Time-Space Clustering of Date at Birth in Childhood-Onset Diabetes

GISELA G. DAHLQUIST, MD
BENGT A.J. KÄLLEN, MD

OBJECTIVE — To investigate whether there was a temporal and geographical clustering of time of birth for infants with childhood-onset diabetes.

RESEARCH DESIGN AND METHODS — The nationwide Swedish Childhood Diabetes Registry, which ascertains 99% of children with recent-onset diabetes (0–14 years), was linked with the Swedish Medical Birth Registry. Clustering of 3,725 patients as to place and time of birth was studied compared with the general population. For each municipality (and in the three large cities of Sweden for each parish), the observed number of patients was compared with the expected number calculated from the average total rate and the number of births in that municipality. Clustering in time of birth within municipality was analyzed using a modification of a set technique by Chen (14).

RESULTS — There was no consistent variability in diabetes risk by calendar birth month, but for specific years, the risk varied during the year. When geographic localization for place of birth was studied on a municipality level, four municipalities showed a statistically significant case excess while one would have been expected by chance. When we looked for clusters in both time and space for date of birth, clearly more clusters than expected were identified ($P < 0.01$). Of the total of 198 primary clusters, 42 included three or more patients being born in the same municipality within an unlikely short period always <2 years.

CONCLUSIONS — This is the first study indicating a clustering according to place and time of birth for later risk to develop type I diabetes. Such a phenomenon would agree with the hypothesis that infections in early life, including fetal infections, can increase the risk for diabetes.

Almost all cases of childhood-onset diabetes are of the autoimmune type (type I). Follow-up studies of first-degree relatives of type I diabetic probands have shown that signs of autoimmunity such as antibodies to islet cells or β-cells are detectable long before the onset of clinical disease (1–5). In childhood-onset cases, exposures that may initiate the autoimmunity may be found already during fetal and perinatal life. The association between maternal-child blood group incompatibility and other perinatal events and risk for childhood-onset diabetes supports this (6).

Viruses have long been suggested to be of importance in the etiology of type I diabetes, mainly on the basis of animal experiments (7). In humans, the most compelling evidence for an association between virus exposure and diabetes was the finding of a very high prevalence of diabetes (30%) in a cohort of children who were followed-up because of the rubella embryopathy syndrome (8). Similarly, a case report suggested an association between congenital cytomegalovirus infection and type I diabetes (9). A recent population-based case-control study also indicated that maternal enteroviral infection during pregnancy is a risk factor for childhood-onset type I diabetes (10). If pre- or perinatal exposures, e.g., for viruses, are associated with an increased risk for childhood-onset diabetes, a clustering of birth dates would be detectable in time and/or space. Also exposures in early postnatal life could give such clustering.

In the present study we have used nationwide Swedish registers to study temporal and geographical clustering of time of birth for infants with childhood-onset diabetes.

RESEARCH DESIGN AND METHODS

— The Swedish Childhood Diabetes Registry records all children in whom type I diabetes was diagnosed before the age of 15 and after 1 July 1977 (11). This registry has a coverage of 99% (11,12), and children born between 1 January 1973 and 31 December 1990 with a diabetes diagnosis before 31 December 1992 were included. The Swedish Medical Birth Registry (13) records all births in Sweden since 1 January 1973. The coverage is ~99%.

The two registries were linked using the unique personal identification number each citizen gets shortly after birth. In this way, the place of birth of each child was determined, defined as county or municipality (and for the three largest cities, parish).

Linkage failed in some cases. Among a total of 3,862 children in the Diabetes Registry, 3,725 matched (96.4%). The explanation for nonmatching is absent identification numbers in either register. Sometimes the hospital reporting to the Childhood Diabetes Registry refused to give the complete identification of the child (57 cases) or the child or the identification number is missing in the Medical Birth Registry (80 cases). The latter is obtained by linkage with the official birth statistics using the maternal identification number; if this is incorrect or missing, no linkage will be obtained.

The following analyses were performed.

1. Temporal changes according to time of birth. The rate of registered infants according to year and month of birth was calculated based on the total number of children born and the number of children in the linked diabetes file born that month.

