PERGAMON

Social Science & Medicine 57 (2003) 551–560

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Geographic analysis of diabetes prevalence in an urban area

Chris Green[a,*], Robert D. Hoppa[b], T. Kue Young[c], J.F. Blanchard[a]

[a] *Epidemiology Unit, Manitoba Health, Room 4050, 300 Carleton Street, Winnipeg, Man., Canada R3B 3M9*
[b] *Department of Anthropology, University of Manitoba, Winnipeg, Man., Canada R3T 5V5*
[c] *Department of Community Health Sciences, University of Manitoba, 750 Bannatyne Avenue, Winnipeg, Man., Canada R3E 0W3*

Abstract

The objective of this research is to identify the sociodemographic, environmental, and lifestyle factors associated with the geographic variability of Diabetes Mellitus (DM) prevalence in the City of Winnipeg, Manitoba in Canada. An ecological regression study design was employed for this purpose. The study population included all prevalent cases of DM in 1998 for Winnipeg. Predictor and outcome data were aggregated for analysis using two methods. First, the spatial scan statistic was used to aggregate study data into highly probable diabetes prevalence clusters. Secondly, predictor and outcome data were aggregated to existing administrative health areas. Analysis of variance and spatial and non-spatial linear regression techniques were used to explore the relationship between predictor and outcome variables. The results of the two methods of data aggregation on regression results were compared. Mapping and statistical analysis revealed substantial clustering and small-area variations in the prevalence of DM in the City of Winnipeg. The observed variations were associated with variations in socioeconomic, environmental and lifestyle characteristics of the population. The two methods of data aggregation used in the study generated very similar results in terms of identifying the geographic location of DM clusters and of the population characteristics ecologically correlated to those clusters. High rates of DM prevalence are strongly correlated with indicators of low socioeconomic status, poor environmental quality and poor lifestyles. This analysis further illustrates what a useful tool the spatial scan statistic can be when used in conjunction with ecological regression to explore the etiology of chronic disease.

Keywords: Diabetes epidemiology; Spatial analysis; Spatial scan statistic; Spatial regression; Racial covariate; Canada

1. Introduction

Type 2 Diabetes Mellitus (T2DM) is one of the most common non-communicable diseases in the world today (Amos, McCarty, & Zimmet, 1997). It is projected that the number of cases of T2DM around the world will increase rapidly over the next 25 years, from 154 million estimated cases in 2000 to 300 million cases in 2025 (King, Aubert, & Herman, 1998). There is great debate about the cause of the T2DM epidemic (Swinburn, 1996). Although there is general consensus that T2DM has both genetic and social roots, there is little consensus on the relative contribution of these factors (Carter, Pugh, & Monterrosa, 1996; Fujimoto, 1996; Haffner, 1998; Hales & Barker, 1992; Hales, Desai, & Ozanne, 1997; McDermott, 1998; Ozanne & Hales, 1998).

This study used two spatial techniques to explore the geographic variability of Diabetes Mellitus (DM) prevalence in the City of Winnipeg, Manitoba. Since 95% of all cases of DM are estimated to be T2DM (Harris, 1995), DM prevalence was used in this study as a proxy for T2DM. A common problem in geographic epidemiology is that observed rates, especially in low incidence or prevalence situations, can often be artefacts of the areal geographic units to which individual events are aggregated for analysis. This can have the effect of rendering invisible small geographic areas that have

*Corresponding author. Tel.: +1-204-774-7144; fax: +1-204-948-2190.

E-mail addresses: chrisgreen@mb.sympatico.ca (C. Green), hoppard@Ms.UManitoba.CA (R.D. Hoppa), tkyoung@Ms.UManitoba.CA (T.K. Young), james_blanchard@umanitoba.ca (J.F. Blanchard).

0277-9536/03/$ - see front matter
PII: S0277-9536(02)00380-5

significantly elevated rates of disease (Meade & Earickson, 2000). The spatial scan statistic, the first geographic method used in this study, avoids this problem by iteratively creating a number of statistically significant high- and low-rate cluster areas from small geographic regions. The spatial scan has been used in a number of recent studies to identify spatial clusters of cancer (Hjalmars & Gustafsson, 1999; Hjalmars, Kulldorff, Gustafsson, & Nagarwalla, 1996; Hjalmars, Kulldorff, Wahlqvist, & Lannering, 1999; Jemal, Devesa, Kulldorff, Hayes, & Fraumeni, 2000; Kulldorff, Athas, Feurer, Miller, & Key, 1998a; Kulldorff, Feuer, Miller, & Freedman, 1997; Kulldorff & Nagarwalla, 1995), child-hood mortality (Sankoh, Ye, Sauerborn, Muller, & Becher, 2001), aviation crashes (Grabowski & Li, 2001) and acute respiratory disease in cattle(Norstrom, Pfeiffer, & Jarp, 1999). This spatial scan method is compared to the more traditional approach of aggregating event data to pre-existing and large geographic administrative areas.

Materials and methods

Study setting

The study was conducted in the City of Winnipeg, Manitoba, Canada. Winnipeg has a population of 632,000 and is the only large metropolitan city in the province. Over the past 20 years, Winnipeg has experienced significant social and physical deterioration of its central core and downtown area, paired with rapid growth of its peripheral suburbs. The majority of the population historically have been of European descent. However, because of in-migration from rural communities and natural population increases, an increasing percentage of the population is of Aboriginal descent. Manitoba has a universal health insurance plan and all residents of the province are eligible to receive health care services without cost.

