See discussions, stats, and author profiles for this publication at: [https://www.researchgate.net/publication/23711358](https://www.researchgate.net/publication/23711358_An_area-level_model_of_vehicle-pedestrian_injury_collisions_with_implications_for_land_use_and_transportation_planning?enrichId=rgreq-71322df86e1bba18920e56c5f7f0b3b0-XXX&enrichSource=Y292ZXJQYWdlOzIzNzExMzU4O0FTOjc0ODM4NzYxNTkwNzg0MEAxNTU1NDQwOTIwMDcw&el=1_x_2&_esc=publicationCoverPdf)

[An area-level model of vehicle-pedestrian injury collisions with implications](https://www.researchgate.net/publication/23711358_An_area-level_model_of_vehicle-pedestrian_injury_collisions_with_implications_for_land_use_and_transportation_planning?enrichId=rgreq-71322df86e1bba18920e56c5f7f0b3b0-XXX&enrichSource=Y292ZXJQYWdlOzIzNzExMzU4O0FTOjc0ODM4NzYxNTkwNzg0MEAxNTU1NDQwOTIwMDcw&el=1_x_3&_esc=publicationCoverPdf) for land use and transportation planning

Article in Accident; analysis and prevention · February 2009 DOI: 10.1016/j.aap.2008.10.001 · Source: PubMed

Provided for non-commercial research and education use. Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

Accident Analysis and Prevention 41 (2009) 137–145

Accident Analysis and Prevention

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

An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning

Megan Wier^{a,∗}, June Weintraub^a, Elizabeth H. Humphreys^{a, 1}, Edmund Seto^b, Rajiv Bhatia^c

^a *San Francisco Department of Public Health, Environmental Health Section, Program on Health, Equity and Sustainability, 1390 Market St., Ste. 910, San Francisco, CA 94102, United States*

^b *University of California - Berkeley, School of Public Health, University Hall, Berkeley, CA 94720, United States*

^c *San Francisco Department of Public Health, Environmental Health Section, 1390 Market St., Ste. 822, San Francisco, CA 94102, United States*

article info

Article history: Received 19 April 2008 Received in revised form 8 September 2008 Accepted 9 October 2008

Keywords: Pedestrian injury Traffic collisions Land use Transportation planning Spatial analyses

A B S T R A C T

There is growing awareness among urban planning, public health, and transportation professionals that design decisions and investments that promote walking can be beneficial for human and ecological health. Planners need practical tools to consider the impact of development on pedestrian safety, a key requirement for the promotion of walking. Simple bivariate models have been used to predict changes in vehicle-pedestrian injury collisions based on changes in traffic volume. We describe the development of a multivariate, area-level regression model of vehicle-pedestrian injury collisions based on environmental and population data in 176 San Francisco, California census tracts. Predictor variables examined included street, land use, and population characteristics, including commute behaviors. The final model explained approximately 72% of the systematic variation in census-tract vehicle-pedestrian injury collisions and included measures of traffic volume, arterial streets without transit, land area, proportion of land area zoned for neighborhood commercial and residential-neighborhood commercial uses, employee and resident populations, proportion of people living in poverty and proportion aged 65 and older. We have begun to apply this model to predict area-level change in vehicle-pedestrian injury collisions associated with land use development and transportation planning decisions.

Published by Elsevier Ltd.

1. Introduction

1.1. Pedestrian safety and planning

In the 20th century, pedestrian needs were rare priorities in urban and transportation planning (Frumkin et al., 2004). Yet, environments that support walking can benefit human health by reducing motor vehicle collisions, motor vehicle-related noise and air pollution, and increasing physical activity and social cohesion (Cavill, 2001; Ewing, 2006; Leyden, 2003; Lavizzo-Mourney and McGinnis, 2003). To achieve walkable communities, planning professionals need practical tools to assess and mitigate the impact of development on pedestrian safety, including vehicle-pedestrian collisions.

0001-4575/\$ – see front matter. Published by Elsevier Ltd. doi:10.1016/j.aap.2008.10.001

Traffic collisions are a major cause of mortality in the United States (Mokdad et al., 2004), and the leading cause of death for persons aged 4–34 (Subramanian, 2006). Nationwide, pedestrians account for 11% of motor vehicle collision fatalities, with approximately 4700 pedestrian deaths in 2006 (NHTSA, 2006a). 15% of those people killed while walking died in California (NHTSA, 2006b).

Among California cities, San Francisco has historically had the highest per capita vehicle-pedestrian injury collision rate (STPP and California Walks, 2002). In stark contrast with the national figure of 11%, pedestrians account for half of San Francisco traffic deaths, with 13 fatalities and 726 non-fatal vehicle-pedestrian collisions in 2006. Pedestrian injuries and fatalities in San Francisco have declined over the last decade, attributed to intersection and mid-block pedestrian safety countermeasures, traffic calming, law enforcement, and improved planning efforts. Still, San Francisco's injury rate remains approximately 100/year/100,000 population (CCSF MTA, 2007; U.S. Census Bureau, 2000) or over five times the Healthy People 2010 national target of no greater than 19 pedestrian injuries/year/100,000 people; San Francisco's fatal injury rate of 2/year/100,000 is twice the national target (US DHHS, 2000).

