# Determinants of the geographical distribution of endemic giardiasis in Ontario, Canada: a spatial modelling approach

A. ODOI[1]*, S. W. MARTIN[1], P. MICHEL[3], J. HOLT[4], D. MIDDLETON[5]
AND J. WILSON[1,2]

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

## SUMMARY

Giardiasis surveillance data as well as drinking water, socioeconomic and land-use data were used in spatial regression models to investigate determinants of the geographic distribution of endemic giardiasis in southern Ontario. Higher giardiasis rates were observed in areas using surface water [rate ratio (RR) 2·36, 95% CI 1·38–4·05] and in rural areas (RR 1·79, 95% CI 1·32–2·37). Lower rates were observed in areas using filtered water (RR 0·55, 95% CI 0·42–0·94) and in those with high median income (RR 0·62, 95% CI 0·42–0·92). Chlorination of drinking water, cattle density and intensity of manure application on farmland were not significant determinants. The study shows that waterborne transmission plays an important role in giardiasis distribution in southern Ontario and that well-collected routine surveillance data could be useful for investigation of disease determinants and identification of high-risk communities. This information is useful in guiding decisions on control strategies.

## INTRODUCTION

*Giardia lamblia* is the most frequently identified intestinal parasite in North America [1–3]. Its waterborne transmission is well documented in Canada and the United States and it is an increasingly important public health concern [4–10]. However, there are information gaps that have to be addressed in order to put control strategies in place. For instance, many studies of risk factors of human giardiasis have involved outbreaks [11–14]. Information from these outbreaks may over-emphasize the importance of risk factors associated with the outbreaks due to more concentrated exposure of defined population groups. In contrast, very few studies have investigated the determinants of endemic giardiasis [15–17]. In addition, few studies have investigated the determinants of variations in its geographical distribution. This has important implications for public-health efforts to control the disease since it is possible that different risk factors may be important in different geographical areas. For instance, land-use factors may be more important in rural than urban areas. Therefore, identification of these factors would guide decisions on control strategies applied in different areas depending on the most important risk factors at play.

Due to the cost of specific longitudinal studies, most epidemiological studies investigating determinants

* Author for correspondence.
(Email: aodoi@uoguelph.ca)

of disease have been restricted to relatively small samples of individuals in limited geographical areas. Consequently, the data obtained from such studies have limitations in the target population to which the results can be inferred, not to mention the difficulty of comparing results from studies carried out in different regions during different periods of time. These limitations could be overcome by using existing databases that cover wider areas and larger target populations, as alternatives to traditional and more costly longitudinal epidemiological studies [18]. Since the mid-1980s, many countries, including Canada, have initiated surveillance for giardiasis. These surveillance systems can be important sources of data on regional variations of disease incidence and may be useful for the detection of high-risk communities and investigation of potential risk factors. However, for the surveillance data to be useful for these purposes, it is imperative that their quality is consistent across the areas under study.

Analysis of surveillance data to investigate the determinants of disease distributions inevitably implies that data across different areas have to be studied. Such analyses face problems of spatial autocorrelation arising from the fact that geographically neighbouring areas tend to have similar disease rates because of the non-discrete nature of the populations and risk factors that do not necessarily segregate along arbitrary geographical boundaries. Hence, usual statistical techniques are not appropriate for these data as they may lead to biased parameter estimates or invalid standard errors depending on the nature of spatial autocorrelation [19]. Proper analyses of these data require use of spatial modelling techniques that adjust for the spatial dependence inherent in the data.

The objective of this study was to investigate and identify determinants of the geographical distribution of giardiasis cases reported to the Reportable Disease Information System (RDIS) between 1990 and 1998 in southern Ontario. Using spatial regression models, drinking-water characteristics were investigated as disease determinants while adjusting for socioeconomic and land-use factors.

## METHODOLOGY

### Study area and data collection

The study area, with a population of 9·98 million residents (approximately 93% of Ontario's total population), included the area of Ontario south of latitude 46·25.

### Water data

A database containing water attribute and distribution information from 257 public water works (PWWs) in southern Ontario was obtained from the Ontario Ministry of the Environment and Energy (MOE) and individual PWWs. A PWW was defined as 'any water works capable of supplying water at a rate greater than 50 000 litres per day or a water works that supplies water for domestic purposes and serves more than five private residences' [20]. The PWW attribute data included water sources, treatment regimes and population served. Maps showing water distribution areas of each PWW were obtained from municipalities, townships, cities and towns. Data from CanMap Desktop Mapping Technologies (Markham, Ontario, Canada) (which contain streets, bodies of water, and municipal and provincial boundaries) were used as visual guides for on-screen digitizing of all water distribution areas [21].

