Diabetologia (2006) 49: 900–904
DOI 10.1007/s00125-006-0208-5

SHORT COMMUNICATION

R. J. Q. McNally · R. G. Feltbower · L. Parker · H. J. Bodansky · F. Campbell · P. A. McKinney

Space–time clustering analyses of type 1 diabetes among 0- to 29-year-olds in Yorkshire, UK

Received: 16 August 2005 / Accepted: 3 January 2006 / Published online: 24 March 2006

Abstract *Aims/hypothesis:* Infections have been suggested to play a role in the aetiology of type 1 diabetes. The presence of space–time clustering is consistent with the notion of an environmental component in disease aetiology, possibly linked to infections. We tested for evidence of space–time clustering among children and young adults under 30 years of age using data from a population-based register in Yorkshire, UK. *Subjects and methods:* Two data sets of children and young people diagnosed with type 1 diabetes were analysed: (1) children aged 0–14 years and resident in Yorkshire during 1978–2002; (2) those aged 15–29 years and resident in West Yorkshire during 1991–2002. Tests for space–time interactions between cases were applied. Addresses at diagnosis were geo-coded and used as the basis for the analyses. *Results:* The study analysed 3,019 type 1 diabetic patients in the 0–14 years age group and 989 patients in the 15–29 years group. Statistically significant space–time clustering based on place and time of diagnosis was detected both for the 10–14-year-olds (p=0.04) and for the 15–19-year-olds (p=0.01). *Conclusions/interpretation:* Previous studies of clustering of type 1 diabetes have generally been restricted to childhood. Our results from a data set that includes teenagers and young adults show that space–time clustering was limited to young people aged 10–19 years. This finding is consistent with an aetiology involving late exposure to infection. However, the question of whether this is directly diabetogenic or unmasks latent diabetes cannot be addressed by this methodology.

Keywords Aetiology · Children · Diabetes · Infection · Space–time clustering

Abbreviations *O*: number of pairs of cases observed to be in close proximity · *E*: number of pairs of cases expected to be in close proximity · *S*: the magnitude of the excess (or deficit) · NN: nearest neighbour

R. J. Q. McNally · L. Parker
School of Clinical Medical Sciences (Child Health),
University of Newcastle upon Tyne,
Newcastle upon Tyne, UK

R. J. Q. McNally
School of Population and Health Sciences,
University of Newcastle upon Tyne,
Newcastle upon Tyne, UK

R. G. Feltbower · P. A. McKinney
Paediatric Epidemiology Group,
Centre for Epidemiology and Biostatistics,
University of Leeds,
Leeds, UK

H. J. Bodansky
The General Infirmary at Leeds,
Leeds Teaching Hospitals NHS Trust,
Leeds, UK

F. Campbell
St James's University Hospital,
Leeds Teaching Hospitals NHS Trust,
Leeds, UK

R. J. Q. McNally (✉)
Sir James Spence Institute,
Level 4, Royal Victoria Infirmary, Queen Victoria Road,
Newcastle upon Tyne, NE1 4LP, UK
e-mail: Richard.McNally@ncl.ac.uk
Tel.: +44-191-202-3029
Fax: +44-191-202-3060

Introduction

The aetiology of type 1 diabetes in children and young adults is not well understood, although both genetic and environmental factors are likely to be involved [1]. A possible role for infections and immunological responses has been identified [2]. Higher incidence has been observed in areas of lower population density or greater affluence [3], which may reflect differing patterns of exposure to infections between communities and is consistent with the hygiene hypothesis [4]. This predicts higher disease rates in areas where children's exposure to infections might be lower.

Space–time clustering occurs when excess numbers of cases are observed within small geographical locations for limited periods of time: this could arise from exposure to an

environmental (especially infectious) agent. The present study aimed to test predictions of space–time clustering in 3,019 patients aged 0–14 years and 989 patients aged 15–29 years using rigorous statistical methods. It is the first study to describe space–time clustering among 15–29-year-olds, and to address the effects of sex and population density.