Motor vehicles and pedestrians are two necessary component causes of vehicle-pedestrian injury collisions. San Francisco is a

[∗] Corresponding author. Tel.: +1 415 252 3972; fax: +1 415 252 3964.

E-mail addresses: Megan.Wier@sfdph.org (M. Wier), June.Weintraub@sfdph.org (J. Weintraub), elizahumphreys@gmail.com (E.H. Humphreys), seto@berkeley.edu (E. Seto),

Rajiv.Bhatia@sfdph.org (R. Bhatia).

^{.&}lt;br>1 Present address: Marin Community Clinic, 250 Bon Air Road, Greenbrae, CA 94904, United States.

relatively dense, urban city, with approximately 776,000 residents and over 250,000 additional non-resident employees. By 2025, residential and job growth are expected to increase trips to, from, and within San Francisco by 12% (SFCTA, 2004). Of the projected 5 million trips in 2025, 3.3 million will be within San Francisco and over 50% of those are estimated to be auto trips. Both the relatively high frequency of pedestrian injuries and fatalities and the projected growth in San Francisco's traffic and population underscore the need to prioritize pedestrian safety needs in land use and transportation planning processes.

Currently, limited planning tools are available to evaluate the impacts of land use planning on pedestrian safety conditions. The Pedestrian and Bicycle Crash Analysis Tool software identifies pre-crash actions that lead to collisions, and links them to potential mitigation strategies (PBCAT, 2007). Crossroads software (Crossroads, 2007) and zonal analysis (USDOT, 1998) identify collision patterns and areas with high densities of pedestrian injuries.

Tools for prospectively forecasting the impacts of transportation and land use development on future vehicle-pedestrian collisions would complement the above methods for assessing existing collision patterns. To be useful in a planning context, a vehiclepedestrian injury collision forecasting model needs to be based on available or routinely produced data, provide meaningful, easily interpreted, robust estimates, and be applicable in diverse areas to routine land use and transportation planning decisions. We are not aware of any vehicle-pedestrian injury collision forecasting tools in general use by planners for environmental or health impact assessments.

Empirically, increases in road facility vehicle volume increase the probability of vehicle-pedestrian conflicts on that facility (Lee and Abdel-Aty, 2005). A simple way to forecast change in vehiclepedestrian collisions associated with change in vehicle volume is by applying a *road safety function*—which describes the relationship between traffic volume and collisions. The following power function (1.1) is an empirically supported parametric form of a road safety function, where AADT = Average Annual Daily Traffic:

 \varDelta (%), vehicle-pedestrian collisions

$$
= \left[\left(\frac{\text{Future AADT}}{\text{Baseline AADT}} \right)^{\beta} - 1 \right] \times 100 \tag{1.1}
$$

Typically β < 1, and empirical evidence suggests that 0.5 is a reasonable parameter (Lee and Abdel-Aty, 2005). At β = 0.5, vehiclepedestrian collisions are forecasted to increase proportional to the square root of AADT, with a 50% increase in AADT predicting a 22% increase in collisions. Fig. 1 graphically illustrates the relationship

Fig. 1. Vehicle-pedestrian injury collision increases associated with traffic volume increases: power function and San Francisco final model predictions.

between change in vehicle volume and change in the number of collisions as β varies. Applying this power function (1.1) to estimate collision increases associated with traffic volume changes due to area-level development is more challenging and requires simplifying assumptions, including: (1) development does not affect pedestrian flow and behavior; (2) development does not implement pedestrian safety countermeasures; and (3) AADT changes at intersections or street segments selected for evaluation are reasonable surrogates for changes at adjacent area roadways. (We included an example application in the Appendix A.)

As vehicle volume is not the only variable mediating the impacts of development on vehicle-pedestrian injury collisions, a multivariate area-level model might more robustly predict related change in collisions. In this paper, we describe our development of a contextspecific regression model for forecasting vehicle-pedestrian injury collisions that includes local traffic volume and environmental and area-level population determinants associated with vehiclepedestrian injury collisions.

1.2. Area-level predictors of vehicle-pedestrian collisions

Fig. 2 describes the conceptual framework that informed our model development. Specifically, we sought to understand how an area's built environmental context – street and land use characteristics – as well as compositional factors, including resident and employee population size, population characteristics and travel behaviors, predict the area-level distribution of vehicle-pedestrian injury collisions. Vehicle-pedestrian injury collisions are also associated with a number of individual-level factors including age, alcohol consumption, and other driver or pedestrian behaviors (Laflamme and Diderichsen, 2000; Ryb et al., 2007; Wazana et

Fig. 2. Conceptual framework: an area-level model of vehicle-pedestrian injury collisions.

al., 1997). In the typical study of individual-level determinants, the environmental context of the injury is viewed as a "given" (Christoffel and Gallagher, 1999); however, individual behaviors occur in and are influenced by the environment, which is the focus of our research.

