The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections

Luis F. Miranda-Moreno
Assistant Professor
Department of Civil Engineering and
Applied Mechanics
McGill University
Phone: 514-398-6589
E-mail: luis.miranda-moreno@mcgill.ca

Patrick Morency
Montreal Department of Public Health
Montreal Health and Social Service Agency
Phone: 514-528-2400 ext.: 3327
E-mail: pmorency@santepub-mtl.qc.ca

Ahmed M. El-Geneidy
Assistant Professor
School of Urban Planning
McGill University
Montreal, Quebec H3A 2K6
Phone: 514-398-8741
E-mail: ahmed.elgeneidy@mcgill.ca

For Citation please use: Miranda-Moreno, L., Morency, P., & El-Geneidy, A. (2011). The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention,* 43(5), 1624–1634.

Abstract

This paper studies the influence of built environment (BE) - including land use types, road network connectivity, transit supply and demographic characteristics - on pedestrian activity and pedestrian-vehicle collision occurrence. For this purpose, a two-equation modeling framework is proposed to investigate the effect of built environment on both pedestrian activity and vehicle-pedestrian collision frequency at signalized intersections. Using accident data of ambulance services in the City of Montreal, the applicability of our framework is illustrated. Different model settings were attempted as part of a model sensitivity analysis. Among other results, it was found that the BE in the proximity of an intersection has a powerful association with pedestrian activity but a small direct effect on pedestrian-vehicle collision frequency. This suggests that the impact of BE is mainly mediated through pedestrian activity. In other words, strategies that encourage densification, mix of land uses and increase in transit supply will increase pedestrian activity and may indirectly, with no supplementary safety strategies, increase the total number of injured pedestrians. Our results show that a 30% reduction in the traffic volume would reduce the total number of injured pedestrians by 35% and the average risk of pedestrian collision by 50% at the intersections under analysis, which is in accordance with previous research. Major arterials are found to have a double negative effect on pedestrian safety. They are positively linked to traffic but negatively associated with pedestrian activity. The proposed framework is useful for the identification of effective pedestrian safety actions, the prediction of pedestrian volumes and the appropriate design of new urban developments that encourage walking.

1. INTRODUCTION

Non-motorized transportation (NMT) such as walking is essential for the development of sustainable transportation systems, whether for short trips, access and egress to/from motorized modes (especially transit) or recreation. As recognized in the literature, NMT can offer substantial benefits in countries like the Unites States and Canada by reducing automobile traffic, reducing energy consumption and emissions, and improving health and fitness (Dora, 1999; Woodcock et al., 2007; Krizek, 2007). Furthermore, an increase in the provision and use of public transit is likely to increase walking and thus the number of pedestrians (Vuchic, 1999; Besser and Dennenberg, 2005; Woodcock et al., 2007).

In spite of the importance and benefits of walking, road facilities in urban areas are still a significant source of harm to pedestrians. Every year, a large number of pedestrians are killed or seriously injured in crashes involving motor vehicles. In the United States, nearly 4,800 pedestrians died (representing approximately 11% of the total road fatalities) while about 70,000 suffered nonfatal injuries each year from 2003 to 2006, most of them being injured in urban settings (NHTSA, 2008). In Canada, 400 pedestrians per year were killed comprising approximately 13% of total road user fatalities in the period 2002 - 2006. In addition, around 6,000 per year were seriously injured in this period (Transport Canada, 2007). To address this problem, local government and urban transportation agencies, not only in Canada but also in other countries around the world, have identified the safety and mobility of pedestrians as high priorities. To this end, investments are constantly allocated through different safety improvement programs.

The development of cost-effective safety improvement programs requires modeling tools to guide decision makers. In the past decade, although considerable research effort has been directed towards addressing road safety issues of motorized modes, relatively little has been directed towards NMT. Significant knowledge and methodology gaps still remain. First, there is a lack of empirical studies that have simultaneously investigated the complex relationships between built environment (BE), pedestrian activity and accident occurrence at the micro-level, in particular at the level of urban intersections[1]. To our knowledge, these relationships have been separately investigated in previous studies, yet none have

[1] A high proportion of vehicle-pedestrian collisions take place at intersections. In Montreal, they represent about 63% of the total pedestrian-vehicle collisions. As defined later, BE is represented by land use, demographic characteristics, transit facilities and road network facilities in the proximity of an intersection.

looked at all of them together. Previous research has mainly focused on the relationship between vehicle-pedestrian collision frequency and traffic and pedestrian volumes (Garder, 2004; Harwood et al., 2008; Lyon and Persaud, 2003; Elvik, 2009). Second, the relationship between BE and pedestrian safety has been investigated primarily at the area-wide level (Lascala et al., 2000; Priyantha et al., 2006; Graham and Glaister, 2003; Sebert Kuhlmann et al., 2009; Wier et al., 2009). Few studies have investigated the relationship between BE and pedestrian activity at intersections (Pulugurtha and Repake, 2008; Schneider et al., 2009). Although these two recent studies have investigated the link between pedestrian activity and BE, they have involved a relatively small sample of intersections. Third, most studies concern US urban areas. Transferability of US evidence to the Canadian context may not be adequate given socio-cultural, urban form and mobility pattern differences (Pucher and Buehler, 2006).

Accordingly the aim of this paper is two-fold:

1) To propose a modeling approach to investigate the relationship between the built environment and both pedestrian activity and collision frequency at signalized intersections.
2) To develop models to identify the main factors associated with both pedestrian activity and collision frequency. The predictive capacity of the developed models is also evaluated as part of this objective.

The paper is structured in several sections. Section 2 offers a literature review on pedestrian safety with particular focus on safety analysis at intersections. Section 3 presents a conceptual framework and model formulation to analyse the interaction between built environment, pedestrian exposure and accident occurrence. Section 4 describes the empirical data used in this research. Section 5 presents the results of the pedestrian activity and collision frequency models; public policy implications and model validation are also part of this section. Finally, section 6 includes the final conclusions and directions for future work.

2. LITERATURE REVIEW

In recent years, a large number of studies have been published addressing different issues related to the development of collision prediction models (Miaou et al., 2003; Miranda-Moreno et al., 2005; Persaud and Lyon, 2007). Relative to the amount of research devoted to motor vehicle-only collisions, few studies have dealt with pedestrian collision occurrence at signalized intersections. Among these studies, we can refer to Brüde and Larsson (1993), Lyon and Persaud (2003), Shankar et al. (2003), Lee and Abdel-Aty (2005) and Harwood, et al. (2008). In this literature, one can observe that past studies on vehicle-pedestrian collision occurrence at intersections have mostly looked at the effect of traffic and pedestrian flows on collision frequency, with few studies incorporating BE and geometric design characteristics in their analyses (Lee and Abdel-Aty, 2005; Harwood et al., 2008). The effect of BE on accident occurrence has been mostly studied at the area-wide (e.g., census tract) level (Graham and Glaister, 2003; Wier et al.,2009). In particular, some area-wide studies have been restricted to the study of children safety (Joly et al., 1991; Lascala et al., 2004; Clifton and Kreamer-Fults, 2007). Some others concentrated on alcohol related collisions and their relationships to BE (Lascala, 2001). In the literature, previous works have studied the direct link between BE and pedestrian collision frequency without specifying whether BE patterns influence collision frequency by directly affecting pedestrian activity, pedestrian accident frequency or both. In area-level studies, attributes such as population density, land use mix and/or commute travel patterns are often included in multivariate models, but, most often, they are treated as any other contributing factor of pedestrian collisions instead of attributes related to pedestrian activity.

To explain the number of pedestrian collisions, pedestrian volume, referred to here as pedestrian activity, is a key element. In the literature, several possible definitions of pedestrian activity can be found depending on the data available and the unit of analysis. In general, measurements of pedestrian activity can be classified as either site-specific (e.g., intersections) or area-wide (e.g., census tract).[2] For urban intersections, pedestrian exposure can be determined using (1) raw pedestrian volumes observed during one or more periods of time (e.g. morning peak, noon, and/or afternoon); or (2) pedestrian volumes estimated according to prediction methods such as Space Syntax or BE models. Some estimation methods and alternative measures of pedestrian exposure have been suggested in the literature (Jonah and Engel, 1983; Greene-Roesel et al., 2007; Lasarre et al., 2007; Pulugurtha and Repake, 2008; Schneider et al., 2009).

