Accident Analysis and Prevention 50 (2013) 304–311

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

A spatial and temporal analysis of child pedestrian crashes in Santiago, Chile

Carola A. Blazquez*, Marcela S. Celis

Universidad Andres Bello, Department of Engineering Science, Sazie 2315, Piso 2, Santiago, Chile

ARTICLE INFO

Article history:
Received 20 January 2012
Received in revised form 19 April 2012
Accepted 1 May 2012

Keywords:
Kernel
Spatial autocorrelation
Crashes
Child pedestrian

ABSTRACT

This paper presents a spatial and temporal analysis of child pedestrian crash data in Santiago, Chile during the period 2000–2008. First, this study identified seven critical areas with high child pedestrian crash risk employing kernel density estimation, and subsequently, statistically significant clusters of the main attributes associated to these crashes in each critical area were determined in a geographic information systems environment. Moran's *I* index test identified a positive spatial autocorrelation on crash contributing factors, time of day, straight road sections and intersections, and roads without traffic signs within the critical areas during the studied period, whereas a random spatial pattern was identified for crashes related to the age attribute. No statistical significance in the spatial relationship was obtained in child pedestrian crashes with respect to gender, weekday, and month of the year. The results from this research aid in determining the areas in which enhanced school-age child pedestrian safety is required by developing and implementing effective enforcement, educational, and engineering preventive measures.

1. Introduction

The World Health Organization (WHO) and UNICEF published a report on child injury prevention in 2008, which statistically indicates that child death involved in road traffic crashes corresponds to 2% of total annual fatalities worldwide. In 2004, approximately 262,000 children were killed in pedestrian crashes representing 30% of the total road traffic crash casualties. Currently, crashes are ranked ninth globally among the leading causes of disability, and the ranking is projected to rise to third by 2020 (WHO, 2008). Furthermore, this report predicts that crashes will be the main cause of death in children under the age of 18 by 2030.

Road traffic crashes among children are one of the most serious epidemic problems in developing countries due to rapid motorization and other factors (Nantulya and Reich, 2002). In Chile, traumatism derived from pedestrian vehicle crashes is the leading cause of death among children over one year of age. Although, statistics indicate that the total number of injured children has diminished in recent years, this value is still high when compared to developed countries producing significant social and economic consequences to the families, nation, and health organizations (Romero, 2007). On average, approximately 15% of the total number of crashes that occur in Chile each year are pedestrian crashes, and over 33% of these crashes involve children, and 40% of whom suffer serious and less serious injuries or death. Thus, there is a need for a spatial and temporal analysis of child pedestrian crashes in Chile, and subsequent effective measures to prevent these crashes.

Various studies have used Geographic Information Systems (GIS) as a tool for data management and spatial analysis (e.g., patterns and clusterization) of road traffic crashes, which some combine with statistical models to determine the relationship between crashes and causal factors (Levine et al., 1995; Ladron de Guevara et al., 2004; Ng et al., 2002; Flahaut et al., 2003; Meliker et al., 2004; Erdogan et al., 2008; Erdogan, 2009; Gundogdu, 2010; Anderson, 2009; Steenberghen et al., 2004). Additional studies have addressed the spatial problem of pedestrian crashes (Schneider et al., 2004; Kim and Yamashita, 2005; Truong and Somenahalli, 2011), and other researches have focused on pedestrian crashes that involve children. For example, some studies have employed GIS technology to locate zones with high child pedestrian severity and risk near schools (Banos and Huguenin-Richard, 2000; Austin et al., 1997; Miller, 2000; Clifton and Kreamer-Fults, 2007), or within high density residential areas (Lightstone et al., 2001; Dissanayake et al., 2009).

The aforementioned studies focus on child pedestrian crashes that occurred in urban zones of Europe and North America. No previous research has analyzed pedestrian crashes (particularly with the involvement of children) both spatial and temporally, and that portrays the Chilean reality. Therefore, this paper presents, to the authors' knowledge, the first results ever published for child pedestrian crashes in Chile using KDE and Moran's *I* spatial autocorrelation statistics in a GIS environment over a period of time. The results of this study may be utilized as a supportive tool for decision making and appropriate allocation of resources

* Corresponding author. Tel.: +56 2 661 8644; fax: +56 2 770 3076.
E-mail address: cblazquez@unab.cl (C.A. Blazquez).

0001-4575/$ – see front matter
http://dx.doi.org/10.1016/j.aap.2012.05.001

for safety enhancements in road infrastructure, and pedestrian education offered to children and parents to increase awareness and usage of warning signs and crossing signals.

The first objective of this study is to identify critical areas with a high likelihood of child pedestrian crash occurrence using kernel density estimation (KDE) in a GIS environment. Many studies have employed the KDE technique to analyze road traffic crashes due to its simple implementation and easy understanding (Banos and Huguenin-Richard, 2000; Flahaut et al., 2003; Steenberghen et al., 2004; Schneider et al., 2004; Pulugurtha et al., 2007; Erdogan et al., 2008; Xie and Yan, 2008; Anderson, 2009; Kuo et al., 2011). The second objective of this study is to perform a spatial correlation test using Moran's *I* index to assess whether the pattern of child pedestrian crashes had an average tendency to cluster in space and time within each of the identified critical areas, and whether a spatial dependence of these patterns exists relative to the main crash attributes (e.g., crash contributing factors, age, gender, time of day, month, road type, and traffic signs).

