

Research

Open Access

## Geographical and temporal distribution of human giardiasis in Ontario, Canada

Agricola Odoi\*<sup>1</sup>, S Wayne Martin<sup>1</sup>, Pascal Michel<sup>2</sup>, John Holt<sup>3</sup>, Dean Middleton<sup>4</sup> and Jeff Wilson<sup>1,5</sup>

Address: <sup>1</sup>Department of Population Medicine, University of Guelph, Guelph, Ontario, Canada, N1G 2W1, <sup>2</sup>Université de Montreal, CP 5000, St-Hyacinthe, Québec, Canada, <sup>3</sup>Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada, <sup>4</sup>Ministry of Health and Long-Term Care, Ontario, Canada and <sup>5</sup>Division of Enteric, Foodborne and Waterborne Diseases, Population and Public Health Branch, Health Canada

Email: Agricola Odoi\* - aodoi@uoguelph.ca; S Wayne Martin - swmartin@uoguelph.ca; Pascal Michel - pascal\_michel@hc-sc.gc.ca; John Holt - jholt@msnet.mathstat.uoguelph.ca; Dean Middleton - Dean.Middleton@moh.gov.on.ca; Jeff Wilson - jwilson@uoguelph.ca

\* Corresponding author

Published: 11 August 2003

Received: 03 May 2003

International Journal of Health Geographics 2003, 2:5

Accepted: 11 August 2003

This article is available from: <http://www.ij-healthgeographics.com/content/2/1/5>

© 2003 Odoi et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

**Background:** *Giardia* is the most frequently identified intestinal parasite in North America. Although information on geographical distribution of giardiasis is critical in identifying communities at high risk, little has been done in this area. Therefore, the objective of this study was to investigate the geographical and temporal distribution of human giardiasis in Ontario in order to identify possible high risk areas and seasons. Two spatial scales of analyses and two disease measures were used with a view to identifying the best of each in assessing geographical patterns of giardiasis in Ontario. Global Moran's I and Moran Local Indicators of Spatial Associations were used to test for evidence of global and local spatial clustering, respectively.

**Results:** There were seasonal patterns with summer peaks and a significant ( $P < 0.001$ ) decreasing temporal trend. Significant ( $P < 0.05$ ) global spatial clustering of high rates was observed at the Census Sub-division spatial scale but not at the Census Division scale. The Census Sub-division scale was a better scale of analyses but required spatial empirical Bayesian smoothing of the rates. A number of areas with significant local clustering of giardiasis rates were identified.

**Conclusions:** The study identified spatial and temporal patterns in giardiasis distribution. This information is important in guiding decisions on disease control strategies. The study also showed that there is benefit in performing spatial analyses at more than one spatial scale to assess geographical patterns in disease distribution and that smoothing of disease rates for mapping in small areas enhances visualization of spatial patterns.

### Background

*Giardia lamblia*, a flagellated enteric protozoan parasite, is the etiologic agent of human giardiasis whose infection

may result in asymptomatic cyst passage, acute diarrhea or chronic disease accompanied with malabsorption and weight loss [1]. The parasite is a common cause of both

endemic and epidemic disease worldwide and is the most frequently identified intestinal parasite in the United States and Canada [2–4]. However, there are information gaps that need to be addressed to better assist in the development of policy guidelines specific for these infections. For instance, few studies have assessed the geographical and temporal distribution of giardiasis with a view to identifying high risk areas and seasons. This information is important in guiding decisions on resource allocation for disease control purposes.

When performing investigations of spatial disease patterns, two questions usually arise: what is the appropriate spatial scale (areal unit) to use in order to identify existing spatial patterns [5,6]; and what is the best disease measure? [7–11]. In relation to the best spatial scale, it has been reported that spatial patterns in disease distribution may change with a change of the spatial scale of analysis, a phenomenon called the Modifiable Areal Unit Problem (MAUP) [5]. The presence of the MAUP makes it necessary to perform spatial analyses at more than one areal unit to minimize its effects. In the current study, all spatial analyses were performed at two spatial scales to alleviate this problem and to identify the appropriate spatial scale to use for mapping giardiasis in Ontario, Canada.

Regarding the disease measure to use for mapping, standardized Mortality or Morbidity Ratios have often been used in disease maps [12,13]. However, when investigating diseases patterns in relatively small areas, use of these measures have certain limitations. Since regional disease rates have a non-constant population base, it implies that mapping and statistical comparisons of rates with very different variances have to be done. Rates arising from areas of low population usually have higher variances and are therefore more unstable than those from areas with high population [7]. Often, these low population areas represent rural communities that occupy large geographical extents. Consequently, the rates from these areas draw undue visual impression while being the least reliable due to higher variances [14]. In this study, we used spatial empirical Bayesian smoothed rates to alleviate or minimize this limitation of standardized rates.

With the above issues in mind, the objectives of this study were to: (a) describe the geographical and temporal distribution of giardiasis cases reported to a surveillance system in Ontario in order to identify areas with unusually high rates; (b) identify the appropriate spatial scale for giardiasis mapping in Ontario. The paper reports on both temporal and spatial patterns of giardiasis distribution in Ontario as well as results of the investigation of evidence of spatial clustering of the disease. Results of investigations of the best spatial scale and disease measure for mapping are also presented.

## Methods

### **Data Collection and manipulation**

The study was conducted in southern Ontario, an area that included approximately 10 million people. All spatial analyses were performed at the census division (CD) ( $n = 40$ ) and the census sub-divisions (CSDs) ( $n = 644$ ) spatial scales. Data on giardiasis cases reported from January 1990 to December 1998 were extracted from the Reportable Disease Information Systems surveillance database [15]. Cases of giardiasis in the database were defined as persons with clinically compatible signs and symptoms and either with an epidemiological link to two or more laboratory confirmed cases, or demonstrating trophozoites or cysts in stool or small bowel specimen. Diagnostic test used was microscopic examination of stool samples.

The personal identifiers of the patients were deleted before the database was released to the investigators. However, the postal code (PC) of residence of the each of the cases was available and was used as the spatial location identifier of the cases. Analyses were first performed with all cases (22,496). Those cases whose risk settings were either hospital (43 cases), local camping (607), vacation (49) or travel (4,766) were then excluded, since their sources of infection likely were outside their respective places of residence, and the analyses repeated to assess their impact on the results.

A Postal Code Conversion File (PCCF) was obtained from Statistics Canada [16]. This file contained all valid PCs and the latitudes and longitudes of their centroids. The PCCF was merged to the disease file using the PC identifier to aid mapping of the spatial distribution of giardiasis cases. The 1991 Canadian census data [17] supplied the standard population distribution for direct standardization of rates, whereas the 1996 Canadian census [18] provided the denominator data for calculation of area specific disease rates.

### **Statistical Analyses**

#### *Creation of spatial weights, rate standardization and smoothing, and temporal analyses*

Since the results of spatial analyses are dependent on the spatial weights used, a number of different row standardized weights (boundary, adjusted boundary, and inverse distance spatial weights) were created at both the CD and CSD spatial scales in ArcView GIS [19] and SpaceStat software [20].

Direct age and sex standardized rates (STDRATES) were computed in STATA [21]. The standardized rates were expressed as the number of cases per 100,000 person-years. Spatial empirical Bayesian (SEB) smoothing was performed at the CSD spatial resolution using inverse

distance weights in SpaceStat [20]. The CSDs were considered as small areas because most of them had 20 or fewer giardiasis cases [22]. Using SEB smoothing, disease rates from low population areas located in regions with no evidence of spatial patterns, are shrunk towards the global mean [7,9]. However, in regions of the map where there exists clear evidence of spatial patterns in disease rates, the less reliable estimates are drawn towards a local rather than a global mean [9]. Therefore the resulting SEB estimators represent a compromise between the overall mean rate, and the local mean of the rates of the nearby areas. The method also uses information from the entire map to make estimates of the rates in areas with missing data.