Despite the fact that pedestrian volumes are an essential element in road safety analysis; few transportation agencies collect pedestrian data from a large number of sites on a regular basis. Among other reasons, this is due to the fact that site-specific pedestrian count studies are expensive and time-consuming[3]. To address this issue, a simple and efficient method would be to develop predictive built-environment models based on a sample of intersections in an urban area (Pulugurtha and Repake, 2008; Schneider et al., 2009). The goal of this paper is to estimate pedestrian activity based on built-environment attributes in the proximity of an intersection or a crosswalk. However, very little empirical evidence exists in the literature following this approach. One of the few studies is the recent work done by Pulugurtha and Repaka (2008), which develops a model to measure pedestrian activity using data collected for a sample of signalized intersections (176 intersections) in the city of Charlotte, North Carolina. Based on a standard regression analysis, they found that population, total employment, urban residential area, and the number of transit stops are statistically significant variables related to pedestrian

[2] At the area-wide level (e.g. census tract, neighbourhood, municipality, etc.), several measures have been proposed, such as observed pedestrian volumes at a sample of sites, and estimated pedestrian volumes based on travel surveys or network analysis models, for instance. An extensive literature review of previous studies is provided by Greene-Roesel (2007).

[3] Alternative methods have also been proposed to collect pedestrian activity data at the site-specific level (e.g. signalized intersections or pedestrian crossings). A common method is to obtain pedestrian volumes from manual counts taken directly in the field for one or several rush hours. With new automatic data collection methods (pedestrian sensors or video camaras), expansion factors can be defined to obtain daily flows (Schneider et al. 2009). A review of these alternative counting strategies and pedestrian detection technologies has been carried out by Greene-Roesel et al., (2007).

activity. However, the study was conducted in a midsized North American city with a relatively low population density and uses a relatively small sample of intersections. A similar and recent study is the one by Schneider et al., (2009). Using a small sample of 50 intersections in Alameda, CA, these authors linked pedestrian activity at intersections to a variety of surrounding land uses, transportation system attributes, and neighbourhood socioeconomic characteristics. Similar to the Pulugurtha and Repaka (2008), some significant variables include total population, employment, number of commercial retail properties, and the presence of a regional transit station in the proximity of an intersection. However, the results of this study are limited by the fact that pedestrian counts were obtained during few hours, and the sample size was small. Moreover, despite the important contributions of these two studies, they did not validate the predictive capacity of the developed models.

3. CONCEPTUAL FRAMEWORK

For a given intersection, a conceptual framework showing the potential relationships between built environment, pedestrian activity and the burden of road injury is presented in **Figure 1**. This conceptual framework is inspired and supported by previous research (Feng, 2001; Harwood et al., 2008; Clifton et al., 2009; Elvik, 2009; Ewing and Dumbaugh, 2009). The elements of this framework and their relationships are discussed as follows:

- ***Built environment (BE)*****:** This is represented by land use characteristics (employment density, commercial, parks, number of schools, etc.), demographics (population density, children, seniors population, etc.), transit supply (kilometers of bus lanes, number of transit stops, the presence of metro stations, etc.) and road network characteristics (kilometers of streets and major roads, number of intersections, speed limits, etc.), all in the vicinity of an intersection – i.e., in a buffer zone of 50 mto 150 m. As shown in **Figure 1**, pedestrian activity is expected to be associated directly to BE. However, the direct effect of BE on collision frequency is less clear, since its major impact seems to be through pedestrian activity and traffic volume.

- ***Risk exposure***: This is a function of pedestrian activity, traffic volume and motor-vehicle operating speeds. In the literature, pedestrian and traffic volume have been identified as the main determinants of pedestrian-vehicle collision frequency, while motor-vehicle speed is one of the main contributing factors associated with severity.

  Past studies generally show that daily traffic volume (measured usually as average annual daily traffic or ADDT), is the main determinant of pedestrian collision frequency. Most published studies agree that the risk faced by pedestrians with respect to traffic volume is non-linear, and that individual pedestrian risk increases as the number of motor vehicles goes up (e.g., Elvik, 2009). The elasticities reported in the literature range between 0.2 and 1.2 (Harwood et al., 2008; Elvik, 2009; Leden, 2002). Moreover, pedestrian activity (also referred to as pedestrian volume during a period of time), has also been identified as one of the most influential factors in predicting pedestrian-vehicle collisions (Harwood et al. 2008). Among others, Lyon and Persaud (2003), Leden (2002) and Harwood, et al (2008), have found a statistically significant and positive relationship between pedestrian activity and collision frequency at different types of intersections. More importantly, past studies also suggest that the risks faced by pedestrians are highly non-linear and that as the number of pedestrians increases, the risk faced by each pedestrian decreases – this is refereed as the "safety in numbers" effect (Leden, 2002; Jacobsen, 2003). The literature reports elasticity estimates between 0.3 and 0.7 for the pedestrian volume and collision frequency relationship – Geyer et al. (2006); Harwood et al. (2008); Elvik, (2009).

- ***Geometric design:*** Few studies have investigated the effects of geometric design attributes – including road width, number of lanes, presence of marked pedestrian crosswalks, presence of median, types of turn restrictions, etc. - on collision frequency at intersections. For instance, Harwood et al. (2008) have completed comprehensive studies of pedestrian accident occurrence and the effect of some geometric attributes. In their study, they found that after controlling for pedestrian and traffic volumes, the number of lanes and the presence of raised medians had significant effects on collision frequency. The effect of curb parking was also documented by Box (2004). In addition, previous case-control studies have showed that on-street parking is more common in the environments of injured child pedestrians (e.g., Agran al., 1996). It is worth noting that geometric design may also have an effect on speeds and in turn, on severity. For instance, road width or traffic calming measures should be associated with operating speeds (King et al., 2003, Ewing and Dumbaugh, 2009; Tester et al., 2004).

Insert Figure 1 here

As illustrated in **Figure 1**, there are several dimensions to studying the linkages between BE, risk exposure and pedestrian safety. In this research, we narrow the analysis to focus on the relationship between BE, pedestrian activity, and pedestrian-vehicle collision frequency. The gap we wish to address is the absence of studies exploring the effect of BE on both pedestrian activity and collision occurrence at intersections. To study the mechanisms by which BE characteristics link directly to pedestrian activity and pedestrian collision frequency, a two-equation modeling framework is proposed. For a given intersection, the proposed two-equation model is then defined as:

$$\text{(i)} \quad P = f(\boldsymbol{\alpha}; \boldsymbol{BE}, \varepsilon_{1i})$$

$$\text{(ii)} \quad \theta = f(\boldsymbol{\beta}; V, P, \boldsymbol{x}, \varepsilon_{2i}) \qquad \textbf{(1)}$$

where:

P - pedestrian activity (daily or peak-hours average) at a given intersection

θ - pedestrian accident frequency (in yearly-base period)

BE - built environment attributes in the proximity of an intersection including: land use types (L), road network connectivity (R), transit supply (T) and demographic characteristics (D),

V	- motor vehicle volumes (daily average)
$\boldsymbol{x}$	- geometric design characteristics
$\boldsymbol{\alpha}, \boldsymbol{\beta}$	- vectors of regression parameters
ε_{1i} , ε_{2i}	- model error representing unobserved heterogeneities

In the first equation, pedestrian activity (P) is assumed to be directly associated to *BE* attributes (*L, R, T* and *D*) and unobserved variations (ε_{1i}). In the second equation, the number of accidents in an intersection during a given period of time depends mainly on vehicle traffic volume (V), pedestrian activity (P), intersection geometry and unobserved variations (ε_{2i}). Note that model errors can be independent or dependent - a distinction which is tested later. In addition, *L, R, T* and *D* in the proximity of an intersection are assumed to have a direct marginal effect on collision frequency. To test this hypothesis, collision frequency models with and without *BE* variables are compared. As mentioned above, this paper challenges the hypothesis that the effect of *BE* patterns on pedestrian safety at intersections are mainly through the generation of pedestrian activity as an important determinant of risk exposure.

Modeling approach

To represent these two outcomes, P and θ, different model settings are tested. For the pedestrian activity (P in **Eq. 1**), log-linear and standard negative binomial (NB) regression models are used. This corresponds to the fact that pedestrian activity is a positively skewed count variable. For instance, the log-linear model is defined as,

$$ln(P_i) = \boldsymbol{\alpha} \cdot \boldsymbol{BE_i} + \varepsilon_{1i}, \qquad \textbf{(2)}$$

where $ln(P_i)$ is the log of the number of pedestrians counted during a day or several hours at intersection i ($i=1,...,n$), $\boldsymbol{\alpha}$ is a set of regression parameters to be estimated from the data and $\boldsymbol{BE}$ is the set of built environment factors such as land use, demographic characteristics, transit supply and road network connectivity.