2. Geographical distribution of the birthplace of patients over the

From the Department of Pediatrics and Department of Epidemiology and Public Health (G.G.D.), Umeå University, Umeå; and the Tornblad Institute (B.A.J.K.), University of Lund, Lund, Sweden.

Address correspondence and reprint requests to Gisela Dahlquist, MD, Department of Pediatrics, University Hospital, S-901 85 Umeå, Sweden.

Received for publication 18 July 1995 and accepted in revised form 9 November 1995.

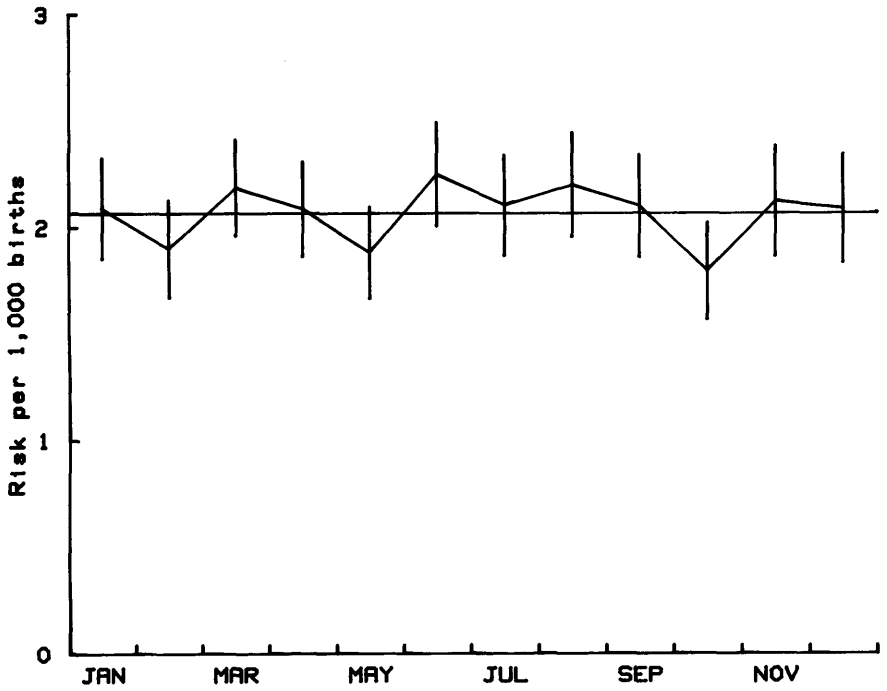

Figure 1—*Average risk to be recorded in the Childhood Diabetes Registry (per 1,000 births) according to month of birth. Vertical bars indicate 95% CI.*

whole time period. For each municipality (and in the three large cities, Stockholm, Gothenburg, and Malmö, for each parish), the observed number of cases was compared with the expected number and calculated from the average total rate and the number of births in that municipality (parish).

3. Clustering in time of birth within municipality (for the three large cities, parish). The set technique introduced by Chen (14) for the monitoring of congenital malformations was used. The technique is based on the following reasoning. A set is defined as all births between two births with the event under study, in this case later childhood diabetes. The first child with diabetes within a municipality also delimits a set. The length of the set will distribute randomly around a mean determined by the probability of the event to occur. If n nonevent births have occurred between the two event births, and the probability for the event for all births is p, then the probability for the set to be randomly short is $P = [1 - (1 - p)^n]$. If the P value is below a certain level, the second case is regarded as the result of a clustering.

In the present situation, however, the value of p depends on the follow-up time. Because all children were followed to the end of 1992, the follow-up time will be determined solely by the date of birth. Therefore, for each year and month of birth, the risk of a child to be registered in the Childhood Diabetes Registry was determined, and this risk was applied to each child born in the set. If a child entered the set with the same date of birth as the child with diabetes, its p value was halved under the assumption that half of the nondiabetic children born that day were born before and half after the diabetic child.

