Applied GIS

an international, refereed, open source journal

ISSN: 1832-5505

URL: http://www.arts.monash.edu.au/ges/research/Gis/public/epress.html

Volume 3, Number 4

April, 2007

CONTENTS:

All papers published during 2007 are part of *Volume* 3.
Each paper constitutes one *Number.*

Hence this paper should be cited as:

Cloutier, M., Apparicio, P. & Thouez, J. (2007) - GIS-based spatial analysis of child pedestrian accidents near primary schools in Montréal, Canada, *Applied GIS*, 3(4): 1-18.

GIS-based spatial analysis of child pedestrian accidents near primary schools in Montréal, Canada

Marie-Soleil Cloutier
Geography Department,
Université de Montréal,
Montreal, Canada
ms.cloutier@umontreal.ca

Philippe Apparicio
INRS UCS,
Montreal, Canada

Jean-Pierre Thouez
Geography Department,
Université de Montréal,
Montreal, Canada

Abstract: In Montréal, Canada, accidents affecting child pedestrians (5 to 14 years old) remained almost constant from 1994 to 1999 despite the great amount of prevention measures. Moreover, the elementary public school environment has been barely taken into account by past and present research on factors affecting the risk of accident even though children attend school most weekdays. We argue here, therefore, that the integration of the local environment into the spatial analysis of child pedestrian accidents could help to reduce them. Accordingly, we have integrated socio-economic and environmental data into a geographic information system in order to perform a geographically weighted regression and results demonstrate that the average network distance separating accident and closest school is less than 500 meters, thereby confirming a relationship of proximity between these two locations. Results also demonstrate the relevance of adding a spatial dimension to the regression model by suggesting that prevention initiatives should take into account the particular nature of each neighbourhood so that more relevant risk factors can be targeted.

Key words: Child pedestrian accidents, spatial analysis, geographically weighted regression, risk factors, geographic information system

1. Introduction

According to the latest *World Health Organization* injury report, road traffic accidents are the primary cause of mortality and morbidity in the 5 to 14 year-old age group within North America, and the vulnerability of these children increases the severity and long-term consequences of their injuries (Krug, 1999; Peden et al., 2004). Since modern societies are concerned about their children, there is a need to explore further this public health preoccupation with local examples. In Montréal, Canada, the number of accidents affecting pedestrians in this age group remained almost constant for the six-year period 1994-1999. Figures available raise concerns about child safety on streets, and according to police reports, between 193 and 228 accidents occurred each year, including thirteen deaths and 152 seriously injured victims (Société de l'assurance automobile du Québec, 2004).

Historically, studies trying to link risk factors, exposure to these various factors and frequency of accidents relied on individual data, but more analyses have been conducted on a territorial basis in the past years. This work arose from the awareness of the spatial interaction existing among accident locations and risk factors. Spatial analysis is helpful in this kind of work because it allows a broader scope of study.

Accidents and their surrounding environment can be visualised on maps and analysed spatially in order to detect locations of high risk (Thil, 2000; Flahaut et al., 2003). This means that additional information can be drawn from existing datasets and outcomes can be focused on preventive environmental approaches - because such initiatives, according to many authors, have been shown to be effective in reducing the risk of accidents more often than pedestrian education programs (Sibert, 1991; Stevenson et al., 1997; Reading et al., 1999; Duperrex et al., 2002). As Roberts et al. (1995) mentioned:

> *While the validity and applicability of any single efficacy estimate is open to question, arguably the most important consideration is that there are studies that have consistently shown a beneficial effect of traffic calming. This cannot be claimed for child pedestrian education programs.*

It is already known that a majority of child-related collisions happens close to home (Kendrick, 1993; Wazana et al., 2000). Other epidemiological studies have shown that accidents involving children occur on the way to school and/or during time periods close to school hours (Joly et al., 1991; Carlin et al., 1997; LaScala et al, 2004). However, because most studies use traditional administrative areas (enumeration areas or census tracks) as the basis of their analysis, the school environment and the fact that children attend school on most weekdays are not sufficiently taken into account.

The objective of this paper is to integrate the elementary public schools environment into the evaluation of child-pedestrian accidents. This means that the study period is limited to the different school calendars (weekdays of the school years - September to June) and that the proximity between accidents and schools is considered through the construction of 'catchment areas'.

The latter was done because by putting the school at the center of a space of proximity, we do not study the problem starting from the usual administrative sectors, but starting from the immediate school surroundings, the real center of attraction of children during weekdays. Using such 'catchment areas' instead of conventional statistical areas has three advantages. Firstly, the areas better reflect the daily reality of pedestrians in terms of proximity to school. Secondly, the results provide local information for each school separately. Thirdly, the 'catchment areas' are used as a common space for the integration of variables that are available at different scales. Note that the variables chosen represent factors that are known from past research to affect the risk of child pedestrian accidents.

2. Data collection

The study area is the island of Montréal, Canada's second largest city with 1.8 million people on the island, expanding to almost 3.4 million within the census metropolitan agglomeration (Figure 1) (*Statistics Canada*, 2002).

