Clustering of Childhood IDDM

Links with age and place of residence

GRAHAM R. LAW, PHD
PATRICIA A. MCKINNEY, PHD
ANTHONY STAINES, MRCPI
RHYS WILLIAMS, FFPHM
MIKE KELLY, PHD
FREDA ALEXANDER, PHD
ESTELLE GILMAN, PHD
H. JONATHAN BODANSKY, MD

OBJECTIVE — To improve understanding of the etiology of IDDM by analyzing spatial and space-time distribution of the incidence in children.

RESEARCH DESIGN AND METHODS — Statistical tests to detect clustering were applied to a population-based register of 1,490 children (aged 0–16 years) with IDDM in Yorkshire, northern England. The Knox test analyzed clustering in space and time, and the Potthoff-Whittinghill test quantified spatial differences in incidence between small-area census units (electoral wards). The Potthoff-Whittinghill test was conditioned for childhood population density and deprivation (Townsend index).

RESULTS — Both tests demonstrated clustering of IDDM in Yorkshire children. Space-time and spatial clustering is strongest in the younger children (0–4 and 5–9 years of age), even after conditioning for known associations. Clustering was more common in the county of Humberside during the years 1982–1985 and in wards of low population density (<0.26 0- to 16-year-old subjects per hectare).

CONCLUSIONS — The study revealed a nonrandom space-time distribution of IDDM in children not accounted for by known covarying demographic factors. The Potthoff-Whittinghill test has not previously been applied to childhood IDDM. The new finding of strong clustering in young children is consistent with early exposure, possibly in utero, to infectious agents or localized environmental sources.

The etiology of IDDM involves both genetic susceptibility and environmental exposures, with the former necessary but not sufficient for disease development. Involvement of an environmental component is demonstrated by a concordance rate of 13–40% for monozygotic twins (1,2). The nongenetic determinants of IDDM remain unknown, but recent studies have shown a nonrandom geographic distribution of IDDM (3,4), suggesting that exposures that vary geographically are important. Rates may vary in relation to levels of deprivation, urban/rural locations, and population density (5–9). A tendency for cases to aggregate, known as clustering, indicates an infectious origin or localized or point sources as the environmental cause of a disease. Interest in clustering of childhood leukemias has prompted extensive development of robust statistical tests applicable to distributions of rare chronic diseases (10). The population-based Yorkshire Register of Childhood Diabetes collects cases of childhood IDDM, and the aim of the present study was to seek evidence of spatial and space-time clustering.

From the Paediatric Epidemiology Group (G.R.L., P.A.M., M.K.), Centre for Health Services Research and the Division of Public Health (A.S., R.W.), Nuffield Institute for Health, University of Leeds, Leeds; the Department of Public Health Services (F.A.), The Medical School, University of Edinburgh, Edinburgh; the Department of Public Health and Epidemiology (E.G.), Institute of Public and Environmental Health, University of Birmingham, Birmingham; and The Diabetes Centre (H.J.B.), Leeds General Infirmary, Leeds, U.K.

Address correspondence and reprint requests to Patricia A. McKinney, PhD, Paediatric Epidemiology Group, 32 Hyde Terrace, Leeds LS2 9LN, U.K. E-mail: trisham@epid.leeds.ac.uk.

Received for publication 18 July 1996 and accepted in revised form 4 December 1996.

RESEARCH DESIGN AND METHODS

Study population

The Yorkshire Register of Childhood Diabetes records demographic details for incident cases of 0- to 16-year-old subjects diagnosed with IDDM while resident within the three counties of North Yorkshire, West Yorkshire, and Humberside in the north of England (Fig. 1). The register is 97% complete for the years 1978–1990 (11) and holds information on 1,490 children.

The alphanumeric postal code of each address at the time of diagnosis was validated using the U.K. Postcode Address File, which also assigned each case a map grid reference (to 100 m) and a small-area census unit (an electoral ward). The date of diagnosis was assigned as the date of the first insulin injection.

