Title
Association between Roadway Intersection Characteristics and Pedestrian Crash Risk in Alameda County, California

Permalink
https://escholarship.org/uc/item/0d48w4gz

Authors
Schneider, Robert J.
Diogenes, Mara Chagas
Arnold, Lindsay S.
et al.

Publication Date
2010

Peer reviewed
Association between Roadway Intersection Characteristics and Pedestrian Crash Risk in Alameda County, California

Robert J. Schneider, Mara Chagas Diogenes, Lindsay S. Arnold, Vanvisa Attaset, Julia Griswold and David R. Ragland, SafeTREC

2010
RR-2010-8
http://www.safetrec.berkeley.edu
ABSTRACT

Each year from 1998 to 2007, an average of approximately 4,800 pedestrians were killed and 71,000 pedestrians were injured in United States traffic crashes. Because many pedestrian crashes occur at roadway intersections, it is important to understand the intersection characteristics that are associated with pedestrian crash risk. This study uses detailed pedestrian crash data and pedestrian volume estimates to analyze pedestrian crash risk at 81 intersections along arterial and collector roadways in Alameda County, California. The analysis compares pedestrian crash rates (crashes per 10,000,000 pedestrian crossings) with intersection characteristics. In addition, more than 30 variables were considered for developing a statistical model of the number of pedestrian crashes reported at each study intersection from 1998 to 2007. After accounting for pedestrian and motor vehicle volume at each intersection, negative binomial regression shows that there were significantly more pedestrian crashes at intersections with more right-turn-only lanes, more non-residential driveways within 50 feet (15 m), more commercial properties within 0.1 miles (161 m), and a greater percentage of residents within 0.25 miles (402 m) who are younger than age 18. Raised medians on both intersecting streets were associated with lower numbers of pedestrian crashes. These results, viewed in combination with other research findings, can be used by practitioners to design safer intersections for pedestrians. This exploratory study also provides a methodological framework for future pedestrian safety studies.
INTRODUCTION

Each year from 1998 to 2007, an average of approximately 4,800 pedestrians were killed and 71,000 pedestrians were injured in United States traffic crashes. During this 10-year period, pedestrians represented approximately 11.5 percent of all traffic fatality victims (1). Pedestrian safety is a critical issue in the United States, but pedestrian fatalities are even more common in many developing countries. International research shows that 30 to 50 percent of traffic fatality victims in India, Kenya, and Brazil are pedestrians (2,3,4).

Roadway intersections are critical locations for pedestrian safety. A study of pedestrian crashes in six states found that the most common location for fatal and injury pedestrian crashes was within 50 feet of an intersection (5). From 1998 and 2007, there were 6,289 pedestrian crashes in Alameda County, California. A pedestrian crash was defined as any reported traffic crash involving a vehicle and one or more pedestrians. Of these crashes, 3,525 (56.1 percent) were at or within 50 feet of a roadway intersection. Therefore, intersections are important locations to study pedestrian safety.

Purpose

The purpose of this study is to identify how specific roadway intersection characteristics (such as number of lanes, roadway crossing width, and traffic volume) are associated with pedestrian crashes. After accounting for differences in pedestrian and vehicle volumes (measures of pedestrian exposure to crash risk) and surrounding neighborhood variables, which intersection design characteristics are associated with greater numbers of pedestrian injuries and fatalities? The results will help transportation planners, designers, and engineers create safer environments for pedestrians.

This exploratory study is designed to account for pedestrian crash risk using specific pedestrian crash location data and pedestrian exposure estimates. Since extensive background data collection was required to evaluate pedestrian safety at this level of detail, the analysis is based on a relatively small number of intersections in one California county. Therefore, more research will be needed to further explore associations between intersection design factors and pedestrian safety. Nonetheless, this study identifies several key intersection characteristics related to pedestrian crash risk and presents a methodological framework for future pedestrian safety studies.

Factors Associated with Pedestrian Crashes

A significant body of research focuses on factors associated with pedestrian crashes (4). Many studies investigate the relationship between pedestrian safety and roadway design variables (6). The following factors are identified in the literature:

- Pedestrian crossings with more motor vehicle lanes and longer crossing distances can be more dangerous than those that are narrower (6,7,8,9,10).
- Longer traffic signal phases and pedestrian wait times tend to be associated with lower levels of comfort and more pedestrian violations at signalized intersections (8,11).
- Right-turn channelization islands and higher right-turn-on-red motor vehicle volumes are perceived to make intersections less safe for pedestrians (8).
- Higher speed limits are associated with a greater risk of crashes at uncontrolled intersections (10).

In addition, high speed limits and high vehicle speeds increase the risk of severe pedestrian injuries (12,13).

- Street segments with sidewalks tend to have fewer fatal pedestrian crashes than segments without sidewalks (14,15).
- Median islands can help mitigate pedestrian crash risk at midblock locations, uncontrolled crossings, and signalized intersections (6,7,9).
- Mid-block locations with a signalized pedestrian crossing or high-intensity activated crosswalk (HAWK) signal have higher rates of motorists yielding to pedestrians than other types of crossing treatments (16).
Studies have also identified relationships between neighborhood characteristics and pedestrian safety. For example:

- More pedestrian crashes tend to occur in urban areas (12,17).
- Pedestrian crash rates tend to be positively associated with proximity to alcohol sales establishments, bus stops, and median annual neighborhood incomes of less than $25,000 (6).
- The presence of schools, parks, and malls are all associated with a greater frequency of pedestrian crashes (18,19).

One possible explanation for these influences is that many land use characteristics are related to both pedestrian exposure and human behavior. For example, pedestrian activity levels in cities tend to be higher than rural areas. Locations near alcohol sales establishments may have more pedestrians and drivers consuming alcohol.