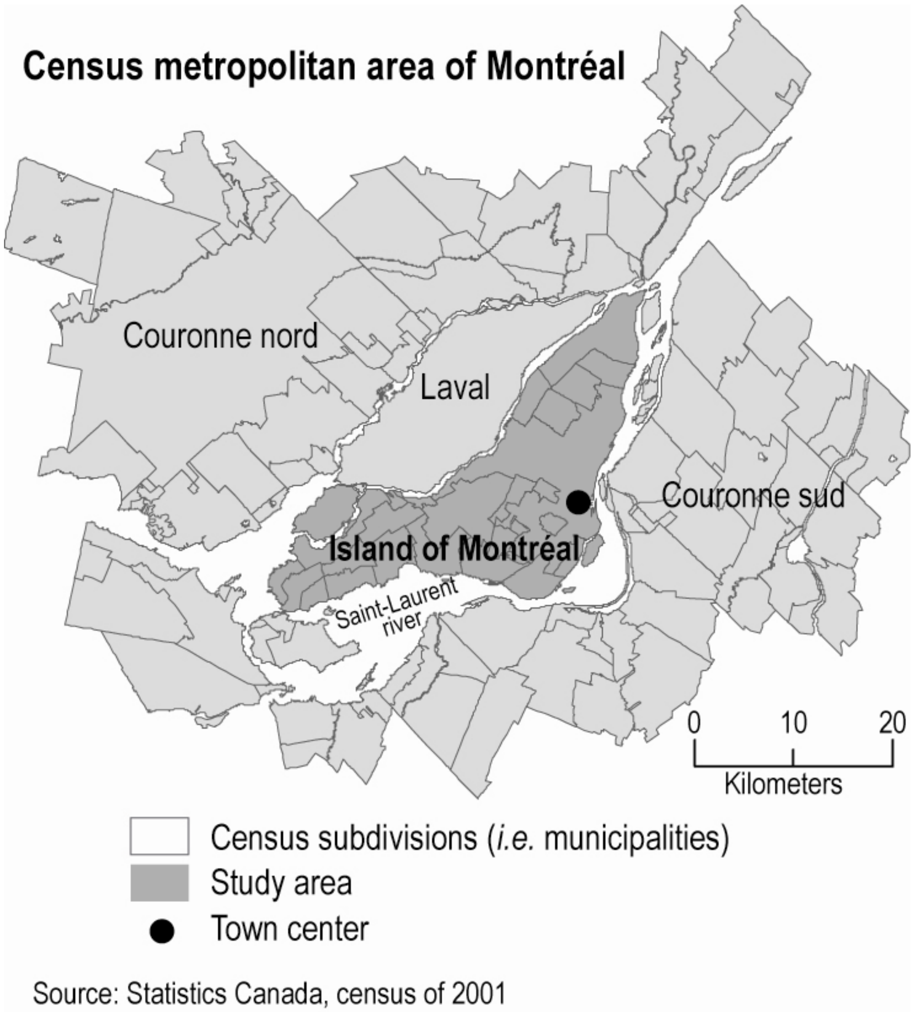


Figure 1 - The study area in the context of Montréal census metropolitan agglomeration

In order to model child pedestrian accident risk around primary schools, three sources of data were used: 1) information on schools; 2) police reports on child pedestrian accidents; 3) characteristics of the built environment around schools.

2.1 School dataset

Data on primary schools were provided by the School Taxation Management Council (*Conseil de gestion de la taxe scolaire de l'Île de Montréal* – CGTSIM). This dataset includes the current address, the enrolment for the year 1999-2000 (September to June) and the linguistic affiliation of each school. Schools that were not open throughout the study period (1994-1999) are excluded, together with specialised ones, for a total of 331 schools included in the study (93.8 per cent of the original dataset).

2.2 Child pedestrian accidents dataset

Data on child-related pedestrian accidents (0-14 years old) for the period 1995-1999 were provided by the Quebec Public Automobile Insurance Society (*Société de l'assurance automobile du Québec – SAAQ*). SAAQ data sets are the only ones that bring together all police accident reports filled on site by police officers. These reports include information on the victims (age, sex, severity of injuries) and on the event itself (date, time, location). This dataset includes events involving any contact on the street between a pedestrian and a car that leads to any type of injury for the pedestrian. Preliminary work was done to allow for a focus on schools: children below the age of five are excluded (318 accidents) as well as all summertime accidents (223), all undated accidents (36), and all those occurring during weekends or national holidays (311 plus 74). This led to the selection of 1335 accidents (58 per cent of the original dataset).

2.3 Built environment dataset

Past studies have identified socio-economic and environmental contexts as major risk factors for child pedestrian accidents. Certain results point out socio-economically disadvantaged areas as 'spots' of higher accident risk (Pless et al., 1987; Dougherty et al., 1990; Hasselberg et al., 2001; Graham et al., 2005) whereas others link higher levels of traffic and higher residential density to greater numbers of victims (Braddock et al., 1994; Roberts et al., 1995; LaScala et al., 2004).

In this study, four different elements were computed for each school in order to characterise the built environment: social deprivation, street network density, major road density (as a proxy for higher traffic) and land-use diversity (entropy index). Data on the built environment came from two main sources. Socio-economic data were extracted from the 2001 population census at the dissemination area level, while land-use and street network data were provided by the City of Montréal Geomatic Division (Statistics Canada, 2002; City of Montréal, 2004).

With regard to the reliability of the data, we should mention that all data included came from primary government sources that create, manage and update them regularly. The CGTSIM updates the school list every year and the five different lists (for each year of the study period) were used to find which school should be kept in our database. The SAAQ, in turn, fill in the accident database every year according to police and coroner reports. Furthermore, the staff agreed to our request to extract child pedestrian accident directly from the main database for a specific period.

The general underestimation of the number of victims in police databases is well documented in the literature. Nevertheless, the monopoly that the SAAQ has in this field led us to affirm that their data are the most reliable and up-to-date source of information regarding accident site location in Montreal (Dhillon et al. 2001). Finally, the road network and the land-use classification were updated by the city's employees regularly and the latest version was given to us at the time we placed a request.

3. Data preparation

3.1 School and accident mapping

The street network database mentioned above was used to map selected schools and accident records according to their location attributes. Accident records where the location information was unreadable because of transcription errors are excluded from the study (104 cases). All the spelling problems were reviewed one by one in order to reduce the number of accidents excluded in the geocoding process. In the end, 331 schools and 1231 accidents have been mapped (92.2 per cent of the selected accidents).

3.2 Defining schools' 'catchment areas'

Catchment areas are created around schools based on network distance. Since accidents usually happen on roads, it is even more important to use the improved accuracy of a road network to create these areas. This approach, which is more precise than Thiessen polygons, uses global image analysis functions such as *Costallocation* and *EucAllocation* to determine catchment areas (Figure 2).

