Application of Three Focused Cluster Detection Methods to Study Geographic Variation in the Incidence of Multiple Sclerosis in Manitoba, Canada

Mahmoud Torabi^a Chris Green^a Nancy Yu^a Ruth Ann Marrie^{a, b}

^aDepartment of Community Health Sciences and ^bDepartment of Internal Medicine, University of Manitoba, Winnipeg, Man., Canada

Key Words

Geographic epidemiology · Multiple sclerosis · Spatial patterns

Abstract

Background: Macroscopic geographic variation in the incidence and prevalence of MS is well-recognized. Microscopic geographic variation in the distribution of MS is also recognized, but less well-studied. Most studies have focused on prevalent cases of MS, although studies of variation in disease incidence are more relevant for developing etiologic hypotheses. We aimed to study geographic variation in the incidence of MS using three different methods. **Methods:** We used population-based administrative (health claims) data to identify 2,290 incident cases of MS in the province of Manitoba, Canada from 1990 to 2006. We applied three focused cluster-detection procedures, including the circular spatial scan statistic (CSS), flexible spatial scan statistic (FSS), and Bayesian disease mapping (BYM), to the dataset. **Results:** The CSS and FSS methods identified 30 and 26 regions as potential clusters, respectively, although the regions identified differed slightly due to the non-circular shape of some regions in Manitoba. The BYM approach identified 37 regions as potential clusters, again with some differences as compared to the other two methods. Twelve regions were identified as potential clusters by all three methods. All methods identified the western part of the city of Winnipeg

as a significant cluster. Using the BYM approach, the incidence of MS was highest among areas of higher socioeconomic status. **Conclusions:** Two methods CSS and FSS only capture geographical variations and are not able to control for confounders at the same time which may lead to misidentification of clusters. However, the BYM method can simultaneously identify geographical variations and control for possible confounders.

© 2014 S. Karger AG, Basel

Introduction

Multiple sclerosis (MS) is a chronic immune-mediated disease of the central nervous system affecting more than 2.5 million people worldwide [1, 2]. The incidence and prevalence of MS are well-recognized to vary across large geographic regions. Relatively few studies have explored the spatial mapping of MS, but identification of spatial clusters of MS may suggest novel etiologic hypotheses for the disease, the etiology of which remains unknown, although it is believed to be a consequence of environmental exposures acting in genetically susceptible individuals [3]. Most of the prior spatial mapping studies have used prevalent rather than incident cases of MS, although incidence is a better measure of the risk of disease [4–7].

Spatial cluster detection methods can be classified into two groups [8]. The first method is a focused approach

where regions with a high rate of disease around a possible cause (i.e., a toxic waste site) are located [9, 10]. Focused cluster detection methods include the circular spatial scan statistic (CSS) [11], flexible spatial scan statistic (FSS) [12] and Bayesian disease mapping (BYM) [13]. Focused tests are used to discover potential clusters in a specific area of interest by testing the null hypothesis that there is no cluster against the alternative hypothesis that a cluster exists. The second method is a non-focused or general cluster detection approach where extra-Poisson variability is used to identify areas with a high number of disease cases [13–15]. Approaches include the Besag and Newell (BN) [16, 17] test and the maximizing excess event test (MEET) [18]. Non-focused tests are used to identify any significant clusters in a region without specifying a region of interest.

We aimed to study geographic variation in the incidence of MS in the province of Manitoba, Canada. Specifically, we compared the performance of three focused cluster-detection procedures, CSS, FSS, and BYM, by applying each approach to a dataset of incident MS cases in the province of Manitoba, Canada.

Materials and Methods

Study Population

We conducted this study in the Canadian province of Manitoba. The population of Manitoba was stable across the study period ranging from 1.13 million people in 1990 to 1.18 million people in 2006 [19]. We used an administrative (health claims) dataset of incident MS cases for the entire province of Manitoba. Manitoba Health provides comprehensive universal health insurance to all residents of Manitoba and maintains a population registry of permanent residents in the province. The population registry captures demographic information (sex, date of birth), and dates of birth and migration into and out of the province. We used the registry to determine the population size and distribution across the province.

