Research

Investigation of clusters of giardiasis using GIS and a spatial scan statistic

Agricola Odoi*[1,2], S Wayne Martin[1], Pascal Michel[3], Dean Middleton[4], John Holt[5] and Jeff Wilson[1,6]

Address: [1]Department of Population Medicine, University of Guelph, Guelph, Ontario, Canada, [2]Hamilton District Health Council, Hamilton, Ontario, Canada, [3]Université de Montreal, CP 5000, St-Hyacinthe, Québec, Canada, [4]Ministry of Health and Long-Term Care, Toronto, Ontario, Canada, [5]Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada and [6]Division of Enteric, Foodborne and Waterborne Diseases, Population and Public Health Branch, Health Canada, Guelph, Ontario, Canada

Email: Agricola Odoi* - aodoi@uoguelph.ca; S Wayne Martin - swmartin@uoguelph.ca; Pascal Michel - pascalm@MAGELLAN.UMontreal.CA; Dean Middleton - Dean.Middleton@moh.gov.on.ca; John Holt - jholt@msnet.mathstat.uoguelph.ca; Jeff Wilson - jwilson@uoguelph.ca

* Corresponding author

Abstract

Background: *Giardia lamblia* is the most frequently identified human intestinal parasite in Canada with prevalence estimates of 4–10%. However, infection rates vary by geographical area and localized 'pockets' of high or low infection rates are thought to exist. Water-borne transmission is one of the major routes of infection. Sources of contamination of drinking water include humans, domestic and wild animals. A previous study in southern Ontario, Canada, indicated a bivariate association between giardiasis rates and livestock density and/or manure use on agricultural land; however these variables were not significant when the variable 'rural' was added to the model. In that study, urban areas were defined as those with a minimum of 1,000 persons and a population density of at least 400 persons per Km^2; all other areas were considered rural. This paper investigates the presence of local giardiasis clusters and considers the extent to which livestock density and manure application on agricultural land might explain the 'rural' effect. A spatial scan statistic was used to identify spatial clusters and geographical correlation analysis was used to explore associations of giardiasis rates with manure application on agricultural land and livestock density.

Results: Significant ($P < 0.05$) high rate spatial clusters were identified in a number of areas. Results also showed significant ($P < 0.05$) associations between giardiasis rates and both livestock density and manure application on agricultural land. However, the associations were observed in only two regions.

Conclusions: There is evidence that giardiasis clusters in space in southern Ontario. However, there is no strong evidence to suggest that either livestock density or manure application on agricultural land plays an important role in the epidemiology of giardiasis in the study area. Therefore these factors do not seem to explain the higher rates of giardiasis reported in rural areas. The spatial scan statistics methodology used in this study has an important potential use in disease surveillance for confirming or refuting cluster alarms.

Introduction

Infection with *Giardia* occurs throughout the world, in humans and over 40 other species of animals, with human prevalence ranging from 1 to over 90% depending on location and age of persons sampled [1-3]. In Canada, *Giardia lamblia* is the most frequently identified human intestinal parasite with prevalence estimates of 4–10% [4,5]. However, even within a country, infection rates vary widely and localized clusters of high or low infection rates exist [1]. In Ontario, Canada, the prevalence of *Giardia* infection is reported to have seasonal patterns suggesting seasonally active risk factors [6,7].

The main mode of infection is ingestion of *Giardia* cysts. This may occur by ingestion of contaminated water, food or by hand-to-mouth transfer of cysts, but water-borne transmission is thought to be one of the major routes of infection. There are a number of sources of water contamination, including domestic and wild animals as well as human sources [8]. In particular, infected cattle have been suspected to be an important source of human giardiasis [9,10]. Concentrations of *Giardia* cysts in water have been found to be significantly associated with the prevalence of the *Giardia* in animals [11]. Moreover, significantly higher concentrations of *Giardia* cysts have been reported in watersheds with cattle activity than in those with no cattle access [12].

A previous study carried out in Ontario using several multi-variable spatial regression techniques found statistically significant bivariate associations between giardiasis rates and both livestock density and manure use on agricultural land [13]. However, these associations were not significant when the variable 'rural' was added to the model. In that study, urban areas were defined as those with a minimum of 1,000 persons and a population density of at least 400 persons per Km^2; all other areas were considered rural [14]. The findings from the above study seem to suggest that different modes of transmission of giardiasis may be important in different local geographical regions. Therefore, this paper investigates the presence of clusters of giardiasis in southern Ontario and explores the extent to which livestock density and manure application on agricultural land might explain the 'rural' effect observed in the above study. If livestock plays a major role in the transmission of giardiasis to humans, then we would expect clusters of high giardiasis rates in areas with high livestock densities. This would have important implications for public health professionals since control strategies would vary depending on the most important risk factors in a particular location.

In this study we use geographical information systems (GIS) and a spatial scan statistic to investigate geographical clusters of human giardiasis reported to a surveillance

system in southern Ontario. The spatial scan statistic implemented in SaTScan software [15] offers several advantages: it corrects for multiple comparisons, adjusts for the heterogeneous population densities among the different areas in the study, detects and identifies the location of the clusters without prior specification of their suspected location or size thereby overcoming pre-selection bias, and the method allows for adjustment for covariates [16,17]. Geographical correlations of giardiasis rates with livestock density and manure application on agricultural land are also explored to assess the extent to which these factors might explain the higher rates of giardiasis reported in rural areas.

Methodology

Study area and data collection

The study included the area of Ontario to the south of latitude 46.25029 (approximately south of the French River). This area is referred to, throughout this article, as southern Ontario. Data on 22,496 cases of human giardiasis reported in southern Ontario from January 1, 1990 to December 31, 1998 were extracted from the Reportable Disease Information System (RDIS) surveillance database. The extracted data contained date of birth, gender, postal code (PC) of residence, date of disease onset, laboratory date, treatment date, date of death and possible sources of infection.

A Postal Code Conversion File (PCCF) containing all valid postal codes (PC) as well as the names of each of the Census Sub-divisions (CSD) (i.e. municipalities) where the PCs were located was obtained from Statistics Canada [14]. This file contained the latitudes and longitudes of the centroids of each of the PCs. The spatial location of each patient was identified as the centroid of their residential PC; this was later aggregated to the CSD level. The use of PC centroids to identify the spatial location of the patients was appropriate for the purposes of this study because all spatial analyses were performed at higher geographical levels than the PC and therefore no errors were introduced by assigning patients to their residential PCs. However, we note that if individuals seek health care in Ontario, their place of residence (which may not necessarily be the place of likely exposure) is recorded in RDIS.