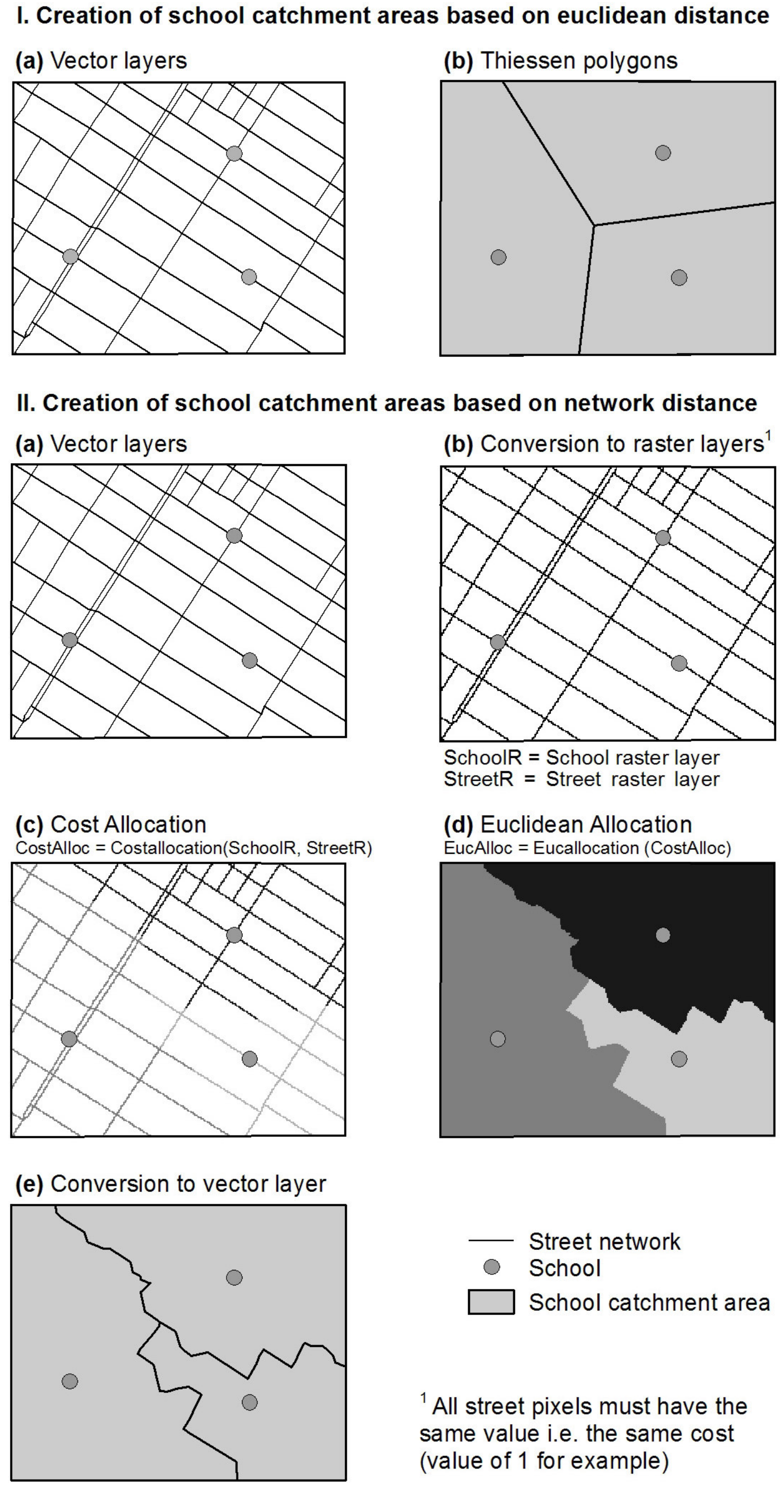


Figure 2 - Defining catchment areas: two approaches

These areas are shaped in *ArcGIS 9.1* using the spatial analyst extension (Environmental Systems Research Institute (ESRI) Inc., 2005). In the end, 327 catchment areas were created because four contain two schools with an identical location (in the same building). By creating these school catchment zones, it becomes possible to allocate each accident location to only one school - the closest - even though there was no way in the data provided to verify whether the child was really attending this specific school.

3.3 Defining the school environment

Once the catchment areas were delimited, four different types of variables were computed to qualify social and built environments surrounding schools. These include population characteristics, social and demographic characteristics, land use diversity and street network characteristics. The process of integrating all sources of data was always the same and was divided in three steps.

First, each geographical unit (census area, road segment, point data, etc.) was clipped (cut) and intersected with the catchment areas to split and extract entities. The resulting entities kept their attributes and were given the unique identifier of the catchment area into which they fell. Second, in the case of polygons with population attributes, the population was fragmented proportionally to the area of each parts of the original polygon. For example, if 50% of the area of polygon *p* falls into the catchment area *c*, then only 50% of the population of polygon *p* was related to catchment area *c*. Third, all the entities were merged according to the catchment area into which they fell and relevant attributes were calculated using different operations. For example, populations were summed and density measures were computed to have one measurement for each catchment area.

3.3.1 Child population

As in many public health researches, a population at risk should be integrated in the model under an assumed positive relationship: the more there are children in a neighbourhood, the higher are the possibilities for child pedestrian accidents. Accordingly, the proportion of children (5 to 14 years) is calculated for each catchment area. This indicator was chosen because we are more interested in the effect of the presence of children in the population as a whole than at the total number of children per catchment area.

3.3.2 Social deprivation index

A social deprivation index was computed to verify the persistence of the relationship illustrated in previous work with the actual accident database. This index includes five variables: the percentage of single-parent families, the unemployment rate, the percentage adults (20 year-old and over) with less than a ninth grade education, the percentage of low income population and the percentage of immigrants having arrived during the period 1996-2001. This information was extracted at the dissemination area level from the most recent Population Census (2001).

The proposed deprivation index represents the sum of these five standardised variables (0 to 1 scale): it varies from minimal deprivation (0) to maximal deprivation (5). This index has been described and analysed in Apparicio et al., 2004). It is important to note here that this variable has an indirect effect on the exposure to risk since socio-economically disadvantaged areas might have more children walking to school, as mentioned in other studies (Roberts et al., 1996; Macpherson et al., 1998). It is not the purpose of the present project to verify this hypothesis again, but the results should be analysed with this relationship in mind.

3.3.3 Land use diversity

The land-use classification has been used here to create an entropy index measuring the land occupational diversity of the school zones. The sixteen categories are: high, medium and low residential density; small retail; shopping center; office space; community service and equipment; public utility; industrial; quarry; landfill site; green space; golf course; cemetery; and rural and vacant space.

Diversity in the present sense goes beyond the use of simple residential density: it integrates all other land uses in an informative index representing a “risk factor”. A highly diverse neighbourhood would attract more people from outside the area, creating flows of traffic different than other neighbourhoods with low diversity. This index is calculated following equation (1) and varies from 0 (homogeneity) to 1 (diversity).

(1) $E_i = -\sum_{i=1}^{n}\left[\left(A_{ij} / A_i\right)\ln\left(A_{ij} / A_i\right)\right] / \ln n$

where:

Ei = Land use diversity of zone *i.*
n = number of land-use;
A_i = area of proximity zone *i;*
A_{ij} = area of land-use *j* in proximity zone *i.*

3.3.4 Traffic

The street network database includes a hierarchical classification of streets (City of Montréal, 2004). This classification allocates street segments to four categories according to the number of lanes and daily traffic: local streets, collector roads, arterial roads and highways. Two variables are constructed from this network (in meters per square kilometre): the road density (all categories), and the main road density (collector and arterial). Note here that highways are excluded because usually pedestrians do not travel along them.

3.4 Defining other school variables

Two variables added to the model are specific to schools: enrolment for a chosen year and linguistic affiliation (English or French). The first one is simply the number of students registered at each school from September 1999 to June 2000 in order to verify the relationship between the size of a school and the number of accident close by.

