Effect of population mixing and socioeconomic status in England and Wales, 1979-85, on lymphoblastic leukaemia in children

C A Stiller, P J Boyle

Abstract

Objectives—To examine the effects of migration, diversity of migrant origins, commuting, and socioeconomic status on the incidence of acute lymphoblastic leukaemia in childhood.

Design—Poisson regression analysis of incidence rates in relation to the variables of interest. Setting—The 403 county districts of England

and Wales during 1979-85.

Subjects-Children aged under 15 years.

Results—There were significant trends in the incidence of lymphoblastic leukaemia at ages 0-4 and 5-9 years with the proportion of children in a district who had recently entered the district. While there was no consistent relation between the proportion of recent incomers in the total population of a district and its incidence rate, the combination of higher migration with greater diversity of origins or distance moved was associated with higher incidence in both age groups. Incidence increased significantly at age 0-4 with the level of employment in a district and at age 5-9 with the proportion of households with access to a car. No significant trends were found with commuting.

Conclusions—The results for level of child migration and diversity of total migration provide evidence of an effect of population mixing on the incidence of childhood leukaemia which is not restricted to areas experiencing the most extreme levels of mixing.

Introduction

In 1988 Kinlen suggested that childhood leukaemia might be a rare response to an as yet unidentified infection or infections.¹ Outbreaks of infection would be most likely when carriers and susceptible persons are brought together by high levels of population mixing—for example, as a result of migration—and in these settings there would be a correspondingly high incidence of leukaemia among children whose immunity is limited. Support has been given to this hypothesis by a series of studies by Kinlen *et al* in which an increased incidence of childhood leukaemia was found in communities experiencing a high degree of population mixing, including rural new towns, towns with large increases in commuting, and areas receiving large numbers of servicemen, migrant construction workers, or wartime evacuees.²

Rapid population growth need not in itself give rise to an abnormal level of mixing. In their study of new towns throughout Britain, Kinlen *et al* differentiated between overspill new towns attracting migrants from a relatively limited set of areas and rural new towns where the origins of the incomers tended to be geographically more diverse.³ Only the rural new towns, where the introduction of a wider range of different types or strains of virus might be expected, had a raised incidence of leukaemia. That study did not include areas of rapid growth which were not designated new towns, however, nor did it differentiate *quantitatively* between areas of rapid population growth that attract most of their incomers from a limited set of origins and those that attract them from diverse origins.

Several studies have indicated that there is a higher risk of childhood leukaemia in general, or of acute lymphoblastic leukaemia in particular, associated with higher socioeconomic status, whether measured at the level of the individual or of area of residence.⁴⁷. Among rural areas supplying relatively large numbers of North Sea oil workers or close to very large construction sites whose workers were housed in special camps and thus subject to a lot of population mixing the excess of childhood leukaemia was most obvious in areas of high social class.⁸

We tested the population mixing hypothesis quantitatively by analysing variations in the incidence of acute lymphoblastic leukaemia in children aged under 15 years at county district level throughout England and Wales according to measures of migration and commuting. Simultaneously, we analysed the effects of various indicators of socioeconomic status.

Subjects and methods

LEUKAEMIA CASES

Registrations for acute lymphoblastic leukaemia in children below the age of 15 years in England and Wales during 1979-85 were taken from the population based National Registry of Childhood Tumours.¹⁰ Cases were ascertained from regional cancer registries, the national cancer registry at the Office of Population Censuses and Surveys, regional children's tumour registries, the register of patients treated by members of the United Kingdom Children's Cancer Study Group, entries to the Medical Research Council's leukaemia trials, and death certificates. The diagnoses and dates of diagnosis were verified and, if necessary, amended with information from hospitals, general practitioners, and the organisers of clinical trials.

There were 2035 cases registered during the study period, with histological or haematological confirmation of diagnosis available for all but 18 (1%). During 1979-83, over 99% of incident cases of childhood leukaemia were thought to have been registered,¹¹ and there is no reason to believe that ascertainment deteriorated from this level during 1984-5. The address at diagnosis was postcoded for each case. Census enumeration districts and hence county districts of residence were derived from the postcodes.

POPULATION DATA

Mid-year estimates of the population in age groups 0-4, 5-9, and 10-14 years in each county district and calendar year were obtained from the Office of Population Censuses and Surveys.