The population denominators for the 536 electoral wards were taken from the 1981 U.K. census: mean childhood population for each ward was 6,450. Childhood population density was calculated as the number of children per hectare. The Townsend index of deprivation is a composite score of four census indicators: household overcrowding, car ownership, unemployment, and housing tenure (12).

Tests for clustering

The Knox test was applied to examine space-time clustering (13), using Mantel's modification (14) to calculate a z statistic (standard normal deviate) as a measure of significance. Thresholds were selected to include a range of distances in both space and time that were likely to be interpretable while remaining within the resolution of the data. The space and time thresholds were 1, 2, 4, 6, 10, 15, and 20 km and 90, 180, 360, 540, 720, and 900 days, respectively.

Comparison of the observed and expected cases occurring in a geographic-area unit can provide evidence for spatial clustering. The electoral wards in the U.K. differ in size, and classic methods of comparing observed to expected numbers using Poisson goodness-of-fit tests are not appropriate. Potthoff and Whittinghill (15) developed a powerful test (16) that assumes a minimum cluster is two cases and deter-

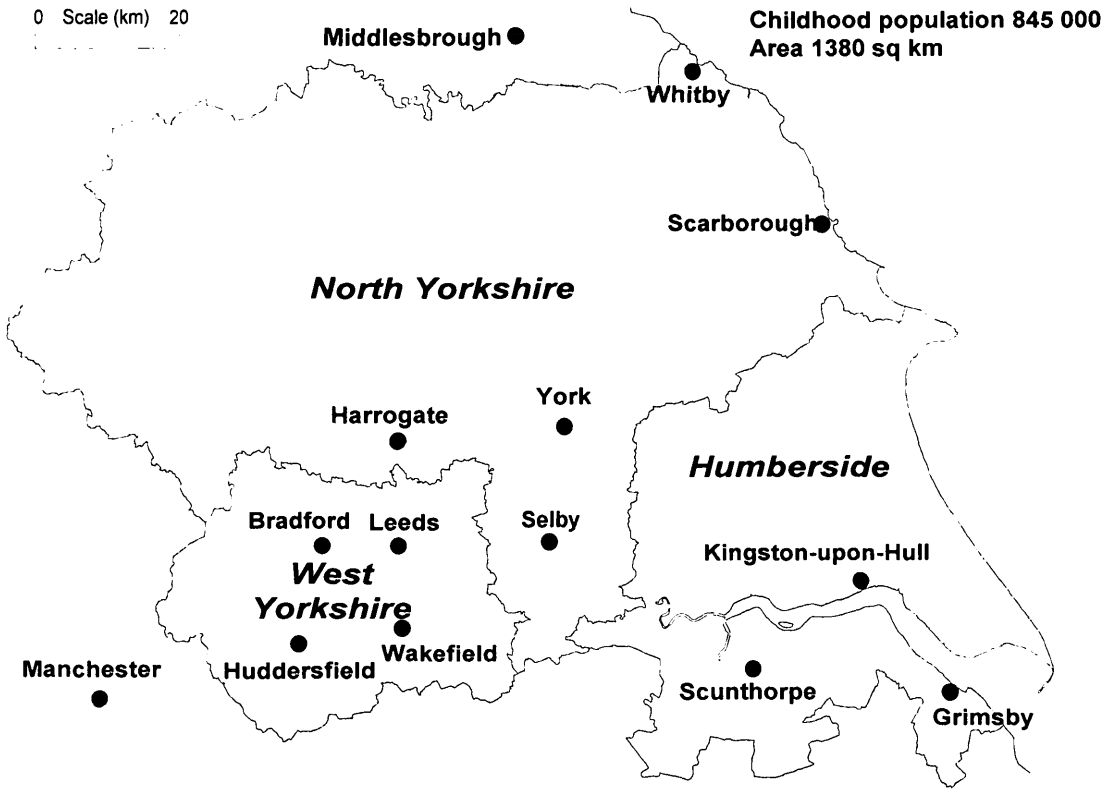

Figure 1—*Yorkshire Register of Childhood Diabetes study area.*

mines whether pairs occur more frequently than expected in a random distribution. A test score (standard normal deviate) (17) indicates whether the excess is significant. The expected numbers were calculated by applying the age- and sex-specific incidence rates for the whole area to each small unit. To adjust for known variation in incidence associated with demographic variables, conditioning was imposed on the expected number of cases based on a relative risk derived from the incidence rates in quartiles of childhood population density and the Townsend deprivation index.