The second variable is the school language. Because of the English and French heritage of Montréal, 85 (26 per cent) of primary schools are English speaking while the rest are French speaking. Since the English ones are less numerous, their “registration” territory is larger. Consequently, we can hypothesize that the school language is significant because it has an influence on the mode of transportation of children: English-speaking schools have students coming from farther away, and thus fewer walking to school. Integrating school language as a variable in the model permits to taken into account this exposure to risk (in terms of walking or not).

In the end, eight variables are included in our analysis and some of them have been transformed in order to normalise their distribution (the operation is given in brackets):

1. number of accidents per school zone (square root) as the dependent variable;
2. school language — 1: English-speaking school, 0: French-speaking school;
3. school enrolment (log);
4. proportion of children — five to 14 years-old;
5. social deprivation index;
6. road network density;

7. main road density (log); and
8. entropy index.

All variables are assumed to have a positive relationship with the number of accident except the school language. English-speaking schools should have fewer accidents.

4. Methodology

The methodology is based on three types of treatment: 1) point pattern analysis to describe the spatial distribution of child pedestrian accidents in Montréal; 2) a multiple regression model to explain globally the number of accidents; and 3) a geographically weighted regression model to show spatial variations in the relation between the number of accidents and selected explanatory variables.

4.1 Point pattern analysis

In order to detect departures from spatial randomness within the distribution of accidents, several classical measures are used to describe the accident sites distribution:

- mean center (a central tendency measure) (eq. 2);
- standard distance and standard deviational ellipse, which is a useful way to graphically represent the dispersion of points on the map (eq. 3);
- nearest neighbour index (a point pattern detector measure) (eq. 4); and
- density mapping (Lee et al. 2001).

Density mapping is a good way to illustrate where points are concentrated. In a kernel density calculation as the one performed here, points that fall within the search radius are added together and then divided by the search area size to get a density value for each cell of an output matrix. The kernel feature means that points lying near the center of a raster cell's search area are weighted more heavily than those lying near the edge, giving a smoother result (McCoy et al., 2002).

(eq. 2) $$\left(\bar{x}_{mc}, \bar{y}_{mc}\right) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n} \right)$$

(eq. 3) $$SD = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}_{mc}\right)^2 + \sum_{i=1}^{n}\left(y_i - \bar{y}_{mc}\right)^2}{n}}$$

where: $\bar{x}_{mc}$ and $\bar{y}_{mc}$ = geographic coordinates of the mean center.
x_i and y_i = geographic coordinates of point *i*;
n = number of points;
SD = standard distance

(eq. 4) $$R = \frac{r_{obs}}{r_{\exp}} = \frac{\sum_{i=1}^{n} d_i / n}{\frac{1}{2\sqrt{n / A}}}$$

where: d_i = distance between point *i* and its nearest neighbours;
n = number of points;
A = area of the referenced zone.

4.2 Regression model and geographically weighted regression (GWR)

A multivariate regression model was computed to evaluate the link between selected variables and frequency of accidents. Such a global model tries to fit a regression equation in order to predict a dependent variable according to the values of *p* independent variables (eq. 5).

(eq. 5) $$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i$$

where: y_i = dependent variable

β_0 = intercept

β_j = regression coefficient for the independent variable *j*

ε_i = error

However, this kind of model, largely used social sciences is not always appropriate with spatial data for two reasons. Firstly, it does not take into account spatial autocorrelation among the dependent and independent variables. Secondly, it cannot capture the non-stationary component of the relationship. This latter term refers to the fact that the process we are trying to investigate might not be constant over space: the measurement of the relationship depends in part on where the measurement is taken (Bailey et al., 1995; Anselin et al., 1998; Fotheringham et al., 2002).

It is now possible to overcome these two drawbacks with the help of a geographically weighted regression (GWR) proposed by Fotheringham et al. (2002). GWR is an extension of the global model where parameters vary in space according to the geographic coordinates (u_i, v_i) of point *i* (eq. 6). In fact, GWR calculates a regression equation for each geographic observation and provides local values for R^2, β_0, β_k, Student *T*. GWR local equations are based on a least squares method combined with a weight matrix *W(i)* with values decreasing according to a distance *d* between *i* and *j* (eq. 7): the closer *i* and *j* are the higher will be the weighting factor (Fotheringham et al., 2002).

(eq. 6) $$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{p} \beta_j(u_i, v_i) x_{ij} + \varepsilon_i$$

(eq. 7) $$\hat{\beta}_i = (X^T W(i) X)^{-1} X^T W(i) y \text{ with } W(i) = \begin{pmatrix} w_{i1} & 0 & \ldots\ldots & 0 \\ 0 & w_{i2} & \ldots\ldots & 0 \\ . & . & \ldots\ldots & . \\ . & . & \ldots\ldots & . \\ 0 & 0 & 0 & w_{in} \end{pmatrix}$$

Two kernel functions are proposed in the model to calculate the weight matrix - a Gaussian function (eq. 8) and a Bi-square function (eq. 9). In the first case, the weighting is 1 when *i* and *j* coincide in space and the weighting of the rest of the data decreases according to a Gaussian curve. In the second function, the weighting of the data is continuous (near-Gaussian) up to a distance *b* (bandwidth) from the regression point while the rest of the data (beyond *b*) are given a weight of 0. This bandwidth *b* can be user-defined (in terms of a distance or a number of neighbours) or it can be estimated through a cross-validation (CV) method (eq. 10) that minimises the CV value through the *golden section search technique* developed by Greig (1980). Fotheringham et al. (2002) propose other methods such as the minimisation of the AIC (*Akaike Information Criterion*), the BIC (*Bayesian Information Criterion*) or the SIC (*Schwartz Information Criterion*) to optimise the bandwidth.

In this study, we decided to automate the bandwidth according to the Gaussian Kernel weight function using the cross-validation (CV) method. The *b* value is optimised at 4.7 kilometres using this method. It is worth noting that we obtained similar results with the AIC optimisation but not with a user-defined bandwidth (for example with bandwidth of 5, 7 or 10 kilometres).

(eq. 8) $w_{ij} = \exp\left[-\frac{1}{2}(-d_{ij}/b)^2\right]$

(eq. 9) $w_{ij} = \left[1-(d_{ij}/b)^2\right]^2$ if $d_{ij} < b$

= 0 otherwise

(eq. 10) $CV = \sum_{i=1}^{n}(y_i - \hat{y}_{i \neq i}(b))^2$

The GWR model used for this paper assumed normality, a distribution not usually found in researches on road traffic safety - it is rather the Poisson distribution that characterised accidents, considered as “rare” events. However, since the accidents are aggregated in school zones, the dataset follows a distribution between the normal and the binomial model. Accordingly, all the variables have been tested and transformed if needed to follow normality before the model was calculated. It is worth noting that the newly released version of the GWR software allows the choice between three different models: normal, logistic and Poisson. Indeed, GWR models can be used to study other road safety data with the Poisson distribution.

