ARTICLE IN PRESS

PERGAMON

Accident Analysis and Prevention xxx (2003) xxx–xxx

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

An accident waiting to happen: a spatial approach to proactive pedestrian planning

Robert J. Schneider*, Rhonda M. Ryznar[1], Asad J. Khattak[1]

Department of City and Regional Planning, University of North Carolina at Chapel Hill, CB 3140 New East Bldg., Chapel Hill, NC 27599-3140, USA

Received 14 May 2001; received in revised form 21 January 2002; accepted 13 February 2002

Abstract

There are about 75,000 pedestrian crashes in the United States each year. Approximately 5000 of these crashes are fatal, accounting for 12% of all roadway deaths. On college campuses, pedestrian exposure and crash-risk can be quite high. Therefore, we analyzed pedestrian crashes on the campus of the University of North Carolina at Chapel Hill (UNC) as a test case for our spatially-oriented prototype tool that combines perceived-risk (survey) data with police-reported crash data to obtain a more complete picture of pedestrian crash-risk. We use spatial analysis techniques combined with regression models to understand factors associated with risk. The spatial analysis is based on comparing two distributions, i.e. the locations of perceived-risk with police-reported crash locations. The differences between the two distributions are statistically significant, implying that certain locations on campus are perceived as dangerous, though pedestrian crashes have not yet occurred there, and there are actual locations of police-reported crashes that are not perceived to be dangerous by pedestrians or drivers. Furthermore, we estimate negative binomial regression models to combine pedestrian and automobile exposure with roadway characteristics and spatial/land use information. The models show that high exposure, incomplete sidewalks and high crosswalk density are associated with greater observed and perceived pedestrian crash-risk. Additionally, we found that people perceive a lower risk near university libraries, stadiums, and academic buildings, despite the occurrence of crashes.

Keywords: Pedestrian; Safety; Survey; Spatial Analysis; Geographic Information Systems; Negative binomial model

1. Introduction

On average, a pedestrian injury occurs every 6 min and a pedestrian fatality occurs every 107 min in the United States (NHTSA, 2000). The 4906 pedestrian fatalities in 1999 accounted for approximately 12% of highway deaths that year (BTS, 1999). Yet, this fact may inspire little local action until a serious pedestrian injury or fatality occurs in a neighborhood, in a specific commercial area, or on a college campus. One of the 4906 pedestrian fatalities of 1999 occurred in November on the campus of the University of North Carolina at Chapel Hill (UNC) when a post-doc was struck by an automobile while trying to cross Manning Drive near the hospital. As a result, student rallies were held, a series of pedestrian safety articles ran in local papers, and a Pedestrian Safety Committee, with members from UNC, the Town of Chapel Hill, and the North Carolina Department of Transportation was formed. Yet, this response was reactive rather than proactive. Though no fatalities had occurred to focus attention on pedestrian safety during the previous 5 years, the 57 pedestrian crashes that had been reported on campus could have indicated safety problems. In addition, pedestrians and drivers who used the UNC Campus roadways each day could have suggested locations where pedestrian safety problems existed. To prevent future fatalities, *proactive* methods are needed to identify where pedestrian problems exist and what types of factors are related to pedestrian crash-risk on campuses, in neighborhoods, in commercial areas, and all other areas where people walk.

1.1. Objective

This study has two objectives. First, to determine if perception data can add important information for a proactive approach to crash avoidance. This is accomplished by first demonstrating that there is a difference between the locations

* Corresponding author. Present address: Toole Design Group, 535 Main Street, Suite 211, Laurel, MD 20707, USA. Tel.: +1-301-362-1600; fax: +1-301-362-1780.
E-mail addresses: rschneider@tooledesign.com (R.J. Schneider), rhonda.ryznar@tufts.edu (R.M. Ryznar), khattak@email.unc.edu (A.J. Khattak).
[1] Tel.: +1-919-962-4760; fax: +1-919-962-5206.

0001-4575/03/$ – see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0001-4575(02)00149-5

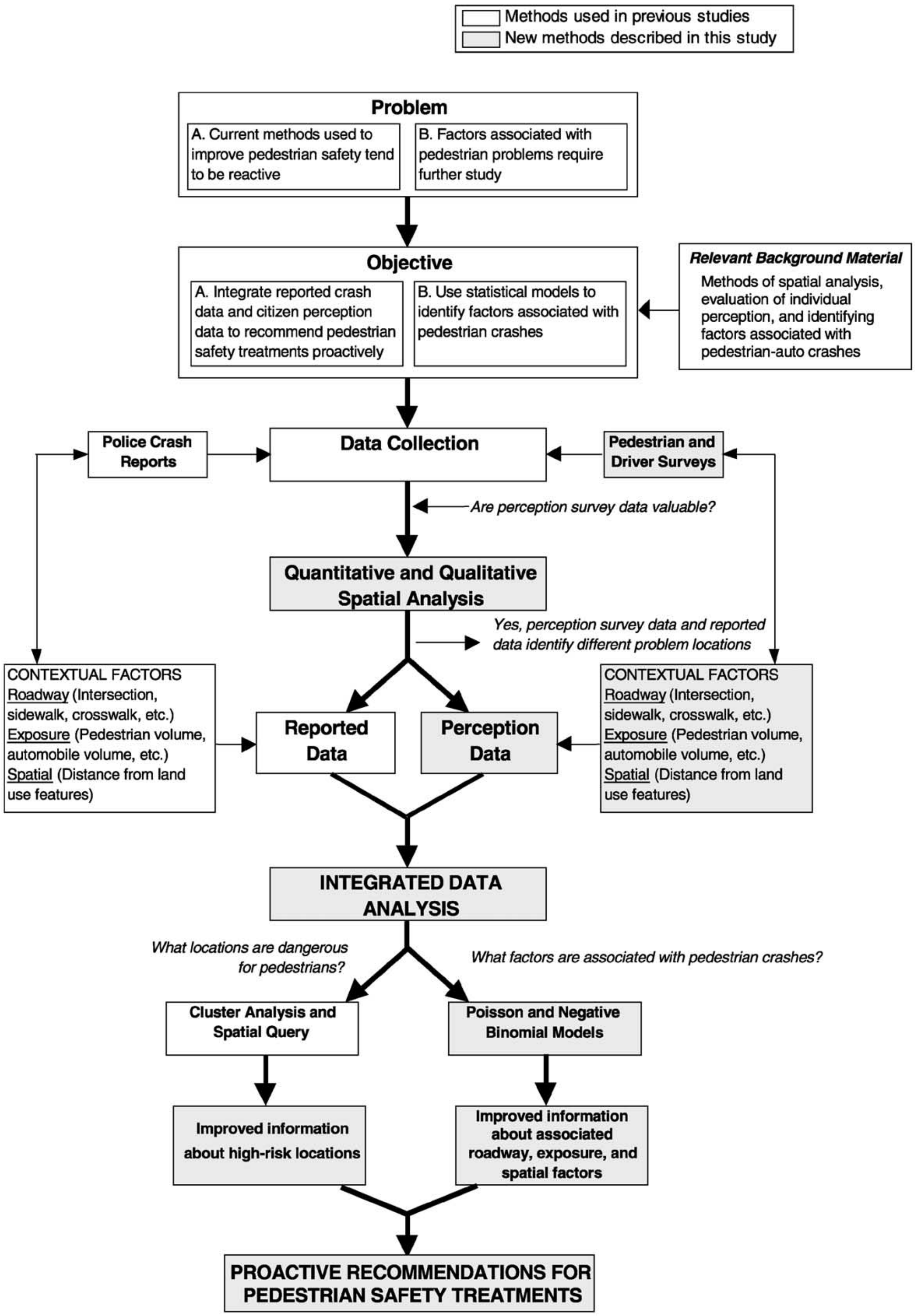

Fig. 1. Outline of method.

of perceived high-risk areas and the locations of actual reported crash data. If the two point patterns are from two different spatial distributions then we may safely assume that perception data could add some important information about the environmental factors associated with pedestrian-auto crashes, such as roadway features, exposure, etc. It is likely that the locations of both actual and perceived pedestrian risk depend on a combination of physical/environmental, as well as, individual factors. Physical/environmental factors include the presence of sidewalks, traffic, and roadway crossings, while individual factors include the ability to judge distance and speed, visual capabilities, and the physical ability to move quickly and change direction.

Our second objective is then to analyze the factors from both geographic distributions through a regression analysis. Our analysis focuses on several specific, measurable physical/environmental factors that can be changed through engineering, education, and enforcement policies to improve pedestrian conditions (prevent crashes and improve perceptions of safety). Therefore, the methodology outlined in this paper could be used by decision makers as a tool to select appropriate pedestrian crash countermeasures at key locations (Fig. 1). By including the factors from perceived high-risk locations we are taking a proactive approach that might avoid an accident "waiting to happen."

2. Previous safety research on perception and spatial analysis

A number of studies have used mapped data to analyze transportation safety problems, but relatively few have incorporated perception data and spatial statistics to show differences between actual and perceived-risk. The literature addressing these topics is reviewed below and summarized in Table 1.

Most studies use only police-reported crash data, yet many suggest that perception may also affect pedestrian crash-risk (Schneider et al., 2001; Karim, 1992; Duncan et al., 1999; Austin et al., 1995; Butchart et al., 2000; Landis et al., 2001). Butchart et al. use perception to examine the implications for crash prevention. They find that people believe inadequate signage, inadequate traffic lights, and alcohol involvement to be associated with pedestrian crashes and that increased enforcement and traffic calming can help prevent these crashes. Landis et al. ask pedestrians to rate their level of comfort on roadway segments to develop a Pedestrian Level of Service Model. They find that pedestrians feel greater comfort when a sidewalk is present, when the sidewalk is wider and further from traffic, when on-street parking or a line of trees is present, and when there are lower vehicle speeds and volumes.