MEASURES OF POPULATION MIXING AND SOCIOECONOMIC STATUS

We calculated for each district the proportions of all residents and of children aged under 15 who had been resident outside the district one year previously from the small area statistics from the 1981 census.

We also distinguished quantitively between places where population increases have been large but the incomers have tended to come from a limited set of origins and those where the incomers came from a diverse set of origins. A value was assigned to each of the 403 local authority districts on the basis of the diversity of origins that the incomers to that district came from by using data extracted from the special migration statistics

Childhood Cancer Research Group, , University of Oxford, Oxford OX2 6HJ C A Stiller, epidemiologist

School of Geography, University of Leeds, Leeds LS2 9JT P J Boyle, *lecturer*

Correspondence to: Mr Stiller.

BMJ 1996;313:1297-300

from the 1981 census. This statistic, based on a form of Shannon's entropy (derived from information theory),¹² measures the extent to which some total population (in our case, the total incomers to a destination district) is distributed among its component parts (the origin districts). Thus, for each destination we can measure the Shannon's entropy (H_i) as:

$$H_j = \sum_{ij} \log(1/p_{ij})$$

i=1

where p_{ii} is the number of migrants coming from each origin area i and terminating in area j as a proportion of the total incomers to area j. This value lies on a scale which ranges from zero to log (n-1). A low value indicates that the incomers to area j came predominantly from a limited number of places whereas a high value indicates that they tended to come from a more diverse range of areas. As it is plausible that the importance of population mixing would also increase if the incomers tended to come from distant areas, where viruses are more likely to differ from those in the destination areas, we additionally calculated the natural logarithm of the mean distance between districts for all incomers to each district. No information was available, however, on the previous district of residence of child migrants in the 1981 census.

From the special workplace statistics from the 1981 census we calculated the total proportion of the economically active population in each district who were commuting to another district and the proportion commuting to another district by public transport. We also calculated the proportion in the armed forces for each district as they form a particularly mobile group and a previous study had found an association between mortality from childhood leukaemia and the presence of large military camps.¹³

The same three measures of socioeconomic status of districts derived from the small area statistics were used as for the previous national study of incidence during 1969-83.6 These were proportion of economically active men who were working ("employment"); proportion of households with car ("car owner"); and proportion of households that were owner occupied ("housing tenure"). Areas where these proportions were high were taken to be areas of high socioeconomic status. As large scale sales of local authority housing did not begin until 1981, this will have had only a small effect on the housing tenure variable. Again, as in the earlier study, a summary "socioeconomic score" was defined as the average of the standardised deviates for each of the three variables. The centiles of the three variables and the values of the socioeconomic score in the present study relate to England and Wales alone and therefore differ from those in the previous study which also included Scotland.

STATISTICAL METHODS

As in the previous national study of socioeconomic factors, possible relations between incidence rates and the variables listed above were investigated by Poisson regression methods by using the GLIM statistical package. Throughout the analyses, a multiplicative model was used—that is, $\log(m/p) = b_0 + b_1x_1 + b_2x_2 + ... + b_kx_k$ —where m is the number of cases in a specified age group and area, p is the population at risk, and the $x_i(i = 1, 2...k)$ are the values of k variables that may affect the incidence rate, with the coefficients $b_0, b_1, ..., b_k$ being estimated as part of the analysis.

Each variable was analysed in two ways. The first was to consider it as a continuous variable with the aim of investigating trends in incidence with that variable. The second was to consider it as a categorical variable, with districts classified as those below the 10th centile, between the 10th and 90th centiles, and above the 90th centile with respect to the variable in question. The aim here was to investigate the possibility that while variations in a factor within a "normal" range might have no discernible effect on incidence, threshold effects may be of importance.¹⁴

The significance of the improved goodness of fit obtained by adding variables to the Poisson regression model was tested by the reduction in scaled deviance, which approximately follows a χ^2 distribution.

Results

The national annual incidence rates per million for acute lymphoblastic leukaemia were 48.7 (SE 1.5) at age 0-4, 26.6 (1.1) at age 5-9, and 15.7 (0.8) at age 10-14. When the Poisson model was fitted for uniform incidence across all districts, there was significant variation between districts at ages 0-4 (deviance 472.9 on 402 df; P<0.01) and 5-9 (465.7; P<0.02) but not at age 10-14 (404.5; P>0.4).