The 1996 Canadian population census [18] provided the denominators for calculation of spatial empirical Bayesian smoothed giardiasis rates. Land-use data were extracted from the 1996 census of agriculture [19].

Data manipulations

To add geographical co-ordinates to the giardiasis data, the latter were merged to the PCCF using the PC identifier. Internal consistency of the disease data was evaluated by checking fields for implausible values, screening for

duplicate dates (birth dates, date of episode, laboratory dates, treatment dates and date of death) and checking for chronological plausibility of events. Duplicate records were identified and removed. Cases for which the risk settings (or probable sources of infections) were hospital (43 cases), local camping (607 cases), local vacation (49 cases) or travel (4,766) were excluded from the analyses.

Cattle density was calculated as: (1) numbers of cattle per hectare of agricultural land and (2) numbers of cattle per hectare of pasture land. The two measures of cattle density were selected so as to explore possible differences in their propensity to contaminate drinking water. Similar computations were performed for livestock density. Livestock was defined as the total number of cattle, pigs, sheep and goats. Intensity of use of animal manure was calculated as the percentage of agricultural land on which manure was applied. The calculation included all types of animal manure applied using a variety of application techniques such as solid spreaders, liquid spreaders and irrigation systems.

Statistical and geographical analyses
(a) Detection and identification of giardiasis clusters
A spatial scan statistic implemented in a software program, SaTScan, was used to test for the presence of giardiasis spatial clusters and to identify their approximate locations [15,20-24]. The theory behind the spatial scan statistic is a generalization of a test proposed by Turnbull and coworkers [25]. The statistic uses a circular window of variable radius that moves across the map. The radius of the window varies from 0 up to a specified maximum value. As the window of the statistic moves across the map, it defines a set of different neighboring CSDs. If the window contains the centroid of a CSD, the whole CSD is included in the window [24]. The cluster assessment is performed by comparing the number of cases within the window with the number expected if cases are randomly distributed in space. The test of significance of the identified clusters is based on a likelihood ratio test [22] whose p-value is obtained through Monte Carlo testing [26].

Identification of spatial high or low rate clusters was done under the Poisson probability model assumption using a maximum spatial cluster size of 5% of the total population. For statistical inference, 999 Monte Carlo replications were performed. The null hypothesis of no clusters was rejected when the simulated p-value was less than or equal to 0.05 for the primary cluster and 0.1 for the secondary clusters since the latter have conservative p-values.

To get a better understanding of the disease distribution, the spatial distribution of the clusters was compared to the distribution of spatial empirical Bayesian smoothed giardiasis rates. The latter were chosen for comparison due

to the fact that the census sub-divisions are small areas and therefore have unstable rates [7]. For details on how these were computed, please refer to an earlier article [7].

(b) Cartographic and GIS manipulations
All GIS manipulations and cartographic displays were performed in ArcView GIS [27]. The geographical distributions of agricultural and land-use factors including livestock densities and the percentage of agricultural land on which manure was applied were also mapped. Jenk's optimization classification method was used to determine the critical intervals for livestock density and manure-use maps [28]. This method identifies the critical intervals using a statistical formula which identifies groupings and patterns inherent in the data. Since it uses patterns inherent in the data, it minimizes the sum of the variance within each of the classes and therefore attains a goodness of variance fit of 0.91, and it maximizes information about both the map area and the parameter being mapped resulting in a very efficient map [28]. Visual comparisons of the spatial patterns in the distribution of the above potential risk factors and distribution of giardiasis clusters were then made.

(c) Geographical correlation analyses
Since data on livestock density and manure application on agricultural land were available at the Census Consolidated Sub-divisions (CCSD) spatial level, this geographical unit was used as the unit of analysis for the assessment of correlations between the land-use variables (livestock density and manure application on agricultural land) and giardiasis rates. As a first step, global Spearman's rank correlation coefficients between giardiasis rates and land-use factors (cattle and livestock densities as well as manure application) were calculated for all areas. For health planning purposes, the Ontario Ministry of Health and Long-Term Care has sub-divided the province into seven health planning regions. In the second step, Spearman's correlation analyses were repeated for each of these health planning regions in order to assess geographical differences in associations between giardiasis rates and the land-use factors.

Results
Distribution of high rate spatial giardiasis clusters
A map showing the distribution of the counties in southern Ontario is presented in Figure 1 and the distribution of spatial empirical Bayesian smoothed rates of giardiasis is shown in Figure 2. Visual examination of Figures 1 and 2 reveals the presence of potential clusters in areas surrounding Georgian Bay, as well as in Waterloo, Wellington, Oxford, Lanark, Frontenac and Lennox and Addington counties. Formal cluster analysis identified several clusters with high giardiasis rates (Table 1 and Figure 3). The largest cluster (cluster 3) was in areas

Table 1: Significant high rate giardiasis spatial clusters in southern Ontario, 1990–98

Cluster[a]	No. of Cases in cluster	Expected No. of cases	Annual cases per 100,000 person-yrs	Cluster RR[b]	No. of CSDs[c]	P-value
1	936	491	3.7	1.91	1	0.001
2	908	502	3.5	1.81	1	0.001
3	1160	759	3	1.53	124	0.001
4	1102	744	2.9	1.48	60	0.001
5	340	166	4	2.05	1	0.001
6	177	84	4.1	2.12	1	0.001
7	439	290	2.9	1.51	5	0.001
8	23	3	13.5	7	1	0.001
9	16	1	22.3	11.5	1	0.001
10	33	8	7.4	4.34	2	0.001
11	751	586	2.5	1.23	3	0.001
12	21	3	12.2	6.3	1	0.001
13	25	6	7.8	4	1	0.001
14	103	56	3.5	1.83	11	0.001
15	14	2	13.2	6.81	1	0.001
16	21	5	8.2	4.23	1	0.001
17	23	8	5.9	3.06	1	0.001
18	276	212	2.5	1.3	14	0.001
19	96	61	3.1	1.59	2	0.056
20	19	6.3	5.8	3	1	0.082

[a] See Figure 3 for the spatial distribution of the geographically large clusters [b] Rate ratio [c] Census Sub-divisions

surrounding Georgian Bay including the Bruce Peninsula on the west and Parry Sound, Muskoka, Haliburton and Nipissing on the East. The second largest cluster (cluster 4) included parts of the following counties: Huron, Grey, Bruce, Wellington, Perth, Oxford and Waterloo (Figures 1 and 3). A number of the significant clusters presented in Table 1 are not visible in the map (Figure 3) due to their geographically small sizes.

Eight clusters with significantly low giardiasis rates were also identified (Table 2 and Figure 4). The primary low rate cluster (cluster 1) was in Middlesex county. Cluster 2 was principally in York Regional Municipality whereas cluster 3 included parts of Halton and Peel counties. Cluster 8 included Stormont, Dundas and Glengarry United Counties, Prescott and Russell United Counties as well as Ottawa-Carleton Regional Municipality.

