RESEARCH REPORT

Geographical mapping of type 1 diabetes in children and adolescents in south east Sweden

U Samuelsson, O Löfman

J Epidemiol Community Health 2004;**58**:388–392. doi: 10.1136/jech.2002.004135

See end of article for
authors' affiliations

Correspondence to:
Dr U Samuelsson, Division
of Paediatrics, Department
of Molecular and Clinical
Medicine, Linköping
University, S-581 85
Linköping, Sweden;
ulf.samuelsson@lio.se

Accepted for publication
19 June 2003

Study objective: As earlier studies have shown space-time clusters at onset of type 1 diabetes in the south east region of Sweden we investigated if there also has been any geographical clusters of diabetes in this region.

Design: The place of residence (coordinates) at the time of diagnosis were geocoded in a geographical information system (GIS). All children diagnosed with type 1 diabetes up to 16 years of age at diagnosis between 1977–1995 were included. The population at risk was obtained directly from the population registry for the respective years and geographical area levels.

Setting: South east region of Sweden containing 5 counties, 49 municipalities, and 525 parishes.

Main results: A significant geographical variation in incidence rate were found between the municipalities ($p < 0.001$) but not between the counties. The variation became somewhat weaker when excluding the six largest municipalities ($p < 0.02$). In municipalities with increased risk ($> 35.1/100\ 000$) the major contribution comes from children in age group 6–10 years of age at diagnosis. There were no obvious differences between the age groups in municipalities with decreased risk ($< 20.1/100\ 000$). Boys and girls had about the same degree of geographical variation.

Conclusions: Apart from chance, the most probable explanation for the geographical variation in the risk for children and adolescents to develop type 1 diabetes between the municipalities in the region is that local environmental factors play a part in the process leading to the disease.

During the years several studies have found spatial clustering at the time of diagnosis in childhood diabetes.^{1–10} We,¹¹ as well as others,^{8, 12} have also reported indications of space-time clustering at diagnosis. Also evidence of birth date time and space clustering of children who later develop diabetes has been shown.¹³

The studies regarding spatial clustering have shown that the incidence is higher in rural than urban areas^{4, 6, 8–10} and, as seen in a study from Scotland,² higher incidence in highlands than areas in the lowlands. Patterson and coworkers⁸ made further analysis and found decreased incidence in areas with highest population density. Children in such areas are more likely to have early infections than areas with low density. These findings therefore imply a protective role of early infections through stimulation of the immune system in agreement with the thesis by Karlsson.¹⁴

In the south east part of Sweden we have earlier been able to show that children sometimes are diagnosed in space-time clusters¹¹ and that there is a difference in birth pattern in children who develop diabetes compared with the background population.¹⁵

Using the same population based material as in the previous study¹⁵ we investigated if there also have been any geographical or spatial clusters of diabetes in children in this part of Sweden.

METHODS

In the Swedish healthcare system, all children with diabetes up to 16 years of age are diagnosed and treated at paediatric clinics. From the seven paediatric clinics in the south east part of Sweden all 1144 children with age at onset of type 1 diabetes less than 16 years and diagnosed between 1977 and 1995 were included. The catchment areas of these clinics form a contiguous geographical area of 42 356 km², and on average 1 290 000 inhabitants. Information regarding date of birth, sex, residential address (at the time of diagnosis), and

date of diagnosis (= first insulin injection) was abstracted from medical records by nurses or doctors in the diabetes teams at each clinic. As all cases are locally registered we can be sure that the study includes all diabetic cases in the region. To validate this further we compared our figures with the reports to the Swedish Diabetes Registry. This register is estimated to have completeness close to 100%.^{13, 16, 17} We found a complete accordance. During the study period the mean annual number of children aged 0–15 years was 259 565.

Geographical units

The south east region of Sweden has five counties (Blekinge, Jönköping, Kalmar, Kronoberg, and Östergötland), 49 municipalities, and 525 parishes. Large parts of this region are predominantly rural. Each county has one or two paediatric clinics and, related to Swedish circumstances, one or two major cities (above 50–60 000 inhabitants, all ages included).

