

## Short Communication

# Space-time clustering analyses of childhood acute lymphoblastic leukaemia by immunophenotype

RJQ McNally<sup>\*,1</sup>, FE Alexander<sup>2</sup> and JM Birch<sup>1</sup>

<sup>1</sup>Cancer Research UK Paediatric & Familial Cancer Research Group, Standcliffe, Central Manchester and Manchester Children's University Hospitals NHS Trust, Hospital Road, Manchester M27 4HA, UK; <sup>2</sup>Department of Public Health Sciences, The University of Edinburgh Medical School, Teviot Place, Edinburgh EH8 9AG, UK

Space-time clustering analyses of acute lymphoblastic leukaemia in children, by immunophenotype, were carried out using a population-based registry. Significant evidence was found of space-time clustering for cases of the precursor B-cell sub-type, in the childhood peak, based on time and location at birth.

*British Journal of Cancer* (2002) **87**, 513–515. doi:10.1038/sj.bjc.6600498 www.bjcancer.com

© 2002 Cancer Research UK

**Keywords:** acute lymphoblastic leukaemia; children; aetiology; infection; immunophenotype; space-time clustering

We have previously reported significant evidence of space-time clustering in childhood leukaemia, diagnosed during 1954–1985, and based on place of birth and time of diagnosis, particularly among cases of acute lymphoblastic leukaemia [ALL] aged 0–4 years (Birch *et al*, 2000). The finding of space-time clustering is consistent with the involvement of infectious agents in the aetiology. There have been several hypotheses, supported by indirect epidemiological evidence, proposing such a role for infections. Kinlen (1988) demonstrated that situations of unusual rural population mixing, which promoted contacts between susceptible and infected individuals, are associated with excesses of ALL, whilst Greaves (1988) has proposed that delayed exposure to common infections would lead to a greater chance of the occurrence of precursor B-cell (common) ALL. Smith (1997) suggested that the relevant infectious exposures occur *in utero*. Our previous findings were consistent with Greaves' model and we hypothesised that the clustering cases would mainly involve precursor B-cell ALL. It has not been possible previously to test this hypothesis directly, because no population-based data on ALL, by immunophenotype, were available over a sufficiently long time period. We have now been able to perform such an analysis, using cases of ALL included in the Manchester Children's Tumour Registry (MCTR), during 1980–2001, for which full data on immunophenotype are available. The aim of this new study was to test the prediction that space-time clustering is concentrated among cases of the precursor B-cell sub-type.

## PATIENTS AND METHODS

All cases, diagnosed with ALL between the 1st January 1980 and 31st December 2001, and registered by the MCTR were analysed. 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 kilometre. The

following diagnostic groups were specified a priori for analysis: (1) all precursor B-cell (common) ALL; (2) precursor B-cell ALL, aged 18–54 months; (3) all non-precursor B-cell ALL; (4) non-precursor B-cell ALL, aged 18–54 months; (5) total ALL.

There are four possible space-time interactions between: (1) times and places of diagnosis; (2) times and places of birth; (3) time of diagnosis and place of birth; and (4) time of birth and place of diagnosis. The interpretation of these interactions will depend on the extent of migration between birth and diagnosis among cases (Birch *et al*, 2000).

Knox (1964) 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 ( $N=40$  for birth locations and 41 for diagnosis locations).

As previously discussed, two problems are apparent with the Knox test: boundary problems and the arbitrariness of the thresholds chosen (McNally *et al*, 2002). To overcome these a second order procedure based on K-functions (Diggle *et al*, 1995) is used, employing both geographical distance and nearest neighbour (NN) approaches as described above.

Data were also analysed by examining clustering pairs which contained at least one male case ('male: any') and clustering pairs which contained at least one female case ('female: any'). Using internal methods, addresses were classified as being located in a more densely populated area, or a less densely populated area (McNally *et al*, 2002). Analysis by population density was undertaken by considering clustering pairs which contained at least one case from a more densely populated area ('more densely populated: any') and clustering pairs which contained at least one case from a less densely populated area ('less densely populated: any'). The

\*Correspondence: Dr Richard JQ McNally;

E-mail: richard.mcnelly@man.ac.uk

Received 9 April 2002; revised 27 May 2002; accepted 29 May 2002

methods are given in more detail in Birch *et al* (2000) and McNally *et al* (2002).