Data sources

Sociodemographic data including self-reported Aboriginal status were obtained from the 1996 Census Canada microdata files. Data on the quality of the physical and social environment in 1999 were obtained from the City of Winnipeg. Data on smoking rates and DM prevalence data for 1998 were obtained from the Manitoba Health Epidemiology Unit. Definitions for the sociodemographic, lifestyle and environmental predictor variables are included in Table 2. The methodology used to generate population-based DM prevalence estimates has been described previously (Blanchard et al, 1996). This method used the standardized case definition of either one hospital visit or two physician visits with a DM diagnosis (ICD 250) within a 2-year period in order to generate DM incidence and prevalence estimates from hospital and physician claims data. This methodology was unable, however, to distinguish between T2DM and DM. Hospital and physician claims data were available for all residents of Manitoba since the entire population is covered by a universal health care program. Population denominator data were obtained from the Manitoba Health population registry of all citizens insured for health services in the province. All data were aggregated initially to the neighborhood level ($n=230$) using the geocoding functionality within the GIS software Arc-View 3.2 (Environmental Systems Research Institute, 1999).

Spatial methods

Two methods were used to explore the geographic variability and clustering of DM within the City of Winnipeg and to identify, using ecological methods, the social and environmental factors associated with variability in DM. The first is the spatial scan method and the second is the pre-existing regions method. Linear regression and analysis of variance were used with both these methods to identify the socioeconomic, environmental and lifestyle factors ecologically associated with this variability. These factors included self-reported Aboriginal status, education, income, family structure, unemployment, housing conditions, crime and smoking rates.

Spatial scan method:

The spatial scan statistic was used to test for the presence of clusters of DM and to identify their approximate location. The open domain software Statscan distributed by the National Cancer Institute was employed for this purpose (Kulldorff, Rand, Gherman, Williams, & DeFrancesco, 1998b). The spatial scan statistic, which works by aggregating together the unique combinations of small-area geographies which have a high probability of being clusters is an especially powerful tool to use in low-prevalence and low-incidence situations. Traditional epidemiological approaches which require that rare events be aggregated to pre-existing higher-level geographies can often mask the existence of real clusters. The statistic assumes the number of cases in each geographic region to be Poisson distributed. The method tests the null hypothesis that within any age and gender group, the risk of having DM is the same as in all regions combined. This means that the expected age and gender-adjusted prevalence rate is constant over the whole area.

The spatial scan statistic places a circular window of varying size on the map surface and allows its center to move so that at any given position and size, the window includes different sets of adjacent neighborhoods. If the

window contains the centroid of a neighborhood, then the whole neighborhood is included in the window. As the window is placed at each neighborhood centroid, its radius is varied continuously from zero up to a maximum radius which never includes more than 50% of the total population. The method creates a large number of distinct circular windows, each containing a distinct set of adjacent neighborhoods, and each a possible candidate for containing a cluster of prevalent diabetes cases. For each window, the method uses a Monte Carlo simulation to test the null hypothesis that there is an elevated risk of DM prevalence. The Statscan software allows any number of covariates to be implemented into the model, and calculates indirectly standardized rates. Details of how the likelihood function is maximized over all windows under the Poisson assumption have been described elsewhere (Kulldorff et al., 1997).

In this study, the Statscan software was applied to the 230 Winnipeg neighborhoods in order to generate possible clusters of DM prevalence. Age and gender were applied as covariates. Two iterations were undertaken. The first iteration used the default setting within the Statscan software which maximizes the cluster size at 50% of the total study population. The second iteration set the maximum generated cluster size at 10% of the total study population. Smaller maximum cluster sizes result in a larger number of smaller clusters with more extreme values. The Monte Carlo simulation used to test significance was set at 999 iterations. The software was set to generate both high and low clusters. Only statistically significant clusters were retained for analysis. Non-cluster areas were aggregated together into one cluster area and assumed to have a relative risk of 1.0. Clusters were initially mapped using Arc-View 3.2 (Environmental Systems Research Institute, 1999) in order to identify their physical location. Social and environmental predictor variables were then aggregated to the cluster areas in order to identify their possible relationship to DM prevalence. Where appropriate, analysis of variance and non-spatial and spatial linear regression techniques were used to formally explore the relationship between predictor variables and DM prevalence.

Pre-existing regions method

In the second method, DM prevalence and predictor data were further aggregated to the 23 Health regions used to organize the delivery of services within the City of Winnipeg. DM prevalence was directly standardized by age and gender to the 1998 Winnipeg population. Chloropleth maps of all variables were generated to visually examine their spatial distribution. Spatial clustering of all variables were assessed using the Moran's I statistic. Non-spatial and spatial regression techniques were used to explore the relationship between predictor variables and the standardized DM prevalence.

Analysis of variance and regression analysis

In both the spatial scan and pre-existing regions methods, variables were log transformed when necessary in order to ensure that the regression assumptions of normality and heteroskedascity were not violated. Regression analyses were undertaken using the S-PLUS Spatial Statistics extension for Arc-View 3.2 (Mathsoft, 1999). A simultaneous autoregressive model was used for spatial regressions. Analysis of variance was undertaken using NCSS (Hintz, 2000).

Results

In Winnipeg in 1998 there were 29,885 prevalent cases of DM, resulting in an overall DM prevalence rate of 47.3 cases/1000 population. Prevalence rates of DM were observed to be higher in men than in women and to increase rapidly in the age 65 and over age group.