Karim finds that campus road users are able to rank specific locations correctly according to crash frequency. Yet, perceiving high-risk areas may not be so easy or reliable. Austin, Tight, and Kirby report that parents of school children misjudge locations with high crash-risk along their children's walking routes to school. This finding is supported by Duncan, Khattak, and Hughes who showed that distinguishing between crash and non-crash sites involving "walking along roadway" crashes is difficult, even for safety experts. In a pre-cursor study, Schneider et al. found mixed results. Four clusters of police-reported pedestrian crashes were identified on the UNC Campus. Campus pedestrians and drivers perceived only two of these locations as dangerous, while they believed that there was high pedestrian crash-risk at two additional locations where no crashes had been reported recently.

Discrepancies among these studies could be the result of Karim giving survey participants a finite set of locations to choose from, Austin, Tight, and Kirby analyzing dangerous locations from open-ended responses, and Schneider et al. allowing survey participants to mark any location on a study area map.

The relationship between risk taking and driving behavior has also been explored (Näätänen and Summala, 1976; Wilde, 1994; Burns and Wilde, 1995; Hino et al., 1996; Summala, 1996; Ward and Wilde, 1996; Wilde et al., 1998). Much of this mildly research supports Risk Homeostasis Theory, which states that people try to maintain a target level of risk that maximizes the difference between the perceived benefits and perceived costs of a behavioral choice. For example, when lateral sight distances were increased at a railroad crossing, drivers perceived less risk and therefore traveled at faster speeds approaching the crossing (Ward and Wilde, 1996).

To further explore the relationship between actual and perceived-risk of pedestrian crashes while incorporating environmental effects, this study also draws on new technologies and methods of spatial analysis. A variety of studies in recent years have addressed the issue of safety analysis using Geographic Information Systems (GIS). Specific uses of GIS range from simple crash plotting, or geocoding, crash locations (Levine et al., 1995; Kim et al., 1995; Hank Mohle and Associates, 1996; Peled et al., 1996; Chu et al., 1999; Miller, 2000; NCCGIA, 2000), to spatial queries that identify the locations of broadside, head-on, and rear end auto crashes (Hank Mohle and Associates, 1996) or that show a geocoded accident database can generate separate plots for bicycle crashes and pedestrian crashes (NCCGIA, 2000), to using GIS to identify spatial clusters of crashes through kernel density estimation (Levine et al., 1995; Kim et al., 1995; NCCGIA, 2000; Schneider et al., 2001).

Levine et al. compare the distribution of crashes at different times during the day and week, with different levels of injury severity, and with and without alcohol involvement, and take spatial analysis one step further by generating standard deviational ellipses showing the spatial distribution of crashes with different characteristics. The mapping of these ellipses allows the difference between the point pattern of one crash type and the point pattern of another crash type to be tested for statistical significance.

Table 1
Summary of literature

Author(s)	Overall findings	Methodology used	Key findings/recommendations
Studies using GIS spatial analysis			
Braddock et al. (1994)	GIS can provide information that is useful for understanding child pedestrian crashes	Crash plotting; cluster analysis; query by date and time of crash, environmental conditions, pedestrian/driver age and sex, injury severity ($N = 358$)	107 of 358 crash sites (30%) were located within two major clusters; Crashes in one cluster tended to involve children that were older and more likely to be seriously or fatally injured
Chu et al. (1999)	GIS can be designed to identify and rank crash locations for the most effective use of safety funds	Identification of high accident locations through clustering techniques; query by crash type, injury severity, time of day	GIS is a tool that can increase productivity in data collection, integrate crash data with roadway and traffic data, and display results graphically to enhance agencies' decision-making capabilities
Hank Mohle and Associates (1996)	GIS can generate maps of automobile crash rates and help analyze crash types	Segment analysis by crash rate; query by crash type, time of day, injury severity, ped/bike involvement	Several streets in San Joaquin County, CA were identified as having a higher risk of auto crashes
Kim et al. (1995)	GIS tools can help identify causes of crashes and characteristics of individuals involved	Moped crash plotting ($N = 240$)	Spatial patterns help identify intersections, roadways, and districts where improving signalization, widening roadways, or changing traffic speeds can reduce moped crashes
Levine et al. (1995)	Spatial analysis can help identify patterns of different crash types and different levels of injury severity	Crash plotting; spatial query by time of day, day of week, alcohol involvement, injury severity, number of vehicles involved, type of impact; calculation of nearest neighbor index and standard deviational ellipse ($N = 19{,}208$)	Employment density is a predictor of crash concentration; Residential locations and alcohol crashes are correlated; Crashes with injuries and fatalities were more widely-dispersed than other crashes
McMahon (1999)	GIS can be used to analyze pedestrian crash-risk	Comparison of walking along roadway crash and non-crash sites using binary logistic regression ($N = 141$); cluster analysis comparing 3 neighborhoods	Older neighborhoods and neighborhoods with more single parent families and unemployment were more likely to be crash sites
Miller (2000)	GIS is a useful tool for analyzing geographic context of crashes	Case study of applying GIS to analyze crash data in central Virginia and literature review	Techniques such as spatial query, segment analysis, buffering, and cluster identification can be used to improve crash analysis
NCCGIA (2000)	GIS methods can be used to make practical decisions	Route planning; cluster analysis to identify high ped/bike crash zones; query by ped/bike involvement	Large search radii can identify clusters over large zones while smaller radii can show crash clusters on road segments and intersections
Peled et al. (1996)	GIS is a useful tool for analyzing geographic context of crashes	Explanation of GIS database for analyzing automobile crashes within a locational context	Maps provide a clear impression of crash distribution and concentrations; Maps present neighboring streets and land uses along with crash locations
Schneider et al. (2001)	Perception data can be used to identify pedestrian problems and help select locations for proactive safety treatments	GIS analysis of pedestrian crash locations reported between 1994 and 1999 ($N = 57$) and locations perceived to be dangerous for pedestrians ($N = 1835$) by 110 drivers and 322 pedestrians on the UNC Campus	Reported pedestrian crashes were concentrated in four locations on campus, while only two of these locations were identified by survey participants; Drivers and pedestrians perceived a high pedestrian risk at two locations that were not identified through reported crash maps

Table 1 (*Continued*)

Author(s)	Overall findings	Methodology used	Key findings/recommendations
Studies incorporating perception			
Austin et al. (1995)	Parents do not perceive dangerous walking locations correctly	Simple analysis of parental survey that identified locations of safety concern on kids' routes to school (N = unreported)	Unsafe locations identified in parental comments did not correspond closely to reported accident locations
Butchart et al. (2000)	Perception of pedestrian injury risk can be used for crash prevention	Survey of households in six neighborhoods in a low income area of Johannesburg, South Africa ($N = 1075$)	Inadequate signage, traffic lights + alcohol involvement were perceived as pedestrian risk factors; Increased enforcement + traffic calming were perceived as preventative measures
Duncan et al. (1999)	Perception of safety is difficult, even for experts	Delphi process in which five pedestrian safety experts attempted to identify crash sites from among those with and without crashes ($N = 147$)	Professionals had difficulty determining crash sites based only on visible physical roadway characteristics
Harrell (1991)	Expectations affect driver detection of pedestrians	Citation of 1985 study by Shinar	Drivers detected pedestrians at night more often when pedestrian encounters were expected
Karim (1992)	Road users are aware of accident-prone locations on campus	Simple statistical analysis of survey responses regarding accident-prone locations (N = unreported)	Road users ranked specific locations correctly in terms of crash frequency on survey
Landis et al. (2001)	The comfort level of pedestrians on roadway segments with different characteristics can be modeled to create a pedestrian level of service	Stepwise multivariate regression estimated effect of sidewalk presence, lateral separation of pedestrians from roadway traffic, driveway frequency, and vehicle mix, volume and speed on pedestrian comfort ($N = 74$ participants, $N = 1250$ observations, $N = 42$ directional roadway segments)	Pedestrian convenience and comfort increase when a sidewalk is present, when the sidewalk is wider and further from traffic, when on-street parking or a line of trees is present, and when there are lower vehicle speeds and volumes
Schneider et al. (2001)	Perception data can be used to identify pedestrian problems and help select locations for proactive safety treatments	GIS analysis of pedestrian crash locations reported between 1994 and 1999 ($N = 57$) and locations perceived to be dangerous for pedestrians ($N = 1835$) by 110 drivers and 322 pedestrians on the UNC Campus	Reported pedestrian crashes were concentrated in four locations on campus, while only two of these locations were identified by survey participants; Drivers and pedestrians perceived a high pedestrian risk at two locations that were not identified through reported crash maps

N: sample size.

2.1. Gaps in the literature

We found no studies that: (a) use quantitative spatial tests to identify clusters of pedestrian crashes; (b) perform inter-distributional spatial tests to show the difference between police-reported crash locations and locations perceived to have a high crash-risk; (c) estimate appropriate rigorous models to identify factors associated with pedestrian crash-risk; or (d) examine the effect of pedestrian exposure on pedestrian crash-risk.

This study attempts to fill some of these gaps by utilizing GIS and spatial analysis capabilities to analyze the spatial distribution of both actual and perceived pedestrian crash locations, and then to determine which environmental factors, including nearby land uses, may have an effect on their occurrence.