### Giardia *and geo-reference data*

Data on cases of giardiasis reported in southern Ontario from January 1990 to December 1998 were extracted from the RDIS database. Among other things, the database had information on date of birth, age, sex, postal code (PC) of residence of the case and date of disease onset. All the personal identifiers including the name and street address were deleted before the database was released to the investigators. Cases were defined as persons with clinically compatible signs and symptoms of giardiasis and with either an epidemiological link to two or more laboratory-confirmed cases, or demonstrating trophozoites or cysts in stool or small bowel specimens.

A 1998 Postal Code Conversion File (PCCF) containing all valid PCs as well as geographical co-ordinates (latitudes and longitudes) of their centroids was obtained from Statistics Canada [22]. The PCCF also contained the names of each of the Census Subdivisions (CSDs) (municipalities) and Census Divisions (CDs) (counties) where the PCs were located. Although the PCs were not the units of analysis, they were used to assign the cases of giardiasis to their respective water distribution areas and CSDs. The water distribution areas were the units of analyses. Each water distribution area is comprised of several PCs.
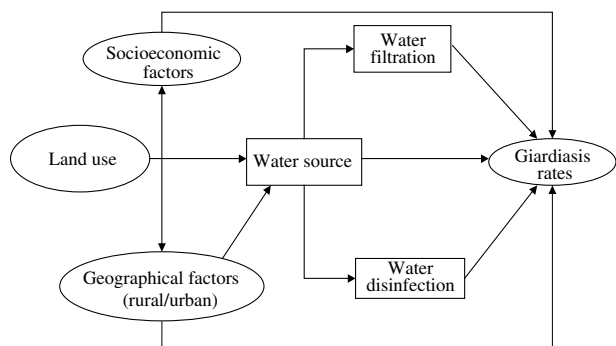
**Fig.** Proposed conceptual model for determinants of giardiasis rates in southern Ontario.

*Socioeconomic and land-use data*

Socioeconomic data were extracted from the 1996 Canadian population census [23]. Although the socioeconomic data were available at the CSD geographical scale, their values were assigned to their respective water distribution areas using the 'spatial join' feature of ArcView GIS [24]. Land-use data were extracted from the 1996 census of agriculture [25].

**Data manipulation**

All data manipulations were done in SAS [26] and ArcView GIS [24]. The RDIS giardiasis database was merged with the PCCF to enable Geographical Information Systems (GIS) manipulations. Since there was no common field that could be used to merge the patient data with the water data, the spatial join feature of ArcView GIS [24] was used. This was achieved using a 'point-in-polygon' join technique whereby the points were centroids of the PCs of residence of the patients and the polygons were the respective water distribution areas supplying water to the respective PCs. This enabled linking of each giardiasis case to their respective water distribution areas. The file resulting from the spatial join was then merged with the land-use and socioeconomic data using the CSD identification.

**Data analysis**

The main potential disease determinants under investigation were the drinking-water characteristics. These were divided into three main variables; water source [ground water (GW) or surface water (SW)]; water filtration (yes or no) and water disinfection (yes or no). The conceptual causal model on which the statistical analyses were based is shown in the Figure. It was believed *a priori* that the source of

Table 1. *Socioeconomic and land-use variables used in the regression analyses*

| Variable description | Units |
| --- | --- |
| Proportion of land where manure was applied using solid spreader | Proportion |
| Proportion of land where manure was applied using irrigation system | Proportion |
| Proportion of land where manure was applied using liquid spreader | Proportion |
| Total proportion of land where manure was applied | Proportion |
| Cattle density | Cattle/km² |
| Density of beef cattle | Cattle/km² |
| Density of dairy cattle | Cattle/km² |
| Density of pigs | Pigs/km² |
| Density of chickens | Chickens/km² |
| Median income | $ (/10 000) |
| Unemployment rate | Per cent |
| Rural areas* | — |

* Urban areas were defined as areas with a minimum of 1000 people and a population density of at least 400/km² [22]. All other areas were considered rural.

water influenced the type of treatment performed since more water systems using SW tended to perform both filtration and disinfection compared to water systems that used GW sources. The list of variables investigated for potential association or confounding with giardiasis rates is shown in Table 1. Socioeconomic factors, rural/urban status and land-use were considered as potential confounders. The socioeconomic variables investigated were unemployment rate and median income. High median income was considered a proxy measure of better quality life. A number of land-use variables were also explored for their potential role as risk factors or confounders in the model. For instance, higher animal density and manure application on agricultural land may be associated with higher levels of contamination of drinking-water sources.