Establishment of a GIS database

The Swedish population registry is upgraded every year. All patients were therefore linked to the registry as recorded on the 31 of December for the year of diabetes onset by use of a personal identification number, unique for all Swedish citizens. The patients were then defined geographically according to their place of residence by matching the population registry to a national property register as reported in a previous study.¹⁸ By this procedure the X and Y coordinates in the national grid for the centroids of the residences were added to the data file and geocoded in a geographical information system (fig 1). The number of cases (the numerator) was then calculated at municipality level by overlay technique using an SQL queering routine that defined all cases within each of the 52 respective municipality polygons. The population at risk, comprising all children

Linking of databases:

Figure 1 The principle of linking databases in the study and the stack of digital maps in the geographical information system used for identification of type 1 diabetes cases and population at risk in certain defined polygons. This could be parish, urban areas, commune, and county respectively dependent of comparison made. The incidences were calculated by extracting the numerator and denominator figures derived from the respective map layer by overlay technique.

0–15 years of age within the same polygons (the denominator) were then defined accordingly and the incidence of type 1 diabetes standardised by five year age intervals calculated for the respective municipality (fig 1).

Geostatistical analysis

Ordinary kriging was used for estimation of spatial distribution of incidence.¹⁹ Kriging is a linear geostatistical interpolation method taken into account spatial autocorrelation using a weighted linear predictor minimising mean² predicting error, spatial autocorrelation was estimated by a semivariogram.²⁰ A semivariogram measures half the average squared difference between pairs of data assuming a covariance between (in this study) incidence values at the respective points as a function of distances. Aggregated incidences calculated from the type 1 diabetes cases in the respective municipalities were converted to a raster map and was used as point values in this study. An isotropic spherical semivariogram model (fig 4) was fitted to an empirical variogram scatter using the aggregated incidence values of the municipalities, assuming the same spatial autocorrelation in each direction. This simplified model thus assumes that the autocorrelation depends only on the distance between two measured points. These weighted incidence values from the semivariogram model were used to smooth the distribution pattern by ordinary kriging and the predicted values plotted in a 3D surface graph (fig 5).

Spatial data were analysed using the software ArcGIS 8.1.

Statistical analysis

In this explorative study we tested differences in incidence rate by χ^2 using the formula:

$(\text{Observed (O)} - \text{expected (E)})^2 / \text{expected}$. Expected values are estimated under the assumption of no geographical variation and are calculated by the formula:

[The overall annual incidence during the study period in the total study area] $\times$ [all children (0–15 years of age) in the county, municipality or parish]. Estimation in subgroups were done by analogy.

The study was approved by the research ethics committee of the faculty of health sciences, Linköping University.

RESULTS

The county of Jönköping had the lowest incidence rate in the region during the study period (22.4/100 000 children 0–15 years of age) (fig 2) and the counties of Blekinge and Kalmar the highest (26.4 each)(NS, compared with Jönköping). Only boys showed a tendency to a more obvious difference in incidence rate between the counties ($p = 0.07$), while girls or the three age groups had no such tendency. We also performed an analysis where we excluded children with diabetes and the background population in the six largest

cities (for example, one city (municipality) in each county apart from Östergötland were two cities (municipalities) were excluded). Three of these cities had more than 100 000 inhabitants and the other three more than 55 000 inhabitants. Those analyses, without the six cities, showed with exception of the girls, weaker χ^2 values and thereby also a weaker tendency for clustering than in the total material.

When comparing the incidence rate between the municipalities in the study region (fig 3) we found significant differences both in total and in some subgroups (table 1). The differences became somewhat weaker when excluding the six largest municipalities (table 1).

In the seven municipalities with increased risk (above 35.1/100 000) the major contribution come from children in age group 6–10 years of age at diagnosis. In contrast there were

Figure 2 The incidence rate during the study period in the five counties in the south east part of Sweden. The figures within each county are the incidence 100 000/person year in children 0–15 years of age at diagnosis.

Figure 3 The incidence rate in the 49 municipalities during the study period.

only minor differences between the age groups in the 11 municipalities with decreased risk (under 20.1/100 000) (table 2).

A descriptive comparison of incidence rate between the parishes shows that even municipalities sometimes may be too big for more careful space-cluster analysis. For instance, the municipality of Ljungby had an incidence of 37.1/100 000 during the study period. When comparing the 19 parishes within this municipality we saw that nine had no cases at all during the study period. Seven of those formed a contiguous area within the municipality. The three parishes with the highest incidence (>60/100 000) were more uniformly spread across the area. This pattern could also be seen in the county of Östergötland and in the part of the county of Jönköping that bordered the county of Östergötland; parishes with no cases or very low incidence formed contiguous areas while parishes with very high incidence were uniformly spread.

The semivariogram model indicated a weak spatial dependency of incidence data, (nugget:34.1; partial sill 49.9) (fig 4). Major range was about 100 km in the model.

The incidence pattern varied with a tendency towards slightly higher values in the central part of the study area as shown by the surface plot (fig 5).