We identified persons with MS by applying a validated case definition to hospital, physician, and prescription claims [20]. We considered the date of the first medical contact for any diagnostic code related to demyelinating disease to be the ‘date of diagnosis’. Incident cases were identified by applying a five-year run-in period. This meant that the first incident case could be identified in 1990. Age-standardized incidence rates were reported by the year of diagnosis, after age-standardization to the 2001 Canadian population. Ecological measures of socioeconomic status and ethnicity (mean household income, proportion of recent immigrants, Jewish ethnicity, visible minority status, and unemployment status) for the MS cases were obtained from the 2001 Canadian Census micro-data files [21]. If the population of a region was over 50,000, that region was considered to be urban. Based on this criterion, only the City of Winnipeg was defined as urban.

The MS cases were geocoded by considering the City of Winnipeg as 230 neighbourhoods (average population during the

study years: 653,100) based on 6-digit postal codes, and considering rural Manitoba as 268 health municipalities (average population during the study years: 500,977) based on the municipal code stated on the health record of MS patients at the time of the first demyelinating disease health claim. The geographical units used in this study were these 498 regions and all the data used in this study were associated with these geographical boundaries. The data were aggregated over the study period. Key requirements for focused spatial cluster detection methods are the observed number of MS cases and the expected number of MS cases or the population size of each region. The CSS, FSS and BYM spatial focused cluster-detection methods are described below.

Circular Spatial Scan Statistic (CSS)

The spatial scan statistic has various uses in epidemiology [22]. The circular spatial scan statistic is used to test the null hypothesis that there is no cluster in region i against the alternative hypothesis that there is a cluster of MS cases in region i based on its j -th nearest neighbours. For each region, the circular spatial scan statistic imposes a circular window S on the region, with the radius of the circle extending from zero to a maximum distance d or a maximum number of regions J to be included in the cluster. Either d or J is predetermined prior to the analyses. Now, $S_{ij}(1, \dots, J)$ defines the window consisting of the $(j - 1)$ -th nearest neighbours to region i . Furthermore, the set of all windows to be scanned by the circular scan statistic is defined by $S_i = \{S_{ij}; i = 1, \dots, m; j = 1, \dots, J\}$. For each circle, a likelihood ratio statistic is calculated based on the number of observed and expected cases of MS inside and outside the circle. Under the null and alternative hypotheses, the likelihood is defined by L_0 and L_i ($i = 1, \dots, m$), respectively. The likelihood ratio statistic is given by

$$\max_i \frac{L_i}{L_0} = \left(\frac{C_i}{Np_i} \right)^{C_i} \left(\frac{N - C_i}{N(1 - p_i)} \right)^{N - C_i} I(C_i > Np_i) \quad (1)$$

where C_i represents the observed number of cases inside a circle, while $(N - C_i)$ represents the observed number of cases outside the circle and p_i represents the probability that a region is inside a circle. The indicator function $I(C_i > Np_i)$ is equal to 1 when $C_i > Np_i$ and 0 otherwise. A possible cluster is defined as a region with a high likelihood ratio statistic [12].

SaTScan [23] or FleXScan [24] software can be used to carry out the CSS method. Usually, 50% of the population at risk is included in a possible cluster; however, we used the FleXScan default, which is $J = 15$ as a maximum number of regions to be included in a potential cluster. As we used aggregate data in our study, the region centroid had to be included in the radius of the circle for the region to be considered part of the circle.

Flexible Spatial Scan Statistic (FSS)

The flexible spatial scan statistic (FSS) is also used to test the null hypothesis that a cluster exists in region i . The FSS method works in the same general way as the CSS method described above. However, with the FSS method, the potential cluster is flexible in shape, although it is still confined to a small neighbourhood around each region. The flexible scan statistic connects adjacent regions to place a window S , with an irregular shape on each region. For any region i , the set of irregularly shaped windows of length j , containing j connected regions including region i , can vary from 1 to J , where J is

the pre-determined maximum length of a cluster. To avoid unlikely cluster shapes, the joined regions are restricted to the subsets of the set of regions i and $(J - 1)$ -the nearest neighbours of region i . Now, $S_2 = \{S_{ij(k)}; i = 1, \dots, m; j = 1, \dots, J; k = 1, \dots, k_{ij}\}$ is the set of all windows to be scanned by the flexible spatial scan statistic. The size of S_2 is much larger than S_1 , which is at most mJ . This comes from the fact that the circular spatial scan statistic only examines J circles for any region i , while the flexible spatial scan statistic examines J circles in addition to all the sets of connected regions whose centroids are found within the J -th largest concentric circle. For the FSS method, the test statistic is based on the likelihood ratio test given in (1), where the circle defined in (1) refers to S_2 instead of S_1 . Similar to the CSS method, the FSS approach considers circles with high likelihood ratio values to be potential regions of disease clusters [13]. The FSS procedure can be implemented using the FlexScan software [24] by applying the FlexScan default, which is a maximum of $J = 15$ regions to be included in a potential cluster.