UNIVARIATE ANALYSIS

Table 1 shows the numbers of cases and incidence rates for the three age groups in the bottom 40, middle 323, and top 40 county districts according to the demographic and socioeconomic variables listed above. Table 2 gives the results of fitting these as continuous and categorical variables, expressed as the reduction in scaled deviance attributable to the variable in question compared with a uniform national incidence.

For age 0-4 there were significant increasing trends in incidence with the proportions of recent incomers in the total population and child population and with the proportion of economically active men in employment. Child migration and employment were also significant when fitted as grouped variables. For age 5-9 there were significant trends with child migration (but not with

Table 1—Numbers of cases and incidence rates per million (SE) in bottom 40, middle 323, and top 40 of country districts classified by demographic and socioeconomic variables

		Age 0-4			Age 5-9			Age 10-14	
Detail	Bottom	Middle	Тор	Bottom	Middle	Тор	Bottom	Middle	Тор
Total migration	191; 43.4 (3.1)	757; 49.5 (1.8)	91; 55.1 (5.8)	112; 24.7 (2.3)	421; 26.6 (1.3)	50; 31.2 (4.4)	82; 15.0 (1.7)	303; 16.0 (0.9)	28; 15.1 (2.9)
Child migration	183; 41.3 (3.1)	784; 50.3 (1.8)	72; 54.1 (6.4)	105; 23.1 (2.3)	426; 26.6 (1.3)	52; 36.9 (5.1)	89; 16.2 (1.7)	306; 16.0 (0.9)	18; 11.0 (2.6)
Total commuting	119; 44.1 (4.0)	797; 48.7 (1.7)	123; 53.9 (4.9)	80; 28.8 (3.2)	438; 26.0 (1.2)	65; 27.9 (3.5)	57: 17.0 (2.3)	311: 15.5 (0.9)	45; 16.0 (2.4)
Public transport		,		, , ,	, , ,	, , ,	, , , ,		., ,
commuting	79; 58.0 (6.5)	799; 47.3 (1.7)	161; 52.0 (4.1)	49; 34.0 (4.9)	467; 26.6 (1.2)	67; 22.8 (2.8)	29; 16.8 (3.1)	331; 15.8 (0.9)	53; 15.2 (2.1)
Employment in									
forces	125; 55.0 (4.9)	831; 47.6 (1.6)	83; 51.5 (5.7)	52; 23.0 (3.2)	476; 26.4 (1.2)	55; 32.7 (4.4)	41; 15.2 (2.4)	345; 16.0 (0.9)	27; 13.8 (2.7)
Employment	138; 40.1 (3.4)	826; 50.6 (1.8)	75; 47.0 (5.4)	86; 24.9 (2.7)	451; 26.8 (1.3)	46; 27.1 (4.0)	72; 17.1 (2.0)	314; 15.7 (0.9)	27; 13.5 (2.6)
With car	162; 47.7 (3.7)	799; 48.9 (1.7)	78; 48.3 (5.4)	74; 22.6 (2.6)	464; 27.4 (1.3)	45; 25.9 (3.9)	63; 15.7 (2.0)	316; 15.7 (0.9)	34; 16.5 (2.8)
Housing tenure	138; 46.2 (3.9)	830; 49.6 (1.7)	71; 43.1 (5.1)	63; 22.1 (2.8)	470; 27.1 (1.2)	50; 28.6 (4.0)	59; 17.0 (2.2)	322; 15.6 (0.9)	32; 15.3 (2.7)
score	161; 44.2 (3.5)	807; 50.0 (1.8)	71; 45.1 (5.3)	81; 23.2 (2.6)	457; 27.3 (1.3)	45; 26.4 (3.9)	72; 16.9 (2.0)	313; 15.7 (0.9)	28; 13.8 (2.6)

Table 2—Reductions in scaled deviance (approximately χ^2) obtained by fitting demographic and socioeconomic variables as grouped variables (for bottom 40, middle 323, and top 40) and as continuous variables