If the risk for the ith child is p, then the probability for the set to be randomly short is

$$P = 1 - (1 - p_1) *(1 - p_2)* \cdots *(1 - p_{(n-1)}) * (1 - p_n)$$

A primary cluster was defined as a set that was so short that its P value was <0.05. As 3,715 cases were studied, one would randomly expect 186 such clusters. A second definition of a cluster is if the P value of the set is <0.01, 37 such clusters should be expected to occur by chance.

A table was then prepared in which for each municipality and year of birth, the number of primary clusters, the total number of clusters, and the total number of cases were presented. From this list, final clusters were selected when two or more clusters were close together in time (the same or adjacent years). The final identification of a time-space cluster is thus made from three (or more) patients being born in the same municipality within an unlikely short period always <2 years.

RESULTS

Clustering in time

Figure 1 shows the risk according to month of birth. All children included in the Swedish Childhood Diabetes Registry by 31 December 1992 and born in 1973–1990 are included. There is no significant variability in risk between months; only the October rate is below the average risk, but as 12 months are studied, this may well be random.

Figure 2*A* and *B* shows the risk for each month and year of birth. The overall probability to be recorded in the Childhood Diabetes Registry varies for each birth cohort with a maximum for children born in 1976, a direct result of the different follow-up times and the noninclusion of children born in 1973–1976 and in whom diabetes was diagnosed before 1 July 1977. Within each year, a variability in risk is seen between months, which might reflect time-dependent clusters.

Clustering in space

For each one of the 283 municipalities in Sweden (except the three large cities of Stockholm, Gothenburg, and Malmö), the observed number of diabetes cases was compared with the expected number, estimated from the average risk (2.14 per 1,000 live births surviving the perinatal period). We found four municipalities in which the case excess reached a statistical significance of $P < 0.005$. Randomly only one would be expected, since we have between 200 and 300 municipalities. The four municipalities were Vaggeryd–15 observed among 2,715 births, expected number 5.82, $P = 0.0011$; Vetlanda–26 observed among 5,876 births, expected number 12.60, $P = 0.0006$; Hagfors–17 observed among 2,970 births, expected