Giardiasis rates per water distribution area were normalized using a log transformation. Geometric mean and median giardiasis rates were calculated by water source and treatment factor combination and the rates expressed as the number of cases per 100 000 person-years.

*Creation of spatial weights matrices for spatial analyses*

Five inverse distance spatial weights, with different distance bands, were created in SpaceStat [27].

Distance bands refer to the critical distance within which two areas would be considered to have a spatial relationship. Thus, any two areas which had a distance between their centroids greater than the upper limit of the specified distance band would be considered to have no spatial relationship, otherwise they would have a spatial relationship (i.e. would be neighbours). The distance bands considered were 0–69, 0–126, 0–217, 0–355 and 0–833 Arc Distance Units (ADU). The upper limits of the distance bands were the quantiles of the spatial distances.

The ADU (also called great circle distance unit) refers to the distance calculated from global coordinate systems (latitudes and longitudes). The arc distance, $d_{ij}$, between two locations $i$ and $j$ is calculated as follows:

$$d_{ij} = 3959 \times arc \cos{(\mathrm{Cos}|Y_i - Y_j|} \\ \times \mathrm{Sin}\ X_i \times \mathrm{Sin}\ X_j + \mathrm{Cos}\ X_i \times \mathrm{Cos}\ X_j),$$

where $X$ and $Y$ are the latitudes and longitudes transformed to radians as follows:

$$X = (90 - \mathrm{latitude} \times \Pi)/180,$$
$$Y = \mathrm{longitude} \times \Pi/180.$$

Each of the inverse distance spatial weights was row standardized so that each row of each of the matrices summed to unity. This was necessary to produce meaningful measures of spatial autocorrelation (Moran's $I$) and to restrict the value of the autoregressive coefficient in spatial regression analyses to values in the range of $-1 < \rho < 1$.

### Test for spatial autocorrelations

A permutations approach was used to calculate Moran's $I$ [28] statistics using each of the row-standardized spatial weights matrices for each of the variables in the study. The results were presented in a correlogram.

### Building of the non-spatial models

Model building of the non-spatial ordinary least squares (OLS) regression model was done in SAS [26]. Backwards elimination was used to select the significant ($P < 0.05$) exposure variables (water source, water filtration and disinfection) and their interaction terms. Each of the potential confounders listed in Table 1 was individually entered into the model and

their effect on the parameter estimates of the exposure variables was evaluated. The change in parameter estimates of the key exposure variables was used to identify important confounding variables. If the change was at least 20 %, or if the parameter estimate for the factor was significantly different from zero, the covariate was retained in the model.

The final OLS model from SAS, henceforth referred to as Model A1, was replicated in SpaceStat [27] for further assessment of the OLS assumptions. Model specification diagnostics, involving testing for assumptions of multi-collinearity, normality, homoskedasticity and spatial independence of errors, were then performed. Multi-collinearity was assessed using the multi-collinearity condition number proposed by Belsley and co-workers [29]. Values of the condition number larger than 20 or 30 were considered suspect. The homoskedasticity assumption was tested using the Lagrange Multiplier (LM) test developed by Breusch and Pagan [30]. Tests for independence of errors were performed using Moran's $I$ [28], the LM test [31], and the robust LM test [32] at each of the inverse distance weights.

### Building of the spatial models

Different models were constructed in SpaceStat [27] to adjust for departures from the OLS regression assumptions (mainly spatial dependence of errors). These included robust variance estimation using the jackknife method, denoted as Model A2. A spatial maximum likelihood estimation model (Model B), and a first-order autoregressive moving average (MA) model (Model C) were used to adjust for spatial autocorrelation of the OLS errors. The goodness of fit of each model was assessed using $-2$ log likelihood (deviance), Akaike's Information Criterion (AIC) [33], and Schwarz's Criterion (SC) [34]. The model with the lowest value of these criteria was deemed to be the best-fitting model.