Certain pedestrian and driver factors are associated with pedestrian crash risk (20,21). For example:

- Males tend to be involved in more pedestrian crashes than females (21).
- Children are more likely to be involved in crashes after darting into the street (21).
- Older pedestrians tend to experience more serious and fatal injury crashes when struck by motor vehicles (21).
- Drivers approaching non-signalized crossing locations at higher speeds are less likely to yield to pedestrians in crosswalks (22).
- Studies of pedestrian and driver behavior related to marked crosswalks have found mixed results (23,24).

Many studies have shown that higher motor vehicle and pedestrian volumes are associated with more pedestrian crashes. For example:

- Higher traffic volumes are associated with more pedestrian crashes at intersections and at uncontrolled crossings of arterial and collector roadways (6,9).
- While the total number of pedestrian crashes at a particular location or in a particular community increases as pedestrian volume increases, several studies suggest that this increase tends to be non-linear. All else equal, a location with 100 percent more pedestrians may only have 30 to 60 percent more (rather than 100 percent more) reported crashes or injuries (6,25,26).

Methodologies for Assessing Pedestrian Crash Risk

Researchers have used many different methods to assess pedestrian risk, including crash analysis, behavior analysis, and expert ratings (4,6). This section focuses on studies that have used reported crash data.

Police crash records have been used to identify intersection design, surrounding neighborhood, and pedestrian and driver factors associated with a higher frequency of pedestrian crashes and injuries (12,21,27,28). However, many of these crash-based studies do not account for differences in pedestrian exposure at different locations; some only use population data as a proxy for exposure (4).

Another group of studies have used crash data to develop pedestrian crash prediction models (29). Many of these models are based only on pedestrian and motor vehicle volumes and do not identify other factors associated with pedestrian crash risk (26,30,31).

A relatively small number of researchers have accounted for a combination of exposure and other roadway design factors for predicting pedestrian crashes at specific locations (6). This type of approach is challenging because pedestrian crashes tend to occur relatively infrequently at any particular roadway segment or intersection location (32). In addition, few agencies have pedestrian counts at specific locations (17), and there is little information available for extrapolating short counts to annual or multi-year time periods, which are needed for comparisons with pedestrian crash data (33).
A recent study extrapolated pedestrian volumes from daytime count periods in order to develop a model to predict the annual number of pedestrian crashes at intersections with different traffic volumes and geometric design characteristics (6). The extrapolation method accounted for some temporal and spatial differences in pedestrian volume patterns. However, better estimates could be obtained if more detailed data about pedestrian activity patterns were available. This would include accounting for weekday and weekend pedestrian activity, seasonal changes in pedestrian volumes, and differences in pedestrian activity patterns near certain land uses (e.g., schools, neighborhood commercial areas).

METHODOLOGY
This section describes the process used to analyze the relationship between pedestrian risk and intersection characteristics at a sample of intersections in Alameda County, California. Data used in the analysis include traffic crashes, pedestrian counts, intersection site characteristics, and characteristics of the area surrounding the study intersections.

Study Area
The analysis focused on 81 of the 7,466 intersections along arterial and collector roadways in Alameda County. Alameda County (Census Bureau 2008 estimated population 1.47 million (34)) is part of the San Francisco Bay Metropolitan Region. It contains intersections in urban, suburban, and exurban areas with a variety of designs and variety of pedestrian activity levels. Oakland is the largest city in the county (population 401,000).

Study Intersections
The intersection selection process was designed to capture a range of intersection site characteristics and surrounding neighborhood characteristics (35). The 81 study intersections were spread throughout the county. The selected intersections included:
- 50 intersections with traffic signals and 26 intersections with an uncontrolled mainline roadway (mainline is defined as the roadway with higher automobile traffic volume; cross-street has lower volume).
- 64 intersections with four roadway legs (approaches); 17 intersections with three legs.
- 42 intersections with marked crosswalks on all sides; seven intersections with no marked crosswalks.
- 34 intersections where the mainline roadway had at least five lanes (including right- and left-turn lanes) for pedestrians to cross; 12 intersections where the mainline roadway had only two lanes for pedestrians to cross.
- 54 intersections with at least one left-turn-only lane; 27 intersections with at least one right-turn-only lane; 25 intersections with no designated turning lanes.
- 10 intersections with medians on all roadway legs; 33 intersections with no medians.
- 29 intersections with at least one non-residential driveway within 50 feet (15 m).
- 35 intersections within 0.25 miles (402 m) of at least one elementary, middle, or high school.
- 19 intersections in commercial retail corridors (at least 10 commercial retail properties within 0.1 miles (161 m).
- 10 intersections in employment centers (at least 2,000 jobs within 0.25 miles (402 m)).
- 44 intersections in neighborhoods where the median annual household income was less than $50,000; 13 intersections in neighborhoods with median annual income less than $30,000.
- 30 intersections in neighborhoods where more than 25 percent of residents were younger than age 18.

Definitions of the intersection variables are listed in TABLE 1. Specific intersection data were collected from field observations, high-resolution aerial photography, the U.S. Census, the San Francisco
Pedestrian Crashes

Pedestrian crash data were obtained from the California Highway Patrol Statewide Integrated Traffic Records System (SWITRS) database for the 10-year period between 1998 and 2007. This analysis focuses on crashes that occurred at or within 50 feet (15 m) of each intersection. This definition of an intersection crash is consistent with the measure of pedestrian exposure (e.g., pedestrian crossing volume) explained in the paragraph below. Of the 81 intersections, 36 (44 percent) had experienced at least one pedestrian crash during the study period. Eight intersections had more than five crashes, and one intersection had 10 crashes.
### TABLE 1. Study Variables