3. Description of study area and data collection

Our technique of integrating police-reported crash data with perception survey data was tested on the campus of the University of North Carolina at Chapel Hill, a 300-ha (740-acre) area that is home to over 23,000 students, 2800 faculty, and many other employees (Peterson's, 2000). The geographic distribution of land uses on the UNC Campus

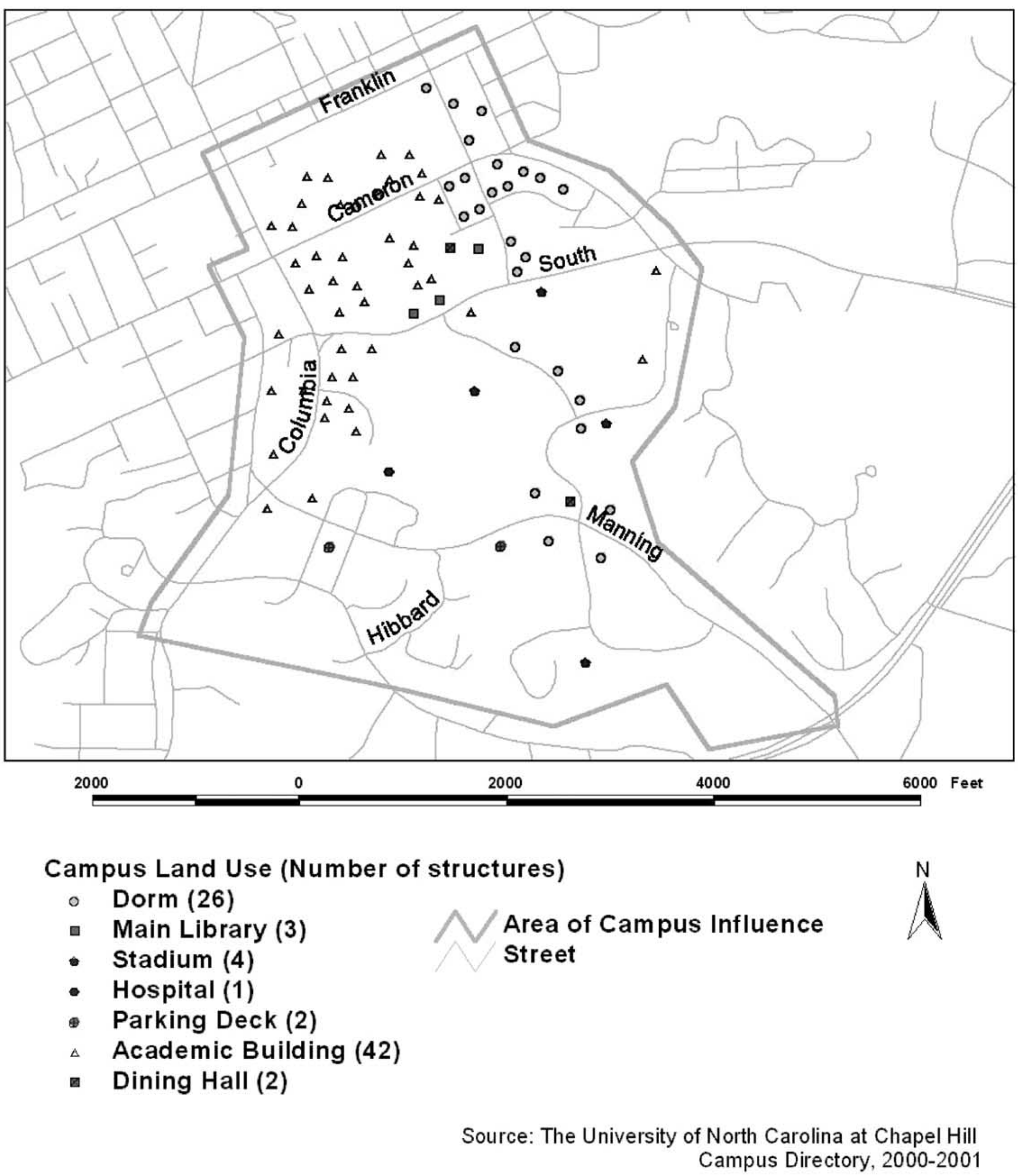

Fig. 2. UNC-Chapel Hill land use (January 2001).

makes the campus an excellent area to study the spatial relationships between pedestrian problems and pedestrian and vehicle flows, development character, and activity destinations. Like many other college campuses, the UNC Campus streets contain a mix of pedestrians, bicycles, and automobiles. Conflicts between pedestrians and vehicles occur as people travel to class, dining, personal recreation, sporting events, and other activities (Fig. 2).

The method developed in this study may improve safety for the 14.5 million students and 2.7 million faculty, staff, and other employees associated with colleges in the United States (more than 6% of the US population) and 90 million

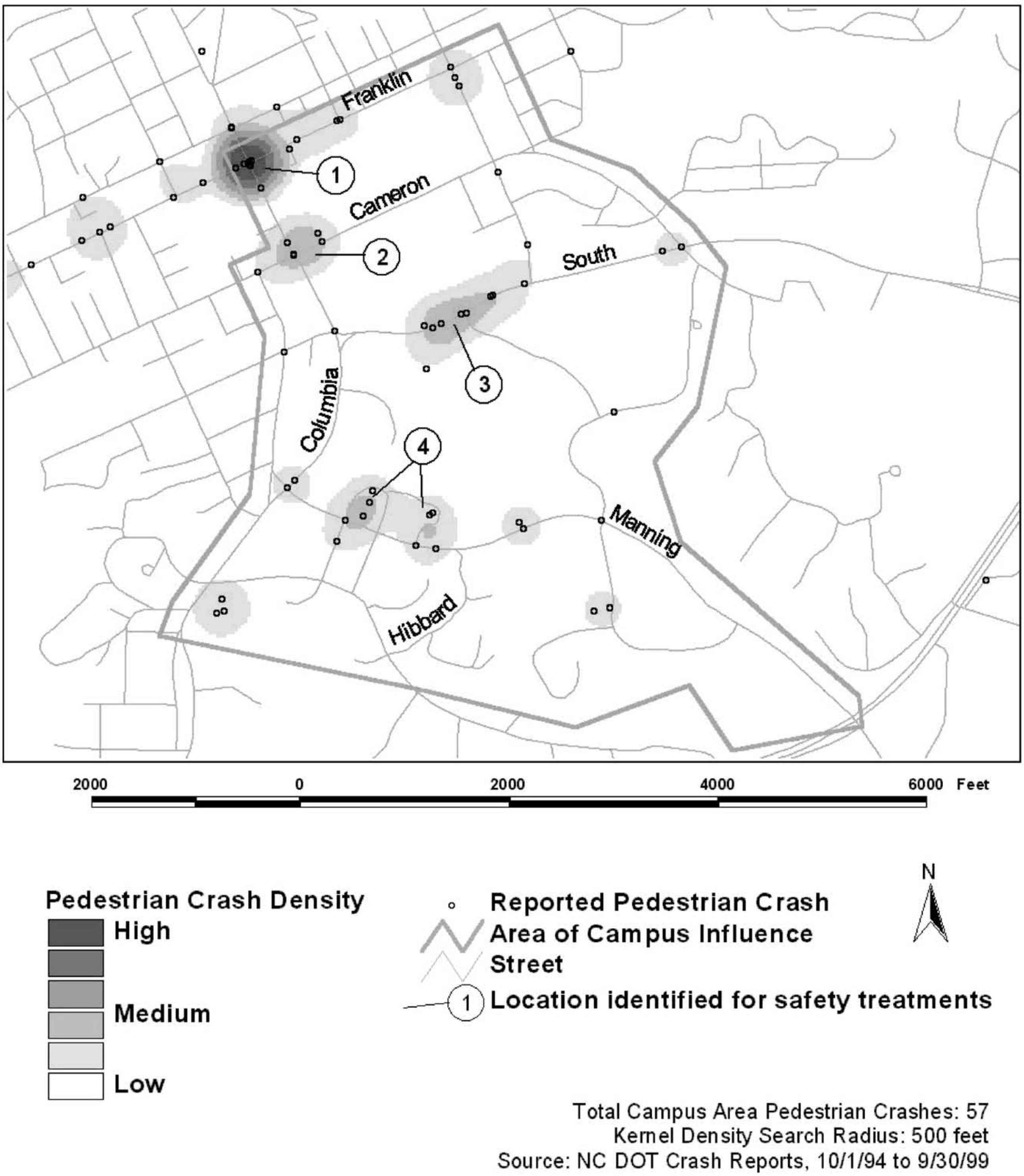

Fig. 3. Reported pedestrian crash density (UNC-Chapel Hill, 1994–1999).

college students and teachers worldwide, many of whom walk on campuses (US Department of Education, 2000). In addition, the method presented here can be applied to neighborhoods, commercial areas, and other locales with large numbers of pedestrians.

3.1. Spatial data

There were 57 police-reported collisions between vehicles and pedestrians on the UNC Campus between October 1994 and September 1999. Reports for each of these crashes were extracted from the state crash database and the crash locations were geocoded on a GIS map with an accuracy of better than 30 m (100 ft) (this was the level of measurement accuracy used in the police reports). Using *CrimeStat*, a spatial statistics software package (Levine, 1999), the data revealed four main clusters of crashes. Fig. 3 shows these clusters and the density of crashes associated with them as determined through kernel estimation.