Distribution of livestock densities and manure use
The median cattle density per Consolidated Census Sub-Division (CCSD) was 2.27 (range: 0.21, 19.3) cattle per hectare of pasture land. The highest cattle density values (3.7 – 19.3 cattle per hectare of pasture land) were observed in the counties of Waterloo Regional Municipality, Wellington, Perth and Oxford counties extending northwards to the counties bordering western Georgian Bay and Lake Huron (Huron, south Bruce and south Grey counties) (Figures 1 and 5). Similar cattle densities were

also observed in Prescott and Russell United counties. High densities (2.5–3.7 cattle per hectare of pasture land) were also recorded for Peel Regional Municipality, Peterborough, Renfrew, and Stormont, Dundas and Glengarry United counties. The lowest cattle density values (0.2 – 0.9 cattle per hectare of pasture land) were recorded in parts of Parry Sound, Muskoka, Haliburton, and a number of central – eastern counties. The spatial distributions of total livestock (cattle, sheep, pigs and goats) followed similar patterns as those of cattle densities with highest densities seen in Perth, Waterloo and the surrounding counties stretching northwards to Bruce county (figure not presented). To the south-west of the study area, only Essex county had low livestock densities. The median livestock density was 3.42 (range: 0.17, 87.9) animals per hectare of pasture land. The spatial distribution of the proportion of agricultural land on which manure was applied followed a similar pattern as the distribution of cattle and total livestock densities (figure not presented).

Correlations between giardiasis rates and land-use factors
Spearman correlation coefficients and their associated p-values for relationships between giardiasis rates and land-use factors are shown in Table 3. When analyses included all geographical regions in the study area, significant but relatively low correlation coefficients were observed only between giardiasis rates and cattle density per hectare of agricultural land ($r = 0.11$; $P = 0.007$) and between
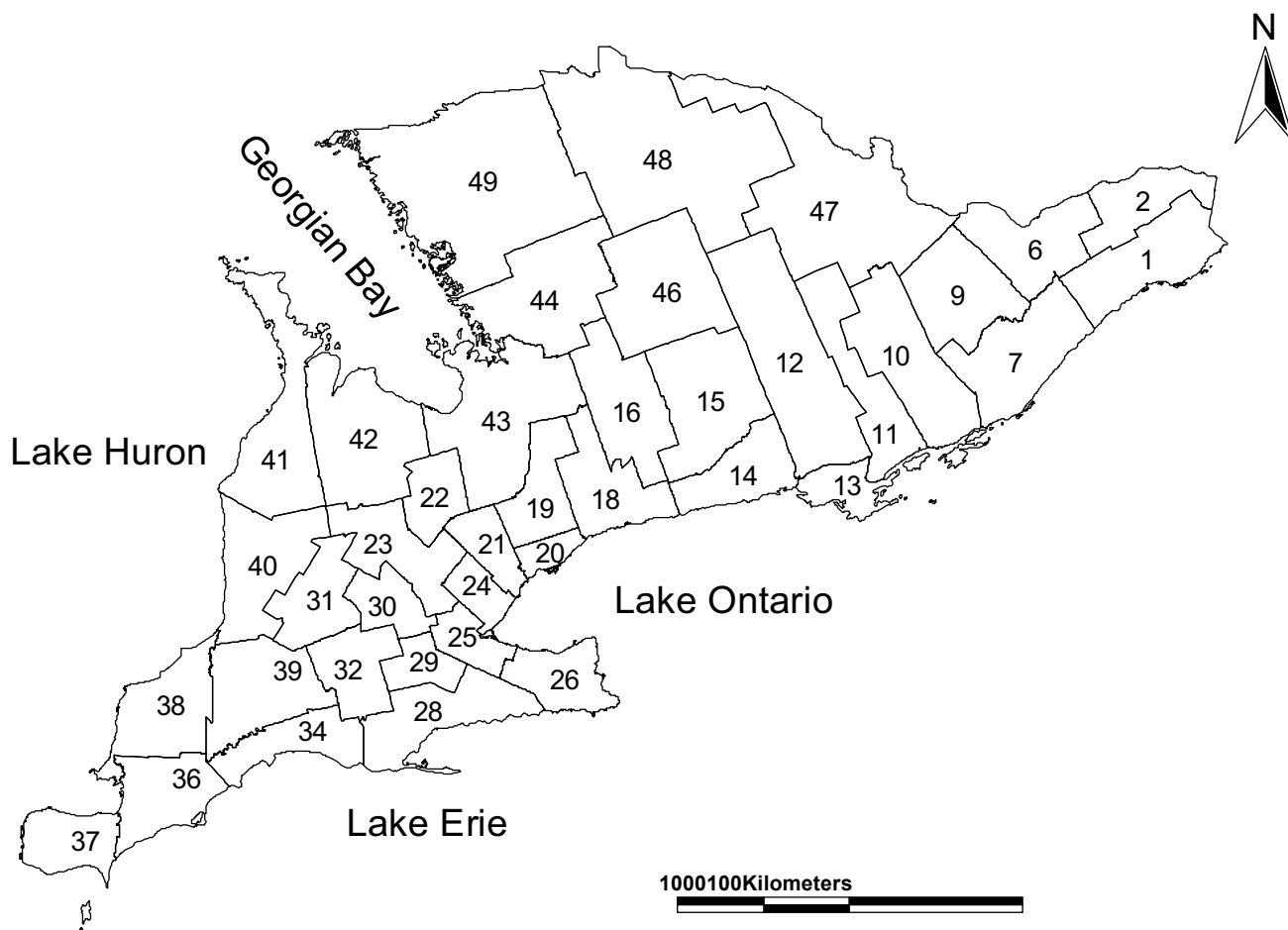
Figure 1
Distribution of counties or census division (CDs) of southern Ontario. A map of southern Ontario showing the distribution of the Census Divisions in southern Ontario. The numbers in each of the census division polygons are the census division identification codes. The identification codes and names of each of the census divisions (or counties) are as follows: (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)