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRASH AND EXPOSURE CHARACTERISTICS</strong></td>
<td></td>
</tr>
<tr>
<td>PedCrash</td>
<td>Number of police-reported pedestrian crashes at or within 50 feet (15 m) of each intersection from 1998 to 2007$^1$</td>
</tr>
<tr>
<td>PedCrossings</td>
<td>Estimated number of pedestrian crossings at the intersection in 10 years$^2$</td>
</tr>
<tr>
<td>lnPedCrossings</td>
<td>Natural logarithm of estimated number of pedestrian crossings at the intersection in 10 years$^2$</td>
</tr>
<tr>
<td>VehicleVolume</td>
<td>Estimated 10-year motor vehicle traffic volume on the mainline roadway passing through the intersection$^3$</td>
</tr>
<tr>
<td>InVehicleVolume</td>
<td>Natural logarithm of estimated 10-year motor vehicle traffic volume on the mainline roadway passing through the intersection$^3$</td>
</tr>
<tr>
<td><strong>INTERSECTION SITE CHARACTERISTICS</strong></td>
<td></td>
</tr>
<tr>
<td>TrafficSignal</td>
<td>Intersection is controlled by a traffic signal (Yes=1, No=0)</td>
</tr>
<tr>
<td>TIntersection</td>
<td>Intersection is a &quot;T&quot; intersection (Yes=1, No=0)$^4$</td>
</tr>
<tr>
<td>NoControl</td>
<td>Traffic on mainline roadway is not controlled by a traffic signal or stop sign (Yes=1, No=0)$^5$</td>
</tr>
<tr>
<td>MainlineWidth</td>
<td>Average curb-to-curb length (feet) of the 2 crosswalks across the mainline road$^5$</td>
</tr>
<tr>
<td>TotalXW</td>
<td>Average number of lanes on mainline approaches to the intersection (including turning lanes)$^6$</td>
</tr>
<tr>
<td>MainlineXW</td>
<td>Proportion of crosswalks across the mainline roadway that are marked (2 marked crosswalks = 1; 1 marked crosswalk = 0.5)$^3$</td>
</tr>
<tr>
<td>MainlineMedian</td>
<td>Proportion of crosswalks across the mainline roadway that have medians (2 legs with medians = 1; 1 leg with median = 0.5)$^6$</td>
</tr>
<tr>
<td>CrossStreetLanes</td>
<td>Average number of lanes on cross-street approaches to the intersection (including turning lanes)$^7$</td>
</tr>
<tr>
<td>CrossStreetXW</td>
<td>Proportion of crosswalks across the cross-street that are marked (2 marked crosswalks = 1; 1 marked crosswalk = 0.5)$^5$</td>
</tr>
<tr>
<td>CrossStreetMedian</td>
<td>Proportion of crosswalks across the cross-street roadway that have medians (2 legs w/ medians = 1; 1 leg w/ median = 0.5)$^6$</td>
</tr>
<tr>
<td>TotalLanes</td>
<td>Total number of lanes (sum of the number of lanes pedestrians are required to cross on all sides of the intersection)</td>
</tr>
<tr>
<td>TotalXW</td>
<td>Total number of marked crosswalks (sum of the number of marked crosswalks on all sides of the intersection)</td>
</tr>
<tr>
<td>PercentXW</td>
<td>Proportion of all crosswalks with a marked crosswalk</td>
</tr>
<tr>
<td>NonResDriveways</td>
<td>Number of non-residential driveways within 50 feet of intersection crosswalks (total of all legs)</td>
</tr>
<tr>
<td>CurbRadius</td>
<td>Curb radius category (&lt;15 feet (&lt;5.7 m)=1, 15-25 feet=2, &gt;25 feet (&gt;=7.62 m)=3)$^3$</td>
</tr>
<tr>
<td>MissingSidewalks</td>
<td>Number of pedestrian approaches with missing sidewalks (a typical 4-leg intersection has 8 pedestrian approaches)</td>
</tr>
<tr>
<td>LeftTurnOnlyLanes</td>
<td>Sum of left-turn-only lanes on all intersection approaches (shared straight-left lanes not included)</td>
</tr>
<tr>
<td>LeftTurnOnlyLanesPresent</td>
<td>One or more left-turn-only lanes present at the intersection (Yes = 1, No = 0)</td>
</tr>
<tr>
<td>RightTurnOnlyLanes</td>
<td>Sum of right-turn-only lanes on all intersection approaches (shared straight-right lanes not included)</td>
</tr>
<tr>
<td>RightTurnOnlyLanesPresent</td>
<td>One or more right-turn-only lanes present at the intersection (Yes = 1, No = 0)</td>
</tr>
<tr>
<td>RightTurnIslands</td>
<td>Separated right turn lanes (with a divider island) on all intersection approaches</td>
</tr>
<tr>
<td><strong>SURROUNDING LAND USE AND TRANSPORTATION SYSTEM CHARACTERISTICS</strong></td>
<td></td>
</tr>
<tr>
<td>TotalPopulation</td>
<td>Total population within 0.25 miles (402 m) from Census Block Groups (2000)</td>
</tr>
<tr>
<td>TotalEmployment</td>
<td>Total number of jobs within 0.25 miles (402 m) from MTC Traffic Analysis Zones (2005)</td>
</tr>
<tr>
<td>CommercialProperties</td>
<td>Number of commercial properties within 0.10 miles (161 m) from Alameda County Assessor's Office parcels (2007)</td>
</tr>
<tr>
<td>Schools</td>
<td>Number of elementary, middle, high, and other schools within 0.10 miles (161 m) from Alameda County parcels (2007)$^7$</td>
</tr>
<tr>
<td>RailStations</td>
<td>Number of regional rail transit stations within 0.10 miles (161 m) from MTC (2008)</td>
</tr>
<tr>
<td>BusStops</td>
<td>Number of bus route stops within 0.10 miles (161 m) from MTC (2008)$^{10}$</td>
</tr>
<tr>
<td>TrailMiles</td>
<td>Total multi-use trail centerline distance (miles) within 0.10 miles (161 m) from MTC (2008)</td>
</tr>
<tr>
<td>StreetMiles</td>
<td>Total street centerline distance (miles) within 0.10 miles (161 m) from Alameda County (2007)</td>
</tr>
<tr>
<td>FreewayPresence</td>
<td>Freeway presence within 0.10 miles (161 m) (Yes = 1, No = 0) from Alameda County (2007)</td>
</tr>
<tr>
<td><strong>SURROUNDING NEIGHBORHOOD SOCIOECONOMIC CHARACTERISTICS</strong></td>
<td></td>
</tr>
<tr>
<td>ProportionMale</td>
<td>Proportion of population within 0.25 miles (402 m) that is male from Census Block Groups (2000)</td>
</tr>
<tr>
<td>ProportionVehicle</td>
<td>Proportion of households within 0.25 miles (402 m) that have no automobile from Census Block Groups (2000)</td>
</tr>
<tr>
<td>MedianIncome</td>
<td>Median income (1999 dollars) of households within 0.25 miles (402 m) from Census Block Groups (2000)$^{11}$</td>
</tr>
<tr>
<td>ProportionUnder18</td>
<td>Proportion of population within 0.25 miles (402 m) that is under 18 years old from Census Block Groups (2000)</td>
</tr>
<tr>
<td>ProportionOver64</td>
<td>Proportion of population within 0.25 miles (402 m) that is over 64 years old from Census Block Groups (2000)</td>
</tr>
</tbody>
</table>