Bayesian Disease Mapping (BYM)

Identifying clusters can also be done via a Bayesian framework using Markov chain Monte Carlo (MCMC) methods [13, 14, 25–27]. Bayesian disease mapping (BYM) can be used to show geographical variation of MS cases across the study region. The BYM method consists of two parts. First, it is assumed that the observed cases of MS follow a Poisson distribution with an area specific parameter $\theta_i N_i$,

$$C_i \sim \text{Poisson}(\theta_i N_i), \quad (2)$$

where the observed cases and population in region i are represented by C_i and N_i , respectively. Now, the second part of the model is given by

$$\log \theta_i = \mu + X_i' \beta + \eta_i + \Phi_i, \quad (3)$$

where θ_i is defined by the incidence rate in region i , μ denotes the overall mean rate across the entire study region and X_i are predictors or covariates with corresponding regression coefficients given by β . Now, η_i is defined as the specified spatial random effects, and Φ_i denotes the unspecified features of region i , which do not incorporate spatial structure. The Φ_i are assumed to follow a Gaussian distribution with zero mean and common variance σ^2_Φ , while the specified spatial random effects η_i are found using an intrinsic conditionally autoregressive model depending on their neighbouring values. Specifically,

$$(\eta_1, \dots, \eta_m)' \sim N(0, \Sigma_\eta), \\ \Sigma_\eta = \sigma^2_\eta D^{-1},$$

where the spatial dispersion parameter is denoted by σ^2_η and D is the neighbourhood matrix with its i -th diagonal element equal to the number of neighbours of the corresponding region, and the off-diagonal elements in each row equal -1 if the corresponding regions are neighbours and zero otherwise [13]. Using vague priors for the parameters, they can be estimated within the Bayesian framework (MCMC) to produce posterior distributions for the parameters in the model. In this study, proper priors are used and the sensitivity of the choice of priors is also investigated. A region is defined as a potential cluster where the lower 95% credible set (equivalent to confidence interval in non-Bayesian analysis) is larger than 10 (per 100,000) [28].

Table 1. Age-standardized annual incidence rates per 100,000 persons by year of diagnosis

Year	Incidence rate	95% CI
1990	14.5	12.2–16.7
1991	10.9	8.9–12.8
1992	12.3	10.3–14.3
1993	12.1	10.0–14.1
1994	10.0	8.2–11.9
1995	8.2	6.5–9.9
1996	8.9	7.2–10.6
1997	9.0	7.3–10.7
1998	14.3	12.0–16.5
1999	15.3	13.0–17.6
2000	19.5	17.0–22.1
2001	15.0	12.8–17.2
2002	13.4	11.3–15.5
2003	14.5	12.4–16.7
2004	12.6	10.6–14.6
2005	12.7	10.7–14.7
2006	9.8	8.0–11.6

In particular, two approaches were taken using the equations (2, 3); one with no predictors and another one with predictors, both with adjustment for age to control for differences in demographic structures. To model the relationship between MS cases and the characteristics of the geographical units, a series of Bayesian Poisson regression models were implemented for each characteristic. A saturated Bayesian Poisson regression model containing multiple predictors was also developed. All models were fitted to individual cases, whereby cases were assigned the ecological characteristics (e.g., average income level) of the geographic unit to which they were geocoded. Variables were categorised using the Jenks natural breaks classification method [29], which attempts to find natural break points in the data when identifying category cut-offs saturated model. We included a variable for the region of residence in the province (Winnipeg urban core; Winnipeg areas outside the urban core [Suburban]; northern Manitoba, south-western Manitoba, south-central Manitoba; and south-eastern Manitoba). Some of the ecological variables such as the proportion of Jewish or visible minorities and urban/rural residence were not included in the saturated model *a priori* to avoid multicollinearity. Potential over-dispersion in the models was managed by incorporating a random variable to capture unspecified variation across small areas. The results are based on Bayesian inference [14] and presented as incidence rate ratios (IRR) and corresponding 95% credible intervals. The WinBUGS software package (MRC Biostatistics unit, Institute of Public Health, London, UK) [30] was used to apply the BYM approach.