A perception survey was also designed to gather additional data about locations where pedestrian safety problems could exist on campus. Four hundred fifty pedestrian surveys were mailed to a random list of students, faculty, and other employees, and 510 driver surveys were mailed to a

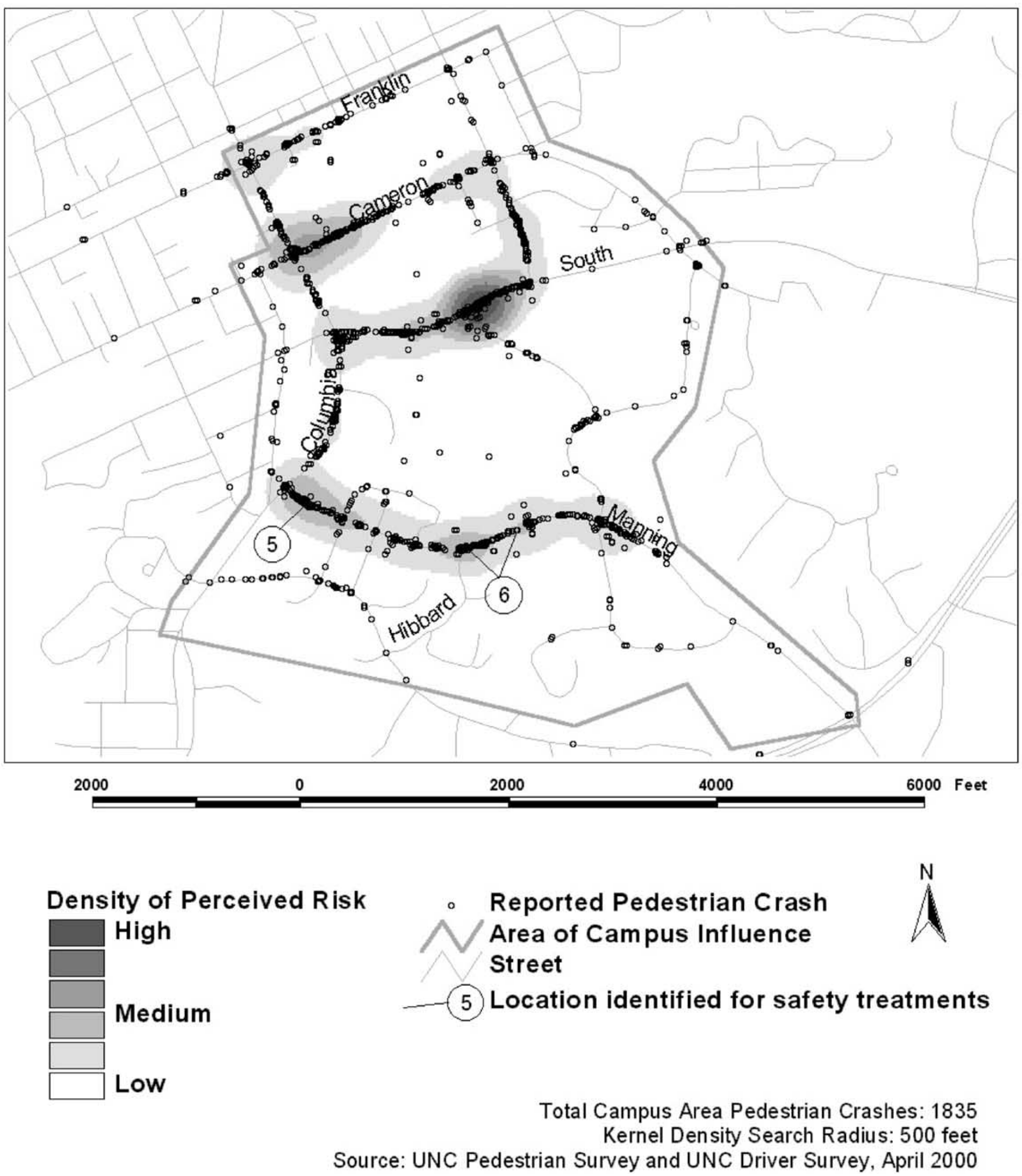

Fig. 4. Perception of pedestrian crash-risk (locations perceived to have a high-risk of pedestrian crashes by pedestrians and drivers on UNC-Chapel Hill Campus).

random list of people with campus parking permits. Over 21% of each group responded. Two hundred fifteen extra pedestrian surveys were given in person at five locations on campus. Though people may have been more likely to mark locations near where the surveys were distributed, pedestrians were over-sampled to ensure that people who walk in all areas of campus were selected. Because it was likely that most pedestrian survey participants had also driven on campus and that most people taking the driver survey had also walked on campus, the pedestrian survey participants were instructed to complete the survey from their perspective as a pedestrian, and driver survey participants were asked to complete the survey from their perspective as a driver. In all, 312 pedestrian and 110 driver surveys were analyzed.

The surveys included two identical maps of the campus area. Participants used the first map to mark the three locations that they believed had the highest risk of pedestrian crashes during daylight. If participants felt that the risk of pedestrian crashes on campus changed at night, they marked the three locations with the highest risk of pedestrian crashes during darkness on a second map. This number of locations was required so that there would be adequate variation in locations if all respondents agreed that one location was the most dangerous point on campus, and that the number of locations would be small enough for respondents to think of and mark on the survey quickly. Yet, some respondents still marked only one or two locations. In all, the 422 pedestrians and drivers provided 1835 locations (data points) on campus that were perceived to have a high-risk of pedestrian crashes (compared to only 57 police-reported crash locations). Twenty-seven off-campus locations were also marked, but these were not used in further analyses. Because there were only minor differences between locations identified during daylight and darkness, the 1835 locations identified in the surveys were combined to create a composite map representing the perception of pedestrian crash-risk on campus (Fig. 4).

4. Reported crashes and risk perception: comparing point distributions

If no difference is found between the spatial distribution of perceived pedestrian risk and the police-reported crash-risk, the survey data would add little value in identifying future crash locations. However, even a visual comparison of Figs. 3 and 4 shows that there are differences between the locations of police-reported crashes and the locations where pedestrians and drivers perceived pedestrian crash problems on the UNC Campus. This visual assessment must be supported by statistical methods to reveal whether the differences between the two point distributions are significant. We use two quantitative techniques (Chi-squared and nearest neighbor cluster analyses) to test the null hypothesis that *the spatial distribution of risk perceptions is not significantly different from the spatial distribution of police-reported pedestrian crash locations*.

4.1. A note about estimating surfaces

We use kernel density estimation to create a probability surface of crashes. Both the kernel density and nearest neighbor cluster analyses assume complete spatial randomness, which implies that there is the potential for reported and perceived pedestrian crash-risk across the entire map surface, i.e. in buildings and in other open spaces away from the roadway network. Though this assumption cannot realistically be made, the purpose of the spatial testing is to show the differences in the actual and perceived-risk distributions. Since the point distributions are naturally constrained to the same roadway network (crashes occurred on roadways, at intersections, and in parking lots, and people marked risky locations on roadways, at intersections, and in parking lots), the crash-risk that is estimated is also constrained along the roadways and, therefore the spatial comparisons presented here are appropriate.

4.2. Chi-squared analysis

The first statistical test used is a Chi-squared test, which examines the relationship *between* points in each distribution (Taylor, 1977; Getis and Boots, 1978). This inter-distributional test shows whether the total counts of police-reported crashes and the points perceived as dangerous on each of the 94 campus roadway network segments or intersections can be classified into similar ranges, and therefore, are distributed similarly. Note that the test does not account for the location of each segment/intersection, but it is an appropriate test to make an initial comparison between both point distributions along the network segments. Each of the 94 campus roadway segments/intersections are classified as having 0, 1, 2, or at least 3 reported crashes. This classification is then compared to the expected number of crashes on each segment/intersection if the crashes were distributed in the same spatial pattern as the perceived dangerous locations.

This parallel classification scheme allows a Chi-squared statistic to be calculated:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_{i_\mathrm{r}} - O_{i_\mathrm{e}})^2}{O_{i_\mathrm{e}}} \qquad (1)$$

where i represents each of the four categories (0, 1, 2, or at least 3 or more police-reported crashes), O_{i_r} is the number of segments/intersections falling into category i according to the number of police-reported crashes on the segment/intersection, and O_{i_e} is the expected number of segments/intersections that would fall into category i if the police-reported crashes were spatially distributed on the roadway segments/intersections in the same manner as the perception survey responses. The test is also performed using the perception data as a base:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_{i_p} - O_{i_e})^2}{O_{i_e}} \quad (2)$$

where i represents each of the five categories (0–29, 30–59, 60–89, 90–119, or 120 or more perception points) and O_{i_p} is the number of segments/intersections falling into category i according to the number of perception points on the segment/intersection.

4.3. Chi-squared results

Both Chi-squared tests showed that the counts of crashes and risk-perception locations on each segment/intersection were distributed into different categories. When the expected number of police-reported crashes on each roadway segment/intersection was based on the distribution of locations perceived to be dangerous, the Chi-squared value was 46.4 with three degrees of freedom (Table 2). Therefore, we are 99.9% confident that there are differences in the manner in which reported and perceived dangerous locations are distributed on the campus roadway intersections and segments. Similarly, a Chi-squared value of 14.5 with four degrees of freedom was generated when the expected number of locations perceived to be dangerous on each segment/intersection was based on the number of police-reported crashes, meaning that the two distributions were significantly different at the 99.0% level.

Though these differences did not account for the spatial arrangement of the segments/intersections, this result showed that some parts of the campus roadway network with larger proportions of police-reported crashes were not matched with a large proportion of perception survey responses (such as the intersection of Franklin Street and Columbia Street) and that specific areas with large proportions of survey responses had few crashes reported in the last 5 years (such as the segment of Manning Drive to the east of its intersection with Columbia Street).

4.4. Nearest neighbor cluster analysis

The second type of spatial comparison, nearest neighbor analysis, is used to identify clusters of points within each spatial distribution. After they are identified, the clusters of one distribution are compared with clusters of the other distribution. This intra-distributional technique uses Euclidean distance to identify sets of points that are clustered more closely than would be expected by random chance. Points are considered clustered when the mean random distance between them is less than a minimum distance based on the standard error of a random distribution:

$$\text{minimum distance} = 0.5\sqrt{\frac{A}{N}} - t\left[\frac{0.26136}{\sqrt{N^2/A}}\right] \quad (3)$$

where A is the area of the study region, N is the number of crash locations ($N = 57$) or locations perceived to be dangerous ($N = 1835$), t is a probability level in the Student's t-distribution, and $[0.26136/(\sqrt{N^2/A})]$ is the standard error

Table 2
Chi-squared test ($N = 94$ segments/intersections)

i, class (number of reported crashes on segment/intersection)	O_{i_r}, number of segments in class i	O_{i_e}, expected number of segments in class i if reported crashes were distributed in same spatial pattern as perception locations
Reported crash data $\left(\chi^2 = \sum_{i=1}^{n}(O_{i_r} - O_{i_e})^2/O_{i_e}\right)$		
$0 \le x < 1$	59	80
$1 \le x < 2$	22	10
$2 \le x < 3$	9	2
$x \ge 3$	4	2
Total	94	94
χ^2	46.4	
d.f.	3	
$\alpha(0.001)$	16.3	
i, class (number of perception locations on segment/intersection)	O_{i_p}, number of segments in class i	O_{i_e}, expected number of segments in class i if perception locations were distributed in same spatial pattern as reported crashes
Perception data $\left(\chi^2 = \sum_{i=1}^{n}(O_{i_p} - O_{i_e})^2/O_{i_e}\right)$		
$0 \le x < 30$	76	59
$30 \le x < 60$	13	22
$60 \le x < 90$	2	9
$90 \le x < 120$	1	2
$x \ge 120$	2	2
Total	94	94
χ^2	14.5	
d.f.	4	
$\alpha(0.01)$	13.3	

distance of a random distribution (Levine, 1999). In other words, for a one-tailed probability, P, fewer than P percent of the points would have nearest neighbor distances less than this lower limit if the point distribution was completely random. For this analysis we use a P-value of 0.01 so that we can be 99% confident that the police-reported and perception location clusters did not occur by random chance. The minimum number of points required to make a cluster is set at three reported crashes and 40 locations perceived to be dangerous.