1) Reported pedestrian crashes were gathered from the California Highway Patrol Statewide Integrated Traffic Records System.

2) Pedestrian crossing volumes are adjusted for control for differences in time of day, day of week, season of year, surrounding land use, and weather when count was taken.


4) "T" intersections are 3-way intersections. Intersections were not considered to be "T" intersections if the fourth approach was a commercial driveway.

5) Mainline roadway is the intersecting roadway with the higher traffic volume; cross-street has the lower traffic volume (estimated).

6) Curb-to-curb length is measured as the shortest possible crossing distance within each crosswalk.

7) Average number of lanes on each mainline approach includes all through, left, and right-turn lanes.

8) Curb radius category reflects the average estimated curb radius of all corners at the intersection.

9) Total schools does not include colleges. There are not enough intersections near colleges to provide conclusive findings about the relationship between college campuses and pedestrian safety.

10) The number of "bus route stops" is the sum of the number of different bus routes servicing each bus stop within a given distance of the intersection (e.g., if 4 routes service a single bus stop, that particular bus stop will be counted 4 times).

11) Median income is calculated as the weighted average of median incomes reported for the census block groups surrounding the intersection. Weights are assigned based on the proportion of the census block group within the specific buffer distance from the intersection.
**Pedestrian Volumes**

The crash risk analysis required an estimate of the total number of pedestrians crossing each intersection during a 10-year period. This pedestrian volume estimate was derived from a combination of manual counts and automated sensor counts. Manual counts were collected during two different two-hour periods at each study intersection during Spring 2008 (50 intersections) and Spring 2009 (31 intersections). One count period was on a weekday afternoon (Tuesday, Wednesday, or Thursday) and one was on a Saturday. Pedestrians were counted each time they crossed a leg of the intersection. This included people crossing within the crosswalk and people crossing the roadway leg up to 50 feet (15 m) from the crosswalk, which corresponded with the proximity measure used to define intersection crashes. A single pedestrian could be counted multiple times if he or she crossed multiple legs of the intersection, since each crossing represented a unique opportunity for conflict with vehicles. For this study, pedestrian volumes at three-leg “T-intersections” only included the number of pedestrians crossing each roadway.

The two manual counts were extrapolated to estimate an annual volume at each intersection. Extrapolation was based on the “typical” Alameda County pedestrian volume pattern, which was calculated from automated sensor counts at 13 locations from April 2008 to May 2009. Adjustment factors were used to account for deviations from the typical pattern depending on nearby land uses (central business district, residential neighborhood, commercial corridor, near multi-use trail, near school), the weather when the count was taken (e.g., rain, clouds, cool temperatures, or warm temperatures), and season (TABLE 2). The weather, land use, and season effects were viewed independently, so more than one adjustment could be made to each count.

Annual estimated volumes were multiplied by 10 to approximate the 10-year pedestrian volume at each study intersection. No data were available to estimate annual increases or decreases in volumes during the 10-year period, so no annual factors were applied. More information about the pedestrian volume extrapolation methods used in Alameda County is provided in other references (36).
TABLE 2. Alameda County Pedestrian Volume Adjustment Factors

<table>
<thead>
<tr>
<th>Land Use Category</th>
<th>Definition</th>
<th>Count Times when Adjustment Factors were Applied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Weekday 12-2 p.m.</td>
</tr>
<tr>
<td>Employment Center</td>
<td>&gt;=2,000 jobs within 0.25 miles (402 m)$^4$</td>
<td>0.83</td>
</tr>
<tr>
<td>Residential Area</td>
<td>&gt;=500 jobs within 0.25 miles (402 m)$^5$ &amp; no commercial retail properties within 0.1 miles (161 m)$^3$</td>
<td>1.37</td>
</tr>
<tr>
<td>Neighborhood Commercial Area</td>
<td>&gt;=10 commercial retail properties within 0.1 miles (161 m)$^3$</td>
<td>0.92</td>
</tr>
<tr>
<td>Near Multi-Use Trail</td>
<td>&gt;=0.5 centerline miles of multi-use trails within 0.25 miles (402 m)$^6$</td>
<td>1.63</td>
</tr>
<tr>
<td>Near School</td>
<td>&gt;=1 elementary, middle, or high school within 0.25 miles (402 m)$^6$</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Weather Adjustment Factors (Counts taken under certain weather conditions were multiplied by these factors to match counts taken during typical Alameda County weather conditions)$^2$