Temporal patterns were displayed by plotting monthly giardiasis rates against month using Microsoft Excel 2000 [23]. Temporal trend was assessed using an ordinary least squares regression model with log of giardiasis rates as the outcome and months as the predictor variable.

#### Computation of measures of spatial clustering

Moran's I was calculated as:

$$I = \frac{N}{S_o} \sum_i \sum_j w_{ij} \frac{(x_i - u)(x_j - u)}{\sum_i (x_i - u)^2}$$

where N is the number of areas (CDs or CSDs),  $w_{ij}$  is the element in the spatial weights matrix corresponding to the observation pair i, j,  $x_i$  and  $x_j$  are observations for areas i and j with mean u, and  $S_o = \sum_i \sum_j w_{ij}$ . Since the weights were row-standardized implying that each row sums to 1,  $S_o = N$ . Both global Moran's I and Moran Local Indicators of Spatial Association (LISA) were computed for giardiasis rates and their statistical significance tested using 999 Monte Carlo permutations. A spatial correlogram for each of the rates was also computed at the CSD spatial scale using the inverse distance spatial weights. Hat matrix and Cook's D values were calculated and mapped to identify areas with high leverage and undue influence on the global Moran's I values, respectively. All computations were performed in SpaceStat [20].

#### Cartographic displays

All cartographic manipulations and displays were done in ArcView GIS [19]. Jenk's optimization classification method was used for determining the critical intervals for mapping of the STDRATES in the choropleth maps. To ease comparisons of the distribution of the STDRATES with the SEB rates, the class intervals of SEB rates were made to match those of the STDRATES maps. Evidence of local clustering of rates was assessed using Moran scatterplots and Moran significance maps.

## Results

### Temporal distribution

The average number of cases per month was  $1.97(\pm 0.51)$ . There were seasonal patterns in the distribution of giardiasis rates with highest peaks in the late summer to early fall. The apparent winter peak in the distribution of giardiasis rates disappeared when a 3-month moving average was computed. There was also a significant ( $P < 0.001$ ) monthly decrease of 0.005% in the rate of giardiasis (Figure 1). The numerical data for Figure 1 has been presented as a table in the Appendix [see Additional file 1].

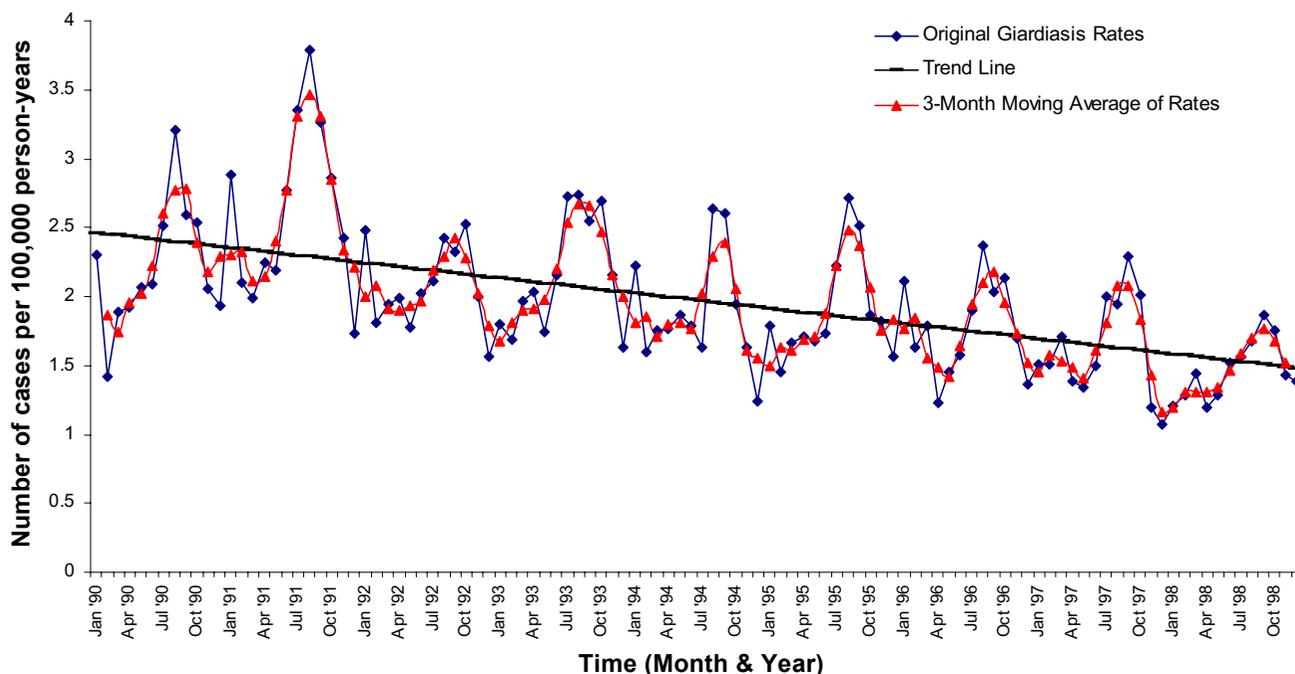
### Spatial distribution

The annual median standardized rate at the CD scale was 2.3 cases/100,000 person-years (range: 0.3 – 3.56). A map of the study area showing the geographical distribution of the 40 CDs is presented in Figure 2 and the description (names) of the CD codes are shown in Table 1. The distribution of age and sex standardized rates at the CD spatial scale is presented Figures 3. Visually, there were high rates (2.81 – 3.56 per 100,000 person-years) in the areas surrounding Georgian Bay and in the counties of Perth, Victoria and Lanark (Figure 3).

The spatial distribution of STDRATES at the CSD spatial scale revealed possible clustering of high rates in the areas of Waterloo and Wellington counties and in the areas surrounding Georgian Bay. There were a number of CSDs with no reported giardiasis cases (Figure 4). In comparison, the SEB smoothed rates produced more distinct patterns of disease distribution making the potential cluster around Wellington and Waterloo counties, those areas to the south and south-eastern borders of Georgian Bay and the areas in eastern Ontario more evident visually (Figures 4 and 5). Most of the areas that had missing (zero rate) values in the STDRATES map had SEB rate estimates.

### Measures of spatial clustering of giardiasis rates

Using both boundary and inverse distance weights, there was non-significant ( $P > 0.05$ ) global spatial auto-correlation of giardiasis rates at the CD spatial resolution (Table 2). However, the results of a global measure of spatial auto-correlation may have been influenced by areas that had high leverage Cook's D values. These CDs were Parry Sound District, Nipissing District county, Middlesex and Halton Regional Municipality. Some counties had significant Moran LISA values (Figure 6). For instance, Wellington county (LISA Moran  $P = 0.032$ ), Durham Regional Municipality ( $P = 0.027$ ) and Oxford county ( $P = 0.037$ ) had significant positive local spatial auto-correlation. All the counties surrounding Georgian Bay, except Simcoe county, were high rate areas surrounded by other high rate areas (high-high).