To model the pedestrian-vehicle collision frequency, count data regression models can be used such as the standard NB model and its extensions. In this case, we assume that the number of pedestrian-vehicle accidents (Y_i) at intersection i over a given time period is Poisson distributed. Then, the standard NB or Poisson/Gamma model for Y_i can be represented as (Winkelmann, 2003),

$$Y_i | \theta_i \sim Poisson(\theta_i),$$

$$\theta_i = \mu_i \cdot exp(\varepsilon_{2i}) \qquad \textbf{(3)}$$

where θ_i is the mean pedestrian accident frequency and $exp(\varepsilon_{2i})$ denotes the multiplicative random effect of the model[4] - which is assumed to follow a Gamma distribution with both parameters equal to $1/\kappa$ and κ being the dispersion parameter. Moreover, $\mu_i = f(\boldsymbol{\beta}; V_i, P_i, \boldsymbol{x}_i)$, where, as defined in **Eq. 1**, V_i and P_i are traffic flow and pedestrian activity, respectively, and are a measure of risk exposure. Here, $\boldsymbol{x_i}$ stands for geometric site-specific attributes and $\boldsymbol{\beta} = (\beta_0, ..., \beta_k)'$ is a vector of regression parameters to be estimated from the data. To capture the nonlinear relationship between the mean number of accidents and traffic/pedestrian flows, alternative functional forms can be specified (e.g., $A_i = \beta_0 V_i^{\beta_1} P_i^{\beta_2}$ could be used for a model only with traffic and pedestrian volumes).

Instead of assuming a Gamma-error term, a Poisson/Lognormal model can be obtained by assuming the model error to be log-normal. Although the NB regression model is the most popular for modeling accident data, a simple extension of this model would allow variability in the dispersion parameter, κ as

[4] This error term represents all unobserved heterogeneities and is assumed to be independent in the NB model.

κ_i, (see e.g., Heydecker and Wu, 2001; Miaou and Lord, 2003; and Miranda-Moreno et al., 2005). Such a model is referred to here as the generalized negative binomial model (GNB). In this model, κ_i is a function of a vector of site attributes, i.e., $\kappa_i = exp(\boldsymbol{w}_i^{'}\boldsymbol{\gamma})$, where $\boldsymbol{w}_i$ is a vector of site-specific attributes which may differ from those included in μ_i and $\boldsymbol{\gamma}$ is a vector of regression parameters. This model assumption may improve the parameter estimates and the fit of the model to the data (Heydecker and Wu, 2001; Miranda-Moreno et al., 2005). The GNB model is applied here in order to structure the unobserved heterogeneities among intersections as a function of some intersection attributes, such as traffic volume. For instance, one may suspect that unobserved factors associated with the pedestrian collision occurrence can have similar patterns structured in clusters, such as intersections situated in the same corridor.

In order to further explore the potential effect of a high frequency of zeros and the presence of clusters (subpopulations) in the collision data, a latent-class NB model (also known as finite mixture models) is also used (Miranda-Moreno, 2006). Latent-class NB models provide a convenient way of identifying latent segments (classes) for which parameters in a specified model differ. Here, for the case with only two classes or subgroups (q=2), we refer to the two-compound NB model. In this case, the probability distribution of Y_i may be simply written as $Y_i|\theta_{iq} \sim \omega Poisson(\theta_{1i}) + (1-\omega)Poisson(\theta_{2i})$, where $0< \omega <1$ gives the proportion of intersections belonging to the class of high collision frequency, and the parameters θ_{1i} and θ_{2i} are the high and low mean accident frequency, respectively – see e.g., Miranda-Moreno(2006). Finally, to test the potential correlation between the error terms (ε_{1i} and ε_{2i}), a Bivariate Poisson model could be also considered, in this case, $cov(\varepsilon_{1i}, \varepsilon_{2i}) \neq 0$, see Karlis and Ntzoufras (2005). Note that the geometric designs of some intersection may influence both collision occurrence and pedestrian flows, e.g. raising medians and crosswalks at a given intersection may reduce collisions and at the same time increase pedestrian activity(Zegeer et al., 2002).

4. DATA ASSEMBLY AND DESCRIPTION

For this research, data collected at 519 signalized intersections in the City of Montreal, Quebec, Canada are used to develop pedestrian activity and collision frequency models at signalized intersections. Data for this analysis is provided by the Montreal Transportation Department (*Direction des transports de Montreal*) and the Montreal Public Health Department (*Direction de Santé Publique de Montreal*). **Figure 2** is a map of Montreal with the studied signalized intersections.

Insert Figure 2 here

Vehicle-pedestrian collision data

For the intersections under analysis, all injured pedestrians for whom an ambulance was sent on the island of Montreal over a five-year period (from 1999 to 2003) are included. Some summary statistics of the vehicle-pedestrian collision frequency are provided in **Table 1**. Urgences-santé (Montreal's ambulance service) is the source of data and has a monopoly in the territory covered - for details see Morency and Cloutier (2006). For every call made to 911 in Montreal, the caller's address is automatically sent to Urgences-santé and the location of the victim is validated over the phone. This location is instantly mapped in a GIS to dispatch and guide an ambulance. As opposed to the police report data usually used in published pedestrian safety research, ambulance service data have less underreporting and misallocation problems. Identical locations were aggregated and the Montreal hierarchical street network (called Geo-base) was used to combine locations within a 15 m radius of an intersection.

Pedestrian and traffic data

For all the studied intersections, pedestrian and traffic volumes were collected by the City of Montreal for the year 2003. Pedestrian volumes were obtained from three different periods: i) peak morning (from

7:30 a.m. to 8:30 a.m.), ii) noon period (from 12:00 to 1:00 p.m.), and ii) peak afternoon (from 4:30 p.m. to 5:30 p.m.). Traffic volumes were also available for the same intersections represented by the average annual daily traffic (AADT). These daily volumes were estimated based on 3-hour traffic counts collected on the same days as the pedestrian data. The data at each intersection were collected on a weekday with normal weather conditions during spring and summer. Some summary statistics of pedestrian and AADT data at the studied intersections are presented in **Table 1.**

Insert Table 1 here

BE: land use, demographics, transit and road network

BE variables in the vicinity of each intersection are generated using Geographic Information Systems (GIS) data obtained from various sources. To account for the impact of buffer dimension, different buffer sizes were tested, including 50, 150, 400 and 600 meters. Because the area under study (central neighbourhoods in Montreal) is more dense, has a rich mix of land uses, and has high transit accessibility, this study uses smaller buffers than those published by Pulugurtha and Repake (2008) and Schneider et al., (2009). The 50-meter buffer was used to find how an intersection's immediate surroundings affected pedestrian activity. A 150-meter buffer examines the effects of characteristics within close proximity to the intersection. The 400-meter and 600-meter buffers served as proxies for how characteristics at a walking distance or neighbourhood scale affect the level of pedestrian activity at a particular intersection. A list of the variables with a short definition is provided in **Table 2**. To extract demographic data, census data at the census tract level were intersected with each buffer generated around the intersections. To classify intersections according to the number of approaches (three-legged versus four-legged intersections), a dummy variable was generated.

Insert Table 2 here

5. MODELING RESULTS

Based on the conceptual model specification defined in **Eq. 1**, a regression modeling analysis is then carried out to investigate the relationship between BE, pedestrian activity and accident occurrence at the signalized intersections. To address the multi-colinearity issue, an exploratory analysis is first carefully done to identify serious problems of correlation between BE factors. Pedestrian activity and accident frequency models are then developed to account for the different modeling issues discussed earlier, such as unobserved heterogeneities, clustering, and correlation between random effects.

Exploratory data analysis

As a first step, an exploratory analysis is conducted to identify high correlations between the different BE attributes. A correlation matrix for each buffer size is generated, including 50, 150, 400 and 600 meters - see **Table 3**. This exercise is also useful for identifying the potential factors related to pedestrian activity and collision frequency. Based on these results, some observations are highlighted:

- The land use variables that are highly correlated with pedestrian activity are commercial area[5], open space, number of jobs and number of schools, all at the 400m buffer. Population and employees living in the buffer are also correlated with pedestrian volumes[6]; however, the two are also highly correlated. For the transit characteristics, presence of metro and bus stops are the ones with the highest correlations at the 150-meter buffer. Among the road network features, local road length, average street length, and proportion of major roads are the variables most correlated to pedestrian activity at the 400-meter buffer.
- Pedestrian volume and AADT are correlated with collision frequency, with a correlation of about 0.4. Other variables, such as commercial area, number of bus stops, and local road length also

[5] with the highest correlation at the 50m buffer (r= 0.51).
[6] with the maximum correlation at the 400m buffer.

show correlations higher than 0.25. Note however that these variables are also highly correlated to pedestrian activity and/or AADT. From this simple analysis, one should be aware that including BE variables along with pedestrian and traffic volumes in the same model may generate multicolinearity problems.

Insert Table 3 here

Pedestrian activity model

The next step is to develop a prediction regression model for pedestrian activity as a function of BE variables, as defined in **Eq. 1**. Log-linear and the standard negative binomial regression models are fitted to the data. Given that the results are very similar, only the log-linear model outcome is then reported.