	Age 0-4		Age	5-9	Age 10-14	
Detail	Grouped (2df)	Linear (1df)	Grouped (2df)	Linear (1df)	Grouped (2df)	Linear (1df)
Total migration	4.1	5.4*	1.8	3.4	0.3	0.2
Child migration	6.8*	5.5*	7.1*	7.8**	2.8	0.8
Total commuting	2.4	0.3	0.8	0.6	0.4	0.0
Public transport						
commuting	3.6	0.8	4.3	1.8	0.1	0.1
Employment in forces	2.5	0.4	3.3	4.6*	0.5	1.4
Employment	6.9*	7.6**	0.4	2.6	1.1	0.6
With car	0.1	1.6	2.4	5.0*	0.0	0.4
Housing tenure	1.8	0.0	2.7	1.6	0.4	0.6
Socioeconomic score	2.5	2.6	1.8	3.8	0.8	0.7

*Significant at 5% level.

**Significant at 1% level.

Table 3—Reductions in scaled deviance (approximately χ^2) obtained by fitting proportion of recent incomers, diversity of incomers' origins, and log mean interdistrict distance moved by incomers as continuous variables

	Degrees	Age (years)			
Detail	freedom	0-4	5-9	10-14	
Total migration	1	5.4*	3.4	0.2	
Total migration and diversity	2	6.8*	10.1**	0.9	
Diversity allowing for migration	1	1.4	6.7**	0.7	
Total migration and distance	2	6.0*	9.7**	1.3	
Distance allowing for migration	1	0.6	6.3*	1.1	

P < 0.05 ** P < 0.01.

total migration), with proportion of workers in the armed forces, and with proportion of households with a car. Only child migration was significant as a grouped variable, and there was a particularly high incidence in districts with the highest proportion of children who had moved into the district within the past year. For age 10-14 there were no significant reductions in scaled deviance by fitting any variable in either model, though there was a suggestion of low incidence in districts with a high proportion of child migrants.

MIGRATION DIVERSITY AND DISTANCE

Table 3 shows the results of fitting the relative entropy due to migration and the log mean distance between districts moved by incoming migrants with allowance for the proportion of migrants in the population. At age 0-4 neither of these additional factors produced a significant reduction in deviance, whereas at age 5-9 they both did. The improvement in fit with total migration and entropy or log mean distance (9.7 or 10.1 on 2 df), however, was less than that obtained by fitting child migration alone (7.8 on 1 df). As explained above, no information was available on the districts of origin of the child migrants.

MULTIVARIATE ANALYSIS

For the age group 0-4 no combination of two variables gave significant reductions in scaled deviance both for the first variable and for the second with allowance for the first. For the continuous variable model applied to the age group 5-9, after we fitted the proportion of workers in the armed forces, we obtained a significant further reduction in scaled deviance by fitting the proportion of child migrants (reduction in deviance 4.1; P<0.05); the proportion in the armed forces, however, did not yield any further significant reduction after first fitting the variable for child migration.

Discussion

Since the publication in 1988 of Kinlen's original finding of an increased incidence of childhood leukaemia in Kirkcaldy district, which included Glenrothes New Town, during the period of rapid expansion of Glenrothes,¹ there have been numerous further studies which have supported his hypothesis of an infectious basis for childhood leukaemia.² ⁹ These studies, however, were all concerned with threshold effects-that is, they focused on incidence rates in areas where an unusually high amount of population mixing had occurred. Our analysis is the first to show a significant relation between population mixing and incidence of childhood leukaemia other than at the extremes of variation.

MIGRATION

Two other studies have investigated the occurrence of childhood leukaemia in relation to population increase. Langford analysed mortality from all childhood leukaemia during 1969-73 and found higher rates in areas that experienced a population increase between 1961 and 1971 of more than 50%, presumably largely due to migration.¹⁴ Rodrigues et al analysed national data for 1979-83 in census tracts classified by percentage change in childhood population between 1971 and 1981.7 They found no overall trend but a significantly raised incidence of lymphocytic leukaemia in census tracts of high socioeconomic status within predominantly rural county districts which had undergone an increase of more than 40% in the child population . The relative importance of incoming migrant children and increased numbers of births within census tracts in producing the increase in the childhood population of these census tracts was unknown. In a small case-control study in northern England, children with leukaemia or non-Hodgkin's lymphoma were slightly more likely to have moved house between birth and diagnosis than their controls who were matched for age and area of birth, but the differences were not significant.15