Subjects and methods

Data on children and young people aged 0–29 years who had been diagnosed with type 1 diabetes up to 31 December 2002 were abstracted from the population-based Yorkshire Register of Diabetes in Children and Young People. The collection of data and its use for research purposes had received appropriate ethical approval. Completeness of ascertainment has been estimated to be more than 98%. The registry has recorded cases of diabetes in 0–14-year-olds for the whole of Yorkshire since 1978 and in 15–29-year-olds for West Yorkshire since 1991 [5]. Of the Yorkshire population, 47% reside in urban areas (>25 persons per hectare). Type 1 diabetes was defined as the diagnosis given by the consultant in charge of the case, dependence on insulin from diagnosis and/or proneness to ketosis.

Table 1 Space–time clustering tests for cases of type 1 diabetes

	Knox test (observed close space–time pairs[a], expected space–time pairs, strength[b], p value[c])		K-function analysis[f] (p value[g])	
Age-group	Geographical distance[d]	NN threshold[e]	Geographical distance[h]	NN threshold[j]
Cases aged 0–14 from Yorkshire and diagnosed during the period 1978–2002				
0–4	O=535; E=515.2 S=3.8% p=0.20	O=239; E=237.8 S=0.5% p=0.48	p=0.10	p=0.18
5–9	O=1,127; E=1,141.7 S=−1.3% p=0.66	O=579; E=583.6 S=−0.8% p=0.57	p=0.45	p=0.66
10–14	O=2,066; E=1,989.2 S=3.9% p=0.04*	O=1,063; E=994.9 S=6.8% p=0.02*	p=0.004*	p=0.04*
0–14	O=10,327; E=10,203.9 S=1.2% p=0.11	O=5,026; E=5,020.1 S=0.1% p=0.47	p=0.12	p=0.61
Cases aged 15–29 from West Yorkshire and diagnosed during the period 1991–2002				
15–19	O=349; E=356.7 S=−2.2% p=0.65	O=244; E=230.2 S=6.0% p=0.19	p=0.20	p=0.01*
20–24	O=429; E=405.7 S=5.7% p=0.13	O=147; E=140.7 S=4.5% p=0.31	p=0.28	p=0.39
25–29	O=552; E=553.4 S=−0.3% p=0.51	O=222; E=210.0 S=5.7% p=0.21	p=0.43	p=0.19
15–29	O=3,963; E=3,920.1 S=1.1% p=0.25	O=1,651; E=1,623.7 S=1.7% p=0.25	p=0.17	p=0.22

*p<0.05

[a]Cases are close in time if dates of diagnosis differ by <1 year

[b]Strength $(S) = \{(\text{Observed} - \text{Expected})/\text{Expected}\} \times 100$ counts of pairs that 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 34th NN for the 0–14-year-olds from Yorkshire and was nearer than the other's 17th NN for the 15–29-year-olds from West Yorkshire

[f]Cases are close in time if dates differ by $<t$ where t is in the range 0.1–1.5 years

[g]p value obtained by simulation (999 runs) with dates of diagnosis randomly re-allocated to the cases in the analysis

[h]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 27–41 for the 0–14-year-olds from Yorkshire and n is in the range 10–24 for the 15–29-year-olds from West Yorkshire

Two separate analyses were undertaken: (1) all patients aged 0–14 years diagnosed between 1 January 1978 and 31 December 2002; and (2) all patients aged 15–29 years diagnosed between 1 January 1991 and 31 December 2002. Analyses were also performed by 5-year age groups to test for age-related differences in susceptibility to a putative aetiological agent.

For each subject, the Easting and the Northing co-ordinates of the address at diagnosis were spatially referenced to the nearest 0.1 km. The following primary aetiological hypotheses were tested: (1) There is space–time clustering of incident cases of diabetes; and (2) space–time heterogeneity of incidence of diabetes is modulated by hormonal factors and therefore is associated with age at diagnosis. For statistically significant groups ($p<0.05$ using the nearest neighbour (NN) threshold version of the *K*-function method) two further hypotheses were tested: (3) space–time clustering of diabetes is associated with either male or female sex—to test for sex-related differences, analyses were performed for clustering pairs that included at least one male case and at least one female case; (4) space–time clustering of diabetes is associated with population density. Analysis by population density was undertaken for clustering pairs of cases that included: at least one case from the 'more densely populated' category and at least one case from the 'less densely

Table 2 Further space–time clustering tests by sex and level of population density