Model B was a correction to Model A1 for spatial effects using each of the row-standardized inverse distance spatial weights matrices specifying the spatial error as the 'Simultaneous autoregressive (SAR) covariance family maximum likelihood approach'. Since the spatial correlogram constructed using the five inverse distance spatial weights revealed evidence of significant spatial autocorrelation at all lag distances, it was necessary to use the full inverse distance spatial weights matrix to capture all the spatial dependence present in the data. This weights

Table 2. *Distribution of giardiasis rates by water source and treatment in southern Ontario* (1990–1998)

| Water type | Covariate pattern | Number of areas | Geometric mean annual rate* | Standard deviation | Median annual rate* |
|---|---|---|---|---|---|
| Surface | Unfiltered/Non-disinfected | 0 | — | — | — |
| | Unfiltered/Disinfected | 10 | 16·9 | 1·5 | 15·7 |
| | Filtered/Non-disinfected | 0 | — | — | — |
| | Filtered/Disinfected | 116 | 5·1 | 1·45 | 5·4 |
| Ground | Unfiltered/Non-disinfected | 13 | 5·6 | 1·42 | 5·2 |
| | Unfiltered/Disinfected | 111 | 4·6 | 1·49 | 5·8 |
| | Filtered/Non-disinfected | 0 | — | — | — |
| | Filtered/Disinfected | 7 | 3·4 | 1·23 | 2·8 |

\* Rate per 100 000 person-years.

matrix also fit the data best. Model B involved fitting the standard regression model with a spatial autoregressive error term:

$$Y = X\beta + \varepsilon,$$
$$\varepsilon = \lambda W\varepsilon + \xi,$$

where, $Y$ is an $N \times 1$ vector of observations on the dependent variable, $X$ is an $N \times K$ matrix of observations of the explanatory variables, $\beta$ is a $k \times 1$ vector of parameter estimates, $\lambda$ is the autoregressive coefficient of the spatial error model, $W\varepsilon$ represents the spatially lagged error term, with $W$ as a row-standardized (standardized to have row sums of unity) proximity matrix and $\varepsilon$ as an $N \times 1$ vector of error terms; $\xi$ is a vector of 'well-behaved' (homoskedastic and independent) error term with mean 0 and a constant variance, $\sigma^2 I$.

The first-order autoregressive MA model (Model C) estimates $\rho$ [the autoregressive coefficient of the spatial MA model], $\gamma$ (the MA component), and $\sigma^2$ the residual variance. Here the spatial moving process in the error term $\varepsilon$ takes the form:

$$\varepsilon = \rho W\xi + \xi,$$

in which the spatial lag pertains to the errors $\xi$ and not $\varepsilon$.

## RESULTS

### Distributions of giardiasis rates by water characteristics

None of the areas used unfiltered and non-disinfected SW (Table 2). Similarly, no areas used filtered and non-disinfected GW. Areas that used filtered and disinfected GW had the lowest giardiasis rates (3 per 100 000 person-years) while those that used unfiltered

and disinfected SW had the highest rates of giardiasis (16·9 per 100 000 person-years).

### Tests for spatial autocorrelation

The spatial correlogram showing results for all variables involved in the final models using different spatial weights is shown in Table 3. There was evidence of significant spatial dependence of all the variables at all the spatial weights, however, the degree of spatial dependence decreased with an increase of distance.

### Non-spatial models (models A1 and A2)

The results of OLS regression (Model A1), and the OLS with jackknife variance estimates (Model A2) are shown in Table 4. The residuals of models A1 and A2 were both normally distributed ($P = 0·1$) and homoskedastic (Breusch–Pagan statistic = 7·9, $P = 0·1$; White statistic = 13·7, $P = 0·2$). Since both OLS assumptions were met, the results of the tests for spatial autocorrelation have meaningful interpretations. Moran's $I$ ($I = 0·024$; $P = 0·004$), the LM (LM = 3·946; $P = 0·047$) and the robust LM (robust LM = 4·924; $P = 0·026$) tests for the spatial error and the MA (Model C) models were all statistically significant indicating the presence of significant spatial dependence in the errors. Thus models B and C were considered valid alternatives to models A1 and A2.

SW and rural areas had positive associations with the log of giardiasis rates in both models A1 and A2 whereas water filtration and median income had negative associations with the log of giardiasis rates (Table 4). In Model A1, SW had a significant positive association [rate ratio (RR) 2·23, 95% confidence

Table 3. *Spatial autocorrelation of the outcome and explanatory variables used in the regression analyses of reported cases of giardiasis in southern Ontario (1990–1998)*