Results

Over the study period, we identified 2,290 incident MS cases, with an average incidence of 12.08 cases per 100,000. Table 1 shows the age-standardized incidence rates by

Fig. 1. Regions identified as potential clusters (a) as shaded regions and order of most likely clusters (b) of MS incidences in the province of Manitoba, Canada, by the CSS method.

year of diagnosis. Most cases were female (1,704, 74.4%), from urban regions (1,553, 67.8%), with an average (SD) age at diagnosis of 40.6 (12.5). The CSS method identified 30 regions as potential clusters (fig. 1a). These 30 regions identified 3 major regions and a couple of minor regions in the City of Winnipeg. The FSS technique identified 26 regions as possible clusters (fig. 2a), but the regions differed slightly from the CSS method. These 26 regions also identified 3 major regions and a couple of minor regions

in the City of Winnipeg, which are slightly different from the CSS method.

The BYM approach was conducted in two ways. First, we did not include any predictors in the model. The BYM method with no covariates included in the model identified 32 regions as possible clusters (fig. 3a). These 32 regions identified one major region in the western part of the City of Winnipeg and some minor regions in the urban core of the City of Winnipeg as potential clusters. Sec-

Fig. 2. Regions identified as potential clusters (a) as shaded regions and order of most likely clusters (b) of MS incidences in the province of Manitoba, Canada, by the FSS method.

ond, we included four predictors (region, annual income, unemployment rate and recent immigrant status) in the model. This approach identified 37 regions as possible clusters (fig. 4a). In addition to some similar regions as identified by the BYM without covariates, the BYM with

predictors also identified some parts of the south western of Manitoba as potential clusters. We noted that 10 regions in western Winnipeg were identified by the all methods CSS, FSS, and BYM (without and with predictors) as potential clusters. Most of the statistically signifi-

Fig. 3. Regions identified as potential clusters (a) as shaded regions of MS incidences and MS incidence rates (b) (per 100,000) in the province of Manitoba, Canada, by the BYM method with no covariates included in the model.

cant clusters were found in the City of Winnipeg. All three methods identified the western part of the city as a statistically significant cluster. The CSS and FSS methods also identified the central part of the city as a single cluster, while the BYM approach identified scattered clusters

in the center of the city. Only the CSS method identified the eastern part of the city as a cluster. Only the BYM method (with predictors) was able to identify some parts of south western Manitoba (outside the City of Winnipeg) as potential clusters.

Fig. 4. Regions identified as potential clusters (a) as shaded regions of MS incidences and MS incidence rates (b) (per 100,000) in the province of Manitoba, Canada, by the BYM method with four predictors (region, annual income, unemployment rate and recent immigrants) included in the model.

Table 2. Univariate Poisson regression analyses of multiple sclerosis incidence rate ratios (IRR) and 95% credible intervals (CI) from 1990 to 2006 in Manitoba, Canada

Parameter	IRR	95% CI
Region 1		
Rural	1.00	–
Urban	2.25	1.34–3.83
Region 2		
Winnipeg non-core	1.00	–
Winnipeg core	1.87	1.22–3.00
Northern Manitoba	2.23	1.42–3.68
South-western Manitoba	2.88	1.85–4.95
South-central Manitoba	2.38	1.28–4.61
South-eastern Manitoba	3.43	1.92–6.42
Annual income		
<\$47,000	1.00	–
\$47,000–\$85,000	1.46	1.26–1.69
>\$85,000	1.49	1.18–1.88
Unemployment rate		
<9%	8.18	4.76–15.04
9%–<19%	5.71	3.22–10.68
≥19%	1.00	–
Recent immigrants		
<9%	1.00	–
9%–<20%	1.32	1.00–1.75
≥20%	1.19	0.85–1.68
Jewish ethnicity*		
<4%	0.81	0.59–1.12
4%–<13%	1.02	0.70–1.45
≥13%	1.00	–
Visible minorities		
<6%	1.13	0.87–1.48
6%–<19%	1.18	0.99–1.46
≥19%	1.00	–

* Percentage of total population reporting Jewish ethnicity.

Figures 1b and 2b also display maps with the order of significant clusters, based on the size of the spatial statistics, for the CSS and FSS approaches, where both methods detected the central and western parts of the City of Winnipeg as the top two potential clusters but with a different order.