giardiasis rates and proportion of agricultural land under manure application ($r = 0.09$; $P = 0.037$). However, when analyses were performed for individual health planning regions (Figure 6) in the study area, more variables were significantly associated with giardiasis rates in some health planning regions. For instance, there was significant correlation between giardiasis rates and cattle density per hectare of pasture land in the Central West Health Planning Region ($r = 0.38$; $P = 0.013$) implying that cattle density might explain approximately 14.4% of the variation in giardiasis distribution in that planning region. Significant but low correlation coefficients were also observed between giardiasis rates and the proportion of agricultural land on which manure was applied in both
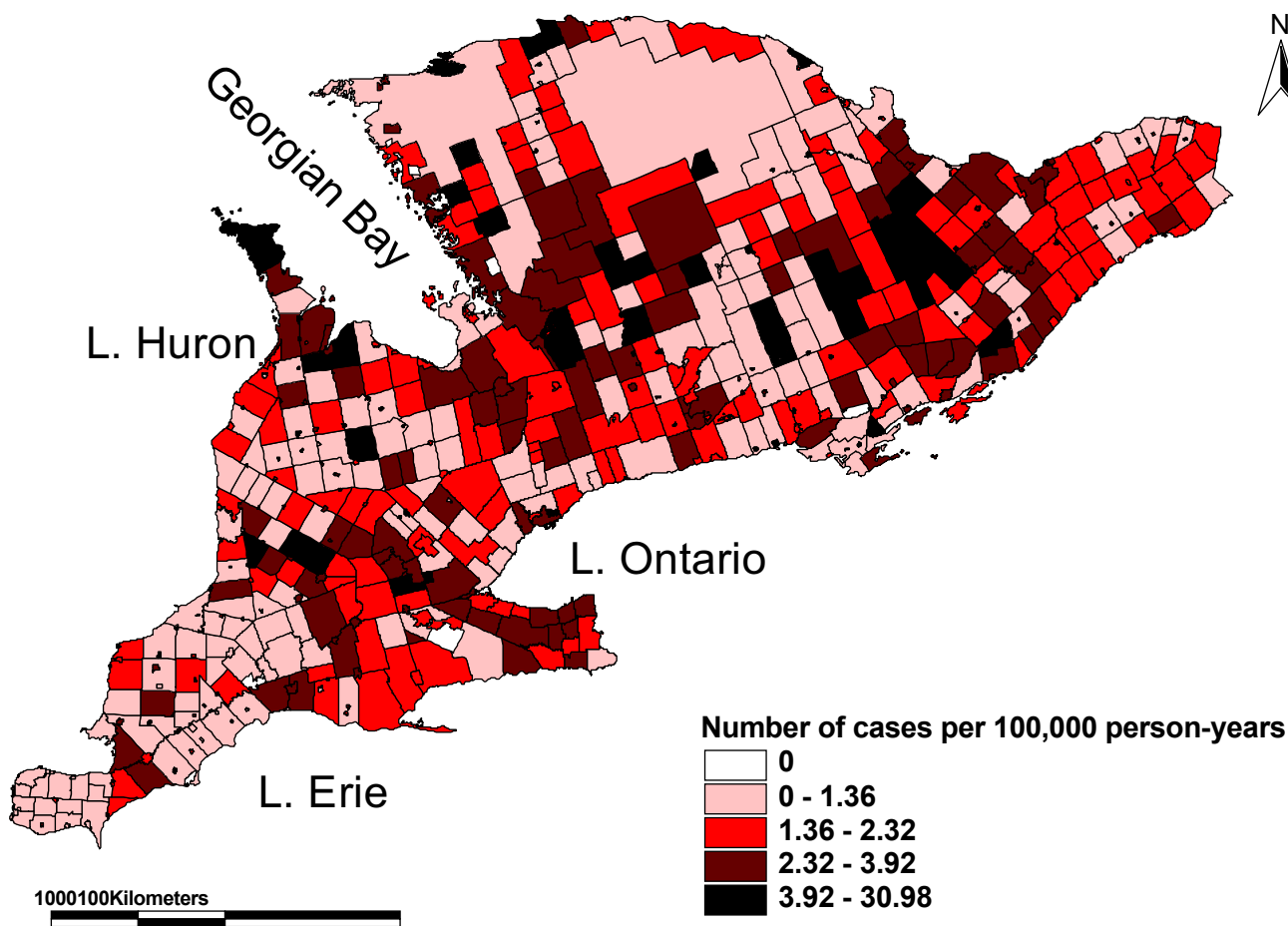
Figure 2
Distribution of spatial empirical Bayesian smoothed giardiasis rates in southern Ontario (1990–98). The light colored areas had the lowest giardiasis rates while the dark areas had the highest rates.

Central West (r = 0.31; P = 0.047) and South West (r = 0.17; P = 0.019) Health Planning Regions. These health planning regions had the highest giardiasis rates and livestock densities as well as some of the significant high rate giardiasis clusters. No significant correlations were observed in the other areas (Table 3).

Discussion
Spatial distribution of giardiasis clusters
The results suggest that there were 'hot-spots' of human giardiasis in a number of areas in southern Ontario. The distribution of the clusters in the Central-west and South-west areas was consistent with a possible involvement of a common risk factor; namely livestock density. However, the distribution of giardiasis clusters in most of the other high risk areas did not coincide with the distribution of

high livestock density and/or manure application suggesting that other modes of transmission might be more important in these areas. It is worth mentioning that although differential reporting biases across the study area can not be totally ruled out as a reason for some of the observed clusters, it is unlikely to be a major problem since reporting of giardiasis is mandatory in Ontario. Due to financial and time constraints this study did not assess reporting biases.