| Variable | Moran's $I$ (s.e.*) $P$ value using IDW69b[a] | Moran's $I$ (s.e.) $P$ value using IDW126[b] | Moran's $I$ (s.e.) $P$ value using IDW217[c] | Moran's $I$ (s.e.) $P$ value using IDW355[d] | Moran's $I$ (s.e.) $P$ value using IDWFULL[e] |
|---|---|---|---|---|---|
| Surface water | 0·489 (0·028) 0·001 | 0·404 (0·020) 0·001 | 0·302 (0·015) 0·001 | 0·238 (0·012) 0·001 | 0·220 (0·011) 0·001 |
| Water filtration | 0·419 (0·028) 0·001 | 0·351 (0·019) 0·001 | 0·260 (0·014) 0·001 | 0·198 (0·012) 0·001 | 0·184 (0·011) 0·001 |
| Water disinfection | 0·176 (0·028) 0·001 | 0·112 (0·020) 0·001 | 0·062 (0·015) 0·008 | 0·055 (0·012) 0·007 | 0·048 (0·006) 0·006 |
| Median income | 0·454 (0·026) 0·001 | 0·301 (0·019) 0·001 | 0·201 (0·014) 0·001 | 0·175 (0·012) 0·001 | 0·163 (0·010) 0·001 |
| Rural areas | 0·167 (0·027) 0·001 | 0·104 (0·019) 0·001 | 0·086 (0·014) 0·001 | 0·076 (0·012) 0·001 | 0·062 (0·010) 0·001 |
| Natural log of giardiasis rates | 0·052 (0·027) 0·025 | 0·037 (0·019) 0·020 | 0·033 (0·014) 0·013 | 0·027 (0·012) 0·012 | 0·024 (0·011) 0·010 |

\* Standard error.
Inverse distance weight (IDW): [a] distance band 0–69; [b] distance band 0–126; [c] distance band 0–217; [d] distance band 0–355; [e] full matrix, i.e. distance band 0–833.

Table 4. *A summary of the competing models for giardiasis rates in southern Ontario (1990–1998)*

| Variable | Ordinary least squares (Model A1) | Jackknife (Model A2) | Spatial error (Model B) | Spatial moving average (Model C) |
|---|---|---|---|---|
| Lambda/rho | — | — | 0·576[a] (0·216) | 0·95[b] |
| Gamma | — | — | — | 0·67[c] (0·14) |
| −2 LL | 756·6 | 756·6 | 752·7 | 746 |
| SC | 784·4 | 784·4 | 780·5 | 757 |
| AIC | 766·6 | 766·6 | 762·7 | 750 |
| **Fixed effects** | **Coefficients (standard errors)** | | | |
| Intercept | −6·26 (0·41) | −6·26 (0·39) | −6·23 (0·45) | −6·23 (0·41) |
| Surface water | 0·80 (0·27) | 0·80 (0·22) | 0·86 (0·28) | 0·82 (0·26) |
| Water filtration | −0·68 (0·27) | −0·68 (0·22) | −0·60 (0·27) | −0·64 (0·27) |
| Median income | −0·42 (0·19) | −0·42 (0·18) | −0·47 (0·20) | −0·47 (0·19) |
| Rural areas | 0·55 (0·15) | 0·55 (0·15) | 0·57 (0·15) | 0·58 (0·15) |

[a] Lambda, autoregressive coefficient (and standard error) of the spatial error model.
[b] Rho, autoregressive coefficient of the moving average model.
[c] Gamma, autoregressive moving average parameter (and its standard error).
−2 LL, −2 *(log likelihood) or deviance; SC, Schwarz Criteria; AIC, Akaike Information Criteria.

interval (CI) 1·32–3·78] with the log of giardiasis rates. The standard error of this variable was adjusted downwards in the jackknife model making the variable more significant and narrowing the 95 % CI (1·45–3·45). The rate of giardiasis was significantly higher in rural areas (RR 1·73, 95 % CI 1·3–2·32) than in urban areas, but significantly lower in areas using filtered water (RR 0·51, 95 % CI 0·3–0·86) than in areas using unfiltered water. The jackknife standard error of water filtration was lower making it more significant and narrowing the 95 % CI around its RR (0·33–0·78). The rest of the variables in the models had similar standard errors in models A1 and A2.

## Spatial models (models B and C)

The spatial models (Table 4) fit the data better than the non-spatial models (A1 and A2) as shown by their −2 LL, SC and AIC; Model C had the best fit. The autoregressive coefficient ($\lambda = 0.58$, 95% CI 0.15–1.0) of Model B was significantly different from zero by both the Wald's and likelihood ratio tests indicating that the spatial error model (Model B) was more efficient than Model A1. Similarly, the spatial autoregressive MA parameter ($\gamma = 0.67$, 95% CI 0.39–0.93) of Model C was significantly different from 0. The parameter estimates in the spatial error models were quite similar to the OLS estimates. Although the standard errors of Model B were adjusted for the spatial dependence of the errors, the inferences on all the parameter estimates remained the same.