The standard deviational ellipse of each nearest neighbor cluster and the number of points they contain are also reported. The standard deviational ellipse measures the dispersion and orientation of the points around the mean center (mean latitude, mean longitude) of the cluster. Equations for drawing standard deviational ellipses are presented by Levine (1999). Similarity between the two distributions is demonstrated when the standard deviational ellipses of the police-reported crash clusters overlap corresponding perceived location cluster ellipses. Yet, if the standard deviational ellipse of a reported crash cluster does not contain the mean center of the closest cluster of perceived locations, we can be 99% confident that the two types of data are identifying different locations with pedestrian problems.

4.5. Nearest neighbor cluster results

Cluster analysis revealed notable differences between concentrations of reported crashes and locations perceived to be dangerous (Fig. 5). Only three police-reported crash clusters were identified at the 99% significance level using the nearest neighbor cluster method. In contrast, there were six clusters of perception locations. The relationship between location and size of the clusters of police-reported crashes and clusters of locations perceived to be dangerous is important to note. At the intersection of Franklin Street and Columbia Street, the center of the cluster of nine reported crashes and center of the perceived location cluster were within 12 m (40 ft) of each other and their standard deviational ellipses overlapped. This means that the perceptions of pedestrians and drivers verified reported pedestrian problems at that intersection. Yet, neither of the other two clusters of reported crashes corresponded closely to a cluster of locations perceived to be dangerous. Though the cluster of five reported crashes on South Road was only 100 m (320 ft) from a cluster of locations perceived to be dangerous on South Road, their standard deviational ellipses did not overlap (Fig. 5). Therefore, we are 99% confident that they refer to different locations of pedestrian problems. The perception cluster is centered on the crosswalk between the Student Union and Student Recreation Center, while the police-reported crash cluster may represent general problems in the vicinity of South Road. The final cluster of four reported crashes was located off Manning Drive near the hospital complex, but its standard deviational ellipse did not overlap with either cluster of locations perceived to be dangerous on Manning Drive. We conclude that the clusters generated by the police-reported crash locations and locations perceived to be dangerous provide further evidence that the two data sets have different spatial characteristics.

4.6. Spatial testing summary

Several other spatial statistics were used to test for significant differences between reported and perceived-risk point patterns. Though they are not reported here, Ripley's K-function (Bailey and Gatrell, 1995; Levine, 1999) and a G-function (Bailey and Gatrell, 1995) also revealed significant differences between the locations of the two distributions.

Overall, the spatial analyses show that some locations of reported pedestrian crashes are identified accurately by perception survey participants, while there is also evidence of a "perception mismatch" between clusters of reported crashes and perceived pedestrian crash-risk. Based on the results of this spatial analysis, there are statistically significant differences between the distribution of police-reported crashes and the distribution of locations perceived to be dangerous, implying that risk-perception data can be valuable in proactive pedestrian planning.

4.7. Applying perception mismatch to crash prevention

The "perception mismatch" between police-reported crash and survey locations has implications for crash prevention. Recommendations can be made for four different types of locations:

(1) *High reported and high perceived-risk*: Top-priority areas for engineering, education, and enforcement treatments.
(2) *High reported and low perceived-risk*: Both engineering and education improvements should be explored. These are locations where there is a physical problem yet people are not aware of the danger. The physical aspect of the problem can be treated with engineering changes, and the awareness aspect can be treated by educating drivers and pedestrians. Note that education may be implemented much more quickly than an engineering treatment, which may have to go through a Capital Improvement Program before being made.
(3) *Low reported and high perceived-risk*: Evaluate these areas to see if pedestrian and driver experience has accurately identified existing and potential problems. If problems are found, it may be possible to treat the sites proactively with appropriate engineering, education, and enforcement measures to prevent future pedestrian crashes.
(4) *Low reported and low perceived-risk*: No countermeasures necessary, but continue normal monitoring.

Note that it is possible that people simply do not use or they walk with extreme caution in areas where there is

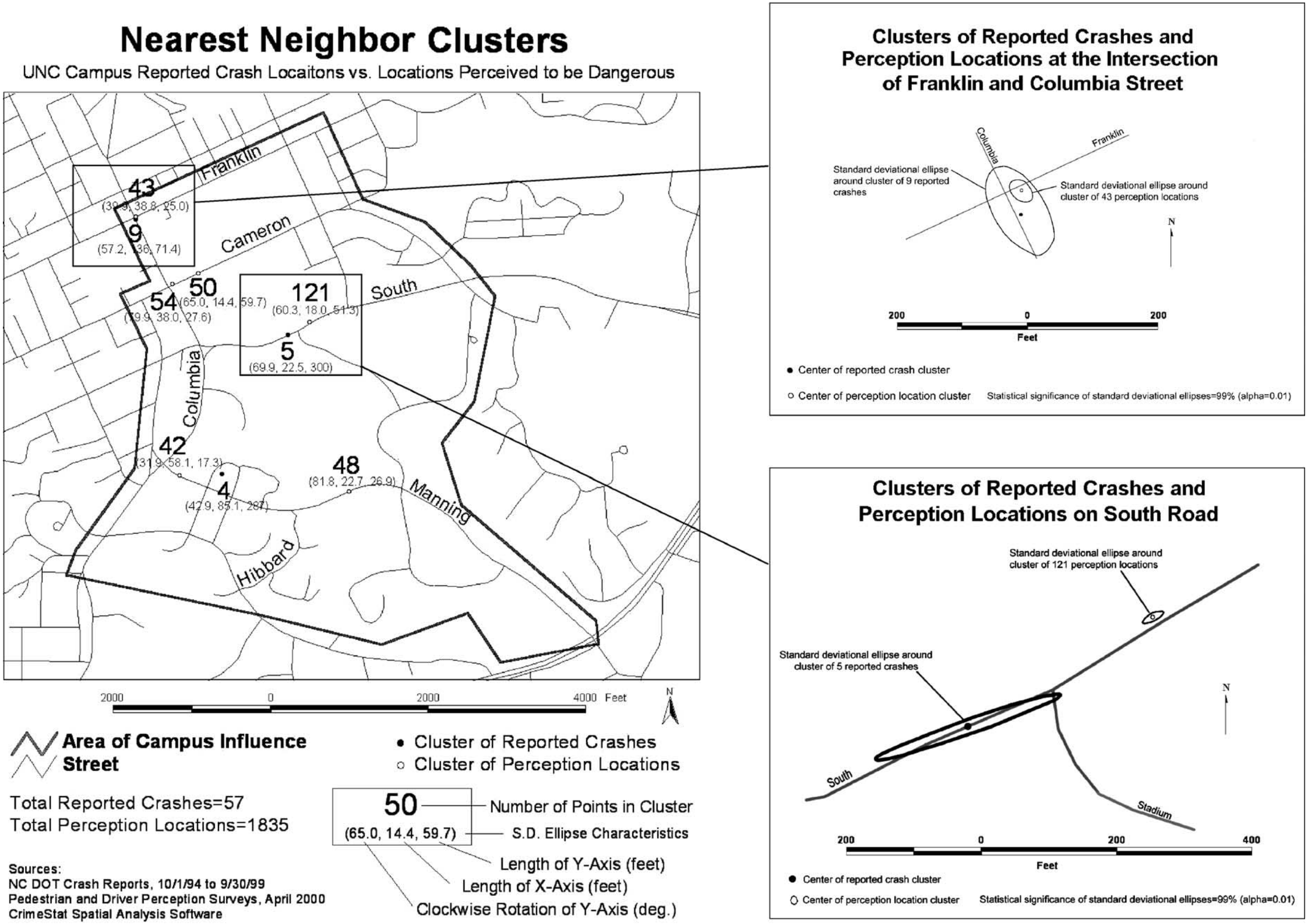

Fig. 5. Cluster analysis. Nearest neighbor clusters (UNC Campus reported crash locations vs. locations perceived to be dangerous).

extreme danger (i.e. high-speed arterials, multi-lane roadways, streets with high traffic and no sidewalks, etc.). We emphasize that our methodology is most useful for situations and areas where there is a moderate perceived crash-risk (common on campuses, in neighborhoods, and in commercial areas). Depending on their ability to confront danger (which may be affected by age, weather, drugs, etc.), pedestrians may walk in these moderate-risk locations. Even pedestrians who perceive danger, but may not take unwarranted risks, can suggest where crashes may occur to others. Further, speeding or reckless drivers who are unaware of pedestrians may cause pedestrian crashes, no matter how careful the pedestrians are when they are in the roadway. Therefore, it is useful to know about the locations where pedestrians perceive danger—it is locations like these where pedestrian crashes could occur in the future if engineering, education, and enforcement treatments are not made.