<table>
<thead>
<tr>
<th>Weather Condition</th>
<th>Definition</th>
<th>Count Times when Adjustment Factors were Applied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Weekday 12-6 p.m.</td>
</tr>
<tr>
<td>Warm</td>
<td>&gt;=80 degrees Fahrenheit (27 degrees Celsius) during first count hour$^7$</td>
<td>1.07</td>
</tr>
<tr>
<td>Cool</td>
<td>&lt;=50 degrees Fahrenheit (10 degrees Celsius) during first count hour$^7$</td>
<td>1.10</td>
</tr>
<tr>
<td>Cloudy</td>
<td>&lt;=0.6 of the expected solar radiation (Langleys per day) during first count hour$^7,8$</td>
<td>1.11</td>
</tr>
<tr>
<td>Rain</td>
<td>&gt;0.01 inch (0.254 mm) of precipitation during either count hour$^7$</td>
<td>1.27</td>
</tr>
</tbody>
</table>

Seasonal Adjustment Factors (Counts taken from April through June were multiplied by these factors to match counts taken in Alameda County during a typical time of the year)$^3$

<table>
<thead>
<tr>
<th>Land Use Category</th>
<th>Definition</th>
<th>Count Times when Adjustment Factors were Applied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All Time Periods</td>
</tr>
<tr>
<td>Employment Center</td>
<td>&gt;=2,000 jobs within 0.25 miles (402 m)$^4$</td>
<td>0.98</td>
</tr>
<tr>
<td>Residential Area</td>
<td>&gt;=500 jobs within 0.25 miles (402 m)$^5$ &amp; no commercial retail properties within 0.1 miles (161 m)$^3$</td>
<td>0.97</td>
</tr>
<tr>
<td>Neighborhood Commercial Area</td>
<td>&gt;=10 commercial retail properties within 0.1 miles (161 m)$^3$</td>
<td>0.98</td>
</tr>
<tr>
<td>Near Multi-Use Trail</td>
<td>&gt;=0.5 centerline miles of multi-use trails within 0.25 miles (402 m)$^6$</td>
<td>0.91</td>
</tr>
<tr>
<td>Near School</td>
<td>&gt;=1 elementary, middle, or high school within 0.25 miles (402 m)$^6$</td>
<td>0.93</td>
</tr>
</tbody>
</table>

1) Land use adjustment factors based on hourly automated sensor counts taken at 13 locations in Alameda County between April 2008 and June 2009.
2) Weather adjustment factors based on hourly automated sensor counts taken at 13 locations in Alameda County between April 2008 and June 2009.
3) Employment center, residential area, neighborhood commercial area, and multi-use trail seasonal adjustment factors based on hourly automated sensor counts taken at 13 locations in Alameda County from April 2008 to June 2009. School seasonal adjustment factor based on hourly automated sensor counts taken at 13 locations in Alameda County from May 2009 to June 2009.
4) Source = Traffic Analysis Zones from San Francisco Bay Area Metropolitan Transportation Commission, 2005
5) Source = Land Use Parcels from Alameda County Tax Assessor's Office, 2007
6) Source = Bay Area Multi-Use Trail Centerlines from San Francisco Bay Area Metropolitan Transportation Commission, 2007
8) Solar radiation measurements from the previous 4 to 10 years at each of the three Alameda County weather stations were used to calculate the expected solar radiation measurement for every hour of the year. The weather condition was determined to be "cloudy" if the ratio of the current measurement was <= 0.6 of the expected solar radiation for that specific hour. The threshold was set at 0.6 to match as closely as possible to field data collectors' subjective determinations of when the weather was "cloudy".

Paper revised from original submittal.
ANALYSIS

This section describes how the data were used to identify relationships between pedestrian crash risk and intersection characteristics.

Preliminary Comparison of Pedestrian Crash Rates and Intersection Characteristics

The number of reported pedestrian crashes during the 10-year study period was divided by the 10-year pedestrian volume to estimate a pedestrian crash rate at each intersection. For the 36 intersections with reported pedestrian crashes, crash rates ranged from 0.70 crashes to 98 crashes per 10 million crossings.

As a preliminary step, the characteristics of the 18 intersections with the highest crash rates (7.2 to 98 crashes per 10 million crossings) were compared with the characteristics of the 18 intersections with the lowest crash rates (0.70 to 7.2 crashes per 10 million crossings). Several variables appeared to be associated with higher crash rates. These variables included intersections with three approaches and the number of right-turn-only lanes, number of left-turn-only lanes, number of right-turn islands, and number of approaches with missing sidewalks at the intersection. Other variables appeared to be associated with lower crash rates, including the number of pedestrian crossings at the intersection as well as the total employment, number of commercial properties, number of rail stations, number of bus stops, presence of a freeway, and proportion of households without a motor vehicle near the intersection.

However, this preliminary approach to analyzing crash rates did not adequately represent intersections that experienced no pedestrian crashes during the study period. In addition, it did not control for correlation among variables. For example, pedestrian volume was correlated ($|\rho| > 0.6$) with total employment, number of bus stops, and number of rail stations, which may explain the preliminary association between these variables and lower crash rates. Not accounting for the correlations among variables was also likely to mask the effects of other variables on pedestrian crash risk. A statistical modeling approach helped address these issues.

Statistical Model of Pedestrian Intersection Crashes

The purpose of modeling was to identify intersection characteristics that had a statistically-significant relationship with the occurrence of pedestrian crashes. The total number of crashes reported at each intersection from 1998 to 2007 was the dependent variable used in the modeling process. Since crashes are count data, a Poisson model was considered. Because the statistical distribution of the number of crashes per intersection did not meet the requirement that the mean be roughly equal to the variance, a negative binomial regression model was used to represent the count data. This is a common modeling approach for traffic crashes (6,37).