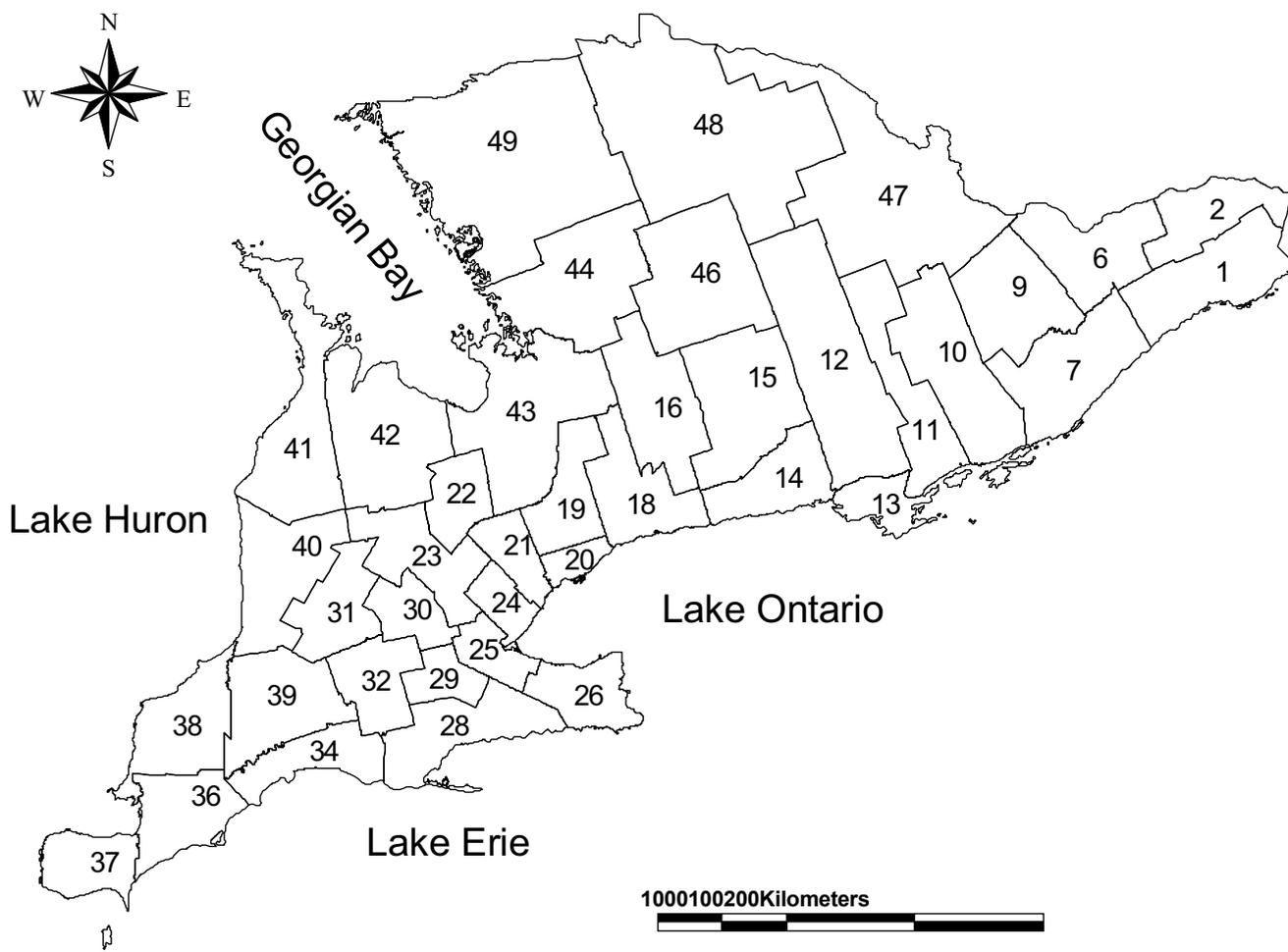
**Figure 1**  
 Monthly distribution of giardiasis rates in southern Ontario. A plot of the monthly rates of giardiasis against time in months. The blue line, with diamond-shaped marks, is the age and sex standardized giardiasis rates. The red line, with triangle marks, is the 3-month moving average of the giardiasis rates. The black straight line is the trend line.

The global Moran test of spatial auto-correlations at the CSD spatial scale using both boundary and inverse distance weights were highly significant ( $P < 0.001$ ). The results of the tests for spatial auto-correlation using inverse distance weights at the CSD spatial scale are shown in Table 3. Significant clustering was observed at only lag 1 in the case of STDRATE, but for both lags 1 and 2 in the case of SEB42 (SEB with distance band 0–42 Arc Distance Units (ADU)) and SEB180 (SEB with distance band 0–180 ADU) (Table 4). Significant values of local Moran's I indicating clustering of high rates were much more evident in the SEB map than STDRATE map (Figures 7 and 8). The visual cluster around the Waterloo Regional Municipality was non-significant. However, significant clustering of high rates were observed in the CSDs in the west and, to a lesser extent, east of Georgian Bay (Figure 7). Significant clustering of high rates (high-high) were also observed in Toronto Metropolitan Municipality when all the cases were included in the analysis. However, this clustering disappeared when cases whose risk setting was traveling, camping or local vacation were excluded from the analysis (Figure not presented).

**Discussion**  
**Geographical and temporal patterns of giardiasis distribution**

The decreasing temporal trend observed in giardiasis distribution could be attributed to better control measures and/or improvement in the quality and/or source of drinking water over the years. The seasonal patterns with highest rates in the summer and early fall have also been reported in other studies [2,24,25] and coincide with increased outdoor activities (camping and swimming). Although there seemed to be winter peaks these disappeared when a 3-month moving average was calculated implying that these peaks may not be important in contrary to results of other studies [26]. The reason for the difference is not clear but may be related to geographical and seasonal variations in cyst contamination of watersheds due to geographical and seasonal differences in patterns of human activity in watersheds.

Although the pattern of high rates in areas surrounding Georgian Bay was visually evident at both the CD and CSD spatial scales it was non-significant at the CD spatial scale, possibly due to edge effect [27,28]. However, a few CSDs in this region had significant clustering of high rates



**Figure 2**  
 Distribution of counties or census division (CDs) of Southern Ontario. A map of southern Ontario showing the distribution of the counties or Census Divisions in the study area. The numbers in the figure represent the unique county identification numbers (codes). The descriptions (names) of the Census Division codes are presented in Table I.