5. Results

5.1 Descriptive statistics

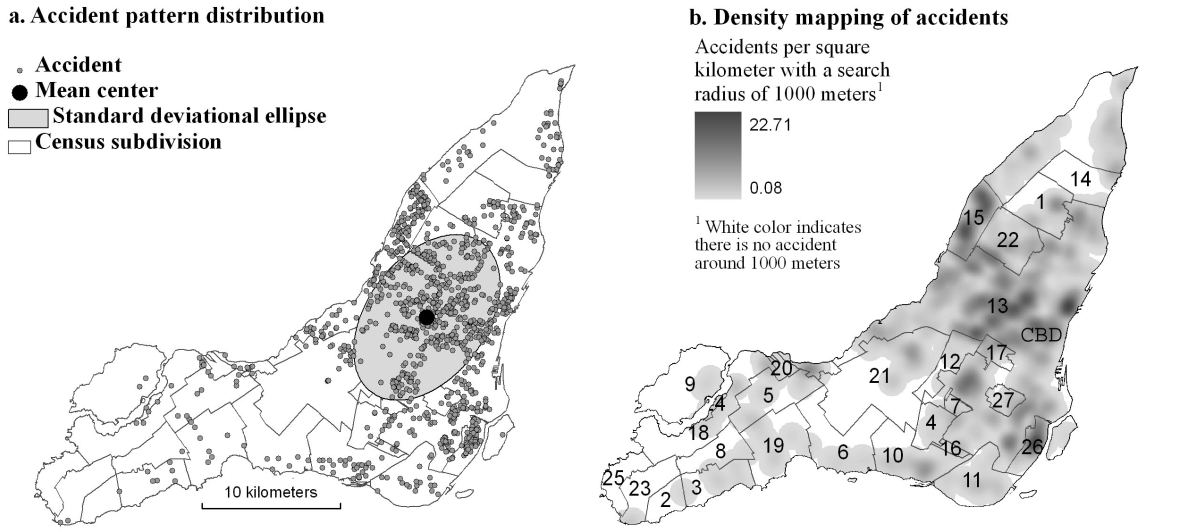
Before analysing the data in depth, we should look at their general portrait. Proportions of the total number of accidents remain almost constant: it fluctuates from 15.5 to 18.5 along the period studied (Table 1). Boys (59.4 per cent) are more involved than girls, as well as children of the 10 to 14 year-old age group (53.4 per cent). Periods of the day with the most accidents correspond to hours when children are not in class but are moving from or to school - in the morning (15.5%); at lunchtime (14.5%); and after class (30.5%). These numbers confirm the relevance of studying schools as a point of concentration for child pedestrian accidents.

Years	***N***	**%**	**Time**	***N***	**%**
1994	228	18.5	**7 to 9 am**	191	15.5
1995	204	16.6	**11 am to 1 pm**	179	14.5
1996	191	15.5	**3 to 5 pm**	376	30.5
1997	211	17.1	**Other**	485	39.5
1998	204	16.6			
1999	193	15.7			
Total	**1231**	**100.0**	**Total**	**1231**	**100.0**

Table 1 - Descriptive statistics

5.2 Spatial distribution of accidents

Figure 3a shows the spatial distribution of accidents involving children on the Island of Montréal between 1994 and 1999. The nearest neighbour index value (0.56) indicates that the accident distribution tends to be clustered. Moreover, the standard deviational ellipse around the mean central point demonstrates that even though accidents occur all over the island, more tend to occur within and close to the ellipse, in addition to two other concentrations of accidents in socio-economically disadvantaged areas -: in Montréal-Nord (north of the ellipse: 16) and in Verdun (south of the ellipse: 27). Finally, the kernel density of accident locations shows a high concentration of accidents in the same areas, areas that correspond to the more densely populated boroughs of the island (Figure 3b).



Point pattern statistics

Distance standard	8008.90 m
Observed average nearest neighbor distance	176.78 m
Excepted average nearest neighbor distance	317.77 m
Nearest neighbor index	0.56
Z score value	29.77

Name of boroughs

1	Anjou	8	Kirkland	15	Montréal-Nord	22	Saint-Léonard
2	Baie-d'Urfé	9	L'Île-Bizard	16	Montréal-Ouest	23	Sainte-Anne-de-Bellevue
3	Beaconsfield	10	Lachine	17	Outremont		
4	Côte-Saint-Luc	11	LaSalle	18	Pierrefonds	24	Sainte-Geneviève
5	Dollard-des-Ormeaux	12	Mont-Royal	19	Pointe-Claire	25	Senneville
6	Dorval	13	Montréal	20	Roxboro	26	Verdun
7	Hampstead	14	Montréal-Est	21	Saint-Laurent	27	Westmount

Figure 3 - Spatial distribution of child pedestrian accidents in Montréal, 1994-1999

5.3 Relationship between accident locations and schools

The mean network distance between each accident and the closest school is 495 meters; and 25% of the accidents occur less than 263 meters from the closest school (Table 2). Comparatively, the mean network distance between each school is 694 meters (minimum: 83 meters; maximum: 4.6 kilometres). These distances demonstrate that accidents happen closer to school and confirm the relevance of the present study. We can even hypothesize that these accidents took place while children were walking to or from school.

Minimum	0.00 m	**Percentiles** (% of accidents at *x* distance of the closest school)	**Distance** (m)
Mean	495.78 m	**5**	46.512
Maximum	3113.45 m	**25**	262.095
Standard Deviation	346.46	**50**	449.513
Skewness	1.47	**75**	656.785
Kurtosis	4.33	**95**	1165.094

Table 2 - Network distance between child pedestrian accidents and the closest school: some descriptive statistics

5.4 Global multivariate regression

A global regression model was computed to verify the statistical relationship between the number of accidents and the selected variables (Table 3). The model explains 36.5 percent of the variance in the number of accidents per catchment area, the deprivation index being the strongest variable. Low variance inflation factor (VIF) values indicate that there is no evidence of high colinearity among variables (Miller, 1990). Four out of seven parameters are statistically significant at the five percent level: the number of accidents near a school is positively related to social deprivation, to main road density and to entropy while it is negatively correlated to school language (English).