It is especially important to identify risky locations proactively for pedestrians because pedestrian injuries are often much more severe when accidents occur. In locations that have experienced no pedestrian crashes, the very first crash can be fatal, as we saw on the UNC Campus. In addition, perceived-risk can be used for more than a superficial examination of personal ability to "read" the safety of the environment. The perception survey and resulting maps can be shared with the public to point out specific locations as being dangerous and in need of attention.

5. Modeling reported and perceived crash-risk

On the UNC Campus, reported crash locations do not correspond closely to perception survey locations near the UNC Hospital and south of Manning Drive; perception survey locations do not correspond closely to police-reported crash locations along parts of Cameron Street and on Columbia Street between South Road and Manning Drive (compare Figs. 3 and 4). The spatial pattern of "perception mismatch" requires further investigation. A richer understanding of this spatial pattern can be obtained by examining the combinations of pedestrian and vehicle volumes, roadway features, and nearby land uses that may be underlying physical/environmental causes of pedestrian crashes and risk perception.

We use Poisson and negative binomial crash models to show the exposure, roadway, and land use factors that are related significantly (at least at the 90% confidence level) to the police-reported and perceived-risk of pedestrian crashes. Incorporating these geographic/land use factors in the regression modeling process demonstrates the connectivity between spatial and multivariate crash analysis.

5.1. Regression model data

The campus roadway network is divided into 38 intersections (nodes, defined as the area within 15 m (50 ft) of the intersection of two or more roadway center lines) and 56 segments (links, defined as lengths of roadway between intersections), each of which are assigned specific pedestrian crash-risk, exposure, roadway, and land use characteristics. These 94 units of analysis will be referred to as segments/intersections in the following sections. By considering the geographic context of the UNC Campus, this study shows that exposure, physical roadway, and *spatial characteristics* are related to pedestrian safety problems.

Three exposure, five roadway, and seven land use characteristics were collected for each segment/intersection from traffic and pedestrian volume maps, field observation, and GIS measurement. Descriptions of these variables and their hypothesized relationships to pedestrian crash-risk (represented by the number of reported crashes or number of perceived locations on each segment/intersection) are presented in Table 3.

5.2. Regression models

The Poisson statistical distribution represents the occurrence of "infrequent events." Therefore, Poisson modeling is appropriate for modeling the number, or count, of actual or perceived crash locations on each campus roadway intersection/segment. However, the Poisson model requires that the mean equal the variance of the count data. The negative binomial regression model can relax the mean–variance equality assumption (Greene, 1997; Cameron and Trivedi, 1998). The difference between the means and variances of reported crashes (mean = 0.606, variance = 1.44) and locations perceived as dangerous (mean = 19.5, variance = 946) is relatively large, favoring the negative binomial model. However, for comparison, both Poisson and negative binomial models are presented.

5.3. Poisson model

The Poisson model uses Y_i to denote the number of crash occurrences or locations perceived to be dangerous for the ith of $N = 94$ campus roadway segments/intersections. This number of crash occurrences on each element can be Poisson distributed with probability density:

$$P[Y_i = y_i] = \frac{e^{\lambda_i}\lambda_i^{y_i}}{y_i!} \tag{4}$$

where λ_i is the segment/intersection i's expected crash frequency; $y_i = 0, 1, 2, \ldots$ (realized value of the crash frequency); and $i = 1, 2, \ldots, N$ and $y_i!$ denotes the factorial of y_i. Note that the mean and variance of Y_i equals λ_i.

To incorporate explanatory variables x_i the parameter λ_i is specified to be

$$\lambda_i = \exp(\beta' x_i) \tag{5}$$

where β' is the vector of estimated parameters; and x_i is the roadway element i's explanatory variables (e.g. pedestrian

Table 3
Campus roadway intersection/segment variables and hypothesized relationships to pedestrian risk

Variable	Description	Hypothesized relationship (assumes "all else being equal, . . . ")
Dependent variables		
NOBSERV	The number of pedestrian crashes reported on each segment/intersection between 10/1/1994 and 9/30/1999	N/A
NPERCEP	The number of locations marked on each segment/intersection by perception survey respondents	N/A
Exposure variables		
LNLENGTH	Natural logarithm of segment/intersection length in feet	Greater segment/intersection length, higher pedestrian exposure, and higher automobile exposure increases risk of pedestrian crashes
LNPED	Natural logarithm of estimated daily pedestrian volume	
LNAUTO	Natural logarithm of estimated daily vehicle volume	
Roadway variables		
INTERSEC	0 = segment; 1 = intersection (within 50 ft of the center of an intersection)	Higher risk on segments than intersections because of mid-block crossings and dangerous crossing behaviors; yet intersection risks might be higher due to more pedestrian-vehicle conflicts
LANES	Number of lanes; intersection = average of all approaches	Higher risk with more lanes because of difficulty judging gaps and longer crossing distance
SIDEWALK	0 = no sidewalk; 1 = sidewalk on one side or incomplete sidewalk; 2 = complete sidewalks on both sides	Lower risk with sidewalks because pedestrians are physically separated from traffic
BUS1000	Number of bus stops per 1000 linear feet	Higher risk with bus stops because pedestrians hurry to and from buses
XWAL1000	Number of marked crosswalks per 1000 linear feet	Lower risk with crosswalks because pedestrians will converge to them and drivers will yield to pedestrians using them
Spatial/land use variables		
DORM	Distance (in feet) from the center of the segment/intersection to nearest of over 20 dorms	Proximity to dorms may increase risk because dorm residents may be younger and use less care when walking
LIBRARY	Distance to nearest major campus library (Wilson, House, or Davis)	Proximity to libraries may decrease risk because library users may be less likely to have been drinking alcohol and may be more cautious as pedestrians
STADIUM	Distance to nearest sports stadium (Carmichael, Boshamer, Kenan, or Smith Center)	Proximity to stadiums may increase risk because people are less familiar with the environment, are often in a hurry, often drink alcohol, and often leave stadium events en masse
HOSPITAL	Distance to UNC Hospital complex	Proximity to hospital may increase risk because many pedestrians are from outside Chapel Hill and unfamiliar with environment
PARKDECK	Distance to nearest campus parking deck (Craige or hospital)	Proximity to parking deck may increase risk because drivers' attention is focused on searching for parking and not on pedestrians
ACADEMIC	Distance to nearest of over 30 academic buildings	Proximity to academic buildings may increase risk because student pedestrians may hurry to class
DINING	Distance to nearest campus cafeteria (Lenoir Hall or Chase Hall)	Proximity to dining halls may increase risk because students' attention may be focused on eating-related activities

exposure and crosswalk presence). Goodness of fit statistics for this model are discussed in Greene (1997).

5.4. Negative binomial model

The negative binomial model relaxes the mean–variance equality assumption of the Poisson model to allow for unexplained randomness in λ_i. This is done by specifying:

$$\ln \lambda_i = \beta' x_i + \varepsilon_i \tag{6}$$

where ε_i is the error term, which can reflect a specification error such as omitted explanatory variables and/or intrinsic randomness. In addition, the negative binomial model assumes that $\exp(\varepsilon_i)$ has a gamma distribution with mean 1 and variance α^2. The derivation of the probability distribution for this model is given in Greene (1997) and Cameron and Trivedi (1998).

Compared with the Poisson model, this model has an additional estimable parameter α, such that

$$\mathrm{Var}[y_i] = E[y_i]\{1 + \alpha E[y_i]\} \tag{7}$$

This is a natural form of overdispersion and the overdispersion rate:

$$\frac{\mathrm{Var}[y_i]}{E[y_i]} = 1 + \alpha E[y_i] \tag{8}$$

The model can be estimated by the standard maximum likelihood method. If α is not statistically different from zero, then the simple Poisson model is more appropriate.

5.5. Regression modeling process

With 57 crashes reported on 94 segments/intersections in 5 years, a single pedestrian crash on one of the segments/intersections has a large influence on the statistical significance of its exposure, roadway, and spatial attributes (there are an average of 0.606 police-reported crashes on each segment/intersection, and 63% of the segments/intersections have no crashes). When using perception data, the average segment/intersection has 19.5 marked survey locations. Therefore, even if a few respondents overestimate the risk of future pedestrian crashes at a location, their individual response will not have a large effect on the accuracy of the perception model. Comparing the factors from these two types of models provides a deeper understanding of the factors that influence pedestrian crash-risk.

The effects of exposure, roadway, and land use (spatial/environmental) factors on reported pedestrian crash locations and perception of crash-risk were hypothesized in Table 3. First, separate exposure, roadway, and spatial models were estimated. Interaction terms were tested in these models, but they did not improve their fit. Then, the variables from the exposure, roadway, and spatial models were used to estimate Poisson and negative binomial combined models. Finally, the most statistically significant variables from the combined models were kept and used to estimate refined models. Results from the refined models are presented below.

5.6. Regression modeling results

Independent variable coefficients, standard errors, significance levels and overall model summary statistics are presented for the four refined models in Table 4. These four refined models include two Poisson (Models 1 and 3) and two negative binomial models (Models 2 and 4) estimating police-reported and perceived-risk frequencies.

The Poisson Model 1 is preferred to Model 2 because it has more significant parameter estimates and α is not statistically significant. However, the negative binomial Model 4 is preferred to Model 3 based on goodness-of-fit and the significance of α. Therefore, Models 1 and 4 are discussed, in this order.

As expected, longer segments/intersections and higher pedestrian volumes are significantly related to higher levels of police-reported crashes, though higher automobile volumes are not statistically significant. In addition, the elasticity of the exposure factors is less than one, indicating that as length or pedestrian volume increase, the number of police-reported crashes on road segments or intersections increase at a decreasing rate. For example, a one percent increase in pedestrian volume increases the number of police-reported crashes by 0.672%.