Spatial scan results

Table 1 shows the results applying the spatial scan statistic. With the maximum cluster size set at 50% of the total study population, two significant ($P<0.001$) clusters were generated with relative risks of 1.3 and 0.84. Fig. 1 shows that the high relative risk cluster is located in the central and northern core of the City of Winnipeg, while the low relative risk cluster is located in the southern suburbs of the City. The high relative risk cluster is the most likely one, with a log likelihood ratio of 291.56.

With the maximum cluster size set at 10% of the total study population, the spatial scan statistic generated 10 significant ($P<0.009$) high and low clusters. Relative risks ranged from 0.69 to 1.45. Fig. 2 shows that the high relative risk clusters are again all located in the central and northern core of the City of Winnipeg, while the low relative risk clusters are located in the southern suburbs of the city. The most likely cluster, with a relative risk of 1.45, and a log likelihood ratio of 282.79 is located in the central core of the city.

Table 2 shows the predictor variables aggregated to cluster areas generated by the spatial scan method when the maximum cluster size was set to 50% of the total population. This table illustrates that high diabetes prevalence is clustering in those areas of the City of Winnipeg which have a high percentage of Aboriginal population, low educational levels, low family income, a high percentage of lone parent families, high levels of unemployment, high numbers of vacant and placarded houses, high levels of crime, and high rates of smoking.

Table 1
DM prevalence analysis, City of Winnipeg, Manitoba, 1998, using the spatial scan statistic

Max. cluster size	Cluster type[a]	Cases	Expected	RR[b]	LLR[c]	*p* value
A 50%	High	7335	5644	1.3	291.56	<0.001
	Non	12782	n/a	1.0	n/a	n/a
	Low	9768	11578	0.84	236.17	<0.001
B 10%	High	4101	2825	1.45	282.79	<0.001
	High	912	778	1.171	11.07	<0.009
	High	2056	1820	1.13	15.64	<0.001
	Non	17527	n/a	1.0	n/a	n/a
	Low	2457	2916	0.84	42.16	<0.001
	Low	677	840	0.81	17.47	<0.001
	Low	1799	236	0.76	78.66	<0.001
	Low	199	272	0.73	11.09	<0.009
	Low	362	509	0.71	23.96	<0.001
	Low	464	659	0.70	33.07	<0.001
	Low	243	353	0.69	19.74	<0 001

[a] Cluster type: High—cluster with relative risk > 1, Non—aggregation of non-clustered population, Low—cluster with relative risk < 1.
[b] RR: Relative risk—Observed DM prevalence/expected DM prevalence.
[c] LLR: Log likelihood ratio.

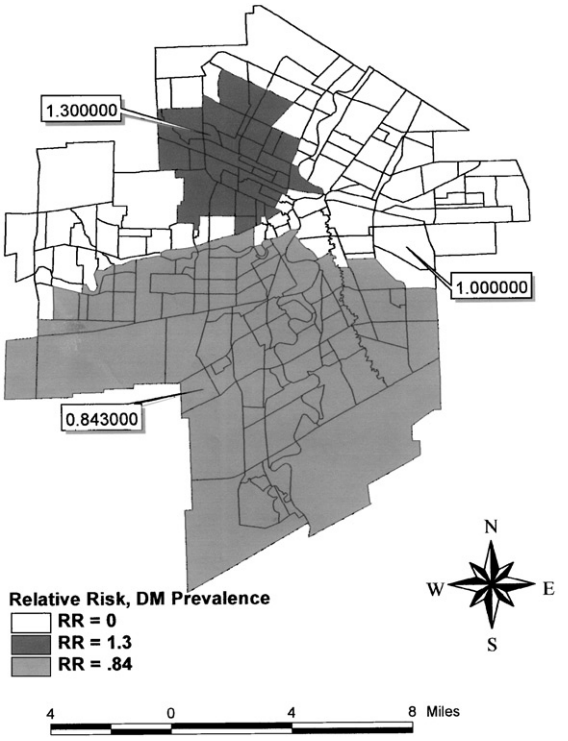

Fig. 1. DM Prevalence Analysis, City of Winnipeg, 1998, using the Spatial Scan Statistic, Maximum Cluster Size set at 50% of Study Population.

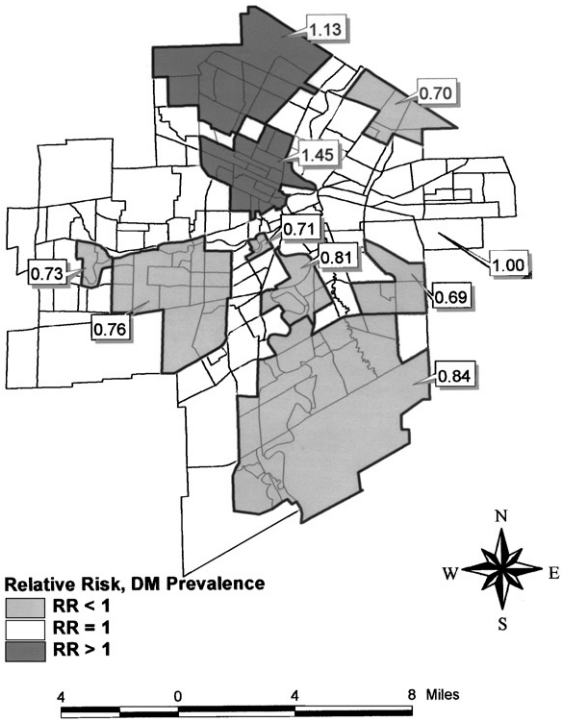

Fig. 2. DM Prevalence Analysis, City of Winnipeg, 1998, using the Spatial Scan Statistic, Maximum Cluster Size set at 10% of Study Population.