Variation in clustering between subgroups

Variation in clustering for different time periods of diagnosis (1978–1981, 1982–1985, 1986–1990), age at diagnosis in years (0–4, 5–9, 10–16), sex, county of residence (Humberside, North Yorkshire, West Yorkshire), and childhood population density (<0.26, 0.26–3.71, or >3.71 0- to16-year-old subjects per hectare) were investigated. The variation in space-time clustering between subgroups was tested using a χ^2 test.

RESULTS — The Knox test for space-time clustering on 1,490 cases showed significantly more pairs of cases diagnosed within 1 km and up to 900 days of each other than would be expected by chance. The most significant excess occurred closer than 1 km and diagnosed within 360 days (z score 2.54, one-sided $P < 0.05$). For all cases, the Potthoff-Whittinghill test (Table 1) showed evidence of spatial clustering, although after conditioning for the effect of population density and deprivation, clustering is not significant. Results for subgroups are shown in Table 1 and summarized in Table 2. Significant spatial and space-time clustering was seen in the younger age-groups (0–4 and 5–9 years), with spatial clustering remaining significant after conditioning for levels of childhood population density and deprivation. Space-time clustering was present in Humberside and during the years 1982–1985, and spatial clustering was seen in low-population-density wards. There was an absence of clustering for boys or girls separately.

CONCLUSIONS — The small-scale geographic distribution of childhood diabetes is poorly described, thereby missing opportunities for identifying environmental etiologic factors. The current study took a high-quality population-based register of children with diabetes to test for space-time and spatial clustering of disease. A recent Swedish study demonstrated space-time clustering in childhood diabetes (3), and analysis of the Yorkshire register tested this hypothesis while extending investigations using recently developed methodologies.

The Knox test clearly showed the onset of IDDM in children and young people clusters in space and time. Significant excesses were present for those living within 1 km over a long time span (90–900 days), with the highest level occurring at 1 km and 360 days. Compared to the Swedish study where the most significant thresholds were 15 km and 7 months (3), the clustering appeared closer both in space and time; however, the Swedish population was from a predominantly rural area and an extended spatial distance would be expected.

The occurrence of diabetes is influenced by genetic susceptibility and cases may therefore cluster in families. Nine sib pairs of cases contributed to the significant results of the Knox test, and removing them from the analysis did not affect the level of statistical significance. Therefore, familial clustering is not able to explain our findings.

Efforts were made to clarify whether the overall clustering was associated with

Table 1—*Results of clustering tests on the register of childhood IDDM in Yorkshire*

Subgroup	Potthoff-Whittinghill (z) Crude	Potthoff-Whittinghill (z) Conditioned	Knox test O/E	Knox test Heterogeneity
Full register	1.92*	1.05	1.12†	—
Age-group at diagnosis				
0–4	2.31*	2.00*	1.56†	$\chi^2 = 6.65$
5–9	5.58*	3.61*	1.02	df = 2
10–16	0.26	0.20	1.19	$P = 0.036$
Year of diagnosis				
1978–1981	0.39	0.61	0.95	$\chi^2 = 8.04$
1982–1985	1.59	1.30	1.19†	df = 2
1986–1990	0.18	−0.02	1.05	$P = 0.018$
Sex				
Male	−0.19	−0.61	1.17	$\chi^2 = 1.57$
Female	0.27	−0.17	1.03	df = 1
				$P = 0.210$
Childhood population density				
Sparse (<0.26)	1.92*	1.30	NA	
Mid	−1.00	−0.96	NA	
Dense (>3.71)	0.61	0.44	NA	
County of residence				
Humberside	0.94	0.66	1.26†	$\chi^2 = 7.74$
North Yorkshire	0.27	−0.03	1.02	df = 2
West Yorkshire	0.17	−0.26	1.04	$P = 0.021$

