$O=85$; $E=78.6$	$I=7.72$	$I=18.20$
	$S=-1.6\%$ ($P=0.56$)	$S=8.2\%$ ($P=0.25$)	($P=0.27$)	($P=0.14$)

[a–j] See Table 1 footnotes.

'younger' or pairs of 'older' cases, by themselves, is difficult to explain and may be a chance finding. Thus, these data do not provide any positive support for the involvement of environmental factors, such as infections in the aetiology of the childhood lymphomas.

However, we have previously reported seasonal variation in the onset of HD [1], which does lend support to a role for environmental factors, particularly infections, but may also include dietary factors. For HD, the lack of space–time clustering together with seasonal variation in onset may be consistent with a hypothesis whereby HD occurs as a very rare response to an ubiquitous environmental factor, such as an infection. Such a mechanism may account for the observed cross-clustering between 'younger' and 'older' HD cases if the latency between exposure and onset of symptoms is very variable. An example of an ubiquitous infection is EBV and we note that the presence of EBV viral DNA in the malignant cells of a substantial proportion of HD cases is usually interpreted as indicative of its aetiological role. However, variable latency is at odds with the interpretation of seasonal variation, but these findings may reflect different sub-groups of cases. Further recent work from the MCTR has shown that markedly higher incidence of the mixed cellularity subtype of HD was associated with greater levels of unemployment and household overcrowding [20]. Such a finding is indicative of a direct role for an environmental exposure, which is most likely to be an infection.

In conclusion, the overall pattern for HD is complex. No space–time clustering was found, but whilst we cannot exclude other environmental factors, we consider that infections remain the most likely aetiological agent. There was also no space–time clustering for NHL and no clear aetiological factors are indicated for childhood NHL.

Acknowledgements

The Manchester Children's Tumour Registry is supported by Cancer Research UK. Jillian M. Birch is Cancer Research UK Professorial Fellow in Paediatric Oncology and Osborn B. Eden is Cancer Research UK Professor of

Paediatric Oncology at the University of Manchester. We thank Mr. D.P. Cairns, Mrs. E.A. Dale, Mrs. D.A. Elliott, Mrs. J.F. Hogg and Mr. C. Nikolaisen for all their hard work on data processing and verification.

References

1. Westerbeek RMC, Blair V, Eden OB, *et al*. Seasonal variations in the onset of childhood leukaemia and lymphoma. *Br J Cancer* 1998, **78**, 119–124.
2. Armstrong AA, Alexander FE, Paes RP, *et al*. Association of Epstein-Barr virus with pediatric Hodgkin's disease. *Am J Pathol* 1993, **142**, 1683–1688.
3. Day NE, Smith PG, Lachet B. The latent period of Burkitt's lymphoma: the evidence from epidemiological clustering. In Lenoir G, O'Conor G, Olweney CLM, eds. *Burkitt's Lymphoma: a Human Cancer Model, IARC Scientific Publication No. 60*. Lyon, IARC, 1985, 187–196.
4. Armstrong AA, Alexander FE, Cartwright RA, *et al*. Epstein-Barr virus and Hodgkin's disease: further evidence for the three disease hypothesis. *Leukemia* 1998, **12**, 1272–1276.
5. Gilman EA, McNally RJQ, Cartwright RA. Space-time clustering of Hodgkin's disease in parts of the UK, 1984–1993. *Leuk Lymphoma* 1999, **36**, 85–100.
6. Alexander FE. Clustering and Hodgkin's disease. *Br J Cancer* 1990, **62**, 708–711.
7. Greenberg RS, Grufferman S, Cole P. An evaluation of space-time clustering in Hodgkin's disease. *J Chron Dis* 1983, **36**, 257–262.
8. Mangoud A, Hillier VF, Leck I, Thomas RW. Space-time interaction in Hodgkin's disease in Greater Manchester. *J Epidemiol Community Health* 1985, **39**, 58–62.
9. Alderson MR, Nayak R. A study of space-time clustering of Hodgkin's disease in the Manchester region. *Br J Prev Soc Med* 1971, **25**, 168–173.
10. Kryscio RJ, Myers MH, Prusiner ST, Heise HW, Christine BW. The space-time distribution of Hodgkin's disease in Connecticut, 1940–69. *J Natl Cancer Inst* 1973, **50**, 1107–1110.
11. Alexander FE, Williams J, McKinney PA, Ricketts TJ, Cartwright RA. A specialist leukaemia/lymphoma registry in the UK. Part 2: clustering of Hodgkin's disease. *Br J Cancer* 1989, **60**, 948–952.
12. Urquhart J, Black R, Buist E. Exploring small area methods. In Elliott P, ed. *Methodology of Enquiries into Disease Clustering*. London, SAHSU, 1989, 41–49.
13. Glaser SL. Spatial clustering of Hodgkin's disease in the San Francisco Bay area. *Am J Epidemiol* 1990, **132**(Suppl.), S167–S177.
14. Birch JM, Alexander FE, Blair V, Eden OB, Taylor GM, McNally RJQ. Space-time clustering patterns in childhood leukaemia support a role for infection. *Br J Cancer* 2000, **82**, 1571–1576.
15. Knox EG. The detection of space-time interactions. *Applied Stats* 1964, **13**, 25–29.
16. McNally RJQ, Cairns DP, Eden OB, *et al*. An infectious aetiology for childhood brain tumours? Evidence from space-time clustering and seasonality analyses. *Br J Cancer* 2002, **86**, 1070–1077.
17. Diggle PJ, Chetwynd AG, Haggvist R, Morris SE. Second-order analysis of space-time clustering. *Stat Methods Med Res* 1995, **4**, 124–136.
18. McNally RJQ, Alexander FE, Birch JM. Space-time clustering of childhood acute lymphoblastic leukaemia by immunophenotype. *Br J Cancer* 2002, **87**, 1267–1274.
19. McNally RJQ, Kelsey AM, Eden OB, Alexander FE, Cairns DP, Birch JM. Space-time clustering patterns in childhood solid tumours other than central nervous system tumours. *Int J Cancer* 2003, **103**, 253–258.
20. McNally RJQ, Alston RD, Cairns DP, Eden OB, Birch JM. Geographical and ecological analyses of childhood acute leukaemias and lymphomas in North West England. *Br J Haematol* 2003, **123**, 60–65.