Analysis of variance was undertaken for all predictor variables. In all cases, the between cluster variance was significant at the $P<0.005$ level.

Regression analysis of predictor variables against the relative risk of the cluster areas generated when the maximum cluster size was set to 10% produced similar

Table 2
Analysis of variance, predictor variables aggregated to spatial scan generated cluster areas for DM prevalence, maximum cluster size set to 50%

Predictor[a]	Clusters			Analysis of variance
	Low cluster (RR[b] = 0.84)	Non-clustered (RR[b] = 1.0)	High cluster (RR[b] = 1.30)	
Aboriginal status	3.8	5.5	16.9	$p<0.005$
Less than grade 9	5.3	8.9	17.3	$p<0.005$
Average family income	62994	50810	37392	$p<0.005$
Lone parent	13.7	16.1	23.6	$p<0.005$
Unemployment	6.1	7.7	14.5	$p<0.005$
Vacant House	0.6	0.8	15.1	$p<0.005$
Crime	56.6	88.5	157.6	$p<0.005$
Smoking	18.1	26.7	35	$p<0.005$
DM cases by cluster	$n=9768$	$n=12,782$	$n=7335$	
Study population by cluster	$n=245528$	$n=258849$	$n=127623$	

[a] Aboriginal status—% of the population reporting aboriginal status; Less than grade 9—% of the population 15 yr + reporting less than grade 9 education; Average family income—average family income; Lone parent—% of families reporting being headed by a lone-parent; Unemployment—% of the population 15 + in the labor force that is unemployed; Vacant house—no. of houses/1000 residential properties that are vacant or placarded; Crime—no. of crimes against property and persons/1000 population; Smoking—% of mothers of newborns smoking on discharge from hospital.
[b] RR: Relative risk—observed DM prevalence/expected DM prevalence.

Table 3
Regression analysis of DM prevalence relative risk vs. predictor variables, using spatial scan generated cluster areas, maximum cluster size set to 10%

Predictor[a]	Non-spatial regression			Spatial regression		
	R[b]	Regression coefficient	p value	Regression coefficient	p value	Residual spatial autocorrelation[c]
Aboriginal Status	0.90	0.034	<0.001	0.0398	<0.001	N.S
Less Than Grade 9	0.97	0.0452	<0.001	0.0456	<0.001	N.S.
Average Family Income	−0.93	−0.0000139	<0.001	−0.000013	<0.001	N.S.
Lone Parent	0.89	0.0342	<0.001	0.0240	<0.001	N.S.
Unemployment	0.93	0.0504	<0.001	0.0558	<0.001	N.S.
Vacant House	0.69	0.0202	<0.001	0.0079	<0.098	$p<0.05$
Crime	0.90	0.00433	<0.001	0.0049	<0.001	N.S.
Smoking	0.88	0.218	<0.001	0.023	<0.001	N.S.

[a] For definitions, refer to footnote a of Table 2.
[b] R Pearsons R.
[c] Spatial autocorrelation of regression residuals. Significance is based upon the Moran's I statistic.

results (Table 3). Non-spatial regression resulted in very high Pearson R values ranging from 0.69 to 0.97, with all regressions significant at the $P<0.001$ level. Education had the greatest predictive value. Spatial regression, which accounts for the spatial clustering of variables in a regression equation, did not appreciably change either the non-spatial regression coefficients or significance levels. With one exception, regression equations generated using spatial regression techniques did not result in any residual spatial autocorrelation, indicating that the regression model was successful in fully accounting for any spatial correlation in the DM prevalence rates.

Pre-existing regions method

When aggregated to the 23 health regions, all variables used in the model were highly spatially clustered. Visual inspection of chloropleth maps showed clustering of DM prevalence in the central core of the City of Winnipeg (Fig. 3), associated with a larger Aboriginal population, low education, low family income, lone parent families, high unemployment, poor housing stock, high crime rates, and high rates of smoking. This visual impression was confirmed by significant Morans I values ($P<0.001$) for all values.

Standardized DM prevalence rates ranged from 37.7/1000 to 78.8/1000 and were significantly different from the mean in all but one region.

Regression of predictor variables against the standardized diabetes prevalence rates for the 23 health regions within the City of Winnipeg generated results similar in strength and direction to the spatial scan analysis. Non-spatial regression again resulted in very high Pearson R values (Table 4). All regression models were significant at the $p<0.001$ level. Unemployment had the greatest predictive value. Spatial regression did not appreciably change either the non-spatial correlation coefficients or significance levels and did not result in any residual spatial autocorrelation.