## RESULTS

The study included 512 cases of ALL, comprising 395 precursor B-cell ALL, 63 T-cell, 36 Null-ALL, 9 B-cell, 1 non-B, non-T cell and eight unclassified. Based on place and time of birth the only statistically significant evidence of space-time clustering was for cases of precursor B-cell ALL, aged 18–54 months ( $P=0.05$ , using the NN threshold version of the K-function method, Table 1a). Examination of cases of precursor B-cell ALL, aged 18–54 months, by gender and population density (Table 1b) shows that the space-time clustering was confined to clustering pairs of cases which contained at least one female case ( $P<0.1$ , using all four methods). There was no association with population density level. There was no evidence for space-time clustering in any group based on time and place of diagnosis, time of diagnosis/place of birth nor on time of birth/place of diagnosis.

## DISCUSSION

Our previous study (Birch *et al*, 2000) showed striking space-time clustering among a proportion of cases of ALL aged 0–4 years, which was based on time of diagnosis/place of birth. In contrast, the present study found limited evidence for clustering based on time of birth/place of birth only. However, even though the results were statistically less significant than in the previous larger study, the clustering for the 18–54 month age group (the childhood peak) of precursor B-cell ALL clearly stood out. This is the first study to demonstrate that space-time clustering for cases of ALL in the childhood incidence peak is specifically due to the precursor B-cell sub-type. The results are consistent with a role for infections, but with exposure occurring pre-natally, or around the time of birth.

The clustering was only apparent among space-time pairs that involved at least one female case. This differs from our previous study, which found an excess of male cases over females involved in space-time pairs. Interestingly, our recent study of incidence

**Table 1** Space-time clustering tests for ALL, in children aged 0–14 years from North West England, based on time and place of birth, and diagnosed during the period 1980–2001

	KNOX TEST [Observed space-time pairs <sup>a</sup> , expected space-time pairs, strength <sup>b</sup> , P-value <sup>c</sup> ]		K-FUNCTION ANALYSIS <sup>f</sup> [Observed integral <sup>g</sup> , P-value <sup>h</sup> ]	
	Geographical Distance <sup>d</sup>	NN Threshold <sup>e</sup>	Geographical Distance <sup>i</sup>	NN Threshold <sup>j</sup>
<i>a. Analyses by Disease group and Age</i>				
Precursor B-cell ALL (aged 0–14 years)	O=212; E=208.0 S=2.0% (P=0.40)	O=106; E=107.4 S=–1.3% (P=0.53)	I=–1.08 (P=0.52)	I=–0.07 (P=0.50)
Precursor B-cell ALL (aged 18–54 months)	O=77; E=70.8 S=8.7% (P=0.25)	O=45; E=35.3 S=27.6% (P=0.06)	I=18.61 (P=0.09)	I=29.76 (P=0.05)
Non-precursor B-cell ALL (aged 0–14 years)	O=17; E=14.2 S=19.8% (P=0.26)	O=6; E=7.5 S=–19.8% (P=0.62)	I=1.25 (P=0.44)	I=–13.02 (P=0.79)
Non-precursor B-cell ALL (aged 18–54 months)	O=1; E=0.9 S=13.3% (P=0.59)	O=1; E=0.8 S=27.5% (P=0.54)	I=–1.78 (P=0.66)	I=0.04 (P=0.42)
All ALL (aged 0–14 years)	O=325; E=332.6 S=–2.3% (P=0.65)	O=167; E=175.8 S=–5.0% (P=0.73)	I=–16.53 (P=0.89)	I=–14.03 (P=0.76)
<i>b. Analyses of precursor B-cell ALL aged 18–54 months by gender and population density</i>				
'Male: any' clustering pairs	O=58; E=59.0 S=–1.7% (P=0.52)	O=32; E=27.8 S=15.3% (P=0.23)	I=5.48 (P=0.41)	I=13.50 (P=0.25)
'Female: any' clustering pairs	O=53; E=43.9 S=20.9% (P=0.098)	O=35; E=23.6 S=48.6% (P=0.016)	I=20.96 (P=0.067)	I=36.83 (P=0.028)
'More densely populated: any' clustering pairs	O=58; E=59.1 S=–1.9% (P=0.52)	O=25; E=21.4 S=17.1% (P=0.24)	I=6.45 (P=0.32)	I=12.69 (P=0.24)
'Less densely populated: any' clustering pairs	O=27; E=24.1 S=12.0% (P=0.30)	O=24; E=21.3 S=12.9% (P=0.30)	I=11.43 (P=0.16)	I=18.28 (P=0.13)