Since relatively low but significant correlation coefficients between giardiasis rates and both livestock densities and manure application were observed but in only two health planning regions (Table 3), it implies that these two factors may not play an important role in the overall epidemiology of giardiasis in southern Ontario. This does not
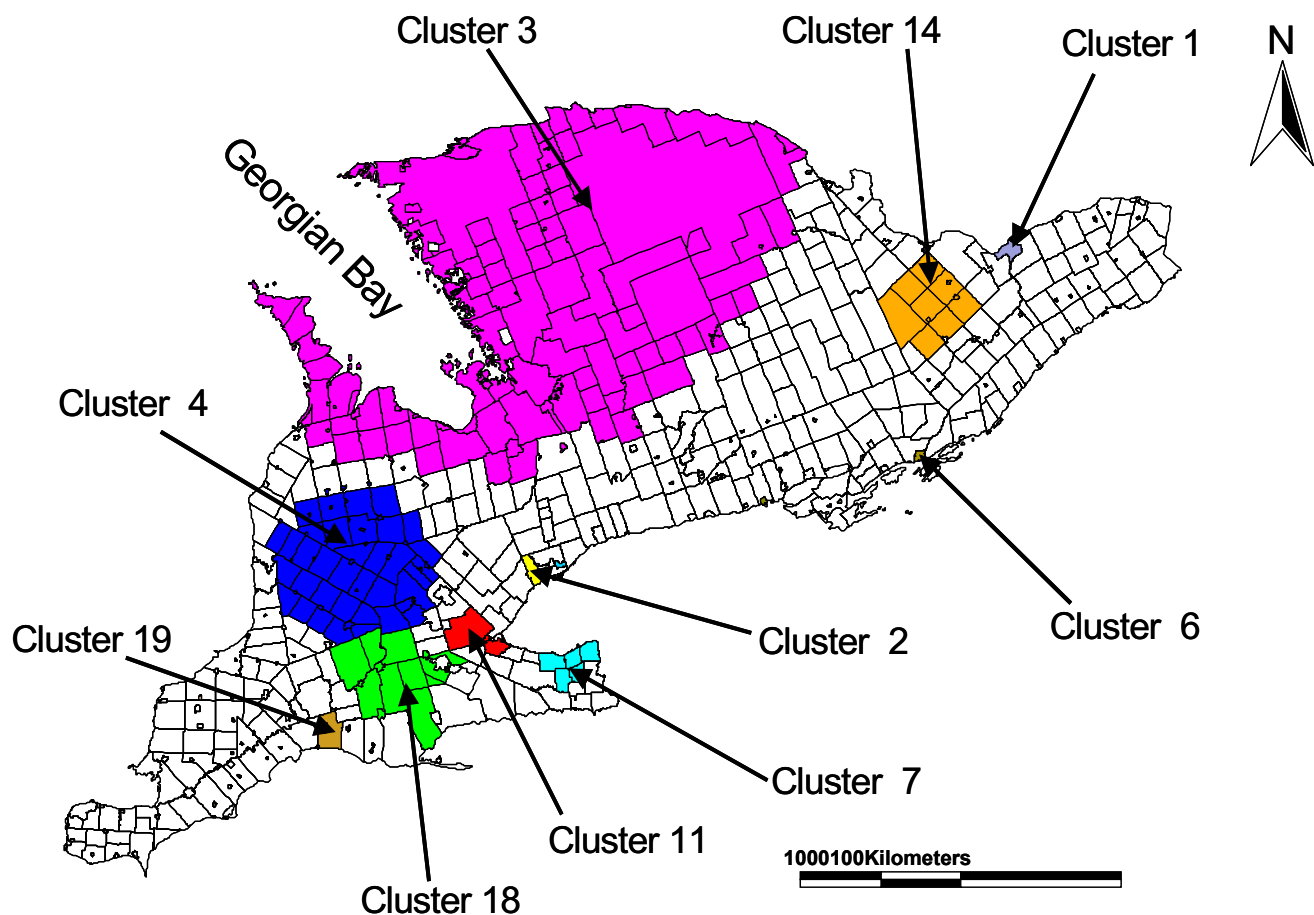
Figure 3
Spatial distribution of significant high rate giardiasis clusters in southern Ontario (1990–98). Only geographically large clusters have been presented in this map. Detailed descriptions of these clusters as well as the geographically smaller clusters (not presented in this figure) are presented in Table 1. Numerical identification of the clusters are in order of their likelihood ratio; the cluster with the highest likelihood ratio is cluster 1 (most likely cluster) while cluster 2 had the second highest likelihood ratio, etc.

Table 2: Significant low rate giardiasis spatial clusters in southern Ontario, 1990–98

Cluster[a]	No. of Cases in cluster	Expected No. of cases	Annual cases per 100,000 person-yrs	Cluster RR[b]	No. of CSDs[c]	P-value
1	102	587	0.3	0.17	15	0.001
2	297	704	0.8	0.42	4	0.001
3	257	542	0.9	0.47	5	0.001
4	275	541	1	0.51	19	0.001
5	494	722	1.3	0.68	6	0.001
6	339	457	1.4	0.742	1	0.001
7	2	20	0.2	0.1	6	0.019
8	105	166	1.2	0.63	14	0.065

[a] See Figure 4 for the spatial distribution of the geographically large clusters [b] Rate ratio [c] Census Sub-divisions
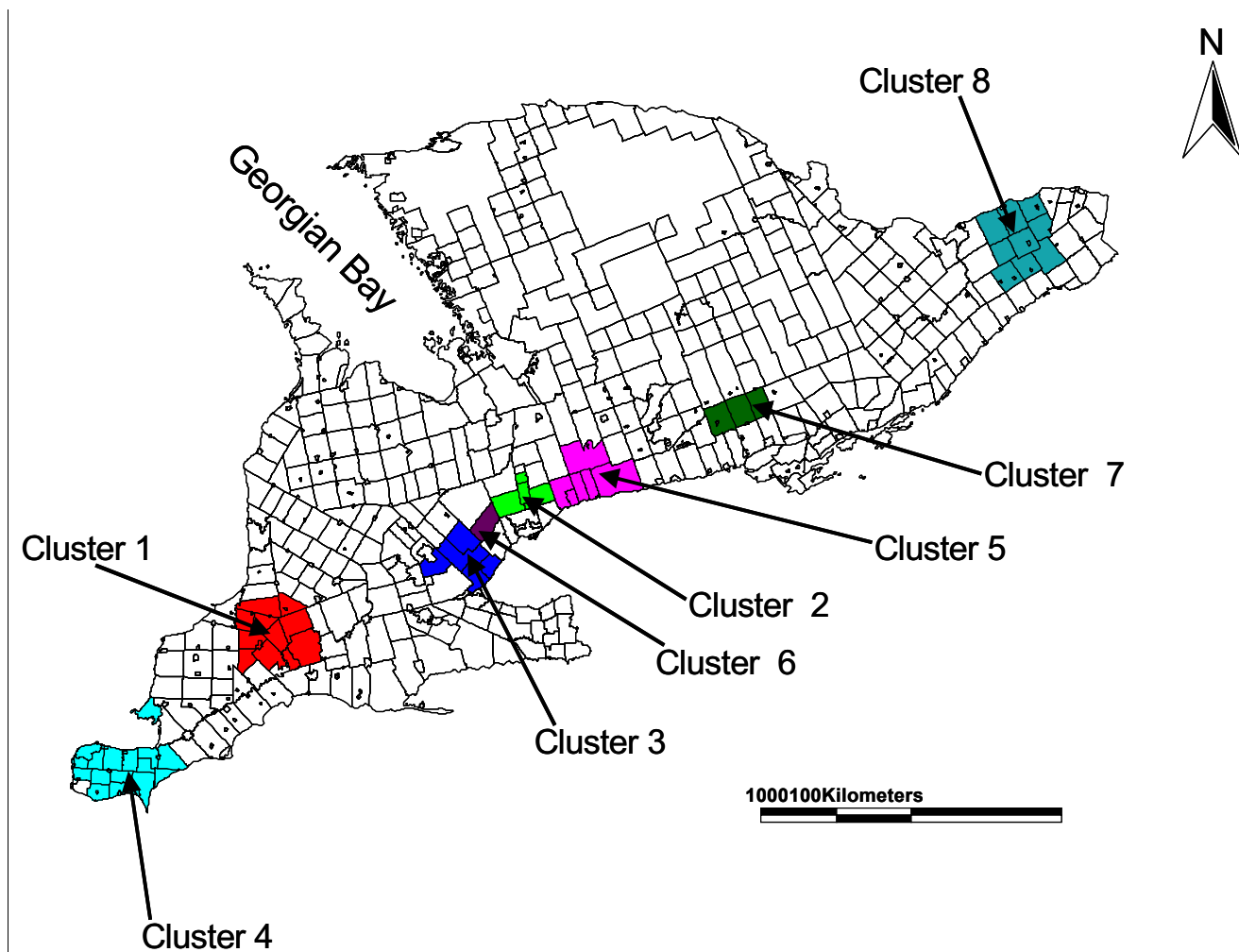
Figure 4
Spatial distribution of significant low rate giardiasis clusters in southern Ontario (1990–98). The numerical identification of the clusters are in order of their likelihood ratio; the cluster with the highest likelihood ratio is cluster 1 (most likely cluster) while cluster 2 had the second highest likelihood ratio, etc. For more detailed cluster information, refer to Table 2.