Different combinations of explanatory variables were tested based on the exploratory analysis reported previously – correlations among explanatory variables were taken into account to avoid problems of multicolinearity. To do so, if there was a high correlation between two explanatory variables (correlation ≥ 0.3), the variable that had a weaker relationship (lower correlation) with pedestrian activity was omitted from the model. In very few cases, best judgment was used[7]. **Table 4** shows the parameter estimates for the final log-linear model. Note that eight variables have statistically significant effects on pedestrian activity, including population density, commercial land use, number of jobs, number of schools, presence of a metro station, number of bus stops, percentage of major arterials, and average street length (as a measure of connectivity). An intersection type indicator was also introduced to the

[7] For instance, the presence of metro stations within 400 meters had a slightly lower correlation value than the number of metro stations within 150 meters. However, the number of metro stops in this buffer was chosen because it had a lower correlation with the number of bus stops. Moreover, to evaluate the goodness-of-fit of each model, the R-squared measure was used.

model to differentiate between three-legged and four-legged intersections. With the exception of the variable representing the proportion of major arterials, all variables have a positive effect on pedestrian activity. The results also show that commercial land use and number of bus stops were related to pedestrian activity as the 50-meter and 150-meter buffers, respectively, with the other six related at the 400-meter buffer. In terms of goodness-of-fit, an R-square of 55% is obtained. These results are in agreement with those obtained in two urban areas in US (Palugurtha and Repaka, 2008; Schneider et al., 2009).

Table 4 also shows the elasticities associated with each explanatory variable calculated at the point of means. From these elasticities, one can see that a 100 percent increase in the population living in the proximity of an intersection (50 meters) is associated with a 34 percent increase in pedestrian activity. A 100 percent augmentation of commercial area or bus stops in the vicinity of an intersection also represents, respectively, a 20 percent and a nearly 37 percent increase in pedestrian activity. The presence of a metro station increases pedestrian activity by 28 percent. Interestingly, a 100 percent increase in the proportion of major arterials is associated with a decrease of 16 percent in pedestrian activity. These results clearly indicate that the implementation of local policies encouraging the density, diversity and design, which are known as the 3 D's – (Cervero and Kocklman, 1997; Boer et al., 2007; Gauvin et al., 2008) and public transit (Besser and Dannenberg, 2005) are expected to increase pedestrian activity (and then the likelihood of pedestrian collisions) at intersections significantly.

Insert Table 4 here

Pedestrian collision frequency models

In the collision occurrence modeling, two issues were taken into account: correlation between pedestrian activity and BE factors, and sensitivity of the results to the model specification. The impact of the correlation problem was investigated by calibrating models with and without BE and the sensitivity was studied by fitting different model settings. This is explained as follows.

As illustrated in **Table 3**, an important issue to take into account is the high level of correlation between the BE and both pedestrian activity and collision frequency – e.g., commercial area is highly correlated with these two outcomes. This makes it difficult to precisely identify and isolate the effect of BE attributes on collision frequency. A simple solution is to fit different collision models with and without BE attributes. Accordingly, the results of a model without BE variables is reported in **Table 5a**. As explanatory variables, this model includes only pedestrian activity, traffic volume (AADT) and a dummy variable for intersection type. Then, by introducing the BE attributes to the previous model, different combinations of variables are tested. The best model of this exercise is reported in the **Table 5b**.

- *Collision frequency model only with pedestrian activity and traffic volume*

From the results in **Table 5.a**, one can see that both pedestrian and traffic volumes are positively and statistically significant, as expected. Based on the NB and GNB models, which are the most common models used in previous studies, the estimated coefficients for pedestrian and traffic volume are 0.45 and 1.15, respectively, for NB, and 0.46 and 1.12, respectively, for GNB model. This suggests that motor vehicle traffic is the major determinant of pedestrian accident frequency. In terms of elasticities, this means that a reduction of 10 percent in the current traffic condition will represent a decrease of 11 percent in the number of pedestrian collisions. For the two-compound NB model, we can see that the effect of traffic flow can vary across the two subpopulations or classes of intersections. For instance, for class 2, the elasticity is approximately 2.4 - this means that traffic flow has a much higher impact in this subgroup.

Overall, results are consistent with previous works in Canada and the US (Elvik, 2009). For instance, in a previous study in Canada, Leden (2002) reported estimates of 0.33 for pedestrian flows and 1.19 for left-turning motor vehicles. In another earlier study in the City of Toronto, Lyon and Persaud (2003) reported estimates ranging from 0.41 to 0.74 for pedestrian volume and from 0.40 to 0.58 for traffic volume. In US cities, Geyer et al. (2006) and Jacobsen (2003) obtained pedestrian coefficients of 0.61 and 0.41, respectively. From these and other results summarized in the recent work of Elvik (2009), it can be observed that the estimates for traffic flow present a larger range of variability (from 0.2 to 1.2), than estimates for pedestrian flows, with estimates generally between 0.3 and 0.7. This variability might be associated to the type of intersections (e.g., three-legged versus four legged intersection), number of years of accident data involved in the analysis, quality of the traffic data, sample size, estimation method, etc. As suspected, the effect of the intersection indicator is also significant, which means that four-leg intersections have a higher frequency of accidents than three-leg intersections.

- *Collision frequency model with BE variables*

From the results in **Table 5.b**, one can observe that some BE variables are statistically significant, including commercial area, and number of bus stops and schools - all other variables are non-significant at the 5% level when other factors are taken into account [8]. However, parameter estimates of pedestrian volume and AADT are substantially reduced with the incorporation of these variables (note however that the dummy variable effect of number of intersection legs stays very similar). For instance, in the NB model, the parameters for pedestrian and traffic volumes go down from 0.46 to 0.26 and 1.1 to 0.90, respectively. Moreover, note that the model fit is only slightly improved when incorporating BE variables. This suggests that most of the observed variability is explained by pedestrian and traffic volumes and that the impact of BE occurs through its association with pedestrian activity. For instance,

---

[8] The negative sign of the number of schools can be related to speed limits and/or some calming measures that may be applied around schools.

as seen in **Table 3**, the number of bus stops and commercial land use are correlated to both pedestrian activity and collision frequency.

- *Sensitivity of the results to the model specification*

In addition to the NB, the GNB and two-compound NB models, as defined in section 3, were applied to the data in order to explore the sensitivity of the results to the model specification. The GNB model was used to test the potential presence of heterogeneities across intersections in the dispersion parameter and the two-compound NB model was applied to explore the presence of clusters or subgroups of intersections. The best outcomes of each model are presented in **Tables 5.a** and **5.b**. From this, we can see that the parameter estimates and the AIC (goodness-of-fit criterion) of the GNB model are very similar to the ones obtained with the NB model. This basically eliminates the potential presence of heterogeneities in the over-dispersion parameter across intersections, meaning that the results are not sensitive to the variability of the dispersion parameters across intersections.

The AIC of the two-compound NB model is very similar to the one obtained with the NB model. However, the results show that the effect of traffic volume is different between the two subgroups –the difference between the two sub-groups being statistically significant. Despite the potential presence of clusters, note that the other model parameters (and the AIC) of the two-compound NB model are very similar to the ones obtained with the NB model. This again shows the stability of the results across model settings. In addition, the potential presence of subgroups of intersections needs to be tested using a larger sample of intersections that includes more site-specific variables, since subgroups might be caused by the omission of important factors.

Finally, to test the potential correlation between the error terms in **Eq. 1** (cov(ε_P, ε_A) $\neq$ 0), a bivariate Poisson regression model was also fitted to the data (Karlis and Ntzoufras, 2005). In this model, the two equations (pedestrian activity and collision frequency) are calibrated simultaneously, assuming dependency between the two model errors. Since non-evidences of endogeneity were identified, the results are not reported in this paper. Although parameter estimates are very similar to those previously reported, more research is needed in this statistical issue.

- *Indirect impact of BE on pedestrian collision occurrence*

To illustrate the indirect impact of BE on pedestrian collision frequency, the elasticities of both the pedestrian activity model and NB collision frequency model without BE factors are used (**Table 4** and **Table 5.a**). From the pedestrian activity model, it can be observed that doubling population density would increase pedestrian activity by 34%, which would represent an increase of (0.34×0.45) 15% on average in the collision frequency, but a reduction of 18% in the individual collision risk, assuming that a "safety in numbers" effect exists. Moreover, the combined effect of land-use and transit strategies (e.g., doubling population and job density, and bus stops in the proximity of an intersection) would increase the collision frequency by around 45% (calculated as the sum of pedestrian activity elasticities times the elasticity of pedestrian activity in the injury frequency model)

Insert Table 5 here.