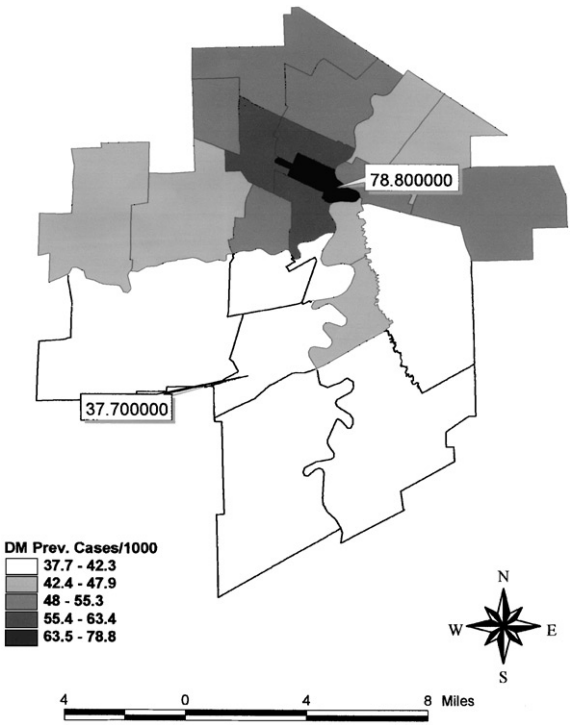

Fig. 3. Standardized Diabetes Prevalence Rates, City of Winnipeg, 1998, by Health Region.

Multiple regression of predictor variables against diabetes prevalence resulted in a model incorporating family income and unemployment, with a Pearsons R value of 0.944. Spatial regression analysis of these predictor variables did not result in any appreciable change in either the regression coefficients or significance levels and did not result in any residual spatial autocorrelation. In this model, diabetes prevalence was positively associated with unemployment and negatively associated with family income. Additional variables could not be incorporated into the model because of the high level of multi-collinearity of predictor variables (Table 5).

Discussion

This study has demonstrated substantial clustering and small-area variations in the prevalence of DM in the City of Winnipeg, and that these variations are associated with variations in socioeconomic, environmental and lifestyle characteristics of the population. This study has also demonstrated that two distinct approaches to spatial analysis, the spatial scan statistic and the pre-existing regions method generate very similar results in terms of identifying the geographic location of DM clusters and of the population characteristics ecologically correlated to those clusters. Finally, our results have shown that when high levels of

Table 4
regression analysis, age standardized DM prevalence rates vs. predictor variables, using existing health boundaries for the City of the Winnipeg, Manitoba

Predictor[a]	Non-spatial regression			Spatial regression		
	R^2[b]	Regression coefficient	p value	Regression coefficient	p value	Residual spatial autocorrelation[c]
Aboriginal Status	0.90	1.009	<0.001	0.969	<0.001	N.S.
Less Than Grade 9	0.90	1.584	<0.001	1.586	<0.001	N.S.
Avg. Family Income	−0.89	−0.0006	<0.001	−0.0006	<0.001	N.S.
Lone Parent	0.86	1.156	<0.001	1.153	<0.001	N.S.
Unemployment	0.92	1.682	<0.001	1.76	<0.001	N.S.
Vacant House	0.74	0.480	<0.001	0.4008	<0.001	N.S.
Crime	0.80	0.110	<0.001	0.1026	<0.001	N.S.
Smoking	0.84	0.729	<0.001	0.766	<0.001	N.S.

[a] For definitions, refer to footnote a of Table 2.
[b] Pearsons R.
[c] Spatial autocorrelation of regression residuals. Significance is based upon the Moran's I statistic.

Table 5
Covariance of predictor variables, Pearsons *R*, using existing health boundaries for the City of Winnipeg

	Abor	LTG9	Income	Lparent	Unemp	Vacant	Crime	Smoking
Abor	1.0	0.90	−0.82	0.94	0.98	0.83	0.86	0.89
LTG9	0.90	1.0	−0.87	0.87	0.90	0.74	0.77	0.86
Income	−0.82	−0.87	1.0	−0.89	−0.86	−0.78	−0.82	−0.87
Lparent	0.94	0.87	−0.89	1.0	0.96	0.85	0.90	0.93
Unemp	0.98	0.90	−0.86	0.96	1.0	0.83	0.90	0.88
Vacant	0.83	0.74	−0.78	0.85	0.84	1.0	0.89	0.79
Crime	0.86	0.77	−0.82	0.90	0.90	0.89	1.0	0.78
Smoking	0.89	0.86	−0.87	0.93	0.87	0.79	0.79	1.0

For definitions, refer to footnote a of Table 2.

non-spatial correlation exist between predictor and dependent variables, spatial regression approaches do not appreciably change the strength or direction of the regression coefficients.

The study has a number of methodological limitations. First, it has relied exclusively on data derived from administrative databases in order to estimate the DM prevalence rates. Since cases have not been individually verified, this approach could result in either an overestimate or underestimate of prevalence rates. However, we have previously studied the accuracy of this approach and found that the specificity is high when compared to local registries of DM (Blanchard et al., 1996). It is possible that some of the small-area variations that we observed are due to variability in health care access and diagnostic practices. This is unlikely, however, since Manitoba provides universal health care to its residents so restricted access to physician and hospital services is not likely.

Secondly, the administrative databases from which the diabetes prevalence rates were derived cannot distinguish between T2DM and Type 1 insulin-dependent diabetes. However, given that it is estimated that approximately 95% of all DM cases are T2DM it is likely that the variability in DM prevalence observed in this study reflects primarily the impact of T2DM (Harris, 1995).

Thirdly, the small number of observations used in regression analysis within both the spatial scan and pre-existing regions approaches means that regression results must be used with some caution. Tests for normality and heteroscadiscity may not have been sensitive to violations of regression assumptions because of the small number of observations. However, given the strength of the direction and significance of the generated correlation coefficients, and their consistency between the two spatial methods, the observed correlations are likely real and significant.