## DISCUSSION

In general, the presence of significant spatial dependence in the errors implies that the OLS regression model is not efficient due to incorrect standard error (S.E.) estimates. However, in this study although there were small changes in the S.E.s estimated in the jackknife and spatial models compared to those of the OLS model, these changes did not result in changes in the significance of the parameter estimates. Therefore, inferences resulting from the OLS model and those of the jackknife and spatial models were the same. This is probably due to the relatively weak spatial autocorrelation of the residuals as evidenced by the value of the Moran's *I*. However, since the spatial autocorrelation of the residuals was statistically significant, it was important to correct for it using spatial models. Furthermore, although the inferences did not change between the OLS and the spatial models, the latter fit the data better and are, therefore, the models of choice for this data.

Land-use factors, pre-chlorination and chlorination of water were not significantly associated with giardiasis rates. However, two drinking-water-related factors (water source and water filtration) were associated with giardiasis rates. Other important covariates were median income and place of residence. Income is a determinant of health in general while place of residence (rural or urban status) was a proxy for general quality of life since in general people living in urban areas tend to have a better quality of life

and easier access to health services than those living in rural areas.

The higher rates of giardiasis in areas using SW has been reported in other studies [5, 15, 35, 36]. In addition, some studies have reported increased giardiasis rates with increase in elevation mainly due to the fact that easily contaminated SWs predominate as drinking-water sources in these areas [17, 37]. In contrast to the results of a study in British Columbia that reported that cyst concentrations were lower in chlorinated water than in raw water [38], there was no significant association between chlorination and giardiasis rates in the current study. However, our descriptive statistics showed that there were generally lower giardiasis rates in areas using disinfected water. Although disinfection was not statistically significant in the multivariable model, it might be biologically important as it could imply that disinfection might reduce cyst viability although the other variables in the model were more important predictors of giardiasis rates. Other authors have reported that although chlorination is effective in removing giardia cysts [39, 40], the organism is quite resistant to chlorine [41, 42]. In fact, there have been reported outbreaks of giardiasis in communities using chlorinated drinking water [12, 43]. The results of the current study and others [14] suggest that communities should not rely upon chlorination alone to protect public water supplies since it may not be effective against giardia cysts.

In agreement with several other reports [4, 17, 44–48] areas where filtration was performed had lower giardiasis rates than those not performing filtration. However, Jephcott et al. [43] reported an outbreak of giardiasis in chlorinated and filtered water in United Kingdom. This was probably a result of the breakdown of the filtration process rather than ineffectiveness of filtration *per se* at removing giardia cysts. Nonetheless, not all filtration systems are equally effective, as higher rates of giardiasis in residents whose areas were supplied by water that underwent mechanical filtration and chlorination have been reported in other studies [6]. Due to unavailability of data, the current study did not assess possible associations with different filtration techniques.

As has been reported elsewhere [14], some communities lack filtered water supplies due to the cost of filtration equipment and the required technical expertise to maintain them. This may be true in rural areas of Ontario, as most of the water systems that

performed filtration were in urban areas. A study of giardiasis in the United States reported higher rates of disease in Vermont than in other states due to the rural nature of Vermont resulting in a higher frequency of contact with sources of infection in the environment, especially water [17]. Alterations to these water systems, such as use of UV treatment and/or micro- and ultra-filtration membranes offer affordable and effective possibilities. Suffice it to say that improved methods of water treatment will be imperative for the continued improvement of drinking-water safety and control of waterborne diseases like giardiasis. Our study showed that rural areas had higher rates even when filtration was in the model, suggesting that probably other lifestyle factors, such as personal hygiene might be important.

Livestock density did not have a significant association with the giardiasis rates in agreement with a report by Chute and co-workers [15]. However, other studies have shown that drinking-water supplies may be contaminated by a variety of domestic and wild animal species [49, 50]. Although manure application in agricultural land had a significant univariate association with giardiasis rates it was not significant in the final model. It seems likely that some communities, especially those in the more rural agricultural areas, will be more at risk of giardiasis in the future as more and more human activities, impacting on land-use practices, continue to increase pressure on the watersheds. It is certainly possible that high livestock densities and greater use of manure on agricultural land may be important determinants of the disease in a few local areas but this association with giardiasis rates is cancelled out at the global scale when all the areas were considered. It is also possible that septic or sewage treatment systems are probable environmental sources of giardia cyst contamination. However, due to time and financial constraints it was not possible to include sewage treatment and disposal data in this investigation.

In accordance with reports of the many studies performed in numerous countries over several years, low median income was a risk factor for giardiasis [51]. The differences in income and health may be associated with differences in human behaviour, living conditions and environment that increase the risk of infection [51].