necessarily contradict reports from other epidemiological studies that showed that cattle and other farm animals play a role in giardiasis transmission to humans [9,10,12]. It is possible that other modes of infection were more important than contamination of drinking water by livestock manure in certain areas of Ontario. For example, parts of cluster 3 were in areas surrounding Georgian Bay, a key holiday destination during the summer, suggesting either possible contamination of drinking water through increased human activities in watersheds or increased contact with water contaminated from other sources. It is also likely that other modes of transmission, such as contamination of water by wild animals and person-to-person transmission as occurs in day care centers, may be

important in different local areas. It is worth noting that although other studies have reported epidemiological evidence implicating cattle and other farm animals in the transmission of giardiasis to humans [9,10,12], to our knowledge there has been no experimental evidence to confirm this.

The cluster detection (SaTScan) methodology
Some of the high rate clusters identified in this study were too large and may be unrealistic when compared to the distribution of the empirical Bayesian rates. This is most likely due to the fact that the borders of the clusters are uncertain and therefore need to be interpreted with caution. The uncertainty of the borders arises because often
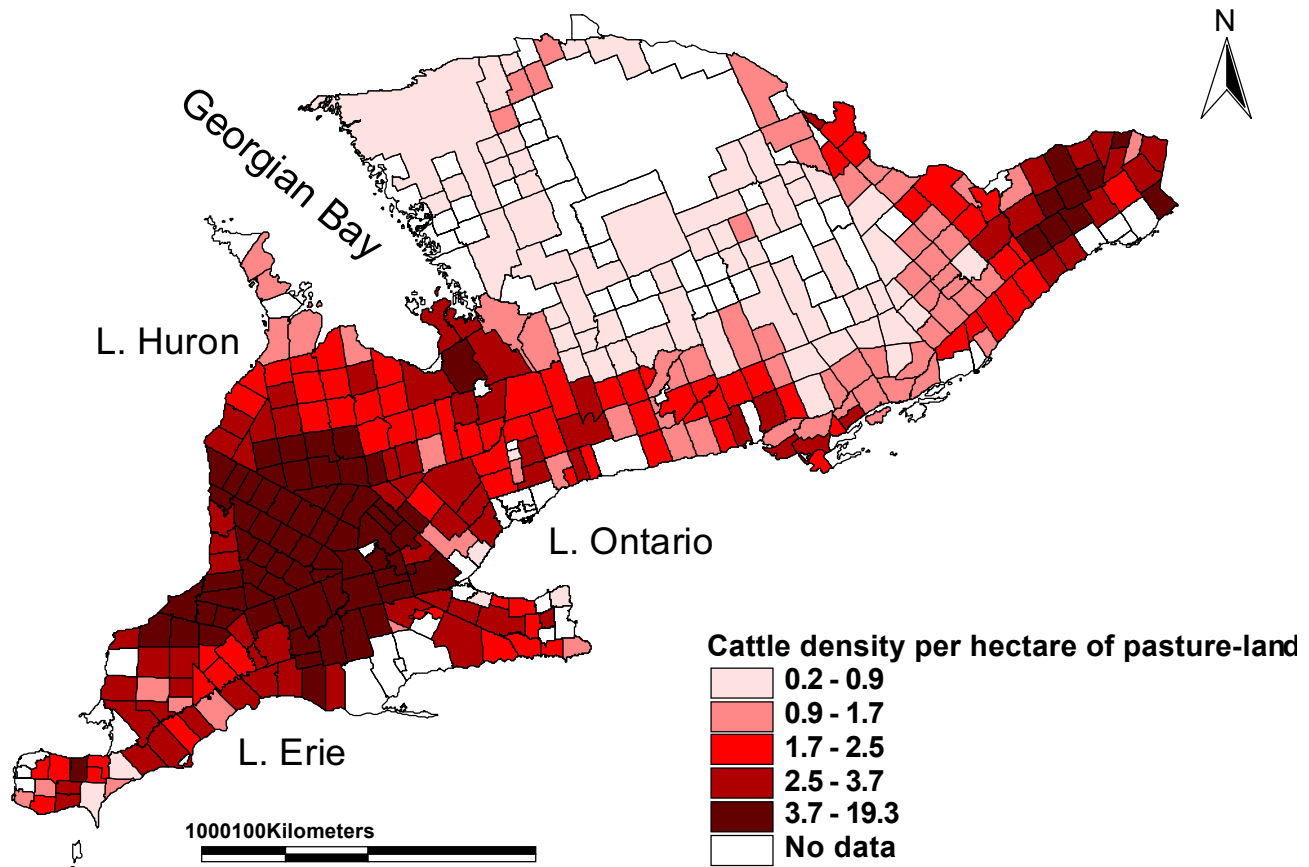
Figure 5
Spatial distribution of cattle density in southern Ontario. The areas with dark shades of red had the highest cattle densities and those with lighter shades had lower cattle densities.