Model validation

To validate the predictive capacity of the previously developed models, the intersection sample was randomly divided into an *exploratory* sample, consisting of 80 percent of the intersections (415 intersections), and a *validation* sample, consisting of 20 percent of the intersections (104 intersections). The exploratory sample was then used to calibrate the pedestrian activity model. This model was then applied to the validation data set to estimate pedestrian volume. Observed versus estimated values were then compared. As a second step, estimated and observed pedestrian volumes in the exploratory sample were used to fit two standard negative binomial (NB) models. According to these two outcomes, the expected collision frequency was computed using the validation sample. The predictive capacity of a model was then determined by comparing the estimated values with the observed collision frequencies. Sample splitting was repeated 100 times to produce 100 sets of exploratory and validation samples (a code in R software was developed for this purpose). To measure prediction accuracy, two measures were used: the Pearson correlation (r_p) and normalized root mean squared deviation (NRMSD)[9]. The results of this validation procedure are reported in **Table 6.** According to these results, the performance of the models seems to be acceptable. For instance, a mean correlation of 0.72 was observed between the predicted and observed pedestrian activity. In addition, the collision frequency predictions based on observed and estimated pedestrian activity are very similar, producing relatively low mean values of NRMSD; mean values of around 0.18 and 0.17 were obtained when using predicted and observed pedestrian activity, respectively (again, lower NRMSD values convey smaller discrepancy between the calculated and observed values). This suggests that when pedestrian activity data is not present, practitioners can use the built environment to predict pedestrian activity and incorporate it in collision prediction models.

Insert Table 6 here.

Some policy implications

Different practical implications can be supported from our analysis:

- As expected, a strong link between BE and pedestrian activity at intersections was confirmed. However, the direct impact of the surrounding BE (ex. land use patterns) on collision frequency seems to be marginal, with traffic volume and pedestrian activity being the main determinants of

[9] Note that the closer to one the correlation is, the better the model predictions are. On the contrary, the closer to zero the NRMSD is, the better the validation results are.

collision occurrence. Although, detailed intersection geometric characteristics were not available (e.g., road width, number of lanes and median presence), their effect was partially captured by the intersection type indicator that was statistically significant. This means that four-leg intersections exhibit a higher pedestrian collision risk than three-leg intersections, after controlling for other factors – this is in accordance with the literature (Harwood et al, 2008).

- Urban policies aimed at increasing population density, land use mix, transit offering and road network connectivity may have the benefit of increasing pedestrian activity. However, their potential indirect effect on pedestrian safety is paradoxical because (1) with no supplementary safety strategies and countermeasures, the total number of injured pedestrians would increase with pedestrian activity and (2) given the fact that the risk of injury to pedestrians is highly non-linear, the risk for each pedestrian would decrease as the number of pedestrians increases. This is obviously assuming the validity of the "safety in number" concept (Jacobsen, 2003). However, these results should be used with caution, since more research is needed on the underlying behavioural and environmental explanatory mechanisms of this concept, as was recently discussed by Bhatia and Wier (2010).

- The greater the number of motor vehicles at intersections, the higher the collision frequency. In addition to their beneficial effect on congestion, noise and emissions, strategies to reduce traffic volume would lower both the individual risk of pedestrian collision and the total number of injured pedestrians. For instance, a reduction of 30% in the traffic volume at each intersection in the area of study would greatly reduce the average pedestrian risk (-50%) and the total number of injured pedestrians at those intersections (-35%).This is shown graphically in **Figure 3**. It is noteworthy that major roads seem to have a double negative effect on pedestrians, being positively associated with traffic volume and negatively related with pedestrian activity. Such results support the idea of retrofitting urban major roads into complete streets (Laplante and

McCann 2008). This is an appropriate strategy, considering that in urban settings like Montreal, the majority of pedestrians are injured on major roads (Morency and Cloutier 2006).

Insert Figure 3 here

CONCLUSIONS

This work aims to understand how the BE affects both pedestrian activity and collision frequency. For this purpose, a modeling framework has been proposed to jointly analyze pedestrian activity and pedestrian-vehicle collision frequency at intersections. This framework is useful for the identification of effective pedestrian safety actions, the prediction of pedestrian volumes and the appropriate design of new urban developments that encourage walking.

Different model settings were tested including the standard NB model, the GNB model (to test for the presence of observed heterogeneities in the dispersion parameter), and the latent class NB model (to explore the presence of sub-groups or clusters of intersections). Overall, the results were consistent across the model specifications for both the collision frequency and pedestrian activity models. Results for the pedestrian activity and collision frequency models have been presented for intersections in Montreal. In accordance with previous studies, the results show that some BE characteristics – including population density, commercial land use, number of jobs, number of schools, presence of metro station, number of bus stops, percentage of major arterials and average street length - have a powerful association with pedestrian activity. The estimated elasticities vary from -0.16 to 0.46; they are all positive with the exception of major arterials. The influence of built environment on pedestrian collisions at intersections, however, seems largely mediated through pedestrian activity and traffic volume, which, based on the results of the NB model, have elasticities of 0.45 and 1.1, respectively. These values fall within the range of those reported in the literature. This study provides some additional evidence that more pedestrian activity and traffic generate more accidents, with traffic volume being the primary cause of collision frequency at intersections. As illustrated in the previous section, a reduction in traffic volume would be associated with great improvements in pedestrian safety. Finally, an original validation procedure measured the predictive capacity of our models.

The proposed two-equation framework has shown the indirect impact of BE on pedestrian collision. For instance, it was estimated that an increase of 100% in population density in the vicinity of an intersection would augment pedestrian activity by 34%. This would represent an increase of 15% in the average number of collisions, but a reduction of 18% in the individual collision risk, assuming that a “safety in numbers” effect (causality) exists.

This study is not without limitations. Our future efforts will concentrate on examining the validity of these findings across a wider spectrum of intersections and using longitudinal or before-after observational studies. The use of cross-sectional data does not take into account the possibility of endogeneity between collision frequency and pedestrian flows. For instance, it is possible that certain geometric design characteristics influence both pedestrian flows and collision occurrence. In other words, the presence of raised medians and crosswalks in an intersection might decrease collisions and attract pedestrians (increasing flows). If this is the case, the estimated relationships between pedestrian activity and number of collisions would only reflect a spurious correlation rather than true causality. To look at this potential issue, a before-after or longitudinal analysis is recommended. In addition, intersection geometric characteristics (ex. road width) should be included. A simultaneous modeling approach will be further explored to investigate this issue. Spatial correlation among outcomes will also be explored to validate these results using a larger sample of intersections.

Acknowledgements

The first author acknowledges the financial support of *Natural Sciences and Engineering Research Council of Canada*. We would like to thank the City of Montreal for providing some of the data used in the analysis. We would also like to add our thanks to Lee Haber, Kevin Manaugh, Sabrina Chan and Tyler Kreider for their help in data preparation. Last but not least we would like to thank the two anonymous reviewers for their feedback on the earlier version of the paper.

References:

Agran, P.F., Winn, D.G., Anderson, C.L., Tran, C., & Del Valle, C.P. (1996). The role of the physical and traffic environment in child pedestrian injuries. Pediatrics, 9:1096-1103.

Bhatia, R., and Wier, M. (2010). "Safety in Numbers" re-examined: Can we make valid or practical inferences from available evidence?. Accid Anal Prev (in press).

Besser, L.M. & Dannenberg, A.L. (2005). Walking to public transit: steps to help meet physical activity recommendations. Am.J.Prev.Med., 29: 273-280.

Boer, R., Zheng, Y., Overton, A., Ridgeway, G.K., & Cohen, D.A. (2007). Neighborhood design and walking trips in ten U.S. metropolitan areas. Am J Prev Med, 32(4): 298-304.

Box, P. (2004). Curb-parking problems: Overview. Journal of Transportation Engineering 130(1), 1-5.

Brüde, U. & Larsson J. (1993). Models for predicting accidents at junctions where pedestrians and cyclists are involved. Accid Anal Prev 25 (5): 499-519.

Cervero, R. & Kockelman, K. (1997). Travel demand and the 3Ds: density, diversity, and design. Transportation Research D, 2(3), 199-219.

Clifton, K. J. & Kreamer-Fults, K. (2007). An examination of the environmental attributes associated with pedestrian-vehicular crashes near public schools. Accid Anal Prev, 39: 708-715.

Clifton, K.J., Burnier, C.V. & Akar, G. (2009). Severity of injury resulting from pedestrian-vehicle crashes: What can we learn from examining the built environment. Transportation Reearch Part D 14(6): 425-436.

Dora, C. (1999). A different route to health:implications of transport policies. BMJ Education and debate 318(7199): 1686-1689.

Elvik, R. (2009). The non-linearity of risk and the promotion of environmentally sustainable transport. Accid Anal Prev 41(4): 849-855.

Ewing, R. & Dumbaugh E. (2009). The built environment and traffic safety: A review of empirical evidence. Journal of Planning Literature 23: 347-367.

Feng, C. (2001). Synthesis of studies on speed safety.Transportation Research Record 1779: 86-92.