DISCUSSION

In the Swedish healthcare system all children up to at least 16 years of age are diagnosed and treated in paediatric departments. The Swedish system of centralised care of children with diabetes means that all cases are reported, and we found complete concordance with the reports to the Swedish Childhood diabetes registry.^{13 16 23} As reported in an earlier study from this region¹¹ there has been no substantial differences in the population shifts between areas in the region. No obvious migration to or from the region has occurred during the study period. Furthermore, there seem to be no large differential changes in the number of children between areas in our region.

Kriging is an interpolator permitting a variety of output maps that is dependent on the model and assumes a stationary stochastic process.²⁰ A certain limitation affects the analysis of the distribution pattern because of the low number of point estimates included in the model and also to the use of aggregated data (area information). This was mainly attributable to the lack of coordinates of individual background population data at the time of the study. An analysis based on original point estimates for case and population data defined at the individual level, would probably yield more accurate estimates.

The explanation for the geographical distribution pattern of cases of type 1 diabetes needs further studies involving the assessment of environmental exposures.

Similar to some other studies^{4 12} we found that boys and girls had about the same degree of geographical variation. Law *et al*¹² found in their study that the age group 0–4 and 5–9 years of age at diagnosis had the most significant spatial clustering. We, in contrast with this saw no evident spatial clustering in the youngest age group. On the other hand we, as well as Law and coworkers, found that age group 6–10 years of age had the most obvious geographical variation.

Most of the studies who have found geographical variation of incidence of type 1 diabetes have also shown that the incidence is higher in rural than urban areas.^{4 6 8 10} When we excluded the most urban areas in the region (that is, the six biggest cities in the region) the significant level regarding spatial clustering decreased, especially for boys and children 6–10 years of age at diagnosis, while girls and the youngest age group had about the same level. This finding is somewhat different to Staines *et al*⁶ who noticed that the effect of population density was weaker for men and for younger children. Also when looking at the communities with the highest incidence in the study region it seems that it is the children in age group 6–10 years who are most responsible for the geographical variation, whereas communities with low incidence have only minor differences between the age groups.

Table 1 Comparisons between the communities in the region. Figures within parentheses are from analysis when the six biggest cities in the region are excluded

	Degrees of freedom (df)	χ^2 Value	p Value
Total	48 (42)	84.87 (62.85)	0.0008 (0.0201)
Boys	48 (42)	66.69 (46.39)	0.04 (0.3)
Girls	48 (42)	67.68 (59.01)	0.032 (0.042)
11–15 years	48 (42)	64.05 (50.07)	0.06 (0.18)
6–10 years	48 (42)	78.84 (56.62)	0.003 (0.13)
0–5 years	48 (42)	48.74 (44.12)	0.44 (0.38)

Table 2 The major contribution for the geographical variation comes from the communities with high incidence, especially from children 6–10 years of age at diagnosis

		df	χ^2 Value	p Value
Communities with increased risk	11–15 years	6	16.75	0.01
	6–10 years	6	27.05	0.000
	0–5 years	6	13.2	0.04
Communities with decreased risk	11–15 years	10	6.9	0.73
	6–10 years	10	13.39	0.20
	0–5 years	10	13.78	0.18

The cause of type 1 diabetes is unknown but genetic susceptibility is an important determinant.^{21–22} Although HLA genotypes are important in explaining global geographical differences in type 1 diabetes, it is reasonable to assume that in our homogenous population where 98% are white who have lived in Sweden for many generations, the families are expected to share a common genetic heritage. It is with this background highly unlikely that variations in the genetic background account for the observed geographical variation.

Environmental factors are, perhaps, more likely to play a part for this variation. Two earlier studies from Sweden found significant variation associated with latitude, but not with population density or with being a coastal or an interior county,²³ the incidence of type 1 diabetes was also higher in counties with less sunshine and colder climates.¹⁷ In this study only one of the seven communities with the highest incidence are coastal. On the other hand most of the communities with the lowest incidence are interior. The climate is about the same in the whole region. If anything, the county in the south (Blekinge, incidence 26.4) has in general less rain and more sunshine than the rest of the region as well as the two communities on the island in the Baltic Sea. Population density as well as urban/rural differences were only briefly investigated but when excluding the most populated communities the level of geographical variation decreased. The analyses regarding the parishes should be interpreted with caution as they in general have small risk population and the observed pattern might therefore be attributable to random variation.

The distribution of high values in the distribution model correlates roughly to the highlands of southern Sweden. Two of the five communities had incidence >35.0/100 000, one had an incidence between 30.1 to 35.0/100 000, and two communities between 25.1 and 30.0/100 000 and year. In previous

report² a similar distribution pattern of type 1 diabetes was reported with higher incidences in highland areas.