*$P < 0.05$ (one-sided); †a subgroup with a significantly high O/E ratio. Childhood population density is tertiles of the number of wards ranked by the childhood (0–16 years) population density (people per hectare).

other factors. Space-time clustering was strongest during the period 1982–1985 and significantly different from that for the time periods before and after. Other European studies have shown differences in the level of clustering over time, such as Sweden during 1981–1985 (3) and Poland from 1982–1985 (18). Temporal changes are not related to increases in incidence, and changes in the genetically susceptible population will not account for space-time clustering of such short duration. Thus, etiology may be related to an environmental factor that varies over time, such as epidemics of infectious diseases.

Table 2—*Summary of significant clustering results*

Subgroup	Clustering: Spatial	Clustering: Space-time
	Entire (U)	Entire
Age-group	0–4, 5–9	0–4
Year of diagnosis	NS	1982–1985
Sex	NS	NS
Childhood population density	Sparse (U)	NS
County of residence	NS	Humberside

U, unconditioned test only.

Clustering in the younger age groups was significant. Other epidemiological characteristics of IDDM have shown different patterns in young children compared with the older age groups; for example, seasonality at presentation appears to be minimal or absent in these very young children (19–21).

Humberside has a 25% excess incidence (22), and attention is now focused on this area as the one in which cases are more likely to cluster compared with the other two counties. An explanation is not readily apparent, but Humberside may experience differential rates of infections, more intense epidemics, or, alternatively, exposure to chemical pollutants dispersed over the area.

Electoral wards were the smallest areas of analysis, and population density varies from <0.01 to 24.7 children per hectare. Transmission of infections is known to be highest in areas of high population density, and the absence of clustering in the densest areas may be a result of early exposure priming the immune system and ensuring that later exposure is unlikely to precipitate an autoimmune response. In sparsely populated areas, an absence of early exposure may result in genetically susceptible children succumbing to later exposure that precipitates the pathogenesis of β-cell destruction. The "late exposure" model has been proposed for poliomyelitis (23), childhood leukemia (24), and Hodgkin's disease (25). Conditions that delay exposure to infections, such as improved household hygiene, may be responsible for the rising incidence in Western countries (26).

In conclusion, evidence of both spatial and space-time clustering has been shown over a lengthy time period (1978–1990). Clustering is prevalent in younger age-groups, during the years 1982–1985, in the county of Humberside, and in wards of low population density. Small-scale clustering suggests that an infectious agent or point source may be involved in the onset or promotion of β-cell destruction leading to IDDM.

Acknowledgments— The register acknowledges financial support from The Children's Research Fund, the Yorkshire Regional Health Authority, and The General Infirmary at Leeds, Special Trustees. Novo Nordisk has provided computer hardware. G.L. was supported by the University of Leeds Scholarship Fund.

The pediatricians, diabetologists, and their staff throughout the former Yorkshire Regional Health Authority are thanked for contributing data to the Yorkshire Register of Childhood Diabetes. The assiduous data collection by Carolyn Stephenson and Helen Lilley is gratefully acknowledged.

References

1. Kaprio J, Tuomilehto J, Koskenvuo K, Romanov K, Reunanen A, Eriksson J, Stengard D, Vaseniemi YA: Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia* 35:1060–1067, 1992
2. Barnett AH, Eff C, Leslie RDG, Pyke DA: Diabetes in identical twins. *Diabetologia* 20:87–93, 1981
3. Samuelsson U, Johansson C, Cartensen J, Ludvigsson J: Space-time clustering in insulin-dependent diabetes mellitus (IDDM) in south-east Sweden. *Int J Epidemiol* 23: 138–142, 1994
4. McKinney PA, Law GR, Bodansky HJ, Staines A, Williams DRR: Geographical mapping of childhood diabetes in the northern English county of Yorkshire. *Diabet Med* 13:734–740, 1996