Table 3: Spearman's rank correlation coefficients for the relationships between giardiasis rates and land-use factors

Health Planning Region	Correlation and (p-value)				
	[a]CDAL	[b]CDPL	[c]LDAL	[d]LDPL	[e]PMA
All Regions	[f]0.11 (0.007)	0.018 (0.673)	[g]0.079 (0.064)	0.0009 (0.983)	[f]0.089 (0.037)
Central West	[f]0.299 (0.05)	[f]0.382 (0.013)	[g]0.262 (0.093)	[f]0.328 (0.034)	[f]0.311 (0.045)
South West	[f]0.167 (0.019)	0.055 (0.440)	0.108 (0.128)	0.026 (0.719)	[f]0.166 (0.019)
North	0.007 (0.969)	0.130 (0.511)	0.033 (0.862)	0.097 (0.607)	0.011 (0.953)
Central East	0.091 (0.41)	0.010 (0.927)	0.085 (0.440)	0.016 (0.882)	0.084 (0.442)
Central South	0.168 (0.422)	0.006 (0.979)	0.119 (0.556)	0.244 (0.261)	0.263 (0.186)
East	0.026 (0.735)	0.029 (0.708)	0.046 (0.560)	0.017 (0.830)	0.037 (0.652)
Toronto	-	-	-	-	-

[a] Cattle density (per hectare of agricultural land) [b] Cattle density (per hectare of pastureland) [c] Livestock density (per hectare of agricultural land) [d] Livestock density (per hectare of pastureland) [e] Proportion of agricultural land on which manure was applied. [f] Significant correlations at 5% significance level [g] Significant correlations at 10% significance level
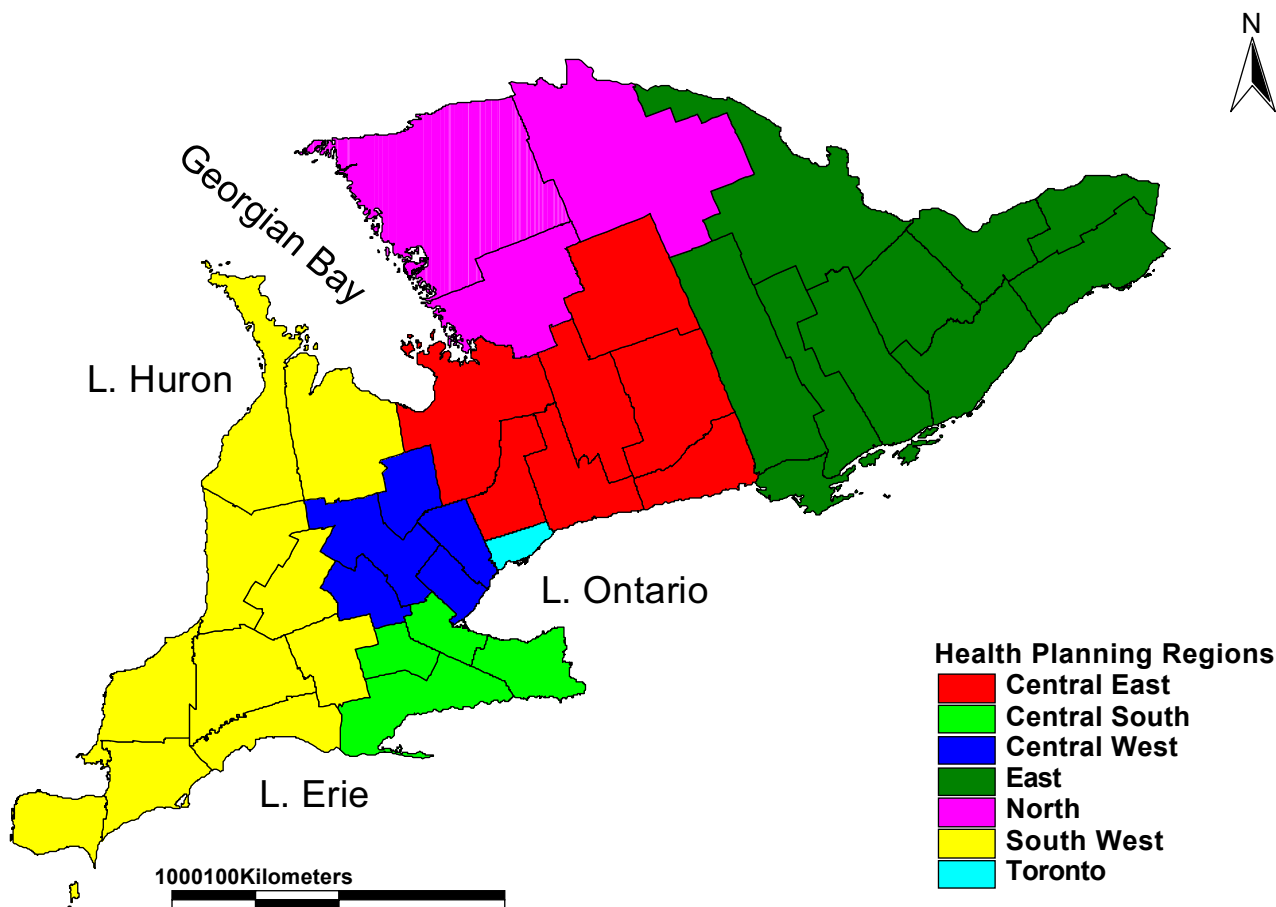
Figure 6
Geographical distribution of health planning regions in southern Ontario

there are many windows that overlap with the potential cluster. These windows may have only slightly lower disease rates and therefore it is probable that some of these areas get erroneously included in the cluster [23]. It is also possible that some secondary clusters in the neighborhood of the primary clusters might actually be parts of the primary cluster [24]. A study by Tango (2000) also reported unrealistically large cluster sizes after applying the procedure on several data sets where many clusters were evident [29]. He observed that this occurred more frequently in the presence of more than one cluster in the study region as opposed to when only one cluster was present. This suggests that the results of cluster analysis should be interpreted together with knowledge of the spatial distribution of rates, especially spatial empirical Bayesian rates [29-31]. It should also be borne in mind that the p-values of the secondary clusters are conservative [24] and therefore they under-estimate their true significance.

For this reason, the significance of the secondary clusters were assessed at the 10% significance level whereas the primary clusters were assessed at the 5% level.

Quite often public health authorities need to respond to demands to investigate potential clusters of different diseases and confirm or refute, with certainty, that a problem exists [32]. However, due to the complexity and cost of rigorous epidemiologic cluster studies, they are usually not able to thoroughly investigate all potential clusters. Therefore, using surveillance data and cluster investigation statistical methodologies, health officials would be able to identify statistically significant clusters and therefore prioritize the clusters that need thorough epidemiological investigations. Systematic use of cluster investigation techniques as part of regular surveillance activities would provide additional intelligence necessary to improve population health. Although we have used the

methodology on retrospective data, it can be applied on prospective data as well and would be a very useful tool for public health epidemiologists in disease surveillance. Moreover, the SaTScan software is free of charge and can be down-loaded from http://www.satscan.org. Lastly, as has been pointed out by Ward and Carpenter [33], publication of these studies will assist epidemiologists and statisticians to address a number of current methodological issues.