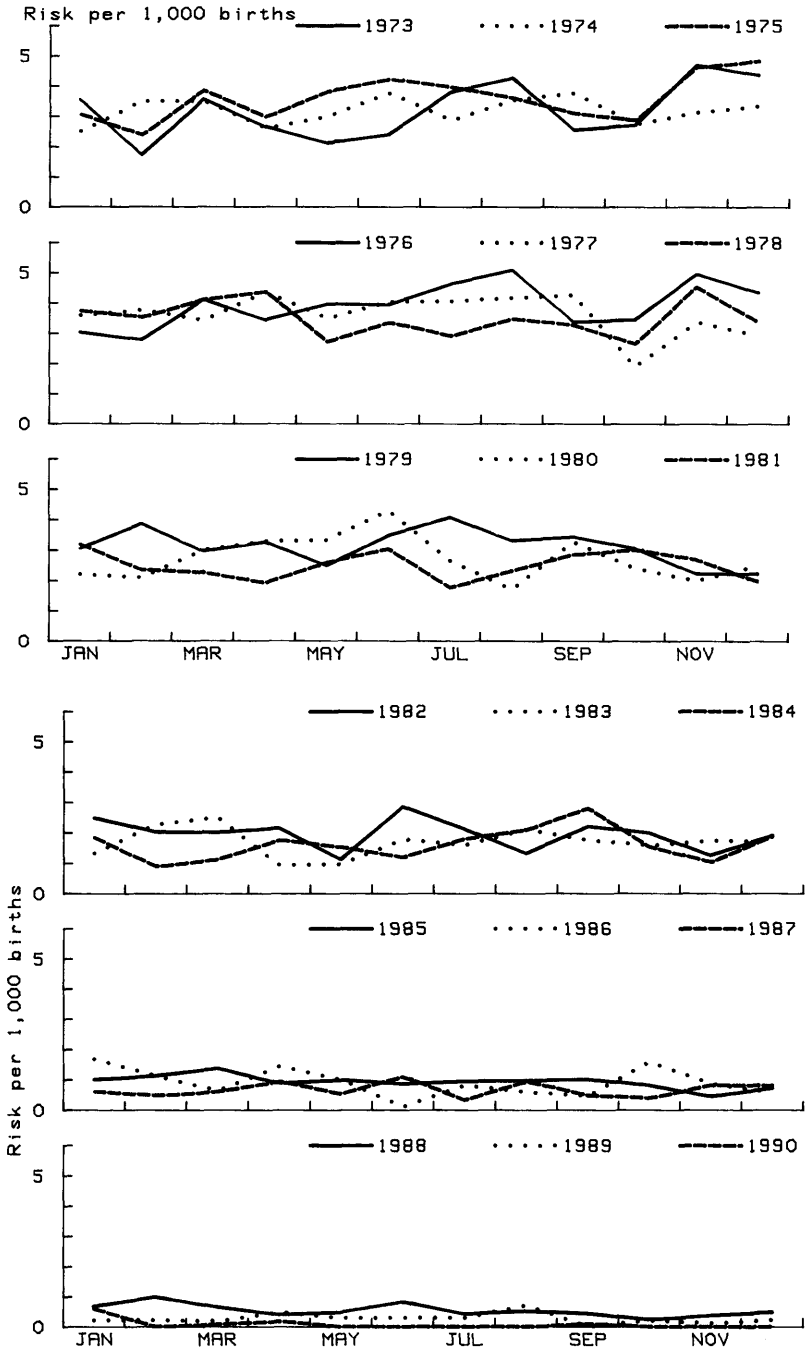

Figure 2—*Risk to be recorded in the Childhood Diabetes Registry (per 1,000 births) according to month of birth each year of birth. A: 1973–1981. B: 1982–1990.*

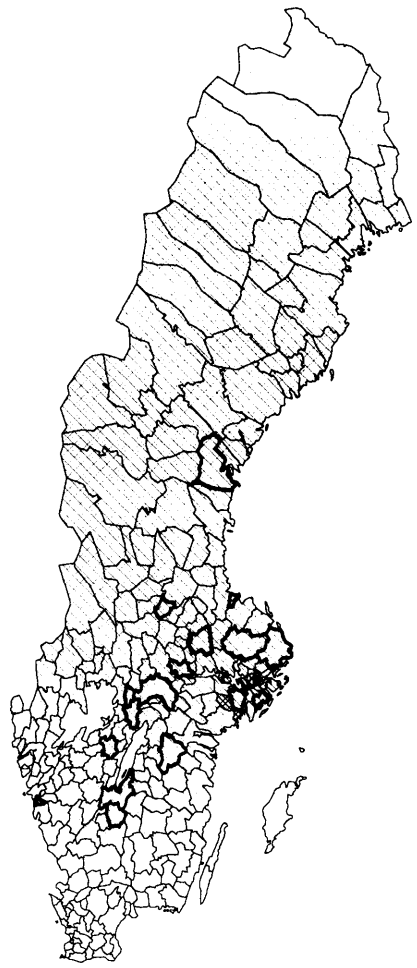

Figure 3—*Map of Sweden with the location of the 42 clusters marked.*

number 6.37, $P = 0.0003$; and Vilhelmina–10 observed among 1,939 births, expected number 4.16, $P = 0.0010$. All four thus show P values around or below 0.001. Taking into consideration multiple testing, we calculated the probability of randomly finding four municipalities among 283 with P values of 0.001, which is small ($P_0(0.238) = 4$, $P = 0.0002$). It is therefore likely that at least some of these four represent true geographic high-risk areas. The four municipalities are located in different parts of Sweden.

For the three large cities, a corresponding tabulation was made on the parish level. Among a total of 77 parishes, only one reached a P value close to 0.01 (11 infants with diabetes among 2,280 born, expected number 4.89, $P = 0.012$). It is rather likely to reach one such P level in 77 samples ($P = 0.61$), and we thus have no evidence for a geographical uneven distribution within the three large cities.