	Knox test (observed close space–time pairs[a], expected space–time pairs, strength[b], *p* value[c])		*K*-function analysis[f] (*p* value[g])	
Age-group	Geographical distance[d]	NN threshold[e]	Geographical distance[h]	NN threshold[j]
Cases aged 10–14 from Yorkshire and diagnosed during the period 1978–2002				
≥1 male case	*O*=1,538; *E*=1,500.4 *S*=2.5% *p*=0.17	*O*=831; *E*=782.0 *S*=6.3% *p*=0.04*	*p*=0.06	*p*=0.07
≥1 female case	*O*=1,539; *E*=1,473.3 *S*=4.5% *p*=0.046*	*O*=744; *E*=697.5 *S*=6.7% *p*=0.04*	*p*=0.003*	*p*=0.04*
≥1 case from more densely populated area	*O*=1,817; *E*=1,751.7 *S*=3.7% *p*=0.06	*O*=670; *E*=631.6 *S*=6.1% *p*=0.07	*p*=0.007*	*p*=0.03*
≥1 case from less densely populated area	*O*=567; *E*=589.6 *S*=−3.8% *p*=0.82	*O*=581; *E*=550.5 *S*=5.6% *p*=0.10	*p*=0.88	*p*=0.46
Cases aged 15–19 from West Yorkshire and diagnosed during the period 1991–2002				
≥1 male case	*O*=258; *E*=282.7 *S*=−8.7% *p*=0.93	*O*=183; *E*=190.1 *S*=−3.8% *p*=0.68	*p*=0.80	*p*=0.28
≥1 female case	*O*=246; *E*=249.9 *S*=−1.6% *p*=0.58	*O*=169; *E*=148.5 *S*=13.8% *p*=0.053	*p*=0.39	*p*=0.001*
≥1 case from more densely populated area	*O*=284; *E*=295.8 *S*=−4.0% *p*=0.74	*O*=95; *E*=93.9 *S*=1.2% *p*=0.47	*p*=0.31	*p*=0.16
≥1 case from less densely populated area	*O*=118; *E*=115.1 *S*=2.5% *p*=0.41	*O*=175; *E*=164.6 *S*=6.3% *p*=0.22	*p*=0.18	*p*=0.03*

*$p<0.05$

[a]Cases are close in time if dates of diagnosis differ by <1 year

[b]Strength $(S) = \{(\text{Observed} - \text{Expected})/\text{Expected}\} \times 100$ counts of pairs that 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 34th NN for the 0–14-year-olds from Yorkshire and was nearer than the other's 17th NN for the 15–29-year-olds from West Yorkshire

[f]Cases are close in time if dates differ by $<t$ where t is in the range 0.1–1.5 years

[g]*p* value obtained by simulation (999 runs) with dates of diagnosis randomly re-allocated to the cases in the analysis

[h]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 27–41 for the 0–14-year-olds from Yorkshire and n is in the range 10–24 for the 15–29-year-olds from West Yorkshire

populated' category. These were defined on the basis of whether the *n*th nearest neighbour was nearer or further away than the median distance for each age group.

Space–time interactions based on time and place of diagnosis were tested. Knox space–time clustering 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 [6]. In the Knox test, a pair of cases was regarded as being in 'close proximity' if they are both diagnosed at addresses that are simultaneously close in space and at times that are close. The number of pairs of cases observed (O) and expected (E) to be in close proximity was obtained. If O exceeded E there was space–time clustering. The magnitude of the excess (or deficit) was estimated by calculating $S = \left[(O-E)/E\right] \times 100$. To adjust for the effect of different population densities, the tests were repeated replacing geographical distance thresholds by distance to the *n*th nearest neighbour, using all locations of all the cases in the data set; n was chosen such that the mean distance was 5 km.

An inherent problem with the Knox test is that thresholds are chosen arbitrarily. A simplification of a second-order procedure based on K-functions was used in the present analyses to partly overcome the problem of arbitrary boundaries [7]. This procedure involved a set of 225 Knox-type calculations where the boundary changed over a pre-specified set of values. Statistical significance was assessed by simulation. The nearest neighbour approach was derived by replacing fixed geographical distances by distances to the *n*th nearest neighbours and this provided the primary result for each analysis.