After accounting for exposure, the results show that when segments/intersections had more marked crosswalks per linear foot, they had a greater number of police-reported pedestrian crashes. Though this result is counter to our

Table 4
Poisson and negative binomial models ($N = 94$)

Independent variable	Observed-risk: coefficient, significance level (P)		Perceived-risk: coefficient, significance level (P)		Mean
	Poisson Model 1:	Negbin Model 2:	Poisson Model 3:	Negbin Model 4:	
Constant	−13.5, (0.0000)	−13.5, (0.0033)	−19.9, (0.0000)	−14.4, (0.0000)	
Exposure factors					
LNLENGTH	0.625, (0.0165)	0.625, (0.0870)	1.05, (0.0000)	0.867, (0.0000)	5.87
LNPED	0.672, (0.0091)	0.667, (0.0771)	1.25, (0.0000)	0.871, (0.0000)	7.09
LNAUTO	0.381, (0.154)	0.383, (0.262)	1.05, (0.0000)	0.796, (0.0000)	8.85
Roadway factors					
INTERSEC	−0.935, (0.129)	−0.932, (0.276)	−0.797, (0.0000)	−0.246, (0.451)	0.404
SIDEWALK	0.0940, (0.851)	0.101, (0.851)	−1.01, (0.0000)	−0.697, (0.0029)	1.61
BUS1000	−0.159, (0.0958)	−0.158, (0.0281)	0.00750, (0.619)	0.0159, (0.774)	1.09
XWAL1000	0.0810, (0.0089)	0.0806, (0.0665)	0.090, (0.0000)	0.0643, (0.0002)	5.29
Spatial factors					
LIBRARY	−0.000841, (0.0685)	−0.000822, (0.103)	0.000218, (0.0063)	0.000316, (0.0884)	1910
ACADEMIC	0.000996, (0.0599)	0.000973, (0.120)	−0.000225, (0.0137)	−0.000348, (0.107)	983
STADIUM	0.000783, (0.0300)	0.000764, (0.0379)	−0.000636, (0.0000)	−0.000644, (0.0001)	1690
α		0.0215, (0.944)		0.332, (0.0000)	
Summary statistics					
log likelihood	−83.0	−83.0	−500	−307	
Restrictive log likelihood	−107	−82.978	−1510	−500	
Goodness of fit R_p^2	0.300		0.799		
Goodness of fit R_d^2	0.361		0.763		
χ^2	48.1, (0.0000)	0.0114, (0.915)	2030, (0.0000)	386, (0.0000)	
	Dependent variable = NOBSERV ($\mu = 0.606$)		Dependent variable = NPERCEP ($\mu = 19.5$)		

expectations, it is consistent with studies by Herms (1972) and Zegeer et al. (2001) who find some locations with marked crosswalks have more pedestrian crashes than control locations without marked crosswalks. However, this result does not necessarily suggest that providing marked crosswalks "causes" pedestrian crashes. Our data does not indicate where the crashes occurred relative to crosswalks. The relation of higher crosswalk density to a higher incidence of pedestrian crashes may result from more dangerous pedestrian or driver behavior in the vicinity of marked crosswalks, weak crosswalk design, or other factors. Further study of crosswalks and related behavior is needed to assess the effectiveness of this treatment.

The number of bus stops was statistically significant in the crash model (90% level), indicating that more bus stops do not create a greater pedestrian crash-risk. The intersection variable was not statistically significant in the model. Finally, the number of lanes was statistically non-significant and dropped from the final model specification.

Among spatial factors, proximity to academic buildings and stadiums is associated with higher crash-risk, whereas proximity to libraries with lower crash-risk. At this point, it is interesting to compare the results of Models 1 and 4. The coefficients for the significant spatial variables in the reported crash model and the perceived-risk model have opposite signs. Therefore, people perceive less risk near some types of land uses where crashes have occurred and perceive more risk near other land-uses where no crashes have occurred. Locations near the main campus libraries have experienced fewer crashes, but they are perceived to have more danger. In contrast, proximity to academic buildings increases police-reported crash-risk but is not perceived to do so. Finally, locations near stadiums are more dangerous according to police-reported crash data than people think. The distance from the segments/intersections to the closest dorm, hospital, parking deck, and dining hall were not included in the refined models because they lacked statistical significance.

The relationship between proximity to certain land uses and crash-risk clearly needs further investigation. Spatial/land use factor results may have implications for the design of college campuses and neighborhoods. Further research could also examine whether a dense traffic network with short blocks and many intersections or dense residential neighborhoods are related to greater actual and perceived pedestrian risk. Our study used segments and intersections as the unit of analysis, so we did not test these relationships.

The results of exposure and roadway factors are quite consistent across Models 1 and 4, increasing our confidence in the appropriateness of behavioral data. Though sidewalk presence was not statistically significant in Model 1, it is significant in Model 4. The negative coefficient indicates that pedestrians and drivers perceived a lower pedestrian crash-risk when more complete sidewalks were provided and more danger when segments/intersections had incomplete sidewalks.

6. Potential biases and other considerations

(1) *Police-reported crashes may have occurred in a different context than is currently perceived by survey participants*: Perceptions of locations with a high-risk of crashes are influenced by a context of exposure, roadway, and spatial factors. Crashes that occurred in 1995 may have occurred when automobile and pedestrian volumes were different, before traffic signal timing was adjusted, or before sidewalks and new buildings were constructed. Unreported pedestrian crashes may also contribute to the spatial mismatch between reported and perceived-risk locations.

(2) *Oversampling at certain locations may create a recency of experience bias*: Distributing surveys at five locations with high pedestrian exposure on campus ensured that input was gathered from a large number of pedestrians in a short amount of time and that people who walk in all areas of campus were represented in the sample. Yet, the 215 pedestrians who took the survey on campus may have remembered areas that they had walked through recently and, therefore, may have been more likely to mark them as dangerous. If the goal of the survey is to gather input from a large number of pedestrians in a short amount of time, then surveys should be given at locations with high pedestrian activity. Yet, if the goal is to obtain a more spatially-representative sample, techniques such as regular, cluster, random, and stratified random sampling should be used.

(3) *The perception survey may have a non-response bias*: About 21% of mail surveys were returned. Though more than 25% of pedestrians agreed to take the survey in-person, the overall survey response rate remained relatively low. Therefore, the survey results may be more likely to reflect the views of people who may be more risk-averse or who care more about the issue of pedestrian safety on campus. These people may be more aware of the exceptional pedestrian problems than people who did not participate, resulting in locations that are commonly viewed as dangerous to be overlooked. Or, their awareness may cause perceived problems to be overstated. Non-response bias could be mitigated by re-mailing the survey or being more assertive when offering the survey in person, though we did not pursue these options in this study.

(4) *The survey results may be influenced by the driver-pedestrian mix, especially if they perceived-risks differently*: Visual inspection of disaggregate perceived-risk maps showed that drivers and pedestrians perceived pedestrian problems at similar locations on the UNC Campus, which allowed the spatial data from both surveys to be combined. Yet, if drivers and pedestrians had not agreed, then this survey would have been biased toward locations with higher pedestrian exposure and locations that pedestrians believed had the greatest

danger (almost three times as many pedestrian surveys were analyzed). In such a case, it would be necessary to generate separate maps and models of pedestrian and driver perception. The perception mismatch between pedestrians and drivers is also important to note because the most dangerous locations may be where one of these two groups do not perceive a high-risk.

(5) *Increasing perception survey sample size may improve perception maps and models*: Increasing sample size can increase the number of perception data points. This can increase the significance of spatial clustering of survey locations, significance of the difference between police-reported and perceived-risk spatial distributions, and significance of the perceived-risk model parameters. However, the sample size could not be increased in this study due to budget constraints.

(6) *The richness of the perception data may depend on the number of locations survey participants are required to identify*: The surveys can be designed to ask respondents to identify any number of dangerous locations, asking participants to mark only one site may result in the selection of a single common location that has received public attention for safety problems. Asking pedestrians and drivers to mark more than one site increases the number of data points that can be used to create risk-perception maps and used to estimate regression models. However, responses by the same person may be correlated. In addition, asking respondents to identify too many dangerous sites lengthens the survey and increases the possibility of identifying low-risk sites as dangerous. Our compromise was to ask for three most dangerous sites.

(7) *The results will depend on the accuracy inherent in the data*: Automobile volumes were estimated to the nearest 1000 vehicles per day from 1997 Chapel Hill traffic data (Town of Chapel Hill, 1999). Pedestrian volumes were estimated to the nearest 50 persons per day based on a campus map of pedestrian flows (Rice, 1998). Known pedestrian volumes at several locations helped guide estimates at other locations. Finally, the GIS distance measurements were made to the nearest 30 meters (100 ft). While the imprecision of each of these measurements may be a problem, the data are reasonably accurate for the purposes of identifying pedestrian risks.

(8) *Maps may be difficult for participants to interpret*: Not all pedestrians and drivers are likely to be familiar with maps of a local area and some groups, such as the elderly, may have a difficult time reading small print. The maps used for this survey were intended to be uncluttered, have several recognizable landmarks, and be easy to read. Yet, printing and mailing costs required the campus map to fit on one page, making building and road labels very small. In addition, study area borders may not have been be clear to participants, resulting in fewer locations marked on the edges of campus.

(9) *Density maps may not reflect the probability of future crashes*: While the kernel density maps presented in Figs. 3 and 4 are visual tools for highlighting concentrations of crashes, they do not show the probability of future crashes or display crash-risk (normalized by pedestrian and vehicle exposure) over a surface. Furthermore, the phenomena of regression-to-the-mean must be recognized. That is, locations where many crashes have occurred recently may regress to the overall mean, which is presumably lower. Furthermore, automobile-pedestrian collisions cannot occur in buildings, parks, or open fields, though density maps may seem to imply this. Possible misinterpretation of density maps is one reason why quantitative spatial tests are needed to identify high-risk locations for appropriate safety countermeasures. However, density maps are still useful for practitioners because it is important to focus on areas where the highest actual and perceived-risk exists, which is often in locations with high pedestrian and vehicle volumes.