Note that the overall relative risk values for the most likely, second most likely, and third most likely significant clusters for the CSS method were 1.52, 1.52, and 2.70, respectively. These values for the FSS method were 1.80, 1.61, and 2.55, respectively. Figures 3b and 4b also depict the choropleth maps of the incidence rate of MS in Manitoba based on the BYM approach with no predictors and multiple predictors respectively, and show that the City of Winnipeg has a high incidence rate of MS. Also, most of southern Manitoba, especially the south-western and

Table 3. Adjusted saturated Poisson regression analysis of multiple sclerosis incidence rate ratios (IRR) with 95% credible intervals (CI)

Parameter	IRR	95% CI
Region		
Winnipeg non-core	1.00	–
Winnipeg core	0.92	0.62–1.34
Northern Manitoba	1.05	0.70–1.60
South-western Manitoba	1.32	0.90–1.97
South-central Manitoba	0.83	0.40–1.68
South-eastern Manitoba	1.00	0.49–1.95
Annual income		
<\$47,000	1.00	–
\$47,000–\$85,000	1.15	1.00–1.34
>\$85,000	1.18	0.94–1.48
Unemployment rate		
<9%	5.95	3.33–11.42
9%–<19%	4.92	2.72–9.46
≥19%	1.00	–
Recent immigrants		
<9%	1.00	–
9%–<20%	1.08	0.86–1.35
≥20%	0.96	0.73–1.27

south-eastern regions, has higher incidence rates than northern Manitoba.

Table 2 shows a series of simple Poisson regression models using the BYM approach by equations (2, 3). The incidence rate of MS was highest among those living in areas with higher incomes and where the unemployment rate was low. Incidence rates were not associated with recent Jewish ethnicity or visible minority areas. The associations with the recent immigrant status were inconsistent. Rates were 225% higher in Winnipeg (urban centre) than in the rest of the province.

The saturated Poisson regression model in table 3 suggested that the highest MS incidence rates occurred in south-western Manitoba (IRR = 1.32; 95% CI: 0.90–1.97) although this did not reach statistical significance. Regions with higher annual incomes and lower unemployment rates experienced higher MS incidence rates. However, there was no association between the MS incidence rates and the recent immigrant status. It is clear from table 3 that some covariates have been adjusted in the saturated model compared to the univariate model in table 2. For instance, the covariate region was significant in the univariate model unlike the saturated model. It shows that if one only uses a univariate model, it may lead to mis-identification of clusters while the saturated model also accounts for other possible covariates.

Discussion and Conclusion

We used three popular approaches in spatial epidemiology to identify potential clusters in MS cases in the Canadian province of Manitoba from 1990 to 2006. The CSS method identified 30 regions of interest, which aggregated into three larger regions in the City of Winnipeg (some parts of western, central, and southern Winnipeg). The FSS method identified 26 regions, some of which overlapped with those identified by the CSS method. When aggregated, the FSS method found that some parts of western, central, and eastern Winnipeg were potential clusters. The BYM methods identified 37 regions (mainly scattered in western, central and southern Winnipeg, and some parts of the south western of Manitoba), again with some similarities and differences to the findings of the other two methods. All three methods identified about 27% (= 10/37) of the same regions as potential clusters of MS cases.

Several studies have explored small area variation in the distribution of MS, with most studies focusing on prevalent cases [5, 6]. A smaller number of studies have evaluated the distribution of incident cases [31], but the study of incident cases is necessary to truly identify clusters. Clusters represent areas where an excess number of new (MS) cases are identified in relation to time or space, and they may indicate real biologic variation or random variation [32]. Identification of such clusters may generate new hypotheses about environmental risk factors. Techniques used to identify clusters have included distribution-free approaches (e.g., spatial scan statistics) and Poisson mixed models, but a few studies have compared the performance of different techniques [7]. An understanding of the differences between these techniques is necessary to interpret findings and select best methods for studies of small area variation.

The three focused spatial cluster-detection approaches (CSS, FSS, and BYM) that we employed had different assumptions. The CSS and FSS methods were distribution free. However, when using these methods, the number of regions to be included in the cluster must be stated *a priori*. The main reason for the observed differences in the findings of the two techniques is due to the non-circular shape of some regions in Manitoba. The FSS method was able to identify those non-circular shaped regions as potential clusters, whereas the CSS procedure could not. On the other hand, the FSS method was unable to detect some circular-shaped regions as potential clusters.

When using the BYM procedure, it is assumed that the number of cases follows a Poisson distribution but the

number of regions to be included in the cluster does not need to be pre-specified.

For the BYM method, a cluster was identified as a region where the lower 95% credible interval of the incidence rate was larger than 10 (per 100,000). However, it is possible to set different decision rules determining a cluster, for instance, the decision rule could be greater than or less than 10 (per 100,000) [33]. However, we divided the regions into five categories based on their incidence rates to have a clear picture of the distribution of incidence rates (e.g., fig. 3b).