Results

The study included 3,019 patients with diabetes aged 0–14 years and diagnosed between 1 January 1978 and 31 December 2002 (1,568 males; 1,451 females; 645 aged 0–4 years; 1,020 aged 5–9 years; 1,354 aged 10–14 years) and 989 patients with diabetes who were aged 15–29 years and had been diagnosed between 1 January 1991 and 31 December 2002 (620 males; 369 females; 358 aged 15–19 years; 286 aged 20–24 years; 582 aged 25–29 years).

For those aged under 15 years, we found that space–time clustering was confined to 10–14-year-old patients (p=0.04 using the NN threshold version of the K-function method). For individuals aged 15–29 years, some evidence for space–time clustering was observed only for patients aged 15–19 years (p=0.01 using the NN threshold version of the K-function method) (Table 1).

Analysis by sex showed that both for patients aged 10–14 years and those aged 15–19 years there was statistically significant clustering only for pairs of cases that included at least one female. Analysis by population density showed that for 10–14-year-old patients there was significant clustering only for pairs of cases that included at least one from a more densely populated area; for patients aged 15–19 years there was significant clustering only for pairs that included at least one from a less densely populated area (Table 2).

Discussion

In this, the largest reported study of space–time clustering in children and young adults with type 1 diabetes, significant evidence of clustering was identified for those aged 10–14 and 15–19 years. The findings are novel because they are not restricted to the childhood age range and they are robust having been performed using well-specified statistical methods and high-quality incidence data.

A previous smaller study of type 1 diabetes in children, aged 0–14 years, from Yorkshire (1978–1990) applying a less sophisticated methodology found clustering limited to 0–4-year-olds [8]. There are two possible explanations for the differences: either, space–time clustering for 0–4-year-olds was limited to the 1978–1990 period, or it is restricted to a very specific geographical and temporal definition of 'close proximity' (<1 km in space and <1 year in time).

Three studies from Sweden and one from Chile have found space–time clustering of type 1 diabetes in children [9–12], whilst a further study from France found no evidence of space–time clustering [13]. Only one of these studies made an adjustment for arbitrary boundaries [11]. The methodology employed in all of the other previous studies involved specifying a number of arbitrary thresholds for 'close in time' and 'close in space'. Tests were repeated using different thresholds but no adjustments were made for multiple testing. The present study partly overcomes the problem of multiple testing that was present in earlier studies and confirms the tendency for cases of childhood type 1 diabetes to cluster in space and time.

The present study found a female preponderance in clustering pairs (especially those aged 15–19 years). This could either have occurred through differential pubertal effects for males and females on the immune system [14] possibly unmasking the disease which is known to have a long latency period or a larger number of susceptible females may have been more exposed to infections via, for example, social contact.

The occurrence of space–time clustering is not consistent with a sustained exposure either geographically or over time and is more consistent with an aetiological agent that displays a temporary occurrence at a number of different locations. Infections would be a highly plausible candidate. Although incidence has been steadily rising in Yorkshire [6], clustering was present throughout the study period.

A role for infections and immunological responses in aetiology is also supported by the findings of a number of recent case-control studies [2, 15–19]. Three ecological studies have found higher incidence associated with lower population density or greater affluence [3, 20, 21] which may reflect differing patterns of exposure to infections between communities, with the numbers and levels of circulating infections likely to be lower in areas which are more affluent and less densely populated. The present

study did not show a clear association with population density.

A number of viruses have been specifically associated with disease onset [22]. However, given the size of the attributable risk estimates from these studies, the majority of cases of type 1 diabetes are unlikely to be triggered by exposure to a single infection or environmental agent. The hygiene hypothesis [4] postulates that exposure to a suite of infections in early life may decrease the subsequent risk of autoimmune conditions such as type 1 diabetes through immune modulation [4].

In summary, we have found evidence of space–time clustering among patients with type 1 diabetes aged 10–14 years, along with a new observation of clustering in 15–19-year-olds; both these results are consistent with an environmental link for disease onset. Although the analysis employs robust statistical techniques, further work is needed to develop a more refined aetiological hypothesis.

Acknowledgements The Yorkshire Diabetes Register in Yorkshire is successful as a result of the support and co-operation of many NHS staff throughout the region; we wish to thank all the medical staff, nurses, medical records officers and other members of the clinical care teams who collaborate with us. We are grateful to C. Stephenson and A. Herbert for data collection and processing. The Register receives funding from the Department of Health. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health.