Garder, P.E. (2004). The impact of speed and other variables on pedestrian safety in Maine. Accid Anal Prev, 36: 533-542.

Gauvin, L., Riva, M., Barnett, T., Richard, L., Craig, C.L., Spivock, M. et al. (2008). Association between neighborhood active living potential and walking. Am J Epidemiol, 167: 944-953.

Geyer, J., Raford, N., Pham, T., & Ragland, D.R. (2006). Safety in numbers. Data from Oakland, California. Transportation Research Record 1982, 150–154.

Graham, D. J. & Glaister S. (2003). Spatial variation in road pedestrian casualties: The role of urban scale, density and land-use mix. Urban Studies 40 (8): 1591-1607.

Greene-Roesel, R., Diogenes, M.C. & Ragland, D.R., (2007). Estimating pedestrian accident exposure: Protocol report (No. UCB-TSC-RR-2007-5). Berkeley: UC Berkeley Traffic Safety Center.

Harwood, D.W., Torbic, D.J. et al. (2008). Pedestrian safety prediction methodolgy. Transportation Research Board NCHRP 17-26.

Heydecker B.G., & Wu, J. (2001) Identification of Sites for Accident Remedial Work by Bayesian Statistical Methods: An Example of Uncertain Inference. Advances in Engineering Software, 32, 859-869.

Jacobsen, P.L. (2003). Safety in numbers: More walkers and bicyclists, safer walking and biking. Injury Prevention 9: 205-209.

Joly, M.F., Foggin, P., & Pless, I. (1991). Socioecological determinants of the risk of accidents in young pedestrians. Rev Epidemiol Sante Publique, 39: 345-351.

Jonah, B. & Engel, G. (1983). Measuring the relative risk of pedestrian accidents. Accid Anal Prev, 15(3): 193-206.

Karlis, D. & Ntzoufras I. (2005). Bivariate Poisson and diagonal inflated bivariate Poisson regression models in R. Journal of Statistical SoftwarE 14 (10): 1-36.

King, M.R., Carnegie, J.A., & Ewing, R. (2003). Pedestrian Safety Through a Raised Median and Redesigned Intersections. Transportation Research Record 1828: 56-66.

Krizek, K.J. (2007). Estimating the economic benefits of bicycling and bicycle facilities: An interpretive review and proposed methods. Essays on Transportation Economics, Springer Publishing: 219-248.

Laplante, J. & McCann B. (2008). Complete streets: We can get there from here. Institute of Transportation Engineers. ITE Journal: 24-28.

Lasarre, S., Papadimitrou E., Yannis G., & Golias J. (2007). Measuring accident risk exposure for pedestrians in different micro-environments. Accid Anal Prev 39: 1226-1238.

Lascala, E.A., Gerber, D., & Gruenewald, P.J. (2000). Demographic and environmental correlates of pedestrian injury collisions: a spatial analysis. Accid Anal Prev, 32: 651-658.

Lascala, E.A., Johnson, F.W., & Gruenewald, P.J. (2001). Neighborhood characteristics of alcohol-related pedestrian injury collisions: a geostatistical analysis. Prev Sci, 2: 123-134.

Lascala, E.A., Gruenewald, P.J., & Johnson, F.W. (2004). An ecological study of the locations of schools and child pedestrian injury collisions. Accid Anal Prev, 36: 569-576.

Leden, L. (2002). Pedestrian risk decrease with pedestrian flow: A case study based on data from signalized intersections in Hamilton, Ontario. Accid Anal Prev 34: 457-464.

Lee, C. & Abdel-Aty M. (2005). Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. Accid Anal Prev 37(4): 775-786.

Lyon, C. & Persaud B. (2003). Pedestrian collision prediction models for urban intersections. Transportation Research Record 1818: 102-107.

Miaou S.P., & Lord, D. (2003) Modeling Traffic Crash-flow Relationships for Intersections: Dispersion Parameter, Functional Form, and Bayes versus Empirical Bayes. Transportation Research Record 1840, 31-40.

Miaou, S.P., Song J.J. & Mallick B.K. (2003). Roadway traffic crash mapping: A space-time modelling approach. Journal of Transportation and Statistics 6 (1): 33-57.

Miranda-Moreno, L.F., Fu L., Saccomano F., & Labbe, A. (2005). Alternative risk models for ranking locations for safety improvement." Journal of the Transportation Research Board 1908: 1-8.

Miranda-Moreno, L. F. (2006). Statistical Models and Methods for Identifying Hazardous Locations for Safety Improvements. PhD thesis. University of Waterloo, Waterloo, Ontario, Canada.

Morency, P. & Cloutier M. (2006). From targeted "black spots" to area-wide pedestrian safety. Injury Prevention 16 (6): 360-364.

National Highway Traffic Safety Administration – NHTSA (2008). National Pedestrian Crash Report - DOT HS 810 968. U.S.Department of Transportation.

Persaud, B. & Lyon C. (2007). Empirical Bayes before-after safety studies: Lessons learned from two decades of experience and future directions. Accident Analysis and Prevention 39 (3): 546-555.

Priyantha Wedagama, D. M., Bird, R.N., & Metcalfe, A.V. (2006). The influence of urban land-use on non-motorised transport casualties. Accid Anal Prev, 38: 1049-1057.

Pucher, J. & Buehler, R. (2006). Why Canadians cycle more than Americans: A comparative analysis of bicycling trends and policies. Transportation Policy, 13, 265-279.

Pulugurtha, S.S. & Repake S.R. (2008). Assessment of models to measure pedestrian activity at signalized intersections. Transportation Research Record 2073: 39-48.

Schneider, R., Arnold L., & Ragland, D. (2009). Pilot Model for Estimating Pedestrian Intersection Crossing Volumes. Journal of the Transportation Research Board, 2140, pp 13-26.

Shankar, V.N., Ulfarrson G.F., Pendyala R., & Nebergall, M.B. (2003). Modeling crashes involving pedestrians and motorized traffic. Safety Science 41 (627-640).

Sebert Kuhlmann, A.K., Brett, J., Thomas, D., & Sain, S.R. (2009). Environmental characteristics associated with pedestrian-motor vehicle collisions in Denver, Colorado. Am J Public Health 99: 1632–1637.

Transport Canada (2007) Canadian Motor Vehicle Traffic Collision Statistics: 2006. Report No. TP3322. Transport Canada, Ottawa, ON.

Tester, J.M., Rutherford, G.W., Wald, Z., & Rutherford, M.W. (2004). A matched case-control study evaluating the effectiveness of speed humps in reducing child pedestrian injuries. Am J Public Health, 94: 646-650.

Vuchic, V. R. 1999. Transportation for livable cities. Center for Urban Policy Research Press, Rutgers University, Brunswick, N.J.

Wier, M., J. Weintraub, et al. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accid Anal Prev 41 (1): 137-145.

Winkelmann, R. (2003) Econometric Analysis of Count Data. Springer, Germany.

Woodcock, J. Banister D., Edwards P, Prentice A. M., & Roberts I. (2007). Energy and transport. Lancet, 370(9592):1078–1088.

Zegeer, C.V., Seiderman, C., et al. (2002). Pedestrian facilities users guide: Providing safety and mobility. No. FHWA-RD-01-102, McLean: Federal Highway Administration.

Figure 1. Conceptual framework - BE, risk exposure and pedestrian safety at intersections