The results relating to migration may be tentatively explained in terms of the age of susceptible members of the population at exposure to the relevant infection. In areas with high levels of inward migration young children would tend to be infected earlier, producing a raised incidence of leukaemia in early childhood. By the age of 10-14 many susceptible children may have already long since been infected, and their incidence of leukaemia is correspondingly low. A similar explanation of immunising doses of infection was suggested for the deficits of leukaemia among older children and adolescents in new towns.3 The raised incidence of leukaemia at age 5-9 in districts with high proportions of child migrants is consistent with the transfer of viruses at school. Interestingly, two studies which used different methods to investigate space-time clustering among children diagnosed during 1966-83 found especially strong clustering at ages 5-9¹⁶ or 4-7,¹⁷ though neither report mentioned the increase in social contacts at school among children of these ages.

COMMUTING

We found no significant relation between commuting and incidence of leukaemia. This may seem inconsistent with the results of an earlier study of Kinlen et al, who found a raised incidence in districts where there had been a large increase in commuting,¹⁸ but districts with high commuting in our study would include those where it had been high for a long time. It is arguable that a change from low to even moderately high levels of commuting could have a greater effect on herd immunity than a sustained high level throughout.

 Incidence of acute lymphoblastic leukaemia in young children increased with the "entropy' attributable to migration into their districts of residence

• Population mixing even at relatively low levels may be important in the aetiology of childhood leukaemia

• The results provide further epidemiological evidence that childhood leukaemia might be a rare response to infection

• Previous studies finding increased incidence in more affluent areas may have been indirectly observing a population mixing effect

SOCIOECONOMIC STATUS

Our results on the relation between socioeconomic status and incidence of acute lymphoblastic leukaemia are similar to those of Draper and colleagues.6 This is not surprising as the two analyses have a large number of cases in common. In multivariate analysis, no two indicators of population mixing and socioeconomic status were of independent significance. The two types of variable are correlated; in particular, areas of low unemployment naturally attract more incoming families and families of higher social class tend to be more mobile. Thus it is impossible to tell the extent to which each of these factors is independently of importance in affecting the incidence of childhood acute lymphoblastic leukaemia. As population mixing is more likely than socioeconomic status to increase rapidly and some of the highest incidence rates for leukaemia have been found in situations of very high increases in levels of mixing, it seems reasonable to conclude that population mixing is of more fundamental importance. Previous studies which have revealed a social class effect on the incidence of childhood leukaemia may in fact have been observing indirectly a population mixing effect.

The present study has provided evidence of an effect of population mixing on the incidence of childhood leukaemia which is not confined to areas experiencing the most extreme levels of mixing. The analysis at district level is nevertheless arguably not sufficiently sensitive and we propose in future to analyse more recent, and hence independent, incidence data at ward level by using flow data from the 1991 census. That census is also the first to provide reliable migrant flow data disaggregated by age group, thus making it possible to investigate specifically the possible effect on incidence of leukaemia of the diversity of origins of incoming children. These further analyses may also enable the effects of population mixing and socioeconomic status to be disentangled more completely.

We are grateful to regional cancer registries, regional children's tumour registries, the UK Children's Cancer Study Group, the Clinical Trial Service Unit, and OPCS for providing notifications of cases of childhood leukaemia. We thank the many people who have worked on the National Registry of Childhood Tumours, particularly Mrs M B Allen, Mr T J Vincent, Dr E M Eatock, Mr M J Loach, Mrs P A Brownbill, Mrs J Williams, Miss A Berry, and Miss A Sabin. We thank Mrs E M Roberts for secretarial assistance. The 1981 census data are Crown Copyright.

Funding: The Childhood Cancer Research Group is supported by the Department of Health and the Scottish Home and Health Department.

Conflict of interest: None.