(10) *Modifications may be required to apply this method to neighborhoods and commercial zones*: When using these techniques in non-campus environments, police-reported crash data can be collected and analyzed in a similar manner as presented in this paper, but the perception survey technique should account for the fact that, unlike a “self-contained” college campus, pedestrians and drivers who use neighborhoods and commercial areas may reside far from the study area. By mailing driver surveys to a random sample of people with parking permits on campus, and administering pedestrian surveys to a random sample of faculty, employees, and students living on and off-campus, a sample was generated for the UNC Campus study. This sample was supplemented with hand-out-hand-back pedestrian surveys given at different locations on campus. A more appropriate technique for a neighborhood or commercial area may be to distribute all surveys in-person at a number of randomly-chosen locations, increasing the likelihood that survey respondents are familiar with the local pedestrian environment.

7. Conclusions

Spatial tests have shown that the distribution of crash-risk perception is different than reported crash-risk. Perception information can help identify locations where an accident is waiting to happen. Therefore, while funds should continue to be allocated to treat locations with reported crash concentrations, this method can be used to justify proactive spending to educate pedestrians and drivers about areas with reported and perceived danger, evaluate risky behaviors and roadway deficiencies at locations with perceived problems, and to make treatments to prevent injuries and fatalities at locations that are overlooked by police reports. Collecting perception data also encourages public participation in the safety analysis process.

The perception survey also increased the amount of data for regression modeling and provided insight into the factors related to pedestrian risk. Models showed that higher pedestrian volume, higher automobile volume, incomplete sidewalks, and fewer bus stops and more marked crosswalks per foot of segment/intersection were correlated with higher risk. Furthermore, perceived-risk was different than police-reported crash-risk in locations that were proximate to campus libraries, academic buildings, and stadiums.

We have developed a prototype tool that can be used to recommend measures for improving pedestrian safety. Due to resource limitations, we did not conduct a before-and-after study, so future researchers should collect crash and injury data over time to document the effectiveness of this method. Mismatched locations that are identified should be studied further to identify dangerous roadway conditions and pedestrian and motorist behavior. If this physical and behavioral site analysis reveals that the pedestrian environment is abnormally dangerous at locations identified from the perception survey but missed by police reports, this prototype tool can be valuable for preventing pedestrian crashes at individual sites.

Future studies can use techniques such as spatial and ordinal regression analysis to examine how perception is related to injury severity. It would be interesting to see if policies aimed at making the pedestrian environment feel safer actually result in fewer, but more severe crashes because pedestrians and drivers have a "false sense of security" and use less caution.

One interesting result of the survey was that while only 28% of drivers felt that they had experienced a "near miss", over 40% of pedestrians felt that they had nearly been struck by a vehicle within the past month. A more in-depth analysis of near miss locations may be incorporated to further explore the differences between how pedestrians and drivers view pedestrian risk.

Acknowledgements

The authors would like to thank Charles V. Zegeer of the UNC Highway Safety Research Center for his input on this project. In addition, we would like to thank all members of the UNC Pedestrian Safety Committee, especially Committee Chair, Chief Derek K. Poarch, Director of Public Safety at UNC, for their cooperation and support. The pedestrian safety improvements that have already been made and are being planned for the campus would not have been possible without the cooperation of the university community, the UNC Highway Safety Research Center, the Town of Chapel Hill, and the North Carolina Department of Transportation. Finally, thanks to the members of the UNC Department of City and Regional Planning for their participation and valuable feedback during pre-testing of the perception survey. We are also grateful to the Carolina Transportation Program for support.

References

Austin, K.P., Tight, M.R., Kirby, H.R., 1995. An Advanced System for the Study of Children's Safety on the Journeys To and From School, vol. 2A, No. 3. VTI Conference No. 4A.

Bailey, T., Gatrell, A., 1995. Interactive Spatial Data Analysis. Wiley, New York.

Braddock, M., Lapidus, G., Cromley, E., Cromley, R., Burke, G., Banco, L., 1994. Using a geographic information system to understand child pedestrian injury. Am. J. Publ. Health 84 (7), 1158–1161.

BTS, 1999. Transportation Statistics Annual Report 1999, BTS99-03. U.S. Department of Transportation, Bureau of Transportation Statistics, Washington, DC.

Butchart, A., Kruger, J., Lekoba, R., 2000. Perceptions of injury causes and solutions in a Johannesburg township: implications for prevention. Social Sci. Med. 50, 331–344.

Cameron, C., Trivedi, P., 1998. Regression Analysis of Count Data. Econometric Society Monograph No. 30, Cambridge University Press, Cambridge.

Chu, Y., Azer, M., Catalanotto, F., Ungar, H., Goodman, L., 1999. Safety/GIS models reviewed and related to long island arterial needs study. In: Proceedings of the ITE International Conference on Enhancing Transportation Safety in the 21st Century.

Duncan, C., Khattak, A., Hughes, R., 1999. The perceived effectiveness of sidewalks and other pedestrian safety treatments as countermeasures for hit-along-roadway crashes. Presented at 81st Annual Transportation Research Board Meeting, on TRB CD-ROM, Washington, DC, 2002.

Getis, A., Boots, B., 1978. Models of Spatial Processes: An Approach to the Study of Point, Line, and Area Patterns. Cambridge University Press, Cambridge.

Greene, W., 1997. Econometric Analysis, second ed. MacMillan, New York.

Greene, W., 1998. LIMDEP. Econometric Software Inc.

Hank Mohle and Associates, 1996. GIS for Small Municipalities. Presentation Materials, OTS Summit.

Harrell, A., 1991. Factors influencing pedestrian cautiousness in crossing streets. J. Social Psychol. 131 (3), 367–372.

Herms, B., 1972. Pedestrian Crosswalk Study: Accidents in Painted and Unpainted Crosswalks. Record No. 406, Transportation Research Board, Washington, DC.

Hino, A., van der Molen, H.H., Wilde, G.J.S., 1996. Risk perception, risk taking, accident involvement and the need for stimulation. Safety Sci. 22 (1–3), 35–48.

Karim, M.R., 1992. Traffic accidents in a university environment. ITE J. 62 (7), 30–34.

Kim, K., Takeyama, D., Nitz, L., 1995. Moped safety in Honolulu Hawaii. J. Safety Res. 26 (3), 177–185.

Landis, B.W., Vattikuti, V.R., Ottenberg, R.M., McLeod, D.S., Guttenplan, M., 2001. Modeling the roadside walking environment: a pedestrian level of service. In: Proceedings of the Papers Presented at Transportation Research Board 80th Annual Meeting.

Levine, N., 1999. CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations. Ned Levine & Associates, Annandale, VA and the National Institute of Justice, Washington, DC, August 1999.

Levine, N., Kim, K., Nitz, L., 1995. Spatial analysis of Honolulu motor vehicle crashes. I. Spatial patterns. Accid. Anal. Prev. 27 (5), 663–674.

McMahon, P., 1999. A Quantitative and Qualitative Analysis of the Factors Contributing to Collisions Between Pedestrians and Vehicles Along Roadway Segments. Masters Project, Department of City and Regional Planning, University of North Carolina at Chapel Hill, NC.

Miller, J.S., 2000. The unique analytic capabilities geographic information systems can offer the traffic safety community. In: Proceedings of the Papers Presented at Transportation Research Board 79th Annual Meeting.

Näätänen, R., Summala, H., 1976. Road User Behavior and Traffic Accidents. North-Holland, Amsterdam.

NHTSA, 2000. Traffic Safety Facts 1999, DOT HS 809-093, U.S. Department of Transportation, National Highway Traffic Safety Administration (NHTSA), National Center for Statistics and Analysis. Online available from 17 November 2000 at <http://www.nhtsa.dot.gov/people/ncsa>.

North Carolina Center for Geographic Information and Analysis (NCCGIA), 2000. Pedestrian and Bicycle Safety Analysis Tools.

Peled, A., Haj-Yehia, B., Hakkert, A.S., 1996. ArcInfo-Based Geographical Information System for Road Safety Analysis and Improvement. http://www.esri.com/library/userconf/proc96/TO50/PAP005/P5.HTM, accessed on 23 January 2000.

Peterson's, 2000. Peterson's Thomson Learning. Peterson's 4 Years Colleges, Lawrenceville, NJ.

Rice, C.A., 1998. University of North Carolina Estimated Pedestrian Flows Map. September 1998.

Schneider, R.J., Khattak, A.J., Zegeer, C.V., 2001. A proactive method of improving pedestrian safety using GIS: example from a college campus. *Transportation Research Record*, 1773, TRB, National Research Council, Washington, DC, pp. 97–107.

Summala, H., 1996. Accident risk and driver behaviour. Safety Sci. 22 (1–3), 103–117.

Taylor, P., 1977. Quantitative Methods in Geography: An Introduction to Spatial Analysis. Houghton Mifflin Company, Boston.

Town of Chapel Hill, 1999. Chapel Hill traffic and pedestrian levels. Draft of long-range plan. Transportation, Section 6. North Carolina Planning Department, NC.

US Department of Education, 2000. Digest of Educational Statistics, 1999. NCES 2000-031 (by Thomas D. Snyder, Production Manager; Charlene M. Hoffman), National Center for Education Statistics, Washington DC.

Ward, N.J., Wilde, G.J.S., 1996. Driver approach behaviour at an unprotected railway crossing before and after enhancement of lateral sight distances: an experimental investigation of a risk perception and behavioural compensation hypothesis. Safety Sci. 22 (1–3), 63–75.

Wilde, G.J.S., 1994. Target Risk: Dealing with the Danger of Death, Disease, and Damage in Everyday Decisions. PDE Publications, Toronto, Ont.

Wilde, G.J.S., Gerszke, D., Paulozza, L., 1998. Risk optimization training and transfer. Transport. Res. Part F 1, 77–93.

Zegeer, C.V., Stewart, J.R., Huang, H., 2001. Safety Effects of Marked Vs. Unmarked Crosswalks at Uncontrolled Locations: Executive Summary and Recommended Guidelines, FHWA-RD-01-075. U.S. Department of Transportation, Federal Highway Administration, March 2002.