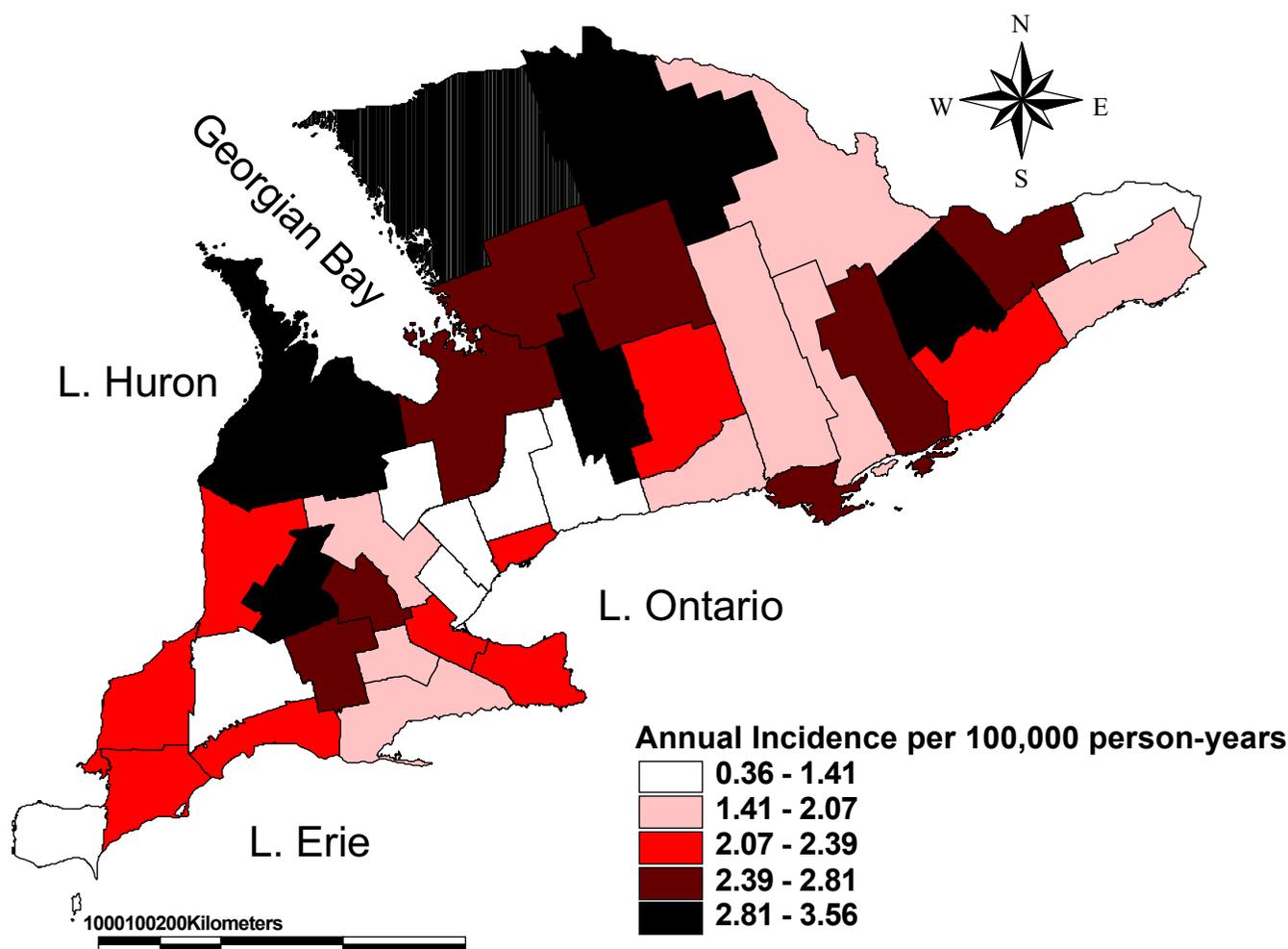
and it was the influence of these CSDs that resulted in the pattern of seemingly high rates at the CD spatial scale. The reason for significant high rate clustering at the CSD spatial scale in areas bordering Georgian Bay could be attributed to increased recreational activities in contaminated water bodies, increased prevalence of infected beavers in these areas or more use of drinking water from surface sources without proper treatment. More detailed studies would need to be carried out to clarify this.

The apparent visual cluster of high rates in the CSDs in Wellington and Waterloo counties was non-significant implying that subjective visual impression of spatial patterns in maps may be misleading. Similar observations have been reported by Walter [29]. Therefore, investiga-

tions should always use both visual and statistical tests to draw conclusions on spatial patterns.

**Smoothing of rates for mapping**

Smoothing of disease rates for mapping is recommended in situations where disease mapping is done in small areas [7,30,31]. In this study a small area was defined as any area with less than 20 cases of giardiasis [22]. Smoothing the rates in these situations was necessary because the rates resulting from these small areas were usually unstable (had high variances) and therefore small variations in the number of cases resulted in dramatic changes in disease rates. It was not necessary to perform smoothing at the CD spatial scale because none of the CDs fitted Cuzick

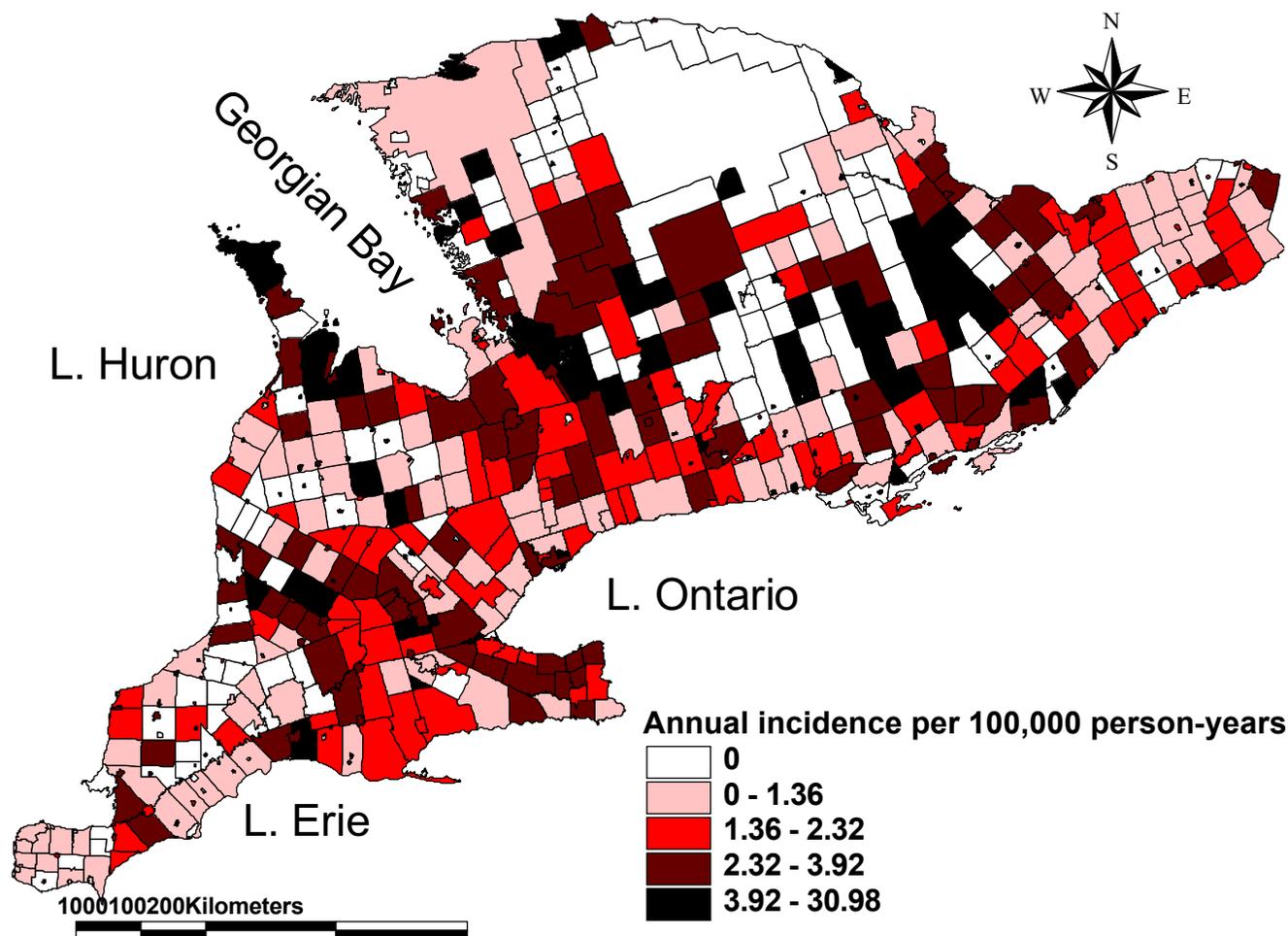


**Figure 3**  
 Distribution of standardized unsmoothed giardiasis rates at Census Division Spatial Scale in southern Ontario. The map shows the distribution of age and sex standardized unsmoothed giardiasis rates at the Census Division spatial scale. The light colored areas had the lowest giardiasis rates while the dark areas had the highest rates.

and Elliott's [22] definition of a small area and therefore their rates were more stable.