Most probably, the environmental factors responsible for the spatial pattern must be more local than common. A study from England that found spatial clustering could demonstrate an association between higher nitrate levels in domestic drinking water and incidence of childhood diabetes.⁵ This exposure is, however, more likely to have an impact on rural populations because agricultural activity sometimes causes contamination of wells by surface water. Some other studies have, however, not been able to confirm this association between nitrate levels and type 1 diabetes.^{24–25} The drinking water supply in our study was explored by a population questionnaire in a part of the region (county of Östergötland, 420 000 inhabitants) and showed that about 15% of the households were supplied by ground water sources from private wells, predominantly in rural areas, and 85% from municipal water work distribution networks.²⁶ Some other studies have found possible relations between low pH levels and low ground water concentrations of zinc and risk of type 1 diabetes.^{27–28} Unfortunately no such data were available at the time of our study.

Other candidates could be variation in exposure to childhood infection between the communities as a consequence of population density or other social factors. We know that some of the communities with the highest incidence have heavy industries (paper mill, metal industry). Interestingly, Dahlquist and Källen found in their study that Vaggeryd was one of the few municipalities in Sweden that had significantly more observed birth than expected of children who later develop diabetes.¹³ In this study Vaggeryd was one of the municipalities that had a significantly higher incidence than the neighbouring municipalities. Vaggeryd has a big paper mill industry.

Figure 4 The semivariogram depicting study area data. The variance of incidences between pairs of points separated by a certain distance is plotted against their respective distances in an empirical semivariogram. To this scattergram a spherical model is fitted. The nugget is the value where the semivariogram model intercepts the y axis and represents a discontinuity at the origin of the y axis caused by microscale variation and measurement errors. The sill is the value at which the semivariogram model attains its highest value on the y axis. (The partial sill is the sill minus the nugget) The range is the distance on the abscissa (the value of the x axis) from origo where this occurs and was used as the search radius for the prediction map.

Figure 5 Surface plot of incidence pattern in study area smoothed by ordinary kriging.

The geographical variation found in this study could, of course, also be attributable to chance, although it is somewhat unlikely. Further investigation is therefore needed and we intend to include information about the children and adolescents who developed diabetes in the region between 1995 and 2001, about 800 children. This will make it possible to do time trends and to subgroup the study population according to age, sex, and to some clinical background parameters.

In conclusion, we found a clear geographical variation in the risk for children and adolescents to develop type 1 diabetes between the municipalities in the south east region of Sweden. Apart from chance, the most probable explanation to this is that local environmental factors play a part in the process leading to the disease.

ACKNOWLEDGEMENTS

We are grateful to the nurses and doctors in charge of diabetes care at the paediatric departments in Jönköping, Kalmar, Karlskrona, Norrköping, Västervik, and Växjö for providing patient data. Special thanks to Professor J Carstensen for statistical advice.

Authors' affiliations

U Samuelsson, Division of Paediatrics, Department of Molecular and Clinical Medicine, University Hospital, Linköping, Sweden

O Löfman, Centre for Public Health Sciences, University Hospital, Linköping

Funding: the study was supported by the Swedish Child Diabetes Foundation (Barndiabetesfonden) and FORSS (Research council in south east Sweden)