5. Barclay RPC, Craig JO, Galloway CAS, Richardson JE, Shepherd RC, Smail PJ: The incidence of diabetes mellitus in certain parts of Scotland. *Scott Med J* 33:237–239, 1988
6. Patterson CC, Carson DS, Hadden DR: Epidemiology of childhood IDDM in Northern Ireland 1989–1994: low incidence in areas with highest population density and most household crowding. *Diabetologia* 389: 1063–1069, 1996
7. Patterson CC, Waugh NR: Urban/rural and deprivational differences in incidence and clustering of childhood diabetes in Scotland. *Int J Epidemiol* 21:108–117, 1992
8. Siemiatycki J, Colle E, Aubert D: The distribution of type 1 (insulin-dependent) diabetes mellitus by age, sex, secular trend, seasonality, time clusters and space-time clusters: evidence from Montreal, 1971–1983. *Am J Epidemiol* 124:545–560, 1986
9. Staines A, Bodansky HJ, McKinney PA, Alexander FE, McNally RJQ, Law GR, Lilley HEB, Stephenson C, Cartwright RA: Small-area variation in the incidence of childhood insulin-dependent diabetes mellitus in Yorkshire UK: links with overcrowding and population density. *Int J Epidemiol*. In press
10. Alexander FE: Clustering and Hodgkin's disease. *Br J Cancer* 62:708–711, 1990
11. Staines A, Bodansky HJ, Liley HEB, Stephenson C, McNally RJQ, Cartwright RA: The epidemiology of diabetes mellitus in the United Kingdom: The Yorkshire Regional Childhood Diabetes Register. *Diabetologia* 36:1282–1287, 1993
12. Phillimore P, Beattie A, Townsend P: Widening inequalities of health in northern England 1981–1991. *BMJ* 308:1125–1128, 1994
13. Knox EG: The detection of space-time interactions. *Appl Statist* 13:25–29, 1964
14. Mantel N: The detection of disease clustering and a generalised regression approach. *Cancer Res* 27:209–220, 1967
15. Potthoff RF, Whittinghill M: Testing for homogeneity II: the Poisson distribution. *Biometrika* 53:183–190, 1966
16. Cartwright RA, Alexander FE, McKinney PA, Ricketts TJ: *Leukaemia and Lymphoma: An Atlas of Distribution Within Areas of England and Wales 1984–88*. London, Leukaemia Research Fund, 1991
17. Alexander FE: Investigations of localised spatial clustering and extra-Poisson variation. In *The Geographical Epidemiology of Childhood Leukaemia and Non-Hodgkin Lymphomas in Great Britain, 1966–83*. Draper G, Ed. London, OPCS, 1991, p. 69–76
18. Rewers M, LaPorte RE, Walczak M: Spatial clustering of insulin-dependent diabetes mellitus in midwestern Poland (Abstract). *Diabetes* 36 (Suppl.):129A, 1987
19. Gamble DR, Taylor KW: Seasonal incidence of diabetes mellitus. *Br Med J* 3:631–633, 1969
20. Gray RS, Duncan LJP, Clarke BF: Seasonal onset in insulin-dependent diabetes mellitus in relation to sex and age at onset. *Diabetologia* 17:29–32, 1979
21. Nystrom L, Dahlquist G, Ostman J, Wall S, Arnqvist H, Blohme G, Lithner F, Littorin B, Schersten B, Wibell L: Risk of developing insulin-dependent diabetes mellitus (IDDM) before 35 years of age: indications of climatological determinants for age at onset. *Int J Epidemiol* 21:352–358, 1992
22. McKinney PA, Law GR, Staines A, Bodansky HJ: Incidence of diabetes in children (Letter). *BMJ* 310:1672, 1995
23. Hillis A: Does insulin-dependent diabetes mellitus (IDDM) follow the poliomyelitis pattern? (Letter). *Am J Epidemiol* 112:452, 1980
24. Greaves MF, Alexander FE: An infectious etiology for common acute lymphoblastic leukaemia in childhood? *Leukaemia* 7:349–360, 1993
25. Gutensohn N, Cole P: Epidemiology of Hodgkin's disease in the young. *Int J Cancer* 19:595–604, 1977
26. Kolb H, Elliott RB: Increasing incidence of IDDM a consequence of improved hygiene? *Diabetologia* 37:729, 1994