Spatial empirical Bayesian smoothing of disease rates was chosen because it utilizes three kinds of information to estimate an area's disease rate: (a) the observed disease events in an area; (b) prior information on the variability of disease rates in the overall map, and (c) information on the disease rates in an area's neighbors since geographically close areas tend to have similar rates of disease [32]. It is worth noting that the standardized rates (STDRATES) ignore both the variability of rates over the entire map and the spatial relationship among neighboring areas [9,30,33]. In areas where there was no apparent local pattern, an area's rate was shrunk towards the global mean.

Consistent with recommendations by Clayton and Kaldor [34], the SEB method smoothed the standardized rates such that in areas where there was a clear evidence of spatial pattern in disease rates, the less reliable estimates were drawn towards a local mean rather than a global one so that after smoothing, the resulting estimators represent a compromise between the overall mean rate, and the local mean of the rates of the nearby areas. This ensures that the final rate estimate is consistent with Tobler's Law (First law of Geography), which states that, 'everything is related to everything else but near things are more related than distant things' [32]. Overall, the SEB smoothing resulted in estimates that were close to the standardized rates when the number of disease events and/or the population at risk in an area were/was large. When the number of disease



**Figure 4**  
 Distribution of standardized unsmoothed giardiasis rates at Census Sub-division Spatial Scale in southern Ontario. The dark areas represent areas with highest rates while the light colored areas had the lowest rates.

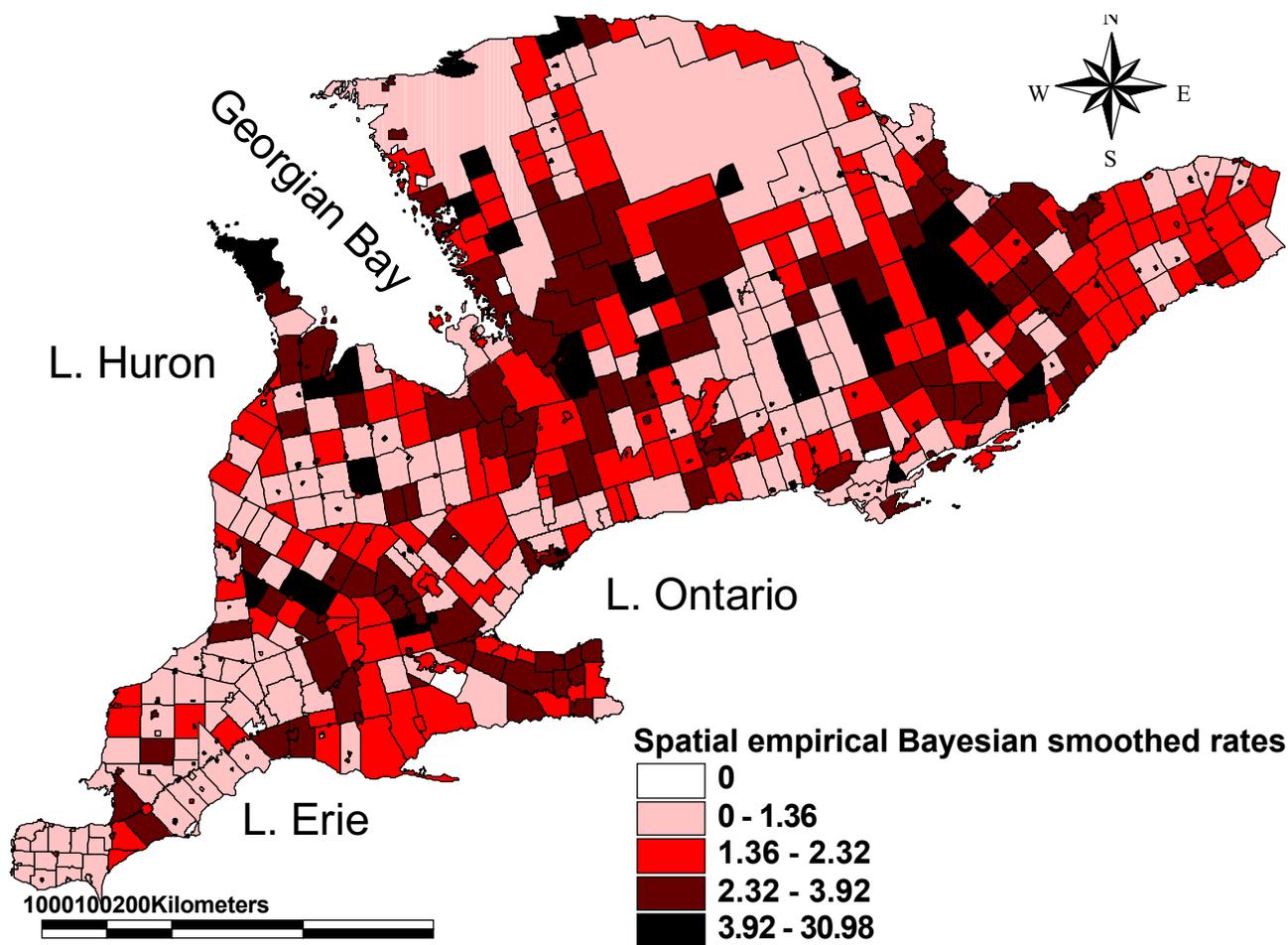
events and/or the population at risk were/was low, prior information on the overall map dominated, thereby shrinking standardized rates towards the overall mean rate. For instance, application of SEB smoothing of the rates from the CSDs in the north of the study area and bordering Nipissing District resulted in shrinking of the rates towards the higher rates of their neighbors.

Standardized rates were sufficient for the CD spatial scale since they were not subject to sampling variation that was encountered at the CSD spatial scale [9,30,33]. In this study, the problem of sampling variation at the CSD spatial scale was aggravated by the fact that areas with low population counts occurred in rural areas occupying large geographical extents, thereby making their visual impact disproportionate. The SEB smoother was important in

reducing this undue 'interest' or attention on these outlying observations.

**Choice of spatial scale of analysis**

Due to the fact that the spatial patterns of disease distribution may change depending on the spatial scale used for mapping, a phenomenon known as the Modifiable Areal Unit Problem (MAUP), all analyses were done at both the CD and CSD in order to identify the appropriate scale for giardiasis mapping. The term MAUP arises from the fact that areal units are not natural but arbitrary constructs that may not necessarily have a relationship with disease distribution [5]. Due to the MAUP there were differences in observed spatial patterns between CD and CSD spatial scales which was quite evident in the areas surrounding Georgian Bay where very high rates in a few CSDs resulted



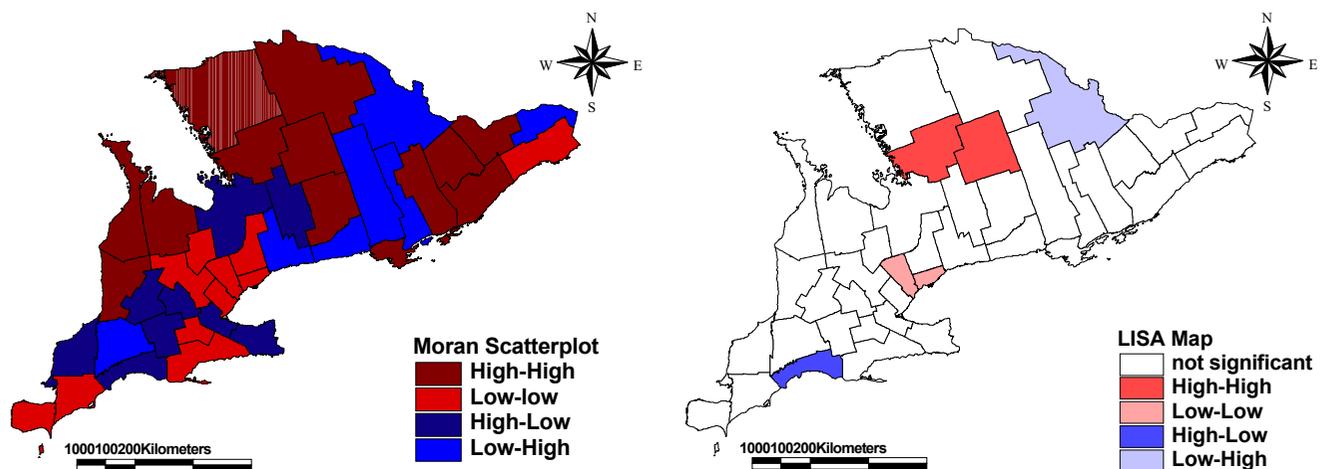
**Figure 5**  
 Distribution of spatial empirical Bayesian smoothed giardiasis rates at Census Sub-division in Southern Ontario. The dark areas represent areas with the highest rates while the light colored areas are those with the lowest rates.