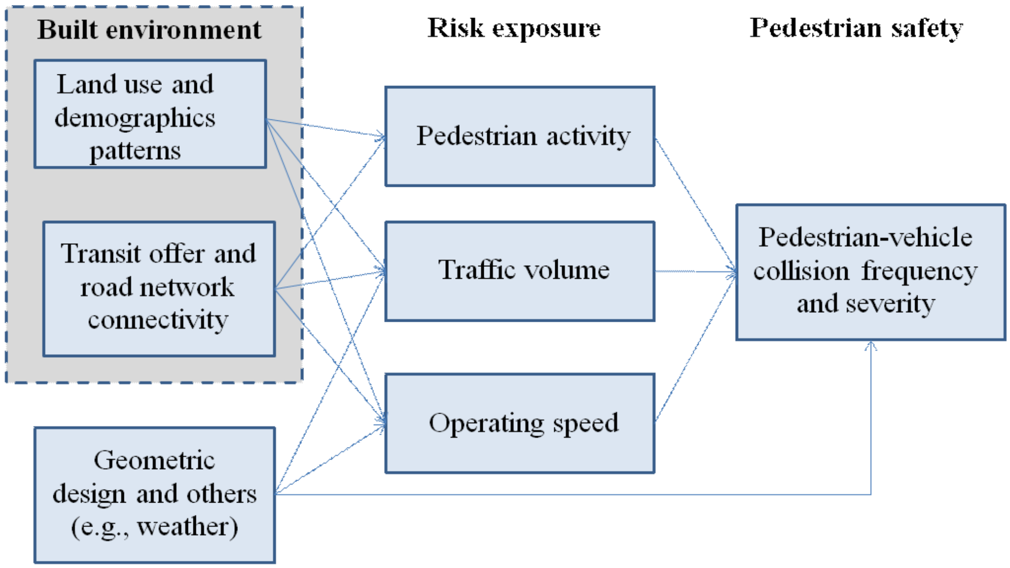

Figure 2. Signalized intersections in the study area

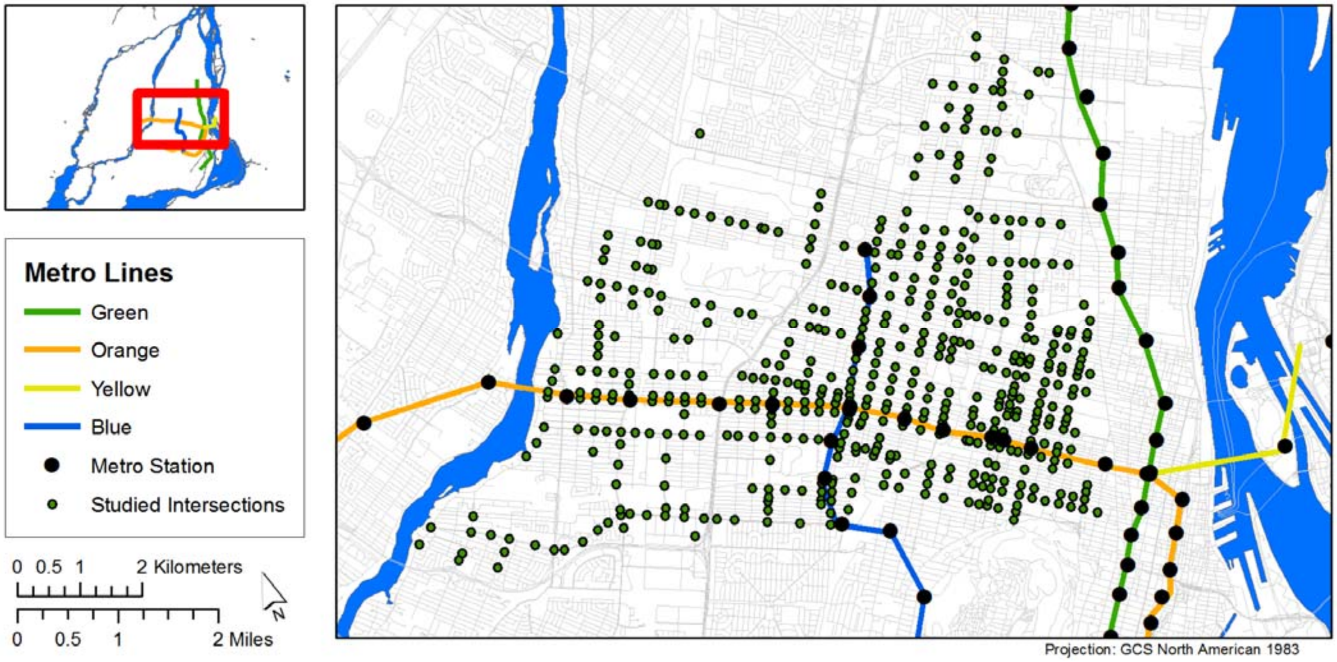

Figure 3. Effect of an AADT reduction on the expected risk of pedestrian collision

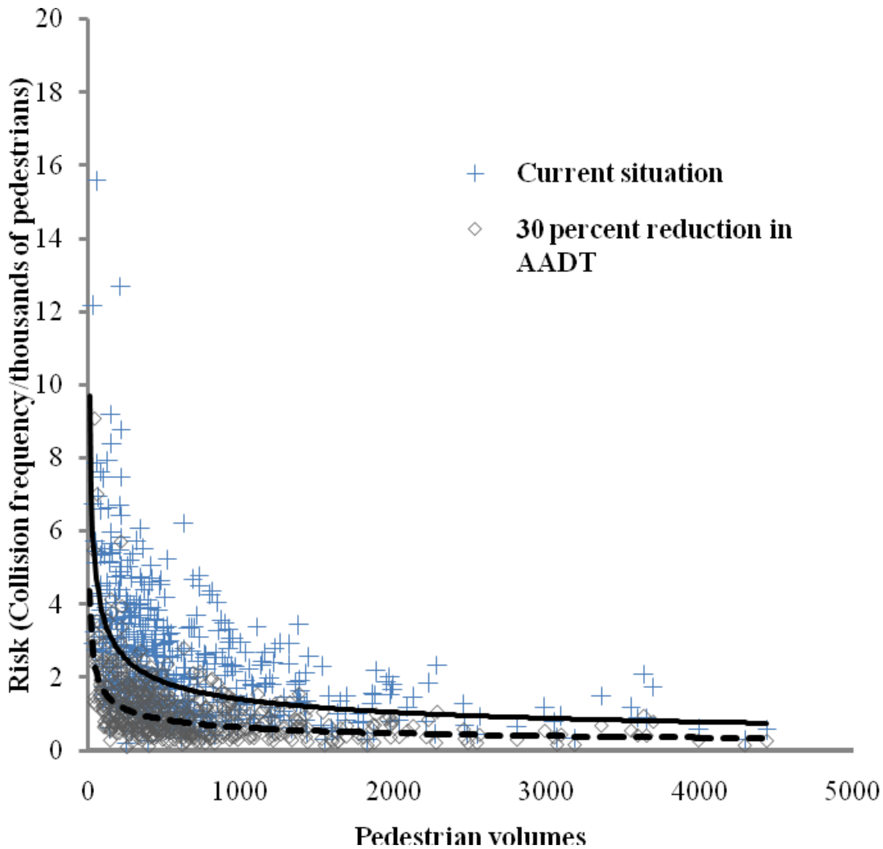

Table 1: Summary statistics of pedestrian, traffic and collision frequency data

Variable	Mean	Std. Dev.	Min	Max
Pedestrian-vehicle collisions	1.2	1.8	0	11
3-hour pedestrian volume	700.8	702.3	10	3,988
AADT	26993.4	13152.5	3236.6	78305.9

Table 2: Built environment variables and some statistics for the 150m buffer

	Variable	Units	Mean+	Std dev.
Land Use (*L*)	Commercial	m^2	4,461.4	5,848.8
	Residential	m^2	41,538.1	15,794.9
	Industrial	m^2	11,359.9	13,510.6
	Parks/Recreational	m^2	4,119.2	8,492.8
	Open Space	m^2	2,128.7	4,815.4
	Government/Institutional	m^2	7,078.6	9,039.4
	Number of jobs	counts	284.5	340.4
	Number of schools	Counts	0.2	0.5
Demographic (*D*)	Population	counts	622.8	270.7
	Workers	counts	389.3	187.9
	Children	counts	122.7	65.6
	Seniors	counts	87.7	40.3
Transit characteristics (*T*)	Presence of metro station	0/1 indicator	0.1	0.3
	Number of bus Stops	Counts	3.4	2.1
	Km of bus route	km	1,047.5	682.3
Road network connectivity (*R*)	Road Length (km)	km	1,282.6	255.2
	Class 1 – Primary Highway	km	12.3	87.0
	Class 2 – Secondary Highway	km	46.6	134.9
	Class 3 – Arterial Road	km	299.0	293.5
	Class 4 – Local Road	km	911.5	321.4
	Number of street segments	Counts	5.7	1.6
	Number of intersections	Counts	6.3	3.2
	Portion of major roads	Sum of class 1 to Class 4 divided by road length	0.3	0.2
	Average Street Length	Road length divided by number of streets	4.5	0.9

+These values are only for the 150 m buffer

Table 3: Correlations between BE and pedestrian activity, traffic and collision frequency