Equation 1 shows the model structure:

$$PedCrashes_i = e^{(\alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + ...)}$$

where:

$PedCrashes_i =$ total number of reported crashes at intersection $i$ from 1998 to 2007

$X_{ij} =$ quantitative measure of each characteristic $j$ associated with intersection $i$

$\beta_j =$ coefficient corresponding to $X_{ij}$ to be determined by negative binomial regression

$\alpha =$ constant to be determined by negative binomial regression
Variables Tested
A variety of model specifications were tested to explore the effects of all of the variables listed in TABLE 1 on intersection pedestrian safety. Each of the model specifications that were considered included pedestrian volume and motor vehicle volume variables plus other explanatory variables. Since there were only 81 intersections available for analysis, it was not desirable to create a model with all possible variables. Several steps were used to narrow the list of variables:

- In order to reduce potential bias due to collinearity, pairs of variables with correlations of $|\rho| > 0.6$ were not included in the same model. The variable that improved the overall model log-likelihood and produced more significant parameter estimates was kept for further testing.
- After estimating a model with all remaining variables, the variables with the least statistically-significant parameter estimates were removed. Then the model was estimated again.
- The variable removal process stopped when all variable parameter estimates had high statistical significance ($p < 0.05$).
- To test for consistency, the process was repeated multiple times by removing variables in a different order.

The natural logarithm form of the exposure variables was tested during this process because previous studies have shown a non-linear association between pedestrian and automobile volumes and pedestrian crashes. Comparing the natural logarithm form with the linear form of these exposure variables in different models showed that the natural logarithm form was a better fit for the data. Several interaction variables were also tested, including the product of pedestrian volume and motor vehicle volume, the quotient of crossing width and number of lanes, and the product of no traffic control and marked crosswalks. These variables did not improve the model.

One variable not included in the final model was the dummy variable representing three-leg intersections. Since the number of intersection approach legs represents a major design characteristic, additional analyses were conducted to determine if this factor had an influence on pedestrian crashes. A separate set of models were estimated using only the 64 intersections that had four legs, and the preferred model from that set had the same variables and similar coefficients to the model with all 81 intersections. Therefore, the model with all intersections was kept as the final model.

RESULTS
The final statistical model suggests that several intersection characteristics have a significant association with pedestrian crashes. This section discusses the preferred pedestrian crash risk model and then addresses specific intersection factors that may contribute to pedestrian crash risk.

Overall Pedestrian Crash Model
The intersection pedestrian crash prediction model is presented in TABLE 3. This model is significantly better than a model based only on constant values, and it has eight explanatory factors that are statistically significant ($p < 0.05$). The model log likelihood statistic is higher than the log likelihood of other alternative models with statistically-significant variables. Equation 2 shows the model formula.
\[ \text{PedCrashes}_i = e^{(-37.3 + 0.577x_1 + 1.50x_2 - 1.37x_3 - 1.24x_4 + 0.425x_5 + 0.286x_6 + 0.0371x_7 + 6.83x_8)} \]

where:

- \( \text{PedCrashes}_i \) = predicted number of reported crashes at any intersection \( i \) during a 10-year period
- \( X_1 \) = lnPedCrossings at intersection \( i \)
- \( X_2 \) = lnVehicleVolume at intersection \( i \)
- \( X_3 \) = MainlineMedian at intersection \( i \)
- \( X_4 \) = CrossStreetMedian at intersection \( i \)
- \( X_5 \) = RightTurnOnlyLanes at intersection \( i \)
- \( X_6 \) = NonResDriveways at intersection \( i \)
- \( X_7 \) = CommercialProperties at intersection \( i \)
- \( X_8 \) = ProportionUnder18 at intersection \( i \)
### TABLE 3. Intersection Pedestrian Crash Model

<table>
<thead>
<tr>
<th>Model Variables</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Z-test</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LnPedCrossings(^2)</td>
<td>0.577</td>
<td>0.162</td>
<td>3.56</td>
<td>0.000</td>
</tr>
<tr>
<td>LnVehicleVolume(^3)</td>
<td>-1.50</td>
<td>0.425</td>
<td>3.54</td>
<td>0.000</td>
</tr>
<tr>
<td>MainlineMedian(^4)</td>
<td>-1.37</td>
<td>0.424</td>
<td>-3.24</td>
<td>0.001</td>
</tr>
<tr>
<td>CrossStreetMedian(^5)</td>
<td>-1.24</td>
<td>0.584</td>
<td>-2.13</td>
<td>0.033</td>
</tr>
<tr>
<td>RightTurnOnlyLanes(^6)</td>
<td>0.425</td>
<td>0.192</td>
<td>2.21</td>
<td>0.027</td>
</tr>
<tr>
<td>NonResDriveways(^7)</td>
<td>0.286</td>
<td>0.125</td>
<td>2.29</td>
<td>0.022</td>
</tr>
<tr>
<td>CommercialProperties(^8)</td>
<td>0.0371</td>
<td>0.0144</td>
<td>2.57</td>
<td>0.010</td>
</tr>
<tr>
<td>ProportionUnder18(^9)</td>
<td>6.83</td>
<td>2.36</td>
<td>2.89</td>
<td>0.004</td>
</tr>
<tr>
<td>Constant</td>
<td>-37.3</td>
<td>8.46</td>
<td>-4.41</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Overall Model**

<table>
<thead>
<tr>
<th>Sample Size (N)</th>
<th>81 intersections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Likelihood</td>
<td>-95.2</td>
</tr>
<tr>
<td>Likelihood Ratio ChiSq (8 df)</td>
<td>55.7</td>
</tr>
<tr>
<td>Probability &gt; ChiSq</td>
<td>0.000</td>
</tr>
<tr>
<td>Pseudo R-Squared(^10)</td>
<td>0.227</td>
</tr>
<tr>
<td>Overdispersion Parameter</td>
<td>0.307</td>
</tr>
<tr>
<td>Significance of Overdispersion(^11)</td>
<td>0.012</td>
</tr>
</tbody>
</table>

1) The dependent variable for the pedestrian crash model is the number of reported crashes at or within 50 feet (15 m) of the center of each study intersection from 1998 to 2007. Overdispersion is modeled using the mean method.