Independent variable	Coefficient value	Standardised coefficient	*t* value	Significance (p)	Variance inflation factor(VIF)
(Intercept)	-0.171		-0.301	0.764	-
Deprivation index	0.563	0.409	8.080	0.000	1.285
Children (% 5-14 years old)*	-0.029	-0.081	-1.472	0.142	1.537
Road network density	0.006	0.023	0.390	0.696	1.791
Main road density (log)	0.223	0.130	2.086	0.038	1.964
Entropy index	0.878	0.110	1.980	0.049	1.547
School enrolment (log)	0.065	0.043	0.920	0.358	1.096
School language **	-0.260	-0.117	-2.450	0.015	1.138

* Proportion per catchment area; ** English speaking: 1, French speaking: 0; Note: R^2 = 0.365; F= 26.15 (p=0.000)

Table 3 - Global multiple regression model for accident frequency in Montréal

5.5 Geographically weighted regression

A geographically weighted regression model was performed with the same variables in order to analyse spatial variability in the model. The *GWR 2.0.5* software is used to compute it (Fotheringham et al, 2002). The Moran's I of residuals from the global regression model was 0.12, revealing low but significant spatial autocorrelation. In the GWR model, Moran's I decreased to 0.022 while the regression coefficient (R^2) increased to 0.561, a great improvement in the model. Accordingly, the results of the ANOVA test performed between the residual of the global multiple regression model (OLS: ordinary least square) and of the GWR model is shown in Table 4 with the conclusion that the improvement between the two models is statistically significant (p=0.001, Fisher value = 2.161).

Source	Sum of squares	Degrees of freedom	Mean sum of squares	test of Fisher (F)
OLS Residuals	196.1	8.00		
GWR improvement	60.4	54.52	1.11	
GWR Residuals	135.7	264.48	0.51	2.16

Table 4 - Analysis of variance (ANOVA) for the GWR model versus the OLS (global) model

Figure 4 shows t-values for each independent variable (see Figure 3 for the name of the municipalities), remembering that values above or below ± 2.56 are significant at the one per cent level. The child population (T2), the road network density (T3) and school enrolment (T6) are not significant in the global model, but the GWR model shows local significance - T2 is significant in the central part of the island; T3 is significant downtown (CBD), and T6 is positively significant in Anjou and Saint-Léonard and negatively significant in Dollard-des-Ormeaux and Pierrefonds. These two areas are characterised differently in terms of school enrolment because the eastern part has a higher proportion of its schools with more than 400 students.

The other variables are also interesting to consider since their significance varies locally. The influence of deprivation (T1), which has the highest coefficient value in the global model, is quite widespread since this coefficient is significant at p=0.01 in 79% of the catchment areas. The t-value is stronger in the core of the city and tends to decline with distance toward the west but stays high along the north-east axis, known to include the most deprived areas of the city.

English-speaking schools (T7) influence the model only in the north-eastern part of the island (Anjou and Montréal-Est, p=0.01; Saint-Léonard, p=0.05). This is probably due to the fact that English-speaking schools are more numerous and regularly distributed across this part of the territory compared to the core of the city, where there are only a few and to the western part, where English is the language of most schools.

The land-use diversity measure (T5) has a high significance in the central south part of the island where entropy values are less than 0.5. These areas are principally occupied by housing but also by major public utilities and services such as cemeteries, parks and railroad tracks.

Finally, the traffic density predictor (T4) is significant in Anjou, Montréal-Nord and Saint-Léonard where there is a lot of arterial traffic for two reasons: incoming cars from off-island suburbs and traffic leaving highways to get in those neighbourhoods.

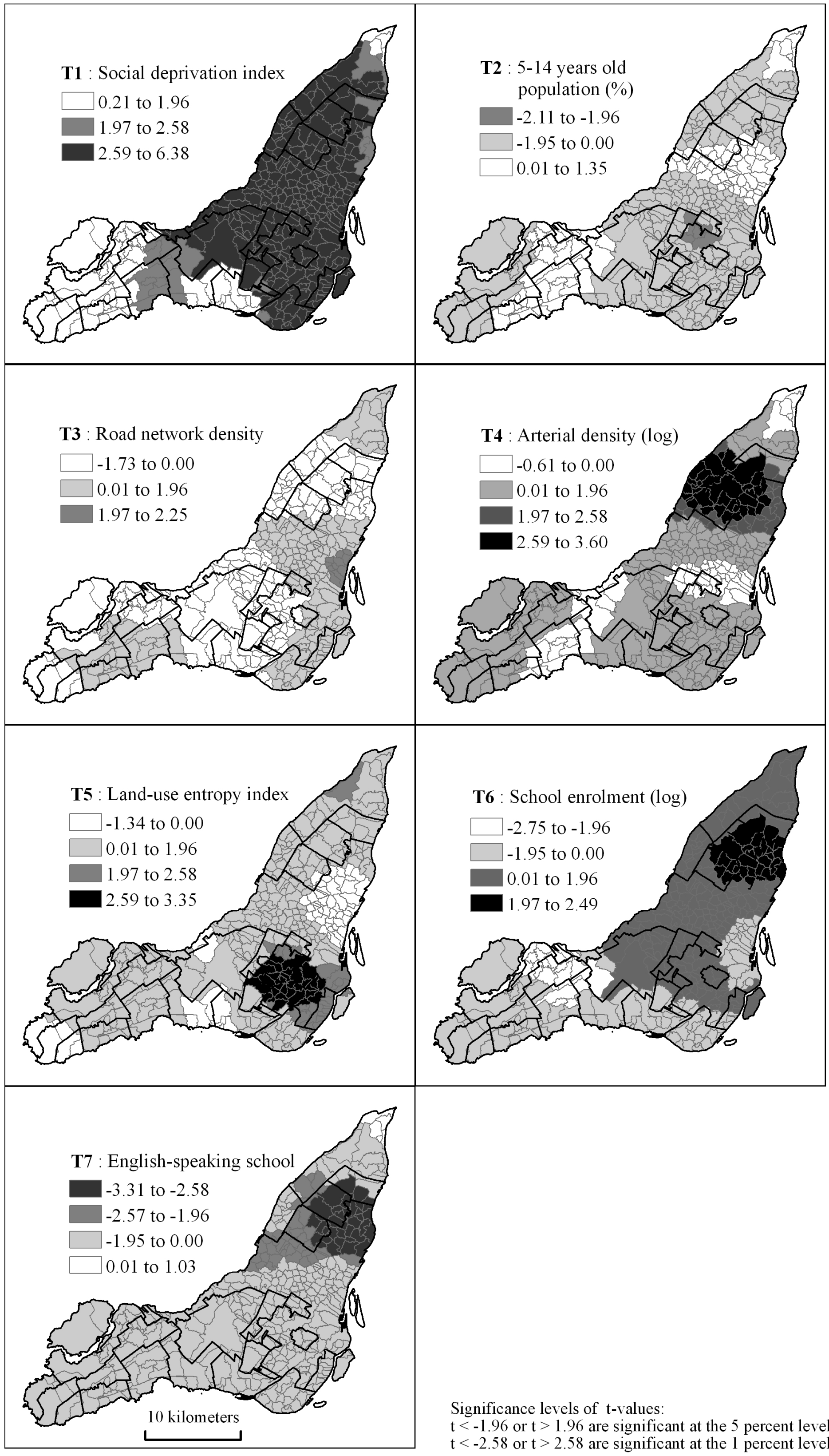


Figure 4 - Local T-values generated by the GWR model

Finally, Figure 5a shows the spatial distribution of the most significant variables - social deprivation (257 areas, 79 per cent of the total number of areas), school enrolment or school language (16 catchment areas each, a total of ten per cent). Figure 5b illustrates the total number of significant variables for each proximity area at the five percent level. Most of the catchment zones have only one or two significant variables (237 areas, 72 per cent), while a smaller proportion is found for three and four variables (23 and 29 areas, 16 per cent). It should also be considered that 38 areas (11 per cent) do not have any significant variables at the five per cent level.