References

1. Rewers M, Norris J, Dabelea D (2004) Epidemiology of type 1 diabetes mellitus. Adv Exp Med Biol 552:219–246
2. EURODIAB Substudy 2 Study Group (2000) Infections and vaccinations as risk factors for childhood type 1 (insulin dependent) diabetes mellitus: a multi-centre case-control investigation. Diabetologia 43:47–53
3. Patterson CC, Carson DJ, Hadden DR (1996) Epidemiology of childhood IDDM in Northern Ireland 1989–1994: low incidence in areas with highest population density and most household overcrowding. Northern Ireland Diabetes Study Group. Diabetologia 39:1063–1069
4. Kolb H, Elliott RB (1994) Increasing incidence of IDDM a consequence of improved hygiene? Diabetologia 37:729
5. Feltbower RG, McKinney PA, Parslow RC, Stephenson CR, Bodansky HJ (2003) Type 1 diabetes in Yorkshire, UK: time trends in 0–14 and 15–29-year-olds, age at onset and age-period-cohort modeling. Diabet Med 20:437–441
6. Knox EG (1964) The detection of space–time interactions. Appl Stat 13:25–30
7. Diggle PJ, Chetwynd AG, Haggkvist R, Morris SE (1995) Second-order analysis of space–time clustering. Stat Methods Med Res 4:124–136
8. Law GR, McKinney PA, Staines A et al (1997) Clustering of childhood IDDM. Links with age and place of residence. Diabetes Care 20:753–756
9. Samuelsson U, Johansson C, Carstensen J, Ludvigsson J (1994) Space–time clustering in insulin-dependent diabetes mellitus (IDDM) in south-east Sweden. Int J Epidemiol 23:138–142
10. Samuelsson U, Catstensen J (2003) Space–time clustering at birth and at diagnosis of type 1 diabetes mellitus in relation to early clinical manifestation. J Pediatr Endocrinol Metab 16:859–867
11. Dahlquist GG, Kallen BA (1996) Time–space clustering of date at birth in childhood-onset diabetes. Diabetes Care 19:328–332
12. Santos J, Carrasco E, Moore A, Perez-Bravo F, Albala C (2001) Incidence rate and spatio-temporal clustering of type 1 diabetes in Santiago, Chile, from 1997 to 1998. Rev Sante Publica 35:96–100
13. Hours M, Siemiatycki J, Fabry J, Francois R (1990) Time clustering and temporospatial regrouping study of cases of juvenile diabetes in the district of Rhone (1960–1980). Rev Epidemiol Santé Publique 38:287–295, [French]
14. Purtilo DT, Sullivan JL (1979) Immunological bases for superior survival of females. Am J Dis Child 133:1251–1253
15. Altobelli E, Petrocelli R, Verrotti A, Valenti M (2003) Infections and risk of type 1 diabetes in childhood: a population-based case-control study. Eur J Epidemiol 18:425–430
16. Carbrera-Rode E, Sarmiento L, Tiberto C et al (2003) Type 1 diabetes-islet associated antibodies in subjects infected by echovirus 16. Diabetologia 46:1348–1353
17. Sipetic S, Vlajinac H, Kocev N, Radmanovic S (2003) The Belgrade childhood diabetes study: association of infections and vaccinations on diabetes in childhood. Ann Epidemiol 13:645–651
18. Salminen KK, Vuorinen T, Oikarinen S et al (2004) Isolation of enterovirus strains from children with preclinical type 1 diabetes. Diabet Med 21:156–164
19. Sipetic SB, Vlajinac HD, Kocev NI, Marinkovic JM, Radmanovic SZ, Bjekic MD (2005) The Belgrade childhood diabetes study: a multivariate analysis of risk determinants for diabetes. Eur J Public Health 15:117–122
20. Patterson CC, Waugh NR (1992) Urban/rural and deprivational differences in incidence and clustering of childhood diabetes in Scotland. Int J Epidemiol 21:108–117
21. Schober E, Rami B, Waldhoer T, Austrian Diabetes Incidence Study Group (2003) Small area variation in childhood diabetes mellitus in Austria: links to population density, 1989 to 1999. J Clin Epidemiol 56:269–273
22. Akerblom HK, Knip M (1998) Putative environmental factors in type 1 diabetes. Diabetes Metab Rev 14:31–67