The methodology presented in this paper was implemented with nine years of child pedestrian crash data for the city of Santiago, Chile. Santiago is the capital of Chile with over 900 km^2 of extension and an estimated population of 6.5 million inhabitants in 2005, representing approximately 40% of the total population of Chile.

2. Methodology

2.1. Identification of critical areas

KDE is a non-parametric method that involves introducing a symmetrical surface over each point feature, assessing the distance from the point to a reference location based on a mathematical function, and subsequently, adding the value of all the surfaces for that reference location (Levine, 2004). Eq. (1) defines KDE for a given set of observations with an unknown probability density function *f*:

$$f(x) = \left(\frac{1}{nh}\right) \sum_{i=1}^{n} \frac{K(x - x_i)}{h} \quad (1)$$

where x_i is the value of the variable *X* at the location *i*, *n* is the total number of locations, *h* is the bandwidth or smoothing parameter, and *K* is the kernel function. This research employed a normal distribution as a kernel function that weighs all points in the study area with near points having more weight than distant points.

The outcome of the KDE depends significantly on the bandwidth and the cell size. According to Steenberghen et al. (2004), the cell size that provides the best concentration patterns of crashes in a dense urban network (such as Santiago) is smaller than the road segments between intersections and large enough to identify crash clustering. Therefore, after a number of tests, a bandwidth or search radius of 1000 m and cell size of 100 m by 100 m was selected for the crash density analysis for the studied period.

2.2. Spatial autocorrelation

Subsequent to the identification of spatial clusters with high child pedestrian crashes over time, a spatial autocorrelation test was performed to determine spatial dependence between the main attributes of these crashes and to detect statistically significant clusters with respect to each of these attributes. If the level of correlation is higher than expected, then neighboring locations have similar values and the spatial autocorrelation is positive. When the level of correlation is lower than expected, high values of the variable are contiguous to low values and the autocorrelation is negative. Thus, the spatial patterns of crash data take into account simultaneously crash locations and their attribute values by measuring for attribute similarity and location proximity into one single index named Moran's *I* (Truong and Somenahalli, 2011). This index is formally expressed by Eq. (2):

$$I = \frac{N\sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^{N}(x_i - \bar{x})^2} \quad \forall i = 1, \ldots, n \wedge \forall j = 1, \ldots, n \quad (2)$$

where w_{ij} are the elements of a spatial binary contiguity matrix with weights representing proximity relationships between location *i* and neighboring location *j*, S_0 is the summation of all elements w_{ij}, x_i is the variable value at a particular location *i*, x_j is the variable value at a another location ($i \neq j$), $\bar{x}$ is the mean of the variable, and *n* is the total number of locations. The values of Moran's *I* index range from −1 to 1, where the former represents a strong negative autocorrelation (i.e., perfect dispersion or clusterization of dissimilar values) and the latter represents a strong positive autocorrelation (i.e., perfect concentration or clusterization of similar values). A value of Moran's *I* near zero indicates a spatially random pattern.

The results from the spatial autocorrelation are always interpreted within the context of its null hypothesis, which states that the attribute being analyzed is randomly distributed among the features in the study area. The statistical significance of the Moran's *I* index is computed with the *Z* score method assuming a standard normal distribution with mean equal to zero and a variance of one. A positive *Z* score for a point indicates that the neighboring features have similar values, whereas a negative *Z* score denotes that the feature is surrounded by dissimilar values (Getis and Ord, 1994).

3. Data description

This study employed the data provided by the Chilean National Road Safety Commission (CONASET), which maintains a database with all road traffic crashes that occur in the country. Of 6078 child pedestrian crashes that arose in Santiago between 2000 and 2008, 5152 (84.8%) were successfully geocoded in a GIS environment, as shown in Fig. 1. This research focuses on pedestrian crashes that involve children between the ages of 5 and 18 and that normally attend school. These school-age children are from now on referred to as "children".

Interactive manual intervention was required to match the crash address information with the spatial database. The remaining percentage of crash locations was not geocoded due to non-existent or incomplete addresses. Currently, police officers fill out a paper form at the scene to report crashes with their main characteristics (e.g., time, date, address, weather, road condition, etc.). Subsequently, this information is manually entered to a computer. Hence, the reported crash data is prone to yield discrepancies in both manual data entry processes.

CONASET classifies road crashes into 42 main contributing factors. This study employed 14 of these factors, which account for 96% of all pedestrian crashes. Table 1 presents the percentages according to injury type for each of the 14 factors, of which 67.2% was the responsibility of the pedestrian, 26.3% was the driver's responsibility, and 6.6% were undetermined causes. Approximately 75% of all child pedestrian crashes were produced by the following factors: Pedestrian crosses road surprisingly or carelessly; imprudence of pedestrian; and driver violates crosswalk. In terms of injury type, over 60% of all children involved in a pedestrian crash during the studied period suffered slight injuries, while only 0.73% of the children were unharmed. In other words, 99.27% of all child pedestrian crashes yielded some degree of physical injury. In addition, 55.8% of child pedestrian crashes that occurred in Santiago during the studied period were boys and 58.1% of all child pedestrian crashes