Conclusions

This study has shown the presence of 'hot-spots' of giardiasis in southern Ontario. The study has also demonstrated that using existing health data, GIS and spatial scan statistics could provide public health officials with additional tools necessary for disease surveillance. The low correlation coefficients between giardiasis rates and both livestock density and manure application in only two health regions implies that livestock density and manure use do not explain most of the higher rates of giardiasis reported in rural areas. More detailed individual level epidemiological investigations need to be carried out in the identified 'hot-spots' to identify the most important determinants of disease distribution and assess the burden of illness due to giardiasis.

Authors' contributions

AO was involved in the conceptualization, research design, execution and write-up of the first draft of the manuscript. SWM and PM collaborated in conceptualization and research design. DM organized and provided the disease data whereas JH, and JW contributed in the research design. All authors were involved in the preparation of the manuscript.

Acknowledgments

References

1. Flanagan PA: **Giardia – diagnosis, clinical course and epidemiology. A review.** *Epidemiol Infect* 1992, **109:**1-22.
2. Kappus KD, Lundgren RGJr, Juranek DD, Roberts JM, Spencer HC: **Intestinal parasitism in the United States: update on a continuing problem.** *Am J Trop Med Hyg* 1994, **50:**705-713.
3. Craun GF: **Waterborne outbreaks of giardiasis. Current status.** In *Giardia and giardiasis: biology, pathogenesis, and epidemiology* Edited by: Erlandsen SL, Meyer EA. New York: Plenum Press; 1984:243-261.
4. Gyorkos T: **Estimation of parasite prevalence based on submissions to provincial laboratories.** *Can J Public Health* 1983, **74:**281-284.
5. Gyorkos T, Meerovitch E, 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:**185-187.
6. Greig JD, Michel P, Wilson JB, Lammerding AM, Majowicz SE, Stratton J, Aramini JJ, Meyers RK, Middleton D, McEwen SA: **A descriptive analysis of giardiasis cases reported in Ontario, 1990–1998.** *Can J Public Health* 2001, **92:**361-365.
7. Odoi A, Martin SW, Michel P, Holt J, Middleton D, Wilson J: **Geographical and temporal distribution of human giardiasis in Ontario, Canada.** *Int J Health Geogr* 2003, **2:**5.
8. Steiner TS, Thielman NM, Guerrant RL: **Protozoal agents: what are the dangers for the public water supply?** *Annu Rev Med* 1997, **48:**329-340.
9. Buret A, denHollander N, Wallis PM, Befus D, Olson ME: **Zoonotic potential of giardiasis in domestic ruminants.** *J Infect Dis* 1990, **162:**231-237.
10. Warburton AR, Jones PH, Bruce J: **Zoonotic transmission of giardiasis: a case control study.** *Commun Dis Rep CDR Rev* 1994, **4:**R32-36.
11. Ongerth JE, Hunter GD, DeWalle FB: **Watershed use and Giardia cyst presence.** *Water Res* 1995, **29:**1295-1299.
12. Ong C, Moorehead W, Ross A, Isaac-Renton J: **Studies of Giardia spp. and Cryptosporidium spp. in two adjacent watersheds.** *Appl Environ Microbiol* 1996, **62:**2798-2805.
13. Odoi A, Martin SW, Michel M, Holt J, Middleton D, Wilson J: **Determinants of the geographical distribution of endemic giardiasis in Ontario, Canada: a spatial modelling approach.** *Epidemiol Infect* in press.
14. Statistics Canada: *Postal code conversion file – May 1999 postal codes. Reference Guidefor catalogue 92F0027XDB. Ottawa, Ontario, Canada* 1999.
15. Kulldorff M, Rand K, Gherman G, Williams G, DeFrancesco D: **Software for the spatial and space-time scan statistics. SaTScan. Version 2.1.** *National Cancer Institute* Bethesda, MD, USA; 1998.
16. Kulldorff M, Nagarwalla N: **Spatial disease clusters: detection and inference.** *Stat Med* 1995, **14:**799-810.
17. Kulldorff M: **Prospective time periodic geographical disease surveillance using a scan statistic.** *J R Stat Soc Ser A* 2001, **164:**61-72.
18. Statistics Canada: *Profile of census divisions and census sub-divisions. Population census of Canada. Ottawa, Ontario, Canada* 1996.
19. Statistics Canada: *Profile of census divisions and census sub-divisions. Census of agriculture. Ottawa, Ontario, Canada* 1996.
20. Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR: **Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico.** *Am J Public Health* 1998, **88:**1377-1380.
21. Kulldorff M: **A spatial scan statistic.** *Commun Stat Theory Methods* 1997, **26:**1481-1496.
22. Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N: **Childhood leukaemia in Sweden: using GIS and a spatial scan statistic for cluster detection.** *Stat Med* 1996, **15:**707-715.
23. Kulldorff M, Feuer EJ, Miller BA, Freedman LS: **Breast cancer clusters in the northeast United States: a geographic analysis.** *Am J Epidemiol* 1997, **146:**161-170.
24. Kulldorff M, Nagarwalla N: **Spatial disease clusters: detection and inference.** *Stat Med* 1995, **14:**799-810.
25. Turnbull BW, Iwano EJ, Burnett WS, Howe HL, Clark LC: **Monitoring for clusters of disease: application to leukemia incidence in upstate New York.** *Am J Epidemiol* 1990, **132:**S136-143.
26. Dwass M: **Modified randomization tests for non-parametric hypothesis.** *Ann Math Statist* 1957, **28:**181-187.
27. ESRI: **Arcview GIS version 3.2.** *Environmental Systems Research Institute, Inc. Redlands, California, USA* 1999.
28. Cliff AD, Haggett P: **Atlas of Disease Distributions: Analytic Approaches to Epidemiological Data.** *Oxford, United Kingdom: Basil Blackwell Ltd* 1988.
29. Tango T: **A test for spatial disease clustering adjusted for multiple testing.** *Stat Med* 2000, **19:**191-204.
30. Rushton G, Lolonis P: **Exploratory spatial analysis of birth defect rates in an urban population.** *Stat Med* 1996, **15:**717-726.
31. Thacker SB, Berkelman RL: **Public health surveillance in the United States.** *Epidemiol Rev* 1988, **10:**164-190.
32. Slovik P: **Perception of risk.** *Science* 1987, **236:**280-285.
33. Ward MP, Carpenter TE: **Techniques for analysis of disease clustering in space and in time in veterinary epidemiology.** *Prev Vet Med* 2000, **45:**257-284.