in high rates of giardiasis for the entire CD in which the high rate CSDs were located. It is therefore not surprising that the results of spatial distribution at the two spatial scales were different; most notable was the rate in Nipissing. Other workers have also reported obtaining different results at different spatial scales [6]. This issue has important implications in resource allocation for disease control since it would complicate decisions as to the exact areas that are at higher risk and therefore need more attention [35]. Analyses at the CD spatial scale resulted in the loss of spatial variability induced by the aggregation process. Other authors have reported similar observations [36]. Therefore the CSD seems to be the more appropriate of the two spatial scales since it achieves a better compromise between detail (resolution) and rate stability conditional on appropriate rate smoothing.

**Clustering of giardiasis in Ontario**

None of the global tests of spatial auto-correlation was significant at the CD spatial scale. However, the global Moran's I may be biased by the fact that the measure is based on an integrated analysis of all data which may not be optimal for the study of rate variations in particular local areas. Therefore, disease excess that was limited to one or two regions may not have a significant impact on the overall measure. Walter [29] reports that unless there exists a sufficient level of regional variation in the disease rate, Moran's I statistic may have limited power. We believe this was the case in this study and therefore, LISA measures are more informative in these cases.

The fact that when the cases infected during traveling and vacation were included in the analysis, Toronto Metropol-



**Figure 6**

Moran scatterplot and LISA maps of giardiasis rates at the Census Division spatial scale. The map on the left represents the distribution of local Moran's I values, at the Census Sub-division spatial scale, of giardiasis rates in southern Ontario. This map shows the four types of spatial association observed in the data (i.e. high-high, low-low, high-low, low-high). These four types of spatial association constitute two forms of spatial autocorrelation: positive and negative spatial associations. Positive spatial associations (i.e. association between similar values) are observed in areas marked high-high (i.e. high rate in an area surrounded by high values of the weighted average rate of the neighboring areas), and low-low (low rate in an area surrounded by low values of the weighted average rate of the neighboring areas). There were also two forms of negative spatial associations (i.e. association between dissimilar values); high-low (high rate in an area surrounded by low values of the weighted average rate of the neighboring areas), and low-high (low rate in an area surrounded by high values of the weighted average rate of the neighboring areas). The areas shaded in red had positive spatial autocorrelation while those shaded in blue had negative spatial autocorrelation of giardiasis rates. The map on the right shows the distribution of significant Moran Local Indicators of Spatial Association (LISA) at the Census Division spatial scale. Significant negative spatial association is shown in blue while significant positive association is colored red. Areas that had non-significant LISA values are blank (no color shading).

itan Municipality became a high rate cluster is an indication that most of the cases in this area got infection elsewhere. Most of the CSDs that had significant local clustering of high rates were in more rural areas possibly due to lower quality of drinking water supply in these areas [37]. These areas are also prime vacation sites for people from Toronto.

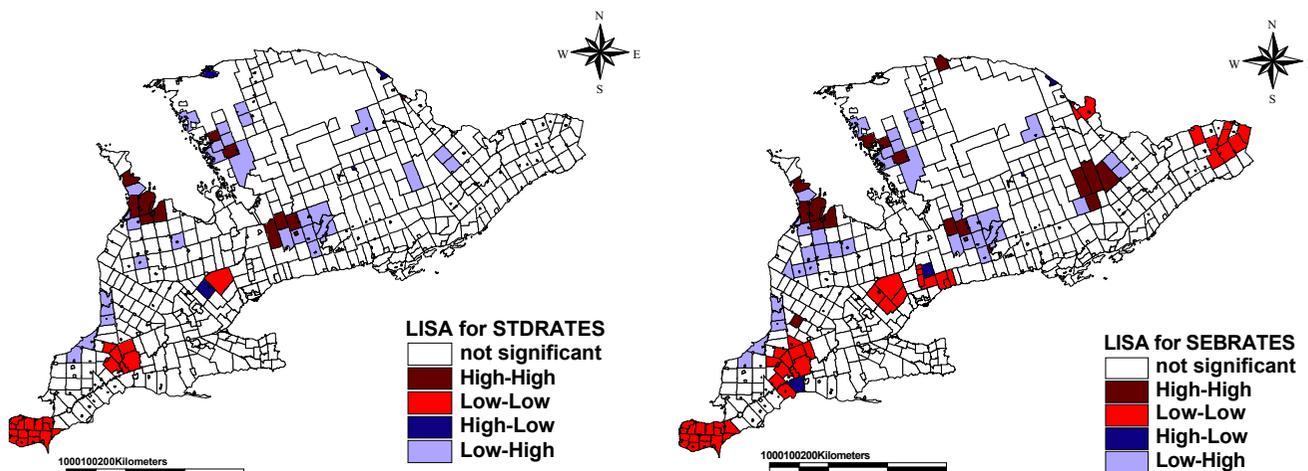
Although the SEB rates enhanced visualization of spatial patterns, the tests of significance resulting from them should be interpreted with caution because they may have a large artefactual component resulting from smoothing [30]. This is evidenced by the fact that all the SEB rates had consistently larger Moran's I values than the STDRATES. Therefore they should only be used for visualization of possible spatial patterns and not for statistical testing of the presence of spatial clustering.

A limitation of the LISA methods is their failure to correct for multiple comparisons while testing for clustering. Some authors have proposed that when using a signifi-

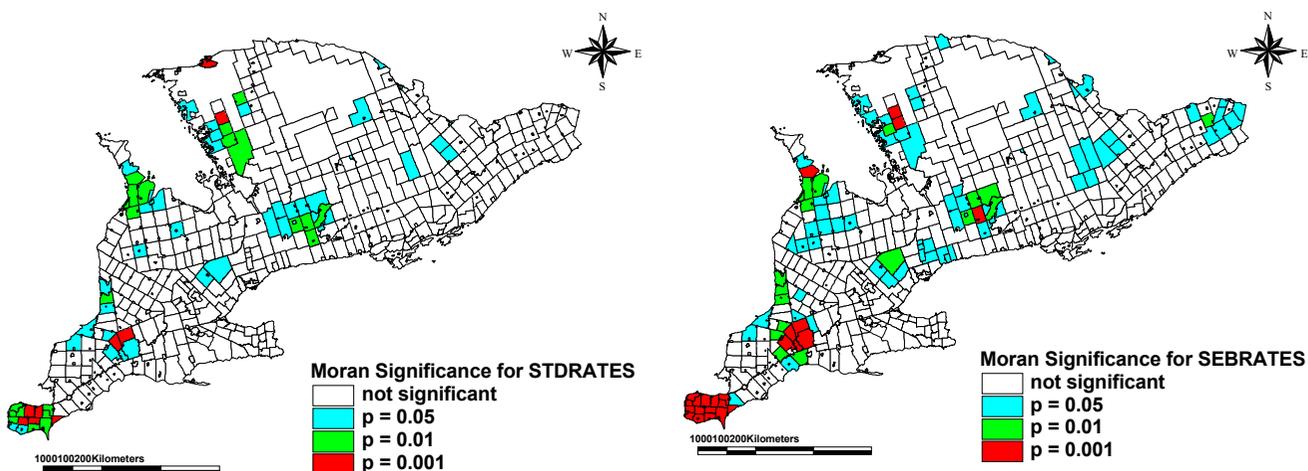
cance level of 0.05 and no correction factor for multiple comparisons, up to 5% of the areas are identified by random chance variation [38]. However, Rothman [39] argues that no adjustments are needed for multiple comparisons since reducing type I error increases type II error thereby reducing the power to detect true statistically significant differences. To our knowledge, there seems to be no consensus on how to adjust for the multiple comparisons when using this kind of methodology [40].