Time and space clustering

Among 3,264 diabetes cases (the three large cities removed), 198 showed primary clusters defined as having a P value of <0.05. The expected number is 163.2, and the excess of 34 primary clusters is statistically significant ($\chi^2 = 7.0$, $P < 0.01$). After scrutiny of all clusters showing the characteristics listed under METHODS (repeated clusters in the same municipality close together in time), 42 final clusters were identified. Table 1 lists these clusters, which were tentatively regarded as true, and their locations are marked in Fig. 3. It should be stressed that the excess of primary clusters (two diabetic infants born too close together in time within a

Table 1—*Listing of 42 clusters identified as tentatively true because of close similarity in time of two clusters in the same municipality*

Municipality	Year(s)	No. of clusters	Cluster strength expressed as *P* value
Haninge	1974, 1975	2	0.023, 0.039
Täby	1979, 1980	2	0.029, 0.006
Södertälje	1977	3	0.030, 0.044, 0.015
Norrtälje	1979, 1980	2	0.026, 0.029
Älvkarleby	1975	2	0.028, 0.024
Uppsala	1974, 1975	2	0.015, 0.038
Linköping	1976, 1979	4	0.048, 0.015, 0.023, 0.029
Vaggeryd	1973	2	0.044, 0.023
Jönköping	1976, 1977	3	0.018, 0.049, 0.014
Mölndal	1979	2	0.042, 0.034
Skövde	1985, 1986	2	0.009, 0.030
Laxå	1975	2	0.023, 0.028
Degerfors	1973, 1974	2	0.043, 0.011
Örebro	1975, 1976, 1979, 1979	4	0.023, 0.040, 0.044, 0.042
Sala	1978	2	0.034, 0.005
Köping	1989	2	0.006, 0.034
Borlänge	1975, 1976	2	0.027, 0.009
Sundsvall	1974, 1975	2	0.018, 0.032

municipality) is statistically significant; the selection of the 42 final clusters is based on the assumption that two primary clusters close together in time and space are more likely true than two primary clusters separated in time or space.

CONCLUSIONS — This is the first study to show a clustering in both time and space of birth dates of children who later developed diabetes. In a study from Iceland, it was shown that birth dates of diabetic children clustered with time (15), and it was suggested that maternal intake of nitrosamine-rich food during pregnancy was a risk factor for type I diabetes. In our study, we found no general variability in risk with calendar month of birth. Within each year of birth, however, an indicated variability by month of birth could be seen, varying between the years. This may indicate a variability consistent with, for example, viral epidemics. A support for such a hypothesis would be a demonstration of clustering with time and space.

The first method to study time and space clustering was that presented by Knox (16), based on the computing of pair-wise distances of cases in time and space. The definition of a critical distance in space and time is arbitrarily chosen. The background population was not taken into account. This method has been revised repeatedly (17,18), but even so, no account is taken of a possible variability in the background population. The technique used in this study overcomes this problem. It is based on the set technique introduced by Chen (14) for surveillance of birth defects. Because the probability to be recorded in the Swedish Childhood Diabetes Registry will vary between different birth cohorts, we modified the technique, taking into account for each month and year of birth the average risk of a child to be registered.

Because multiple comparisons were made, many random clusters will appear. Using the cutoff of $P < 0.05$ for a clustering, one would expect 163 clusters to arise randomly, and 198 were observed–a statistically significant excess. By scrutiny of the observed possible clusters, 42 were identified, at least some of which were regarded as possibly true.

In a previous study, we tried to identify birth cohort effects on Swedish childhood diabetes patients using Poisson regression modeling (19) but were unable to find any clear cohort effects. In that study, however, birth cohorts were analyzed in 2-year intervals, and clusters occurring by month of the year may thus have been hidden. In a previous study, we found significant geographic variability between counties in the onset rates of childhood diabetes (12). A geographic variability may indicate differences in population genetics or in risk exposures.

The present study for the first time gives evidence of a birth date time and space clustering of children who later develop diabetes. It certainly cannot reveal the mechanisms behind the phenomenon but is in accordance with previous evidence (8–10) that viral epidemics, which may cause fetal infection, can be involved in the initiation of autoimmunity that subsequently can lead to type I diabetes. Other pre- or perinatal or even early postnatal exposures could also cause a clustering of dates of birth.