2) LnPedCrossings = Natural logarithm of estimated number of pedestrian crossings at the intersection in 10 years (adjusted to control for time of day, day of week, season of year, surrounding land use, and weather when count was taken).

3) LnVehicleVolume = Natural logarithm of estimated 10-year motor vehicle traffic volume on the mainline roadway passing through the intersection.

4) MainlineMedian = Proportion of crosswalks across the mainline roadway that have medians (2 medians = 1.0; 1 median = 0.5). Mainline roadway is the intersecting roadway with the higher motor vehicle traffic volume; cross-street has the lower traffic volume (estimated).

5) CrossStreetMedian = Proportion of crosswalks across the cross-street roadway that have medians (2 medians = 1.0; 1 median = 0.5). Cross-street roadway is the intersecting roadway with the lower motor vehicle traffic volume (estimated). Mainline has higher traffic volume (estimated).

6) RightTurnOnlyLanes = Sum of right-turn-only lanes on all intersection approaches (shared straight-right lanes not included).

7) NonResDriveways = Number of non-residential driveways within 50 feet of intersection crosswalks (total of all legs).

8) CommercialProperties = Number of commercial properties within 0.10 miles (161 m).

9) ProportionUnder18 = Proportion of population within 0.25 miles (402 m) that is under 18 years old.

10) The Pseudo R-Squared is not the same type of measure as the R-Squared statistic used in ordinary least squares regression. Pseudo R-Squared is a ratio of log likelihood values. It does not represent the proportion of variance in pedestrian crashes explained by the predictor variables.

11) The significance of the overdispersion parameter is the result of a likelihood ratio chi-square test that this parameter is equal to zero. Since the test statistic is significant (<0.05), it is likely that the pedestrian crash data are over-dispersed and is not sufficiently described by the Poisson distribution. This shows that the Negative Binomial distribution is preferred.
Intersection Characteristics Associated with Pedestrian Crashes

The model shows eight characteristics that have a statistically-significant relationship with pedestrian crash propensity at intersections. This section suggests possible reasons for these relationships. Additional field analysis of pedestrian and motorist interactions at each study site could help provide more evidence to support these statistical relationships.

- After controlling for other factors, intersections with higher pedestrian volumes and mainline motor vehicle volumes tended to have more pedestrian crashes.
- An intersection with 100 percent more pedestrian crossings is expected to have approximately 49 percent more crashes (fewer than 100 percent). This finding is similar to previous studies, which suggest that increasing the number of pedestrians crossing an intersection reduces the risk of any individual pedestrian being injured in a crash, independent of all other changes to the local environment (6,25,26).
- An intersection with 100 percent more mainline traffic volume is expected to have 183 percent more crashes (more than 100 percent). This indicates a stronger positive relationship between motor vehicle volume and pedestrian crashes than was found in previous studies (38). One possible explanation for this result may be related to traffic volumes and congestion levels. As traffic volumes increase towards the capacity of a roadway, traffic speeds tend to decrease, which is expected to result in fewer pedestrian crashes. Previous studies may have included a wide range of congestion levels. However, most of the roadways included in this study operated with relatively little congestion at most times of day (only two mainline roadways had more than 9,000 vehicles per lane per day). Under less-congested conditions, the frequency of pedestrian crashes may increase more rapidly as traffic volume increases (39). This relationship requires further study.
- The proportion of mainline and cross-street legs with medians were both negatively associated with pedestrian crashes in all model alternatives. Medians may offer a refuge for pedestrians in the middle of a roadway crossing and may allow pedestrians to concentrate on crossing one direction of traffic at a time.
- The number of right-turn only lanes at intersections was positively associated with pedestrian crashes. This may indicate that intersections with right-turn lanes tend to have longer crossing distances and a more complex set of interactions between pedestrians and motorists. It could also indicate a tendency for more right-turn-on-red collisions.
- The number of non-residential driveways within 50 feet (15 m) of each intersection was positively associated with pedestrian crashes. This suggests that driveways represent additional conflict points between motor vehicles and pedestrians near the intersection. Drivers may be paying more attention to interactions with other vehicles at the intersection and may not look carefully for pedestrians as they exit driveways across the sidewalk.
- The number of commercial retail properties within 0.1 miles (161 m) of the intersection was positively associated with pedestrian crashes. This may suggest that commercial corridors may have particularly risky interactions between vehicles and pedestrians. Drivers may be concentrating on finding parking spaces or looking for particular stores or restaurants, while pedestrians may be crossing streets between cars or outside of crosswalks to take the most direct route to a store entrance or other destination.
- The percentage of neighborhood residents living within 0.25 miles (402 m) of the intersection who are younger than age 18 was positively associated with pedestrian crashes. This may indicate that neighborhoods with more children have slightly more dangerous pedestrian crossing behavior than other neighborhoods.

The model can be used to understand the relationship between intersection characteristics and pedestrian safety. According to the model equation, as pedestrian volume increases, the expected number of pedestrian crashes increases at a decreasing rate (FIGURE 1A). As pedestrian volume increases, the expected risk of a crash for each individual crossing decreases (FIGURE 1B). Both graphs illustrate that medians can help improve pedestrian safety at intersections.
FIGURE 1. Relationship between Pedestrian Volume, Presence of Medians, and Intersection Safety, as Predicted by the Model Equation

A. Pedestrian Volume, Median Presence, and Predicted Pedestrian Crashes

B. Pedestrian Volume, Median Presence, and Predicted Pedestrian Risk

Notes:
1) Model input variables use the mean values from the 81 study intersections.
2) Assumes medians are present on all mainline and cross-street approaches. All other model input variables use mean values from the 81 study intersections.
3) Assumes no medians are present on any intersection approach. All other model input variables use mean values from the 81 study intersections.
FUTURE RESEARCH

A number of issues should be addressed in future research. These include incorporating additional intersection site variables, controlling for differences in pedestrian and driver behavior, improving the quality of data used in the analysis, and considering non-safety aspects of intersection design.