## Conclusions

The study has shown visual and statistical evidence of spatial clustering as well as seasonal patterns and decreasing temporal trend in the distribution of giardiasis rates in Ontario. The study also showed that the CSD spatial scale was a more appropriate scale than the CD for mapping giardiasis rates in Ontario. However, mapping at the CSD spatial scale requires SEB smoothing. The information obtained from this study is useful in guiding decisions in resource allocation to control the illness and in guiding future studies. Further epidemiological investigations to



**Figure 7**  
 LISA maps of STD RATES and SEBRATES at the Census Sub-division spatial scale. The map on the left shows the four types of significant spatial association observed for the standardized unsmoothed rates (STD RATES). The four types are: (i) high-high (high rate in an area surrounded by high values of the weighted average rate of the neighboring areas), (ii) low-low (low rate in an area surrounded by low values of the weighted average rate of the neighboring areas), (iii) high-low (high rate in an area surrounded by low values of the weighted average rate of the neighboring areas), and (iv) low-high (low rate in an area surrounded by high values of the weighted average rate of the neighboring areas). The map on the right shows the distribution of the above mentioned four types of spatial association observed for the spatial empirical Bayesian smoothed rates (SEBRATES). In both maps the areas shaded in red had significant positive spatial autocorrelation while those shaded in blue had significant negative spatial autocorrelation of giardiasis rates. There was stronger spatial association on the map on the right (i.e. for spatial empirical Bayesian smoothed rates). This suggests that the smoothed rates should only be used for visual comparison of spatial patterns and not for statistical analyses since the smoothing may lead to some artefactual effects on the tests for spatial autocorrelation of the rates.



**Figure 8**  
 Moran significance maps of giardiasis rates at Census Sub-division spatial scale. The map on the left shows the distribution of Moran's I significance levels of tests of spatial autocorrelation on standardized giardiasis rates at the Census Sub-division spatial scale while the one on the right shows distribution of the significance levels of the spatial empirical Bayesian giardiasis rates. The different colour shading represent different significance levels.

**Table 1: Description (names) of the Census Division Codes in southern Ontario**

Census Division Code	Census Division Name
1	Stormont, Dundas and Glengarry United Counties
2	Prescott and Russell United Counties
6	Ottawa-Carleton Regional Municipality
7	Leeds and Grenville United Counties
9	Lanark County
10	Frontenac County
11	Lennox and Addington County
12	Hastings County
13	Prince Edward County
14	Northumberland County
15	Peterborough County
16	Victoria County
18	Durham Regional Municipality
19	York Regional Municipality
20	Toronto Metropolitan Municipality
21	Peel Regional Municipality
22	Dufferin County
23	Wellington County
24	Halton Regional Municipality
25	Hamilton-Wentworth Regional Municipality
26	Niagara Regional Municipality
28	Haldimand-Norfolk Regional Municipality
29	Brant County
30	Waterloo Regional Municipality
31	Perth County
32	Oxford County
34	Elgin County
36	Kent County
37	Essex County
38	Lambton County
39	Middlesex County
40	Huron County
41	Bruce County
42	Grey County
43	Simcoe County
44	Muskoka District Municipality
46	Haliburton County
47	Renfrew County
48	Nipissing District
49	Parry Sound District

**Table 2: Tests of spatial autocorrelation of giardiasis rates at the census division spatial scale.**

Disease Measure	Moran's I <sup>1</sup> (P-Value)
STD Rates <sup>2</sup> (1990–92)	-0.090 (0.262)
STD Rates (1993–95)	-0.045 (0.492)
STD Rates (1996–98)	-0.035 (0.538)
STD Rates (All Years)	-0.089 (0.286)

<sup>1</sup> Inverse distance spatial weight with band width of 0–180 great distance units used for computations <sup>2</sup> Age and sex standardized rates

**Table 3: Tests of spatial autocorrelation of giardiasis rates at the census sub-division spatial scale.**

Disease Measure	Moran's I <sup>1</sup> (P-Value)
STD Rates <sup>2</sup> (1990–92)	0.002 (0.266)
STD Rates (1993–95)	0.028 (0.004)
STD Rates (1996–98)	0.004 (0.173)
STD Rates (1990–98)	0.007 (0.097)
SEB Rates <sup>3</sup> (1990–98)	0.019 (0.008)

<sup>1</sup>Inverse distance spatial weight with band width of 0–180 great distance units used for computations <sup>2</sup>Age and sex standardized rates <sup>3</sup>Spatial empirical Bayesian smoothed rates

**Table 4: Moran's I spatial correlogram of giardiasis rates in southern Ontario**

Spatial Weight	STDRATE <sup>1</sup> Moran's I (S.D.) P-value	SEB42 <sup>2</sup> Moran's I (S.D.) P-value	SEB180 <sup>3</sup> Moran's I (S.D.) P-value
Lag 1	0.035 (0.013) 0.011	0.072 (0.0014) 0.003	0.048 (0.013) 0.004
Lag 2	0.006 (0.010) 0.184	0.015 (0.010) 0.05	0.022 (0.010) 0.011
Lag 3	-0.006 (0.009) 0.283	0.0027 (0.009) 0.32	0.003 (0.009) 0.315
Lag 4	-0.003 (0.008) 0.429	0.0062 (0.008) 0.172	0.004 (0.008) 0.25
Lag 5	-0.004 (0.008) 0.348	-0.002 (0.008) 0.481	-0.007 (0.008) 0.216
Lag 6	-0.002 (0.008) 0.489	0.004 (0.008) 0.238	-0.002 (0.008) 0.484

<sup>1</sup>Age and sex standardized rates <sup>2</sup>Spatial empirical Bayesian smoothed rates using 42 great distance units spatial band <sup>3</sup>Spatial empirical Bayesian smoothed rates using 180 great distance units spatial band

investigate risk factors responsible for the observed patterns would provide more information to guide decisions on control strategies.

### Authors' contributions

The author AO was involved in the conceptualization, research design, data collection, execution and preparation of the manuscript. SWM contributed greatly in the design, execution and manuscript preparation whereas PM, JH, and JW collaborated intensely in the research design and manuscript preparation. DM collaborated in the data collection and manuscript preparation.

### Additional material

#### Additional File 1

Title – Appendix: Numeral data for the figure 1 showing monthly distribution of giardiasis rates in southern Ontario. Description: The first column shows the month; column 2 represents the monthly rates of giardiasis presented as number of cases per 100,000 person-years; in column 3 are the natural logarithms of the rates in column 2; column 4 shows the fitted values of the rates, from a time series model, used to draw the trend line; and the values in column 5 are the 3-month centered moving average of the rates.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1476-072X-2-5-S1.wpd>]

### Acknowledgments

We extend our appreciation to the Ontario Ministry of Health and Long-Term Care for providing the data. This research was jointly funded by Health Canada and Department of Population Medicine, University of Guelph. All the work was performed at the Department of Population Medicine, University of Guelph, Ontario, Canada.