This study is the first that has attempted to investigate determinants of the variations of the distribution of endemic cases of giardiasis in southern Ontario using surveillance data. The information obtained from this study will be useful in guiding decisions on future directions of research and in designing disease-control strategies. In addition, the more surveillance data are used, the more likely the quantity and quality of the data will improve. Increased awareness by public health professionals of the potential uses of the surveillance data will help improve the quality and quantity of these databases in the future.

Ongoing activities to reduce the risk of giardiasis through improvements in public water sources and treatment procedures should continue to be undertaken by health authorities.

## REFERENCES

1. Furness BW, Beach MJ, Roberts JM. Giardiasis surveillance – United States, 1992–1997. Mor Mortal Wkly Rep CDC Surveill Summ 2000; **49**: 1–13.
2. 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.
3. Craun GF. Waterborne outbreaks of giardiasis. Current status. Giardia and giardiasis – biology, pathogenesis, and epidemiology, 1984: 243–261.
4. Naumova EN, Chen JT, Griffiths JK, Matyas BT, Estes-Smargiassi SA, Morris RD. Use of passive surveillance data to study temporal and spatial variation in the incidence of giardiasis and cryptosporidiosis. Public Health Rep 2000; **115**: 436–447.
5. Isaac-Renton JL, Philion JJ. Factors associated with acquiring giardiasis in British Columbia residents. Can J Public Health 1992; **83**: 155–158.
6. Fraser GG, Cooke KR. Endemic giardiasis and municipal water supply. Am J Public Health 1991; **81**: 760–762.
7. 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.

8. Roach PD, Olson ME, Whitley G, Wallis PM. Waterborne giardia cysts and cryptosporidium oocysts in the Yukon, Canada. Appl Environ Microbiol 1993; **59**: 67–73.

9. Levesque B, Rochette L, Levallois P, Barthe C, Gauvin D, Chevalier P. Study of the incidence of giardiasis in Quebec (Canada) and association with drinking water source and quality. Rev Epidemiol Santé Publique 1999; **47**: 403–410.

10. Chauret C, Springthorpe VS, Sattar SA, et al. Detection and survival of protozoan parasites in water in the Ottawa (Canada) region. Protozoan Parasites and Water, 1995: 80–83.

11. Herwaldt BL, Craun GF, Stokes SL, Juranek DD. Waterborne-disease outbreaks, 1989–1990. Mor Mortal Wkly Rep CDC Surveill Summ 1991; **40**: 1–21.

12. Moorehead WP, Guasparini R, Donovan CA, Mathias RG, Cottle R, Baytalan G. Giardiasis outbreak from a chlorinated community water supply. Can J Public Health 1990; **81**: 358–362.

13. Levine WC, Stephenson WT, Craun GF. Waterborne disease outbreaks, 1986–1988. Mor Mortal Wkly Rep CDC Surveill Summ 1990; **39**: 1–13.

14. Kent GP, Greenspan JR, Herndon JL, et al. Epidemic giardiasis caused by a contaminated public water supply. Am J Public Health 1988; **78**: 139–143.

15. Chute CG, Smith RP, Baron JA. Risk factors for endemic giardiasis. Am J Public Health 1987; **77**: 585–587.

16. Greig JD, Michel P, Wilson JB, et al. A descriptive analysis of giardiasis cases reported in Ontario, 1990–1998. Can J Public Health 2001; **92**: 361–365.

17. Birkhead G, Vogt RL. Epidemiologic surveillance for endemic *Giardia lamblia* infection in Vermont. The roles of waterborne and person-to-person transmission. Am J Epidemiol 1989; **129**: 762–768.

18. Kuller LH. The use of existing databases in morbidity and mortality studies [Editorial]. Am J Public Health 1995; **85**: 1198–1199.

19. Anselin L. SpaceStat version 1.80 user's guide. University of Illinois, Urbana–Champaign, Urbana, 1995.

20. Ontario Ministry of the Environment. Ontario Drinking Water Objectives. Revised 1994. Queen's Printer for Ontario, 1999.

21. Odoi A, Aramini J, Majowicz S, et al. The status of drinking water in Ontario, Canada (1992–1999). Can J Public Health 2003; **94**: 417–421.

22. Statistics Canada. Postal code conversion file. May 1999 postal codes. Reference guide for catalogue 92F00227XDB. Ottawa, Ontario, Canada, 1999.