	50 m buffer			150 m buffer			400 m buffer			600 m buffer		
Variables	**A**	**ln(P)**	**ADT**	**A**	**ln(P)**	**ADT**	**A**	**ln(P)**	**ADT**	**A**	**ln(P)**	**ADT**
Pedestrian collisions	1.00			1.00			1.00			1.00		
Log of pedestrian volume	0.36	1.00		0.36	1.00		0.36	1.00		0.36	1.00	
Pedestrian volume	0.37	0.84		0.37	0.84		0.37	0.84		0.37	0.84	
AADT	0.40	0.11	1.00	0.40	0.11	1.00	0.40	0.11	1.00	0.40	0.11	1.00
Commercial	0.32	0.49	0.06	0.24	0.42	0.05	0.11	0.36	0.02	0.07	0.35	-0.02
Residential	-0.07	-0.14	-0.20	0.05	0.01	-0.16	0.10	0.16	-0.13	0.10	0.16	-0.10
Open space	-0.03	-0.23	0.52	-0.07	-0.27	0.43	-0.10	-0.34	0.28	-0.05	-0.27	0.24
Parks/recreational	-0.09	-0.13	-0.06	-0.10	-0.15	-0.03	-0.13	-0.16	-0.06	-0.09	-0.12	-0.07
Industrial	-0.04	-0.08	0.07	-0.02	-0.05	0.09	-0.01	-0.06	0.12	-0.02	-0.03	0.11
Number of schools	0.00	0.00	-0.03	-0.08	0.12	-0.01	0.09	0.31	-0.02	0.09	0.27	-0.03
Number of jobs	0.16	0.26	0.06	0.17	0.33	0.17	0.10	0.36	0.08	0.02	0.20	0.09
Population	0.10	0.38	-0.17	0.12	0.42	-0.17	0.15	0.47	-0.15	0.16	0.47	-0.12
Employees	0.06	0.39	-0.18	0.08	0.42	-0.18	0.11	0.46	-0.15	0.11	0.45	-0.13
Seniors	0.10	0.19	-0.13	0.12	0.20	-0.13	0.18	0.23	-0.07	0.19	0.20	-0.05
Children	0.09	0.24	-0.13	0.12	0.28	-0.11	0.20	0.34	-0.06	0.22	0.35	-0.03
Presence of metro station	0.07	0.17	-0.01	0.20	0.32	0.14	0.08	0.32	0.05	0.04	0.24	0.04
Number of bus stops	0.42	0.23	0.56	0.31	0.37	0.36	0.08	0.34	0.00	0.02	0.32	-0.07
Km of bus routes	0.32	0.29	0.49	0.26	0.30	0.38	0.00	0.18	0.05	-0.08	0.08	0.07
Road Length (km)	0.06	-0.08	0.48	0.08	0.14	0.18	0.07	0.27	-0.01	0.07	0.31	-0.02
Class 1 – Primary Highway	-0.02	-0.15	0.31	-0.02	-0.16	0.31	-0.08	-0.20	0.08	-0.10	-0.21	-0.02
Class 2 – Secondary Highway	0.10	0.07	0.12	0.05	0.10	0.06	0.01	0.17	-0.04	0.01	0.19	0.00
Class 3 – Arterial Road	0.22	0.01	0.52	0.17	-0.01	0.49	0.01	-0.08	0.21	0.01	-0.02	0.09
Class 4 – Local Road	-0.26	-0.03	-0.49	-0.10	0.15	-0.43	0.10	0.38	-0.20	0.11	0.39	-0.14
Number of street segments	-0.04	-0.07	0.05	-0.01	0.02	-0.04	0.00	0.03	0.06	-0.02	0.06	0.08
Number of intersections	-0.02	-0.12	0.33	-0.02	-0.01	0.12	-0.04	0.04	0.03	-0.02	0.12	0.03
Portion of major roads	0.29	0.03	0.60	0.17	-0.06	0.57	-0.06	-0.22	0.25	-0.10	-0.22	0.17
Average street length	0.11	0.00	0.43	-0.10	-0.13	-0.23	-0.07	-0.24	0.06	-0.07	-0.18	0.11
Intersection type	0.16	0.30	-0.05	0.16	0.30	-0.05	0.16	0.30	-0.05	0.16	0.30	-0.05

Table 4. Log-liner model for pedestrian activity

Variables	**Buffer**	**Parameter estimates**		**Elasticities**
Intercept		4.08	***	
Population (in 1000's)	400 m	0.08	**	0.34
Commercial (in 1000's)	50 m	0.19	***	0.20
No. of jobs (in 1000's)	400 m	0.17	***	0.28
No. of schools	400 m	0.15	***	0.20
Metro station	400 m	0.33	***	0.28
No. of bus stops	150 m	0.11	***	0.37
% of major arterials	400 m	-0.71	**	-0.16
Average street length	400 m	0.95	**	0.46
Intersection indicator[+]		0.36	***	0.31
Goodness-of-fit	$R^2 = 0.55$			

[+] 4-legged intersection = 1 and 3-legged intersections = 0
**: Statistically significant at 5%,
***: Statistically significant at 1%

Table 5.a Pedestrian collision frequency model – Short model without BE variables

Variables	Parameter estimates			
	NB	GNB	Two-compound NB	
Mean			***Class 1***	***Class 2***
ln (AADT)	1.15 ***	1.12 ***	0.71 ***	2.43 ***
ln (pedestrian volume)	0.45 ***	0.46 ***	0.46 ***	0.46 ***
Intersection indicator	0.40 ***	0.41 ***	0.39 ***	0.39 ***
Intercept	-14.79 **	-14.59 **	-10.24 **	-28.33 ***
Alpha				
Constant	0.41 ***	9.76 ***	0.51 **	0.00
ln (AADT)	-	-1.03 **	-	-
Elasticities				
ln (AADT)	1.15	1.12	0.71	2.43
ln (pedestrian volume)	0.45	0.46	0.46	0.46
Intersection indicator	0.33	0.34	0.32	0.32
AIC	*1406.9*	*1413.1*	*1408.6*	

++ 4-legged intersection = 1 and 3-legged intersection = 0
***: Statistically significant at 1%,
**: Statistically significant at 5%,
*: Statistically significant at 10%

Table 5.b Pedestrian collision frequency model - Extended model with BE variables

Variables	Parameter estimates							
	NB		GNB		Two-compound NB			
					Class 1		*Class 2*	
Mean								
ln (AADT)	0.90	***	0.90	***	0.51	***	1.99	***
ln (pedestrian volume)	0.26	***	0.28	***	0.26	***	0.26	***
Intersection indicator	0.43	***	0.43	***	0.43	***	0.43	***
Commercial (50 m buffer)$^{\&}$	0.14	**	0.14	**	0.14	**	0.14	**
No. of bus stops (50 m buffer)	0.16	***	0.15	***	0.16	***	0.16	***
No. of schools (150 m buffer)	-0.34	**	-0.23	***	-0.25	***	-0.25	***
Intercept	-11.56	***	-11.66	***	-7.40	***	-23.08	***
Alpha								
Constant	0.32	**	8.44	***	0.39	**	0.00	
ln (AADT)	-		-0.92	**	-		-	
Elasticities								
ln (AADT)	0.91		0.90		0.51		1.99	
ln (pedestrian volume)	0.26		0.28		0.26		0.26	
Intersection indicator	0.35		0.35		0.35		0.35	
Commercial	0.15		0.15		0.15		0.15	
No. bus stops	0.30		0.28		0.29		0.29	
No. schools	-0.08		-0.05		-0.06		-0.06	
AIC	*1381.8*		*1381.4*		*1380.8*			

$^{++}$ 4-legged intersection = 1 and 3-legged intersection = 0, $^{\&}$ in 100's

***: Statistically significant at 1%,

**: Statistically significant at 5%,

*: Statistically significant at 10%

Table 6. Validation model results

Model	Normalized root mean squared deviation (NRMSD)			
	Mean*	Median	Min	Max
Pedestrian activity	0.14	0.140	0.08	0.20
Collision frequency with estimated pedestrian volume	0.18	0.18	0.13	0.24
Collision frequency with observed pedestrian volume	0.17	0.17	0.13	0.23
Model	**Pearson Correlation**			
	Mean	**Median**	**Min**	**Max**
Pedestrian activity	0.72	0.73	0.54	0.88
Collision frequency with estimated pedestrian volume	0.56	0.57	0.36	0.77
Collision frequency with observed pedestrian volume	0.58	0.58	0.42	0.76

*Mean and median values are based on 100 validation random sets

Box, P. (2004). "Curb-parking problems:Overview." Journal of Transportation Engineering **130**.

Elvik, R. (2009). "The non-linearity of risk and the promotion of environmentally sustainable transport." Accident Analysis and Prevention **41**(4): 849-855.

Laplante, J. and B. McCann (2008). "Complete streets: We can get there from here." Institute of Transportation Engineers. ITE Journal: 24-28.

Leden, L. (2002). "Pedestrian risk decrease with pedestrian flow: A case study based on data from signalized intersections in Hamilton, Ontario." Accident Analysis and Prevention **34**: 457-464.

Lyon, C. and B. Persaud (2003). "Pedestrian collision prediction models for urban intersections." Transportation Research Record **1818**: 102-107.

Morency, P. and M. Cloutier (2006). "From targeted "black spots" to area-wide pedestrian safety." Injury Prevention **16**(6): 360-364.

Pulugurtha, S. S. and S. R. Repake (2008). "Assessment of models to measure pedestrian activity at signalized intersections." Transportation Research Record **2073**: 39-48.

Schneider, R., L. Arnold, et al. (2009). A Pilot Model for Estimating Pedestrian Intersection Crossing Volumes. Annual Meeting of the Transportation Research Board. washington DC.