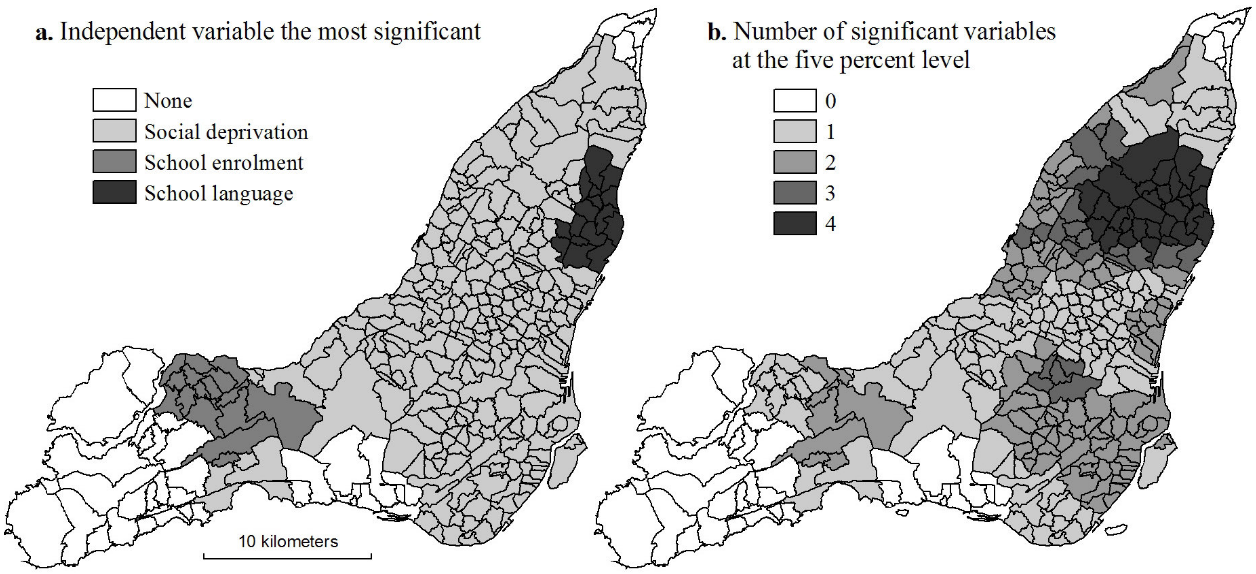


Figure 5 - Significant variables found by the GWR model

6. Discussion

Our descriptive results confirm those of many studies previously done in Montréal and elsewhere. Firstly, peaks of occurrence (hours and months) resemble those noted by Pless et al. (1987), Joly et al. (1991) and DiMaggio and Durkin (2002). Secondly, the spatial distribution of accidents is also similar to that Joly's findings (1991), which pointed to concentrations of accident in the central and south-west neighbourhoods. However, our study, with its analysesof the topic with point pattern methods, offers greater depth. Our conclusions concerning socio-economic status and traffic as factors in the increase of accidents are, furthermore, comparable to earlier research (Dougherty et al., 1990; Roberts et al., 1995; Rao et al., 1997; Posner et al., 2002).

More importantly, the contribution of this paper is to analyse the data in an explicitly spatialised (GIS) framework. This research expands our knowledge of the subject by revealing an association relating the location of child pedestrian accidents and the location of primary schools. GIS techniques allow for the integration of many sources of data in order to model in a more complex way the accident sites through socio-economic, traffic and land diversity variables.

The GWR results demonstrate significant spatial variations in the relationship linking frequency of accident and different risk factors. Even the non-significant variables in the global model contribute to local models. Moreover, the GWR results suggest that prevention initiatives should take into account the particular context of each neighbourhood of the city in order to prioritise interventions on influential risk factors.

These results reinforce the idea that future prevention should address local realities in their intervention planning, especially in poor neighbourhoods. For example, authorities could adapt educational campaigns to the local population in terms of mobility pattern, risk knowledge, etc. instead of following the usual application of traffic safety principles through universal mass campaigns.

Future public policies should also include measures geared to the improvement of neighbourhood road infrastructures since two significant risk factors are directly related to the physical environment (land use and arterial density). Such interventions would have the potential to reduce inequalities in road traffic risk between deprived and other neighbourhoods.

7. Conclusion

The main objective of this paper has been to integrate the school environment into the analysis of child pedestrian accidents. This approach enables schools to be targeted for prevention on the basis of known and new risk factors. Future methodological improvement could center on the choice of scale; Montréal is a large city with many local disparities that are not specifically studied in this project.

To improve further the results at the school level, one might choose only some boroughs, take the schools they contain as representative samples and then analyse in greater depth new elements of the landscape, such as a special road layout for pedestrians. For example, another improvement could be made by integrating more elements in the construction of the catchment areas, such as physical barriers for pedestrians (railways, highways).

Finally, another potential direction for research would be to explore other elements of the urban landscape in order to integrate more risk factors in the analyses. For example, one might examine not only measures of exposure to risk in terms of traffic lights or school-crossing guard programs but also in terms of children's behaviour or mode of transportation to school. The effects of deprivation on the exposure to risk could then be directly tested.

Child pedestrians move within complex spaces and many factors act in these situations to influence whether or not they will be involved in an accident. Knowing more about these factors can help to reduce levels of insecurity among our children.

Acknowledgements

The authors wish to thank the Conseil de gestion de la taxe scolaire de l'île de Montréal (CGTSIM), the Département de Santé publique de Montréal (DSP) and the Police Road Safety Division for the data provided, as well as the Centre de Recherche sur les Transports (CRT) of Université de Montréal for partial funding.