For the BYM method, the model that included predictors drew different conclusions with respect to possible clusters than the model that did not include predictors. This emphasizes the importance of controlling for sources of confounding that may lead to mis-identification of clusters otherwise.

A recent publication using the same dataset also studied geographic variation of MS in Manitoba [34]. That study used an adaptive mean nearest-neighbor smoothing (AMNNS) algorithm [35] to stabilize the rates for geographic areas having fewer than 60 MS cases, and confirmed findings using a spatial scan statistic. Using that approach, higher than expected incidence rates were identified in two areas in the City of Winnipeg (central and south-western) and in rural western Manitoba. The identification of clusters in Winnipeg using AMNNS was consistent with our findings using the BYM approach, however, the BYM approach did not identify rural western Manitoba as a potential cluster. Using Poisson regression models (without incorporating spatial effects) the earlier study found that urban residence, higher income, and lower rates of unemployment were associated with higher incidence rates. The latter two findings were consistent with the findings from the BYM approach with a saturated model. However, there are notable differences between that approach and the BYM approach employed here. For instance, the BYM approach simultaneously accounts for the spatial variation and predictors in the model. The AMNNS method captures only the spatial variation (visually) and cannot control for the predictors at the same time, which may lead to the mis-identification of clusters. Note that the AMNNS method is similar to the methods CSS and FSS in the sense that they descriptively identify potential clusters and are not also able to control for the predictors at the same time.

Geospatial variation in MS incidence and prevalence is well-recognized and likely reflects both variation in underlying genetic susceptibility of the population to MS as

well as exposure to environmental factors. The risk of MS varies among the racial and ethnic groups even when they live in the same regions. On univariate analysis it was found that regions with a higher proportion of recent immigrants had higher incidence rates of MS. A higher socioeconomic status was also strongly associated with increased incidence rates, consistent with prior ecologic analyses in which MS was associated with greater urbanization and higher socioeconomic status [36, 37]. Observations like this contributed to the development of the hygiene hypothesis, which suggests that lack of exposure to infectious agents and loss of symbiotic microorganisms in the gut increase susceptibility to autoimmune diseases such as MS [38]. Although we lacked the data to evaluate whether variation in exposure to infection to the geospatial variation in MS incidence, spatial variation in the rates of infection in Manitoba has been observed [39]. We were also unable to evaluate associations with other putative etiologic factors for MS such as vitamin D insufficiency and toxic exposures [3].

This study has limitations that should be considered. We identified incident cases using administrative data. Administrative data are collected for health system management rather than research purposes, and may underrepresent chronic disease when compared to medical records [40, 41]. However, we used a case definition developed and validated in Manitoba that was subsequently re-validated in Nova Scotia [42]. Because we relied on administrative rather than clinical data to identify the date of incidence cases, this date may not be entirely accurate. However, the diagnostic delay between symptom onset and diagnosis is short and residential mobility rates in Manitoba are low [43]. Further, most migration in North America occurs over short distances to adjacent neighborhoods [44, 45]; therefore, it is unlikely that this affected the geographic mapping of our cases. We assumed that our MS cases are rare to be able to use Poisson model in our BYM method. Our work evaluated the perfor-

mance of focused spatial cluster detection methods rather than non-focused spatial cluster detection methods such as BN and MEET, which have different strengths and weaknesses; these warrant future study. We used census-based (ecologic) measures of socioeconomic status rather than individual level measures. Strengths of the study include the population-based design, the use of incident rather than prevalent MS cases, and the evaluation of multiple cluster detection methods.

The potential clusters of MS appear to be located in the City of Winnipeg, specifically the western and central regions of Winnipeg. As well, the southern part of the province of Manitoba had higher incidence rates than most of northern Manitoba. The BYM method may be the best approach for identifying potential clusters because it allows the investigator to control for potential confounders. Further studies are needed to explore the reason and cause of these increases, especially in broader geographic regions with evenly distributed population density.

Acknowledgements

We would like to thank three referees for constructive comments and suggestions, which led to an improved version of the manuscript. The research of Dr. Torabi was supported by an establishment grant from Manitoba Health Research Council (MHRC). The analyses were performed by Dr. Torabi. All of the authors were involved in the study concept and design; acquisition of data; interpretation of data and critical revision of the manuscript for important intellectual content. This study was funded (in part) by the Health Sciences Centre Foundation and a Don Paty Career Development Award from the MS Society of Canada. The results and conclusions presented are those of the authors. No official endorsement by Manitoba Health is intended or should be inferred.