Fourthly, the ecological approach used in this study has been frequently criticized as being a weak design and commits what is known as the ecological fallacy. The ecological fallacy suggests that it is a mistake to apply characteristics measured at the scale of the population or geographic level to individuals living within those geographies or populations (Morgenstein, 1982, 1995). The ecological design used in this study therefore restricts us to making statements about the characteristics of the populations living in specific geographies. Statements made about individuals living within those geographies can only be made with caution. However, given the arguments by Rose and others (Rose, 1985, 1992; Wilkinson, 1996, 1999) on the primary importance of population and geographic level factors on population health, this study design legitimately provides important clues to the etiology of DM at the population level. It suggests that DM prevalence at the population level is powerfully graded by socioeconomic status, environmental quality, and lifestyle.

Finally, the study used only one lifestyle variable, smoking in mothers of newborns on discharge from hospital, as a proxy for overall lifestyle quality. Given that this variable may be a relatively weak proxy for lifestyle attributes relevant to DM, caution must be taken in concluding that lifestyle is associated with diabetes prevalence at the ecological level. Lifestyle measures more directly related to DM prevalence such as diet, exercise and obesity were not available at the geographic levels required for this study.

The high level of consistency between the results of the spatial scan statistic and the pre-existing regions method in identifying etiological factors associated with DM is encouraging. Previous studies that have used the spatial scan statistic to identify cancer clusters have not attempted to systematically explore possible etiological factors associated with clusters using analysis of variance and linear regression (Hjalmars et al., 1996, 1997, 1999; Jemal et al., 2000; Kulldorff et al., 1998a; Kulldorff and Nagarwalla, 1995; Sankoh et al., 2001; Walsh & Fenster, 1997). This study suggests that the spatial scan statistic in conjunction with analysis of

variance and linear regression may be a useful tool in exploring the etiology of cancer and other chronic diseases.

The relationship observed between DM prevalence and low levels of socioeconmic status, environmental quality and lifestyle at the geographic level is consistent with previous studies (Auslander W F, Haire-Joshu, Houston, & Fisher E B, 1992; Hanis, Chakraborty, Ferrell, & Schull, 1986; Hazuda & Monterrosa, 1992; Hendricks & Haas, 1991; Leonetti, Tsunehara, Wahl, & Fujimoto, 1992; Marshall et al., 1993) . This study provides some of the strongest evidence to date of this relationship, with DM prevalence estimates based upon diabetes prevalence estimates covering the whole population of Winnipeg. Previous studies have not been population based and were often restricted in scope to limited surveys of specific sub-populations.

This study demonstrates that the highest rates of DM are occurring in geographic areas that have the highest concentration of Aboriginal people. It has been hypothesized that populations of Aboriginal, Black, and Mexican American origin are genetically predisposed to develop T2DM supposedly due to the high frequency of the "thrifty gene" in their respective population gene pools. The thrifty gene, it is proposed, conferred an adaptive advantage in historical times of feast and famine. However, in modern conditions of relative plenty, the thrifty gene predisposes individuals to the development of obesity and increased frequency of DM (McDermott, 1998; Neel, 1962, 1982, 1999). In this study, it was observed that the geographic areas with the highest prevalence of DM also had the lowest socioeconomic status, the poorest lifestyles, and the lowest levels of environmental quality. Regression analyses demonstrated that broad neighborhood characteristics such as education and income were more predictive of DM prevalence than Aboriginal status. Once family income and unemployment were used in the regression analysis as predictors, Aboriginal status lost all of its significance as a predictor of DM. This suggests that it may be more the impact of low socioeconomic status that is putting populations at risk of DM in Winnipeg rather than genetic background. This also suggests that population-based studies using race as a covariate need to critically question their use of racial constructs by examining the social and physical circumstances in which particular racially defined groups find themselves. These studies may need to examine whether it is these circumstances which are predisposing these groups to disease and ill-health rather than something inherent in their "race" (King, 1997). There may indeed be a genetic component that confers some variability in DM between individuals, but at the level of the population it appears that larger socioeconomic and environmental factors are more important. Further studies which stratify the analysis by Aboriginal status in order to explore whether the socioeconomic and environmental gradients in DM prevalence observed in this study apply to the non-Aboriginal population alone would add strength to these conclusions.

The high level of multi-collinearity observed between predictor variables also suggests that attempts at disentangling the independent relationships between these variables and DM prevalence may be counter-productive since all predictor variables may in fact be measuring aspects of the same phenomena (Evans & Barer, 1994; Hertzman, Frank, & Evans, 1994; Marmot, 1999; Wilkinson, 1996, 1999). This phenomenon is likely related to social position, access to real life choices, and a sense of personal empowerment. This suggests that ecological studies utilizing socioeconomic predictors should start to locate their analyses within well-developed perspectives on how the social position of particular groups become established, spatially concentrated, reproduced over time, and results in poor health outcomes. The specific pathways by which low social position becomes translated into poor health outcomes through the generalized stress response, poor lifestyle practices, and reduced opportunities are becoming increasingly recognized (Baum, Garofalo, & Yali, 1999; Cohen, 1999; Kawachi, 1999; Lundberg, 1999; McEwen & Seeman, 2001; Pickering, 1999; Williams, 2001).