<sup>a</sup>Cases are close in time if dates of diagnosis/birth differ by less than 1 year. <sup>b</sup>Strength (S)={[(Observed–Expected)/Expected]} × 100 counts of pairs which are close in time and space. <sup>c</sup>I-sided P-value derived from the Poisson distribution. <sup>d</sup>When using geographical distance cases are close in space if their locations are <5 km apart. <sup>e</sup>When using nearest neighbour (NN) thresholds cases are close in space if the locations of one (or both) was nearer than the other's 41<sup>st</sup> NN in the total data set, for diagnosis, or 40<sup>th</sup> NN for birth. <sup>f</sup>Cases are close in time if dates differ by <t where t is in the range 1–15 months. <sup>g</sup> $I = \int R(s,t) ds dt$ , 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  $\leq t$  in time and  $\leq s$  in space,  $K1(s)$ =proportion of pairs whose distance apart is  $\leq s$ , and  $K2(t)$ =proportion of pairs whose distance apart is  $\leq t$ . <sup>h</sup>P-value obtained by simulation (999 runs) with dates of birth randomly re-allocated to the cases in the analysis. <sup>i</sup>Cases are close in space if distances between their locations differ by <s where s is in the range 0.5–7.5 km. <sup>j</sup>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 34–48, for diagnosis, or 33–47 for birth.

trends in ALL (McNally *et al*, 2000), showed that the observed increase was attributable to the precursor B-cell sub-type, and indicated a faster recent rate of increase amongst females.

The present data appear to be more consistent with Smith's hypothesis (Smith, 1997) that the childhood peak is due to an in-utero exposure to infection, rather than with Greaves' hypothesis of delayed exposure to common infections (Greaves, 1988). The data are also consistent with Kinlen's hypothesis of population mixing (Kinlen, 1995), particularly in relation to one study which demonstrated a marked excess of leukaemia in children under 1 year of age suggesting an infection during pregnancy (Kinlen and Hudson, 1991). The differences between males and females may be indicative of an immune modulation hypothesis, with females being more susceptible to leukaemogenic events resulting from exposure to infection. Relative to our earlier findings the current evidence would suggest that there is greater exposure to the aetiological agent at an earlier age, or a greater prevalence of the agent,

or a new aetiological agent. Whichever of these applies we would hypothesise that the agent would be an infection. Finally, it should be stressed that a direct comparison between this study and the previous one is not possible, because data on immunophenotype were not available in the previous analysis.

## ACKNOWLEDGEMENTS

The Manchester Children's Tumour Registry is supported by Cancer Research UK. Jillian M Birch is Cancer Research UK Professorial Fellow at the University of Manchester. We thank Mr DP Cairns, Mrs EA Dale, Mrs DA Elliott, Mrs JF Hogg and Mr C Nikolaisen for all their hard work on data processing and verification.

## REFERENCES

- Birch JM, Alexander FE, Blair V, Eden OB, Taylor GM, McNally RJQ (2000) Space-time clustering patterns in childhood leukaemia support a role for infection. *Br J Cancer* **82**: 1571–1576
- Diggle PJ, Chetwynd AG, Haggvist R, Morris SE (1995) Second-order analysis of space-time clustering. *Stat Methods Med Res* **4**: 124–136
- Greaves MF (1988) Speculations on the cause of childhood acute lymphoblastic leukaemia. *Leukemia* **2**: 120–125
- Kinlen L (1988) Evidence for an infective cause of childhood leukaemia: comparison of a Scottish New Town with nuclear reprocessing sites in Britain. *Lancet* **2**: 1323–1327
- Kinlen LJ, Hudson C (1991) Childhood leukaemia and poliomyelitis in relation to military encampments in England and Wales in the period of national military service, 1950–63. *Br Med J* **303**: 1357–1362
- Kinlen LJ (1995) Epidemiological evidence for an infective basis in childhood leukaemia. *Br J Cancer* **71**: 1–5
- Knox EG (1964) The detection of space-time interactions. *Applied Stats* **13**: 25–29
- McNally RJQ, Birch JM, Taylor GM, Eden OB (2000) Incidence of childhood precursor B-cell acute lymphoblastic leukaemia in north-west England. *Lancet* **356**: 485–486
- McNally RJQ, Cairns DP, Eden OB, Alexander FE, Taylor GM, Kelsey AM, Birch JM (2002) An infectious aetiology for childhood brain tumours? Evidence from space-time clustering and seasonality analyses. *Br J Cancer* **86**: 1070–1077
- Smith M (1997) Considerations on a possible viral etiology for B-precursor acute lymphoblastic leukemia of childhood. *J Immunother* **20**: 89–100