23. Statistics Canada. Profile of census divisions and census sub-divisions. Population census of Canada 1996.

24. ESRI. Arcview GIS version 3.2. Environmental Systems Research Institute, Inc. Redlands, California, 1999.

25. Statistics Canada. Profile of census divisions and census sub-divisions. Census of agriculture, Canada, 1996.

26. SAS Institute. Statistical analysis systems (SAS) version 8.00. SAS Institute Inc., Cary, NC, USA, 1999.

27. Anselin L. SpaceStat version 1.9. Regional Research Institute of West Virginia University, 1995.

28. Moran PAP. The interpretation of statistical maps. J R Statist Soc B 1948; **10**: 243–251.

29. Belsley D, Kuhr E, Welsh R. Regression diagnostics, identifying influential data and sources of collinearity. New York: Wiley, 1980.

30. Breusch T, Pagan A. A simple test for heteroskedasticity and random coefficient variation. Econometrica 1979; **47**: 1287–1294.

31. Burridge P. On the Cliff-Ord test for spatial autocorrelation. J R Statist Soc B 1980; **42**: 107–108.

32. Anselin L. Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. Geogr Anal 1988; **20**: 1–17.

33. Akaike H. Likelihood of a model and information criteria. J Econom 1981; **16**: 3–14.

34. Schwarz G. Estimating the dimension of a model. Ann Statist 1978; **6**: 461–464.

35. Grant DR, Woo PT. Comparative studies of Giardia spp. in small mammals in southern Ontario. II. Host specificity and infectivity of stored cysts. Can J Zool 1978; **56**: 1360–1366.

36. Dennis DT, Smith RP, Welch JJ, et al. Endemic giardiasis in New Hampshire: a case-control study of environmental risks. J Infect Dis 1993; **167**: 1391–1395.

37. Meyer EA, Jarroll EL. Giardiasis. Am J Epidemiol 1980; **111**: 1–12.

38. Isaac-Renton J, Moorehead W, 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.

39. Jarroll EL, Bingham AK, Meyer EA. Effect of chlorine on *Giardia lamblia* cyst viability. Appl Environ Microbiol 1981; **41**: 483–487.

40. Lippy EC. Chlorination to prevent and control waterborne disease. J Am Waterworks Assoc 1986; **78**: 49–52.

41. Gray SF, Gunnell DJ, Peters TJ. Risk factors for giardiasis: a case-control study in Avon and Somerset. Epidemiol Infect 1994; **113**: 95–102.

42. Payment P. Poor efficacy of residual chlorine disinfectant in drinking water to inactivate waterborne pathogens in distribution systems. Can J Microbiol 1999; **45**: 709–715.

43. Jephcott AE, Begg NT, Baker IA. Outbreak of giardiasis associated with mains water in the United Kingdom. Lancet 1986; **1**: 730–732.

44. CDC. Centers for Disease Control and Prevention. Water-related disease outbreaks: annual summary, 1983. Atlanta, GA: Centers for Disease Control, September 1984.

45. Logsdon GS. Comparison of some filtration processes appropriate for Giardia cyst removal. In: Wallis PM, Hammond BR, eds. Advances in Giardia research. Papers from the Calgary Giardia conference held 23–25 February, 1987, Calgary, Alberta, Canada, 1988: 95–102.

46. Logsdon GS, Dewalle FB, Hendricks DW. Filtration as a barrier to passage of cysts in drinking water. In: Erlandsen SL, Meyer EA, eds. Giardia and giardiasis – biology, pathogenesis and epidemiology, 1984: 287–309.

47. Logsdon GS, Symons JM, Hoye Jr. RL, Arozarena MM. Alternative filtration methods for removal of Giardia cysts and cyst models. J Am Waterworks Assoc. 1981; **73**: 111–118.

48. Logsdon GS, Symons JM, Hoye RL. Water filtration techniques for removal of cysts and cyst models. In: Jakubowski W, Hoff JC, eds. Waterborne transmission of giardiasis. Proceedings of Symposium, 18–20 September 1978, Cincinnati, USA. [EPA-600-9-79-001]. 1979: 240–256.

49. Jarroll EL, Hoff JC, Meyer EA. Resistance of cysts to disinfection agents. In: Erlandsen SL, Meyer EA, eds. Giardia and giardiasis – biology, pathogenesis and epidemiology, 1984: 311–328.

50. Davies RB, Hibler CP. Animal reservoirs and cross-species transmission of giardia. In: Jakubowski W, Hoff JC, eds. Waterborne transmission of giardiasis. Proceedings of Symposium, 18–20 September 1978, Cincinnati, USA, 1979: 104–125.

51. Wilkinson RG. Income distribution and life expectancy. Br Med J 1992; **304**: 165–168.