This study raises questions about how we need to better understand the powerful and predictable impact that place has on the health of populations. This study has documented that geographies in the central core of Winnipeg are associated with high levels of DM. These core area neighborhoods are places that seemed to have emerged as gathering places for individuals low on the social hierarchy with few social choices. This has likely occurred as a result of historical, political, and economic forces. The result has been the transformation of the physical and social fabric of these geographies into places of risk with ecological characteristics having strong association with health status. In order to more fully understand how this has happened over time, the unique history of how these high-risk geographies have evolved over time have to be explored more carefully through the use of diverse historical, political, economic and ethnographic methods.

Finally, the results of this study suggest that high rates of DM are tightly embedded within a context of poverty and disempowerment. Population-based prevention programs which focus only on lifestyle modification would likely not be successful. As illustrated by this study, lifestyle quality indicators like smoking are highly correlated with income and education. This suggests that DM prevention programs, to be successful, would require comprehensive policy interventions above and beyond lifestyle modification. These interventions have

to address the socioeconomic resources and opportunities available to individuals.

References

Amos, A. F., McCarty, D. J., & Zimmet, P. (1997). The rising global burden of diabetes and its complications: Estimates and projections to the year 2010. *Diabetes Medicines*, *14*(Suppl 5), S1–85.

Auslander, W. F., Haire-Joshu, D., Houston, C. A., & Fisher, E. B. (1992). Community organization to reduce the risk of non-insulin-dependent diabetes among low-income African-American women. *Ethnology Disease*, *2*, 176–184.

Baum, A., Garofalo, J., & Yali, A. (1999). Socioeconomic status and chronic stress: Does stress account for SES effects on health. *Annals of the New York Academy of Sciences*, *896*, 131–144.

Blanchard, J., et al. (1996). Incidence and prevalence of diabetes in manitoba, 1986–1991. *Diabetes Care*, *19*(8), 807–811.

Carter, J. S., Pugh, J. A., & Monterrosa, A. E. (1996). Non-insulin-dependent diabetes mellitus in minorities in the United States. *Annals of Internal Medicine*, *125*(3), 221–232.

Cohen, S. (1999). Social status and susceptibility to respiratory infections. *Annals of the New York Academy of Sciences*, *896*, 246–253.

Environmental Systems Research Institute. (1999). *Arc-View 3.2.* Redlands, California.

Evans, R., & Barer, M. (1994). *Why are some people healthy and others not: the determinants of health of populations*. New York: Walter de Gruyter Inc.

Fujimoto, W. Y. (1996). Overview of non-insulin-dependent diabetes mellitus (NIDDM) in different population groups. *Diabetes Medicine*, *13*(9 Suppl 6), 7–10.

Grabowski, J. G., & Li, G. (2001). Cluster analysis of pilot fatalities in general aviation crashes, continental United States, 1983-1998. *Academy of Emergency Medicine*, *8*(5), 465–469.

Haffner, S. M. (1998). Epidemiology of type 2 diabetes: Risk factors. *Diabetes Care*, *21*(Suppl 3), C3–6.

Hales, C. N., & Barker, D. J. (1992). Type 2 (non-insulin dependent) diabetes mellitus: The thrifty phenotype hypothesis. *Diabetologia*, *35*, 595–601.

Hales, C. N., Desai, M., & Ozanne, S. E. (1997). The thrifty phenotype hypothesis: How does it look after 5 years? *Diabetes Medicine*, *14*(3), 189–195.

Hanis, C. L., Chakraborty, R., Ferrell, R. E., & Schull, W. J. (1986). Individual admixture estimates: Disease associations and individual risk of diabetes and gallbladder disease among Mexican-Americans in Starr County, Texas. *American Journal of Physical Anthropology*, *70*, 433–441.

Harris, M. (1995). Chapter 1 Summary: Descriptive Epidemiology. In: *Diabetes in America (2nd Ed.)*. National Institutes of Health. (pp. 1–13).

Hazuda, H. P., & Monterrosa, A. E. (1992). Social Class predicts 8-year incidence of diabetes in Mexican Americans and non-Hispanic whites. *Diabetes*, *41*, 179–188.

Hendricks, R. T., & Haas, L. B. (1991). Diabetes in minority populations. *Nurse Practice Forum*, *2*, 199–202.

Hertzman, C., Frank, J., & Evans, R. (1994). Heterogeneities in health status and the determinants of population health. *New York Aldine de Gruyter*, (3), 67–92.

Hintz, J. (2000). *NCSS 2000.* Kaysville, Utah Number Cruncher Statistical Systems 2000.

Hjalmars, U., & Gustafsson, G. (1999). Higher risk for acute childhood lymphoblastic leukaemia in Swedish population centres 1973-94. Swedish Child Leukaemia Group. *British Journal of Cancer*, *79*(1), 30–33.

Hjalmars, U., Kulldorff, M., Gustafsson, G., & Nagarwalla, N. (1996). Childhood leukaemia in Sweden: Using GIS and a spatial scan statistic for cluster detection. *Statistics and Medicine*, *15*(7-9), 707–715.

Hjalmars, U., Kulldorff, M., Wahlqvist, Y., & Lannering, B. (1999). Increased incidence rates but no space–time clustering of childhood astrocytoma in Sweden, 1973–1992: A population-based study of pediatric brain tumors. *Cancer*, *85*(9), 2077–2090.

Jemal, A., Devesa, S., Kulldorff, M., Hayes, R., & Fraumeni, J. (2000). Geographic variation in prostate cancer mortality rates among white males in the United States. *Annals of Epidemiology*, *10*(7), 470–481.

Kawachi, I. (1999). Social capital and community effects on population and individual health. *Annals of the New York Academy of Sciences*, *896*, 120–130.