Additional Intersection Site Variables

Many factors are likely to influence pedestrian crash risk. While this study accounted for a variety of intersection site characteristics, it was not possible to capture all aspects of intersections that may contribute to crashes in specific locations. Crashes may also be related to factors such as traffic signal cycle length, pedestrian signal timing, traffic congestion, and turning traffic speeds. Resources were not available to collect these data, but future research could explore these factors.

Pedestrian and Driver Behaviors

Behaviors such as speeding, yielding, jaywalking, and traffic control violations may have an effect on pedestrian safety. Future studies of intersection design characteristics and pedestrian safety could also attempt to control for differences in pedestrian and driver behaviors around intersections. Some of these behaviors may be different in Alameda County than in other communities.

Data Improvements

The analysis used pedestrian crash data from 1998 to 2007 to identify intersection characteristics associated with pedestrian crash risk. Intersection characteristics were gathered from field observations and aerial images during 2008 and 2009. Since most urbanized areas of Alameda County have been built-out for decades, it is likely that most intersections have changed little during the ten-year study period. A comparison of recent observations with a 1993 aerial photograph showed that five of the study intersections had been expanded during the past 15 years. However, it is not known if or when these changes occurred during the study period.

Given that most of the study intersections had fewer than five reported pedestrian crashes in 10 years, a single pedestrian crash can make a significant difference in crash risk at a particular location. While this analysis has attempted to use the most accurate crash and exposure data available, it is still subject to this limitation. Future studies can reduce the impact of this type of variation by collecting data at more intersections. A larger intersection sample size could also show that more of the variables considered in this study have a statistically-significant association with pedestrian crashes.

Significant effort was made to generate reliable pedestrian exposure data. The 10-year pedestrian volume estimates were extrapolated from two different two-hour counts. While averaging the estimates generated from two manual counts is more reliable than using a single count, these counts are subject to the random variations in pedestrian activity that occur from day-to-day. Five intersections were counted using the same method in both 2008 and 2009. The differences in volumes at these sites between 2008 and 2009 ranged from two percent to 33 percent. It is likely that conducting more manual counts at different times and gathering continuous pedestrian counts over multiple years would provide even more accurate exposure data.

In addition, pedestrian crossing volumes are only one possible measure of pedestrian exposure to crash risk. While this measure is appropriate for this analysis of intersection crashes, other measures could be tested in future studies. These alternative measures could account for crossing distance, crossing time, and size of pedestrian crossing groups (40). Further, crash risk could be analyzed using individual crosswalks at each intersection as the unit of analysis. This would require data showing the specific crosswalk leg where each crash occurred. This information was not available for this study.

There are potential limitations to the secondary data sources used in this analysis. According to a study of pedestrian crash underreporting at eight hospitals in three states, only 56 percent of pedestrian injuries treated in emergency rooms were matched with a corresponding crash records in state police.
crash databases (41). Since emergency rooms tend to treat the most severe injuries, the rate of
underreporting may be even greater for less severe pedestrian crashes.

Traffic volume data were obtained from the California Department of Transportation and local
jurisdictions. While some local traffic volume data were from the last year, other volumes were nearly a
decade old. Ten of the intersections did not have an applicable traffic count available, so mainline traffic
volumes were estimated based on counts from nearby locations. While resources were not available for
counting motor vehicles and improving traffic volume data in this study, this issue should be addressed in
future studies.

Since the model is based on a relatively small number of intersections in one county, it requires
further testing. This should include comparing the model results with data collected in other communities
outside of Alameda County. Variables such as left-turn-only lanes and three-leg intersections showed
associations with greater numbers of pedestrian crashes during some parts of the analysis process, even
though they were not significant in the final Alameda County model. Therefore, these characteristics
warrant further study in other communities.

Other Aspects of the Pedestrian Environment around Intersections
Intersection planning and design should address safety as well as other pedestrian needs. Other factors
that were not evaluated in this study may or may not have a direct association with pedestrian safety, but
they are important for improving the pedestrian environment near intersections. For example, features
such as curb ramps and accessible pedestrian signals are critical for providing accessibility for all
pedestrians, street trees along sidewalks make walking more pleasant, and shorter building setbacks from
the sidewalk can help increase the convenience of pedestrian travel.

CONCLUSION
This study of 81 intersections in Alameda County, California suggests that certain intersection
characteristics are associated with pedestrian crash risk. After controlling for pedestrian and motor
vehicle volumes, more pedestrian crashes occurred at intersections with more right-turn-only lanes, more
nearby non-residential driveway crossings, more nearby commercial properties, and a larger percentage of
children younger than age 18 living near the intersection. Medians were associated with fewer pedestrian
crashes at intersections. While there is a need for additional research, these findings can help inform
intersection design practices and provide safer conditions for pedestrians.
ACKNOWLEDGEMENTS

The authors would like to thank John Bigham, Swati Pande, and other University of California, Berkeley, Safe Transportation Research & Education Center researchers who geocoded the traffic crashes analyzed in this study. In addition, the authors thank all of the volunteers who collected pedestrian counts and students who entered count data into spreadsheets in Spring 2009. This work was partially funded by the California Department of Transportation and the Alameda County Transportation Improvement Authority. Fellowship support was provided by the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program. Although the research described in the paper has been funded in part by the U.S. Environmental Protection Agency's STAR program through grant (FP-91695101-0), it has not been subjected to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.
REFERENCES


TRB 2010 Annual Meeting CD-ROM  
Paper revised from original submittal.


LIST OF TABLES AND FIGURES

TABLES
1. Study Variables
2. Alameda County Pedestrian Volume Adjustment Factors
3. Intersection Pedestrian Crash Model

FIGURES
1. Relationship between Pedestrian Volume, Presence of Medians, and Intersection Safety, as Predicted by the Model Equation