Acknowledgments— This study was supported by grants from the Swedish Medical Research Council (Project 07531), the Nordic Insulin Foundation, and the Swedish Diabetes Association.

References

1. Srikanta S, Ganda OP, Jackson RA, Gleason RE, Kaldany A, Garovoy MR, Milford El, Carpenter CB, Soeldner JS, Eisenbarth GS: Type 1 diabetes mellitus in monozygotic twins: chronic progressive β-cell dysfunction. *Ann Intern Med* 99:320–326, 1983
2. Tarn AC, Thomas JM, Dean BM, Ingram D, Schwartz G, Bottazzo GF, Gale EAM: Predicting insulin-dependent diabetes. *Lancet* i:845–850, 1988
3. Palmer JP, McCulloch DK: Prediction and prevention of IDDM: 1991. *Diabetes* 40: 943–947, 1991
4. Riley WJ, MacLaren NK, Krischer J, Spillar J, Silverstein D, Schatz S, Schwartz J, Malone S, Shaw C, Vadheim C, Rotter J: A prospective study of the development of diabetes in relatives of patients with insulin-dependent diabetes. *N Engl J Med* 323: 1167–1172, 1990
5. Tun RYM, Peakman M, Alviggi L, Hussain MJ, Loss SS, Shattock M, Pyke DA, Bottazzo GF, Vergani D, Leslie RDG: The importance of persistent cellular and humoral immune changes in the pre-diabetic period: a prospective identical twin study. *Br Med J* 308:1063–1068, 1994
6. Dahlquist G, Källén B: Maternal-child blood group incompatibility and other perinatal events increase the risk for early-onset type 1 (insulin-dependent) diabetes mellitus. *Diabetologia* 35:671–675, 1992
7. Yoon JW: Role of viruses in the pathogenesis of IDDM. *Ann Med* 23:437–445, 1991

8. Menser MA, Forrest JM, Bransby RD: Rubella infection and diabetes mellitus. *Lancet* i:57–60, 1978
9. Ward KP, Galloway WH, Auchterlonie JA: Congenital cytomegalovirus and diabetes (Letter). *Lancet* i:497, 1979
10. Dahlquist G, Ivarsson S, Lindberg B, Forsgren M: Maternal enteroviral infection during pregnancy as a risk factor for childhood IDDM: a population-based case-control study. *Diabetes* 44:408–413, 1995
11. Dahlquist G, Gustavsson KH, Holmgren G, Hägglöf B, Larsson Y, Nilsson KO, Samuelson G, Sterling G, Thalone B, Wall S: The incidence of diabetes mellitus in Swedish children 0–14 years of age. *Acta Paediatr Scand* 71:7–14, 1982
12. Dahlquist G, Blom L, Holmgren G, Hägglöf B, Larsson Y, Sterky G, Wall S: The epidemiology of diabetes in Swedish children 0–14 years: a six year prospective study. *Diabetologia* 28:802–808, 1985
13. Cnattingius S, Ericson A, Gunnarskog J, Källén B: A quality study of Medical Birth Registry. *Scand J Soc Med* 80:143–148, 1990
14. Chen R: A surveillance system for congenital malformations. *J Am Stat Assoc* 73: 323–327, 1978
15. Helgasson T, Jonasson MR: Evidence for a food additive as a cause of ketosis-prone diabetes. *Lancet* ii:716–720, 1981
16. Knox EG: The detection of space-time interactions. *Appl Stat* 3:25–29, 1964
17. Mantel N: The detection of disease clustering and generalized regression approach. *Cancer Res* 27:209–220, 1967
18. Siemiatycki J: Mantel's time-space clustering statistic: computing higher moments and a comparison of various data transforms. *J Stat Compat Simul* 7:13–31, 1978
19. Nyström L, Dahlquist G, Rewers M, Wall S: The Swedish childhood diabetes study: an analysis of the temporal variation in diabetic incidence 1978–1987. *Int J Epidemiol* 19:141–146, 1990