Disclosure Statement

The authors declare that they have no conflicts of interest.

References

- 1 Nielsen NM, Jørgensen KT, Bager P, Stenager E, Pedersen BV, Hjalgrim H, Koch-Henriksen N, Frisch M: Socioeconomic factors in childhood and the risk of multiple sclerosis. *Am J Epidemiol* 2013;177:1289–1295.
- 2 Kakalacheva K, Lüneman JD: Environmental triggers of multiple sclerosis. *FEBS Lett* 2011; 585:3724–3729.
- 3 Marrie RA: Environmental risk factors in multiple sclerosis aetiology. *Lancet Neurol* 2004;3:709–718.
- 4 Fromont A, Binquet C, Sauleau EA, Fournel I, Bellisario A, Adnet J, Weill A, Vukusic S, Confavreux C, Debouverie M, Clerc L, Bonithon-Kopp C, Moreau T: Geographic variations of multiple sclerosis in France. *Brain* 2010;133:1889–1899.
- 5 Donnan PT, Parratt JD, Wilson SV, Forbes RB, O'Riordan JL, Swingle RJ: Multiple sclerosis in Tayside, Scotland: detection of clusters using a spatial scan statistic. *Mult Scler* 2005;11:403–408.
- 6 Pugliatti M, Solinas G, Sotgiu S, Castiglia P, Rosati G: Multiple sclerosis distribution in northern Sardinia: spatial cluster analysis of prevalence. *Neurology* 2002;58:277–282.
- 7 Bergamaschi R, Montomoli C, Candeloro E, Monti MC, Cioccale R, Bernardinelli L, Fratino P, Cosi V: Bayesian mapping of multiple sclerosis prevalence in the province of Pavia, northern Italy. *J Neurol Sci* 2006;244:127–131.