### References

- Hill DR: **Giardia Lambli**a. In: *Principles and Practice of Infectious Diseases* 5th edition. Edited by: Mandell GL, Bennette JE, Dolin R. Philadelphia, USA: Churchill Livingstone; 2000:2888-94.
- Furness BW, Beach MJ and Roberts JM: **Giardiasis surveillance – United States, 1992–1997**. *Mor Mortal Wkly Rep CDC Surveill Summ* 2000, **49**:1-13.
- Gyorkos T, Meerovitch E and Prichard R: **Estimates of intestinal parasite prevalence in 1984: report of a 5-year follow-up survey of provincial laboratories**. *Can J Public Health* 1987, **78**:185-7.
- Craun GF: **Waterborne outbreaks of giardiasis. Current status**. *Giardia and giardiasis – biology, pathogenesis, and epidemiology* 1984:243-61.
- Openshaw S: **The modifiable areal unit problem**. *CATMOG No. 38*. Geo Books, Norwich 1984.
- Staines A and Järup L: **Health event data**. In: *Spatial Epidemiology: Methods and Applications* Edited by: Elliott P, Wakefield JC, Best NG, Briggs DG. Great Clarendon Street, Oxford : Oxford University Press; 2000:13-29.
- Clayton D and Bernardinelli L: **Bayesian methods for mapping disease risk**. In: *Geographical and Environmental Epidemiology. Methods for Small-Area Studies* Edited by: Elliott P, Cuzick J, English D, Stern R. Oxford: Oxford University Press; 1997:206-20.
- Heisterkamp SH, Doornbos G and Gankema M: **Disease mapping using empirical Bayes and Bayes methods on mortality statistics in The Netherlands**. *Stat Med* 1993, **12**:1895-913.
- Clayton D and Kaldor J: **Empirical Bayes estimates of age-standardized relative risks for use in disease mapping**. *Biometrics* 1987, **43**:671-81.

10. Gelman A and Price PN: **All maps of parameter estimates are misleading.** *Stat Med* 1999, **18**:3221-34.
11. Lawson A, Bigggeri A, Böhning D, Lesaffre E, Viel JF and Bertollini R: **Disease Mapping and Risk Assessment for Public Health.** West Sussex, England: John Wiley and Sons, Ltd 1999.
12. De Marco R and Zanolin E: **Methods for analyzing and mapping standardized mortality ratios (SMR) for rare causes of death.** *Statistica* 1990, **50**:547-60.
13. Lloyd OL, Wong SL, Yu TS and Wong TW: **Mortality mapping in Hong Kong, 1979-83 and 1984-88: feasibility study and the patterns of cancers.** *Asia Pac J Public Health* 1995, **8**:66-73.
14. Wang K, Vieland V and Huang J: **A Bayesian approach to replication of linkage findings.** *Genet Epidemiol* 1999, **17**(Suppl 1):S749-54.
15. Ontario MOHLTC: **The Ontario Ministry of Health and Long-Term Care.** *Public Health Branch* 1998.
16. Statistics Canada: **Postal Code Conversion File. May 1999 Postal Codes. Reference Guide for Catalogue 92F00227XDB.** Ottawa, Ontario, Canada 1999.
17. Statistics Canada: **Profile of Census Divisions and Census Subdivisions.** *Population Census of Canada* 1991.
18. Statistics Canada: **Profile of Census Divisions and Census Subdivisions.** *Census of Agriculture, Canada* 1996.
19. ESRI: **ArcView GIS. Version 3.2.** Environmental Systems Research Institute, Inc. Redlands, California 1999.
20. Anselin L: **SpaceStat. Version 1.9.** Regional Research Institute of West Virginia University 1995.
21. STATA Corporation: **Intercooled Stata Version 7.0 for Windows 95/98/NT.** Stata Corporation 4905 Lakeway Drive College Station, Texas, USA 1985.
22. Cuzick J and Elliott P: **Small-area studies: purpose and methods.** In: *Geographical and Environmental Epidemiology. Methods for Small-Area Studies* Edited by: Elliott P, Cuzick J, English D, Stern R. Oxford: Oxford University Press; 1997:14-21.
23. Microsoft Corporation: **Microsoft Excel 2000.** Seattle. United States of America 1999.
24. Flanagan PA: **Giardia - diagnosis, clinical course and epidemiology. A review.** *Epidemiol Infect* 1992, **109**:1-22.
25. Daly JJ, Gross MA, McCullough D, McChesney T, Tank SK, Daly EB and Puskarich CL: **Seasonal increase in the incidence of Giardia lamblia in Arkansas.** In: *Advances in Giardia research. Papers from the Calgary Giardia Conference held February 23-25, 1987.* Calgary, Alberta, Canada 1988:71-4.
26. Isaac-Renton J, Moorehead W and Ross A: **Longitudinal studies of Giardia contamination in two community drinking water supplies: cyst levels, parasite viability, and health impact.** *Appl Environ Microbiol* 1996, **62**:47-54.
27. Griffith DA: **The boundary value problem in spatial statistical analysis.** *Journal of Regional Science* 1983, **23**:377-87.
28. Griffith DA and Amrhein CG: **An evaluation of correction techniques for boundary effects in spatial statistical analysis: traditional methods.** *Geographical Analysis* 1985, **15**:352-60.
29. Walter SD: **The analysis of regional patterns in health data. II. The power to detect environmental effects.** *Am J Epidemiol* 1992, **136**:742-59.
30. Bithell JF: **A classification of disease mapping methods.** *Stat Med* 2000, **19**:2203-15.
31. Bernardinelli L and Montomoli C: **Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk.** *Stat Med* 1992, **11**:983-1007.
32. Tobler W: **A computer movie simulating urban growth in the Detroit region.** *Economic Geography* 1970, **46**:234-40.
33. Marshall R: **A review of methods for the analysis of spatial patterns of disease.** *Journal of Royal Statistical Society, Series A* 1991, **154**:421-41.
34. Bernardinelli L, Clayton D and Montomoli C: **Bayesian estimates of disease maps: how important are priors?** *Stat Med* 1995, **14**:2411-31.
35. Atkinson P and Molesworth A: **Geographical analysis of communicable disease data.** In: *Spatial Epidemiology: Methods and Applications* Edited by: Elliott P, Wakefield JC, Best NG, Briggs DG. Great Clarendon Street, Oxford: Oxford University Press.; 2000:251-66.
36. Haining R: **Spatial Statistics and the Analysis of Health Data.** In: *GIS and Health* Edited by: Gatrell AC, Löytönen M. London, UK: Taylor and Francis; 1998.
37. Stirling R, Aramini J, Ellis A, Lim G, Meyers R, Fleury M and Werker D: **Waterborne cryptosporidiosis outbreak, North Battleford, Saskatchewan, Spring 2001.** *Can Commun Dis Rep* 2001, **27**:185-92.
38. Weiss KB and Wagener DK: **Geographic variations in US asthma mortality: small-area analyses of excess mortality, 1981-1985.** *Am J Epidemiol* 1990, **132**:S107-15.
39. Rothman KJ: **No adjustments are needed for multiple comparisons.** *Epidemiology* 1990, **1**:43-6.
40. Thomas DC, Siemiatycki J, Dewar R, Robins J, Goldberg M and Armstrong BG: **The problem of multiple inference in studies designed to generate hypotheses.** *Am J Epidemiol* 1985, **122**:1080-95.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