REFERENCES

- Rytönen M, Ranta J, Tuomilehto J, for the SPAT Study Group and the Finnish Childhood Diabetes Registry Group, *et al.* Bayesian analysis of geographical variation in the incidence of type 1 diabetes in Finland. *Diabetologia* 2001;**44**(suppl 3):B37-44.
- Patterson CC, Smith PG, Webb J, *et al.* Geographical variation in the incidence of diabetes mellitus in Scottish children during the period 1977-1983. *Diabetic Med* 1988;**5**:160-5.
- Cherubini V, Carle F, Gesuita R, *et al.* Large incidence variation of type 1 diabetes in central-southern Italy 1990-1995: lower risk in rural areas. *Diabetologia* 1999;**42**:789-92.
- McKinney PA, Law GR, Bodansky HJ, *et al.* Geographical mapping of childhood diabetes in the northern English county of Yorkshire. *Diabetic Med* 1996;**13**:734-40.
- Parslow RC, McKinney PA, Law GR, *et al.* Incidence of childhood diabetes in Yorkshire, northern England, is associated with nitrate in drinking water: an ecological analysis. *Diabetologia* 1997;**40**:550-6.
- Staines A, Bodansky HJ, McKinney PA, *et al.* Small area variation in the incidence of childhood insulin-dependent mellitus in Yorkshire, UK: links with overcrowding and population density. *Int J Epidemiol* 1997;**26**:1307-13.
- Bodington MJ, Muzulu SI, Burden AC. Spatial clustering in childhood diabetes: evidence of an environmental cause. *Diabetic Med* 1995;**12**:865-7.
- Patterson CC, Carson DJ, Hadden DR. Epidemiology of childhood IDDM in northern Ireland 1989-1994: low incidence in areas with highest population density and most household crowding. *Diabetologia* 1996;**39**:1063-9.
- Patterson CC, Waugh NR. Urban/rural and deprivation differences in incidence and clustering of childhood diabetes in Scotland. *Int J Epidemiol* 1992;**21**:108-17.
- Songini M, Bernardinelli L, Clayton D, *et al.* The Sardinian IDDM study: 1. Epidemiology and geographical distribution of IDDM in Sardinia during 1989 to 1994. *Diabetologia* 1994;**41**:221-7.
- Samuelsson U, Carstensen J, Johansson C, *et al.* Time-space-clustering in insulin-dependent diabetes mellitus (IDDM) in south-east Sweden. *Int J Epidemiol* 1994;**23**:138-42.
- Law GR, McKinney PA, Staines A, *et al.* Clustering of childhood IDDM: links with age and place of residence. *Diabetes Care* 1997;**20**:753-6.
- Dahlquist G, Källen BAJ. Time-space clustering of date at birth in childhood-onset diabetes. *Diabetes Care* 1996;**19**:328-32.
- Karlsson MGE. The importance of cell-mediated immunity for the development of type 1 diabetes. [Thesis number 644]. Linköping: Linköping University, 2000.
- Samuelsson U, Johansson C, Ludvigsson J. Month of birth and risk of developing insulin dependent diabetes in south east Sweden. *Arch Dis Child* 1999;**81**:143-6.
- Dahlquist G, Blom L, Hägglöf B, *et al.* The epidemiology of diabetes in Swedish children 0 - 14 years: a six year prospective study. *Diabetologia* 1985;**28**:802-8.
- Dahlquist G, Mustonen L. Childhood onset diabetes-time trends and climatological factors. *Int J Epidemiol* 1994;**23**:1234-41.
- Kohli S, Noorlind Brage H, Löfman O. Childhood leukaemia in areas with different radon levels: a spatial and temporal analysis using GIS. *J Epidemiol Community Health* 2000;**54**:822-6.
- Cressie N. Geostatistical methods for mapping environmental exposures. In: Elliot P, Wakefield J, Best N, *et al.*, eds. *Spatial epidemiology. Methods and applications*. Oxford: Oxford University Press, 2000.
- Kleinschmidt I, Bagayoko M, Clarke GP, *et al.* A spatial statistical approach to malaria mapping. *Int J Epidemiol* 2000;**29**:355-61.
- Lie BA, Todd JA, Pociot F, *et al.* The predisposition to type 1 diabetes linked to the human leukocyte antigen complex includes at least one non-class II gene. *Am J Hum Genet* 1999;**64**:793-800.
- Eisenbarth GS. Type 1 diabetes: a chronic autoimmune disease. *N Engl J Med* 1986;**314**:1360-8.
- Nyström L, Dahlquist G, Östman J, *et al.* Risk of developing insulin-dependent diabetes mellitus (IDDM) before 35 years of age: indications of climatological determinants for age at onset. *Int J Epidemiol* 1992;**21**:352-8.
- Casu A, Carlini M, Contu A, *et al.* Type 1 diabetes in Sardinia is not linked to nitrate levels in drinking water. *Diabetes Care* 2000;**23**:1043-4.
- Van Maanen JMS, Albering HJ, Van Breda SGJ, *et al.* Nitrate in drinking water and risk of childhood diabetes in the Netherlands. [Letter]. *Diabetes Care* 1999;**22**:1750.
- Centre for Public Health Sciences. *Health and environment, population survey, report 2002*. Linköping: Centre for Public Health Sciences, Linköping University, 2002.
- Haglund B, Ryckenberg K, Selinus O, *et al.* Evidence of a relationship between childhood-onset type 1 diabetes and low groundwater concentration of zinc. *Diabetes Care* 1996;**19**:873-5.
- Stene LC, Hongve D, Magnus P, *et al.* Acidic drinking water and risk of childhood-onset type 1 diabetes. *Diabetes Care* 2002;**25**:1534-8.