King, G. (1997). The race concept in smoking: A review of the research on African Americans. *Social Science & Medicine*, *45*(7), 1075–1087.

King, H., Aubert, R. E., & Herman, W. H. (1998). Global burden of diabetes, 1995–2025: Prevalence, numerical estimates, and projections. *Diabetes Care*, *21*(9), 1414–1431.

Kulldorff, M., Athas, W. F., Feurer, E. J., Miller, B. A., & Key, C. R. (1998a). Evaluating cluster alarms: A space–time scan statistic and brain cancer in Los Alamos, New Mexico. *Am. J. Public Health*, *88*(9), 1377–1380.

Kulldorff, M., Feuer, E. J., Miller, B. A., & Freedman, L. S. (1997). Breast cancer clusters in the northeast United States: A geographic analysis. *American Journal of Epidemiology*, *146*(2), 161–170.

Kulldorff, M., & Nagarwalla, N. (1995). Spatial disease clusters: Detection and inference. *Statistics and Medicine*, *14*(8), 799–810.

Kulldorff, M., Rand, K., Gherman, G., Williams, G., & DeFrancesco, D. (1998b). *SaTScan v2.1: Software for the spatial and space–time scan statistics*. Bethesda, MD: National Cancer Institute.

Leonetti, D., Tsunehara, C. H., Wahl, P. W., & Fujimoto, W. (1992). Educational attainment and risk of non-insulin dependent diabetes or coronary heart disease in Japanese-American men. *Ethnology Disease*, *2*, 326–336.

Lundberg, U. (1999). Stress responses in low-status jobs and their relationship to health risks: Musculoskeletal disorders. *Annals of the New York Academy of Sciences*, *896*, 162–172.

Marmot, M. (1999). Epidemiology of socioeconomic status and health: Are determinants within countries the same as between countries. *Annals of the New York Academy of Sciences*, *896*, 16–29.

Marshall, J. A., Hamman, R. F., Baxter, J., Mayer-Davis, E. J., Fulton, D. L., & Orleans, M. (1993). Ethnic differences in risk factors associated with the prevalence of non-insulin

dependent diabetes mellitus. The San Luis Valley Diabetes Study. *American Journal of Epidemiology*, *137*, 706–718.

Mathsoft, (1999). *S-Plus Spatial Statistics*. Mathsoft.

McDermott, R. (1998). Ethics, epidemiology and the thrifty gene: Biological determinism as a health hazard. *Social Science Medicine*, *47*(9), 1189–1195.

McEwen, B., & Seeman, T. (2001). Protective and damaging effects of mediators of stress: Elaborating and testing the concepts of allostasis and allostatic load. *Annals of the New York Academy of Sciences*, *896*, 30–47.

Meade, M., & Earickson, R., (2000). Scale, spatial analysis, and geographic visualization. In: *Medical Geography*. New York: The Guilford Press. (pp. 438–484).

Morgenstein, H. (1982). Use of ecologic analysis in epidemiologic research. *American Journal of Epidemiology*, *72*(12), 1336–1346.

Morgenstein, H. (1995). Ecologic studies in epidemiology: Concepts, principles and methods. *Annual Review of Public Health*, *16*, 61–81.

Neel, J. V. (1962). Diabetes mellitus: A "thrifty" genotype rendered detrimental by "progress"? *American Journal of Human Genetics*, *14*, 353–362.

Neel, J.V. (1982). *The thrifty genotype revisited*. London, New York: Academic Press, (pp. 283–293).

Neel, J. V. (1999). The "thrifty genotype" in 1998. *Nutrition Reviews*, *57*(5 Part 2), 2–9.

Norstrom, M., Pfeiffer, D. U., & Jarp, J. (1999). A space-time cluster investigation of an outbreak of acute respiratory disease in Norwegian cattle herds. *Preventive Veterinary Medicine*, *47*(1–2), 107–119.

Ozanne, S. E., & Hales, C. N. (1998). Thrifty yes, genetic no. *Diabetologia*, *41*(4), 485–487.

Pickering, T. (1999). Cardiovascular pathways: Socioeconomic status and stress effects on hypertension and cardiovascular function. *Annals of the New York Academy of Sciences*, *896*, 262–277.

Rose, G. (1985). Sick individuals and sick populations. *International Journal of Epidemiology*, *14*(1), 32–38.

Rose, G. (1992). *The strategy of preventive medicine*. Oxford, New York, Tokyo: Oxford University Press.

Sankoh, O. A., Ye, Y., Sauerborn, R., Muller, O., & Becher, H. (2001). Clustering of childhood mortality in rural Burkina Faso. *International Journal of Epidemiology*, *30*(3), 485–492.

Swinburn, B. A. (1996). The thrifty genotype hypothesis: How does it look after 30 years? *Diabetes Medicine*, *13*(8), 695–699.

Walsh, S. J., & Fenster, J. R. (1997). Geographical clustering of mortality from systemic sclerosis in the Southeastern United States, 1981–90. *Journal Rheumatology*, *24*(12), 2348–2352.

Wilkinson, R. (1996). *Introduction: The social economy of health*. London, New York: Routledge (1), (pp. 1–12).

Wilkinson, R. (1999). Health, hierarchy and social anxiety. *Annals of the New York Academy of Sciences*, *896*, 48–62.

Williams, D. (2001). Race, socioeconomic status and health: The added effects of racism and discrimination. *Annals of the New York Academy of Sciences*, *896*, 173–188.