- 8 Torabi M, Rosychuk RJ: An examination of five spatial disease clustering methodologies for the identification of childhood cancer clusters in Alberta, Canada. *Spat Spatiotemporal Epidemiol* 2011;2:321–330.
- 9 Elliott P, Briggs D, Morris S, de Hoogh C, Hurt C, Jensen TK, Maitland I, Richardson S, Wakefield J, Jarup L: Risk of adverse birth outcomes in populations living near landfill sites. *BMJ* 2001;323:363–368.
- 10 Lawson AB, Biggeri A, Williams FLR: A review of modeling approaches in health risk assessment around putative sources; in Lawson AB, Biggeri A, Böhning D, Lesaffre E, Viel J, Bertollini R (eds): *Disease Mapping and Risk Assessment for Public Health*. New York, Wiley, 1999, pp 231–245.
- 11 Kulldorff M: A spatial scan statistics. *Communications in Statistics – Theory and Methods* 1997;26:1481–1496.
- 12 Tango T, Takahashi K: A flexibly shaped spatial scan statistic for detecting clusters. *Int J Health Geogr* 2005;4:1–15.
- 13 Besag JE, York JC, Mollie A: Bayesian image restoration with two applications in spatial statistics (with discussion). *Ann Inst Statist Math* 1991;43:1–59.
- 14 Clayton D, Bernardinelli L: Bayesian methods for mapping disease risk; in Elliott P, Cuzick J, English D, Stern R (eds): *Geographical and Environmental Epidemiology: Methods for Small-Area Studies*. Oxford, Oxford University Press, 1996, pp 205–220.
- 15 Clayton D, Kaldor J: Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987;43:671–681.
- 16 Besag JE, Newell J: The detection of clusters in rare diseases. *J Roy Statist Soc Ser A (JSTOR)* 1991;154:143–155.
- 17 Torabi M, Rosychuk RJ: Spatial event cluster detection using an approximate normal distribution. *Int J Health Geogr* 2008;7:1–22.
- 18 Tango T: A test for spatial disease clustering adjusted for multiple testing. *Stat Med* 2000;19:191–204.
- 19 Manitoba Health: Population report. Winnipeg, MB, Canada: Government of Manitoba, 2013. <http://www.gov.mb.ca/health/population/> (accessed September 13, 2013).
- 20 Marrie RA, Yu N, Blanchard J, Leung S, Elliott L: The rising prevalence and changing age distribution of multiple sclerosis in Manitoba. *Neurology* 2010;74:465–471.
- 21 Statistics Canada: Data Liberation Initiative. Ottawa, ON, Canada: Government of Canada, 2001. <http://www.statcan.gc.ca/dli-idd/dli-idd-eng.htm> (accessed December 5, 2013).
- 22 Fukuda Y, Umezaki M, Nakamura K, Takano T: Variations in social characteristics of spatial disease clusters: examples of colon, lung and breast cancer in Japan. *Int J Health Geogr* 2005;4:1–13.
- 23 Kulldorff M, Rand K, Gherman G, Williams G, DeFrancesco D: SaTScan V2.1: Software for the spatial and space-time scan statistics. National Centre Institute, Bethesda, 1998.
- 24 Takahashi K, Yokoyama T, Tango T: Flex Scan: Software for the Flexible Scan Statistic. National Institute of Public Health, Japan, 2006.
- 25 Bernardinelli L, Montomoli C: Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Stat Med* 1992;11:983–1007.
- 26 Gilks WR, Richardson S, Spiegelhalter DJ (eds): *Markov chain Monte Carlo in practice*. Chapman and Hall/CRC, 1995.
- 27 Cocco E, Sardu C, Massa R, et al: Epidemiology of multiple sclerosis in south-western Sardinia. *Mult Scler* 2011;17:1282–1289.
- 28 Aamodt G, Samuelsen SO, Skrandal A: A simulated study of three methods for detecting disease clusters. *Int J Health Geogr* 2006;5:1–11.
- 29 Pfeiffer DU, Robinson TP, Stevenson M, et al: Spatial visualization; in Pfeiffer DU, Robinson TP (eds): *Spatial Analysis in Epidemiology*. New York, NY: Oxford University Press, 2008, pp 17–31.
- 30 Spiegelhalter D, Thomas A, Best N, Lunn D: WinBUGS version 1.4 User Manual. MRC Biostatistics unit, Institute of Public Health, London, 2004.
- 31 Fromont A, Binquet C, Sauleau E, Fournel I, Despalins R, Rollot F, Weill A, Clerc L, Bonithon-Kopp C, Moreau T: National estimate of multiple sclerosis incidence in France (2001–2007). *Mult Scler* 2012;18:1108–1115.
- 32 Riise T: Cluster studies in multiple sclerosis. *Neurology* 1997;49(2 suppl 2):S27–S32.
- 33 Richardson S, Thomson A, Best N, Elliott P: Interpreting posterior risk estimates in disease-mapping studies. *Environ Health Persp* 2004;112:1016–1025.
- 34 Green C, Yu BN, Marrie RA: Exploring the implications of small-area variation in the incidence of multiple sclerosis. *Am J Epidemiol* 2013;178:1059–1066.
- 35 Anselin L, Lozano N, Kochinsky J: Rate Transformations and Smoothing. Urban-Champaign, IL: Spatial Analysis Laboratory, Department of Geography, University of Illinois, 2006.
- 36 Leibowitz U, Sharon D, Alter M: Geographical considerations in multiple sclerosis. *Brain* 1968;91:37–52.
- 37 Leibowitz U, et al: Epidemiological study of multiple sclerosis in Israel. II. Multiple sclerosis and level of sanitation. *J Neurol Neurosurg Psychiatry* 1966;29:60–68.
- 38 Rook GAW: Hygiene hypothesis and autoimmune diseases. *Clin Rev Allergy Immunol* 2012;42:5–15.
- 39 Green CG, Krause DO, Wylie JL: Spatial analysis of campylobacter infection in the Canadian province of Manitoba. *Int J Health Geogr* 2006;5:2.
- 40 Klabunde CN, Warren JL, Legler JM: Assessing comorbidity using claims data: an overview. *Med Care* 2002;40:26–35.
- 41 Lee DS, Donovan L, Austin PC, et al: Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care* 2005;43:182–188.
- 42 Marrie RA, Fisk JD, Stadnyk KJ, Yu BN, Tremlett H, Wolfson C, Warren S, Bhan V: The incidence and prevalence of multiple sclerosis in Nova Scotia, Canada. *Can J Neurol Sci* 2013;40:824–831.
- 43 Lix LM, Hinds A, Deverteuil G, Robinson JR, Walker J, Roos LL: Residential mobility and severe mental illness: a population-based analysis. *Adm Policy Ment Health* 2006;33:160–171.
- 44 Hurley SE, Reynolds P, Goldberg DE, et al: Residential mobility in the California Teachers Study: implications for geographic differences in disease rates. *Soc Sci Med* 2005;60:1547–1555.
- 45 Rogerson PA, Han D: The effects of migration on the detection of geographic differences in disease risk. *Soc Sci Med* 2002;55:1817–1828.