References

Anselin, L. and D. A. Griffith. 1998. 'Do spatial effects really matter in regression analysis?' *Regional Science Association* 65: 11-34.

Apparicio, P.; Z. Micic and R. Shearmur. 2004. 'Évaluation de l'accessibilité aux supermarchés d'alimentation à Montréal'. *Inédits* 2004-03 www.inrs-ucs.uquebec.ca: 33 pages.

Bailey, T. C. and A. C. Gatrell. (1995) *Interactive spatial data analysis*. New York: Longman Scientific & Technical.

Braddock, M.; G. Lapidus; E. Cromley; R. Cromley; G. Burke and L. Banco. 1994. 'Using a geographic information system to understand child pedestrian injury'. *American Journal of Public Health* 84(7): 1158-61.

Carlin, J. B.; M. R. Stevenson; I. Roberts; C. M. Bennett; A. Gelman and T. Nolan. 1997. 'Walking to school and traffic exposure in Australian children'. *Australian & New Zealand Journal of Public Health* 21(3): 286-92.

City of Montréal. (2004) 'Géobase du réseau routier de l'île de Montréal, Version Juin 2004': Division Géomatique.

Dhillon, P. K.; A. S. Lightstone; C. Peek-Asa and J. F. Kraus. 2001. 'Assessment of hospital and police ascertainment of automobile versus childhood pedestrian and bicyclist collisions'. *Accident Analysis & Prevention* 33(4): 529-37.

DiMaggio, C. and M. Durkin. 2002. 'Child pedestrian injury in an urban setting: descriptive epidemiology'. *Academic Emergency Medicine* 9(1): 54-62.

Dougherty, G.; I. B. Pless and R. Wilkins. 1990. 'Social class and the occurrence of traffic injuries and deaths in urban children'. *Canadian Journal of Public Health*. Revue Canadienne de Sante Publique 81(3): 204-9.

Duperrex, O.; F. Bunn and I. Roberts. 2002. 'Safety education of pedestrians for injury prevention: a systematic review of randomised controlled trials'. *BMJ* 324(7346): 1129.

Environmental Systems Research Institute (ESRI) Inc. (2005) 'ArcInfo-ArcView, Version 9.1': ESRI inc. www.esri.com.

Flahaut, B.; M. Mouchart; E. San Martin and I. Thomas. 2003. 'The local spatial autocorrelation and the kernel method for identifying black zones - A comparative approach'. *Accident Analysis & Prevention* 35(6): 991-1004.

Fotheringham, A. S.; C. Brunsdon and M. Charlton. 2002. *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Wiley.

Graham, D.; S. Glaister and R. Anderson. 2005. 'The effects of area deprivation on the incidence of child and adult pedestrian casualties in England'. *Accident Analysis & Prevention* 37(1): 125-35.

Greig, D. M. 1980. *Optimisation*. London: Longman.

Hasselberg, M.; L. Laflamme and G. R. Weitoft. 2001. 'Socioeconomic differences in road traffic injuries during childhood and youth: a closer look at different kinds of road user.[see comment]'. *Journal of Epidemiology & Community Health* 55(12): 858-62.

Joly, M. F.; P. M. Foggin and I. B. Pless. 1991. 'Geographical and socio-ecological variations of traffic accidents among children'. *Social Science & Medicine* 33(7): 765-9.

Kendrick, D. 1993. 'Prevention of pedestrian accidents'. *Archives of Disease in Childhood* 68(5): 669-72.

Krug, D. E. (1999) 'Injury: A leading cause of the global burden of disease': Geneva: World Health Organisation.

LaScala, E. A.; P. J. Gruenewald and F. W. Johnson. 2004. 'An ecological study of the locations of schools and child pedestrian injury collisions'. *Accident Analysis & Prevention* 36(4): 569-76.

Lee, J. and D. W. S. Wong. 2001. *Statistical Analysis with ArcView GIS.* New York: John Wiley & Sons.

Macpherson, A.; I. Roberts and I. B. Pless. 1998. 'Children's exposure to traffic and pedestrian injuries.[see comment]'. *American Journal of Public Health* 88(12): 1840-3.

McCoy, J. and K. Johnston. 2002. *Using ArcGIS Spatial Analyst*. Redlands, USA: ESRI inc.

Peden, M. M. and World Bank. 2004. *World report on road traffic injury prevention*. Geneva: World Health Organisation.

Pless, B.; R. Verreault; L. Arsenault; J.-Y. Frappier and J. Stulginskas. 1987. 'The epidemiology of road accidents in childhood'. *American Journal of Public Health* 77(3): 358-360.

Posner, J. C.; E. Liao; F. K. Winston; A. Cnaan; K. N. Shaw and D. R. Durbin. 2002. 'Exposure to traffic among urban children injured as pedestrians'. *Injury Prevention* 8(3): 231-5.

Rao, R.; M. Hawkins and B. Guyer. 1997. 'Children's exposure to traffic and risk of pedestrian injury in an urban setting'. *Bulletin of the New York Academy of Medicine* 74(1): 65-80.

Reading, R.; I. H. Langford; R. Haynes and A. Lovett. 1999. 'Accidents to preschool children: comparing family and neighbourhood risk factors'. *Social Science & Medicine* 48(3): 321-30.

Roberts, I.; R. Norton; R. Jackson; R. Dunn and I. Hassall. 1995. 'Effect of environmental factors on risk of injury of child pedestrians by motor vehicles: a case-control study'. *BMJ* 310(6972): 91-4.

Roberts, I.; R. Norton and B. Taua. 1996. 'Child pedestrian injury rates: the importance of "exposure to risk" relating to socioeconomic and ethnic differences, in Auckland, New Zealand'. *Journal of Epidemiology & Community Health* 50(2): 162-5.

Sibert, J. R. 1991. 'Accidents to children: the doctor's role. Education or environmental change?' *Arch Dis Child* 66(7): 890-3.

Société de l'assurance automobile du Québec. (2004) 'Bilan 2003: Accidents, parc automobile, permis de conduire': Québec: SAAQ.

Statistics Canada. (2002) '2001 Canadian Population Census': Ottawa: Statistics Canada.

Stevenson, M. R. and D. A. Sleet. 1997. 'Which Prevention Strategies for Child Pedestrian Injuries? A Review of the Literature'. *International Quarterly of Community Health Education* 13(3): 207-217.

Thill, J. C. 2000. 'Geographic information systems for transportation in perspective'. *Transportation Research Part C-Emerging Technologies* 8(1-6): 3-12.

Wazana, A.; V. L. Rynard; P. Raina; P. Krueger and L. W. Chambers. 2000. 'Are child pedestrians at increased risk of injury on one-way compared to two-way streets?' *Canadian Journal of Public Health* 91(3): 201-206.