- 1 Kinlen L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain. Lancet 1988:ii-1323-7
- 2 Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. Br J Cancer 1995;71:1-5. 3 Kinlen LJ, Clarke K, Hudson C. Evidence from population mixing in Brit-
- ish new towns 1946-85 of an infective basis for childhood leukaemia. ancet 1990;**336**:577-82
- 4 McWhirter WR. The relationship of incidence of childhood lymphoblastic
- Ieukaemia to social class. Br J Cancer 1982;46:640-5.
 Alexander FE, Ricketts TJ, McKinney PA, Cartwright RA. Community lifestyle characteristics and risk of acute lymphoblastic leukaemia in children. Lancet 1990;336:1461-5.
- 6 Draper GJ, Vincent TJ, O'Connor CM, Stiller CA. Socio-economic factors and variations in incidence rates between county districts. In: Draper GI, ed. The geographical epidemiology of childhood leukaemia and non-Hodgkin Image of the second s
- tion to childhood leukaemia and non-Hodgkin lymphomas: an analysis based on small area statistics for census tracts. In: Draper GJ, ed. The geographical epidemiology of childhood leukaemia and non-Hodgkin's lymphomas in Great Britain, 1966-83. London: HMSO, 1991: 47-56. (Studies on Medical and Population Subjects No 53.) 8 Kinlen LJ, O'Brien F, Clarke K, Balkwill A, Matthews F. Rural population
- mixing and childhood leukaemia: effects of the North Sea oil industry in Scotland, including the area near Dounreay nuclear site. BMJ 1993;306:743-48
- 9 Kinlen LJ, Dickson M, Stiller CA. Childhood leukaemia and non-Hodgkin's lymphoma near large rural construction sites, with a com-parison with Sellafield nuclear site. *BMJ* 1995;310: 763-8.
- 10 Stiller CA, Allen MB, Eatock EM. Childhood cancer in Britain. The National Registry of Childhood Tumours and incidence rates 1978-87. Eur J Cancer 1995;31A:2028-34.
- 11 Stiller CA, O'Connor CM, Vincent TJ, Draper GJ. The National Registry of Childhood Tumours and the leukaemia/lymphoma data for 1966-83. In: Draper GJ, ed. The geographical epidemiology of childhood leukaemia and non-Hodgkin's lymphomas in Great Britain, 1966-83. London: HMSO, 1991:7-16. (Studies on Medical and Population Subjects No 53.)
- 12 Shannon CE. A mathematical theory of communication. Bell System Technical Journal 1948; 27:379-423
- 13 Kinlen LJ, Hudson C. Childhood leukaemia and poliomyelitis in relation to military encampments in England and Wales in the period of national military service 1950-63. BMY 1991;303:1357-62.
- 14 Langford I. Childhood leukaemia mortality and population change in England and Wales 1969-73. Soc Sci Med 1991;33:435-40. 15 Alexander FE, McKinney PA, Cartwright RA. Migration patterns of
- children with leukaemia and non-Hodgkin's lymphoma in three areas of northern England. *J Public Health Med* 1993;15: 9-15.
- 16 Alexander FE. Space-time clustering of childhood acute lymphoblastic leu-kaemia: indirect evidence for a transmissible agent. Br J Cancer 1992:65:589-92.
- 17 Knox EG, Gilman E. Leukaemia clusters in Great Britain. 1. Space-time interactions. J Epidemiol Community Health 1992;46:566-72. Kinlen LJ, Hudson CM, Stiller CA. Contacts between adults as evidence
- for an infective origin of childhood leukaemia: an explanation for the s near nuclear establishments in West Berkshire? Br J Cancer 1991:64:549-54.

(Accepted 24 September 1996)

Absence of oats toxicity in adult coeliac disease

Usha Srinivasan, Niamh Leonard, Eileen Jones, Donald D Kasarda, Donald G Weir, Cliona O'Farrelly, Conleth Feighery

Correspondence to: Professor Feighery.

BMJ 1996;313:1300-1

Coeliac disease is a gluten-sensitive disorder characterised by malabsorption and a typical histological lesion. Treatment with a strict gluten-free diet results in complete clinical and histological recovery. The conventional gluten-free diet used to treat coeliac disease proscribes oats cereal as well as wheat, barley, and rye.1 However, the issue of oats toxicity has not been conclusively resolved, and the prohibition of this important cereal deprives patients of a valuable source of fibre and nutrients. The aim of this study was to examine the clinical, histological, and immunological responses of adult patients with coeliac disease to challenge with oats.

Patients, methods, and results

Ten adult patients with coeliac disease in clinical and histological remission were recruited from the coeliac