Previous research on environmental correlates of vehiclepedestrian collisions shows that traffic volume is a significant predictor (Brugge et al., 2002; LaScala et al., 2000; Lee and Abdel-Aty, 2005; Loukaitou-Sideris et al., 2007; Roberts et al., 1995), while injury severity is largely determined by vehicle speed (Ewing, 2006; NHTSA, 1999). Other roadway characteristics associated with pedestrian injuries include street type (e.g., residential, freeway, arterial) and intersection and street design features (e.g., traffic and pedestrians signals, signage, lighting) (Ewing, 2006; Retting et al., 2003). Similarly, the land use type in an area has been associated with vehicle-pedestrian collisions (overall and fatal)—with increases predicted by increasing proportions of land used for commercial, mixed use, park, retail, or community uses (Geyer et al., 2005; Kim et al., 2006; Loukaitou-Sideris et al., 2007; Wedagama et al., 2006).

Pedestrian volumes, at the intersection-level as well as larger geographic regions, are also associated with increased pedestrian injury risk, though individual risk may be attenuated as pedestrian volumes increase (Geyer et al., 2005; Jacobsen, 2003). Actual pedestrian count data is not routinely collected in the United States; however, U.S. Census data on population or commute travel mode data can serve as a surrogate for pedestrian volume (Jacobsen, 2003).

Aside from pedestrian volumes, specific population characteristics can affect vehicle-pedestrian collision risk. Vehicle-pedestrian collisions are a leading cause of injury and death for youth (Walton-Haynes, 2002). Nationally, youth aged 10–20 have the highest population rates of pedestrian (non-fatal) injury at 35 injuries/100,000, well above the overall population rate of 20/100,000 (NHTSA, 2006b). Seniors aged 65 and over actually have non-fatal injury rates slightly lower than the overall population rate (some have speculated due to less pedestrian activity); however, seniors are more likely to die when hit by a vehicle based on national and local data (NHTSA, 2006b; Sciortino and Chiapello, 2005a). The elderly and children take longer to cross a street, increasing their exposure for injury (Demetriades et al., 2004), and children also have less developed cognitive, perceptual, motor and traffic safety skills (Johnson et al., 2004). Lower income children have a higher rate of pedestrian injury than higher income children, though the mechanisms contributing to this disparity – including the physical and social environment – are not well understood (Laflamme and Diderichsen, 2000; Johnson et al., 2004; LaScala et al., 2004).

Findings from many of the above studies may be specific to local contexts, and the resulting findings and risk estimates therefore may not be generalizable. In addition, some of the above studies did not adjust for confounding by important covariates, while others standardized outcome variables by factors we would like to understand as predictors—such as street length or land area.

1.3. Macro-level collision models

Vehicle-pedestrian collisions tend to be dispersed throughout urban areas, and these dispersion patterns are missed by intersection or other micro-level analyses that focus on "black spots" with pre-existing high crash rates (Campbell et al., 2004; Morency and Cloutier, 2006). For example, from 2001 to 2005, eliminating all vehicle-pedestrian injury collisions at the five San Francisco intersections with 10 or more collisions during that period would leave over 98% of the city's vehicle-pedestrian injury collisions unaddressed (CCSF MTA, 2006). However, based on our data review, almost 10% of San Francisco's vehicle-pedestrian injury collisions were concentrated in two of 176 census tracts. A macro-level approach focused on census tracts could inform area-wide community transportation safety planning, and complement micro-level traffic safety mitigation measures such as intersection signalization (Lovegrove and Sayed, 2006).

Transportation researchers have modeled motor vehicle collisions at an area-level using multivariate regression methods, aggregate variables and linked datasets (Hadayeghi et al., 2003; Ladron de Guevara et al., 2004; Lovegrove and Sayed, 2006). Positive associations between collisions and traffic volume or vehicle miles travelled, population density, road network, and area-level socio-demographic characteristics are consistently significant in these macro-level models, which include pedestrian collisions with all motor vehicle collisions. Given potentially different determinants and risk estimates, separate macro-level vehicle-*pedestrian* collision models are warranted. For example, Loukaitou-Sideris et al. (2007) analyzed the spatial distribution of vehicle-pedestrian collisions in Los Angeles, and found pedestrian exposure, traffic, socioeconomic and land use variables were predictive of censustract collision density.

To evaluate and model census-level predictors of vehiclepedestrian injury collisions in San Francisco, we used crosssectional, aggregated data, to (1) describe the distribution of vehicle-pedestrian injury collisions and select environmental and population characteristics in San Francisco census tracts; and (2) estimate the nature and strength of the independent effect of census-tract traffic volume on census-tract vehicle-pedestrian injury collisions, adjusting for covariates. We then discuss the strengths and limitations of this approach and its potential for practical application to predict change in vehicle-pedestrian injury collisions associated with land use development and transportation planning decisions.

2. Methods

This area-level model is based on cross-sectional data for San Francisco, California County, aggregated at the level of the census tract (outlined in Fig. 3).We selected our analytic variables based on the previous literature and our interest in environmental predictors of vehicle-pedestrian injury collisions as detailed in Fig. 2.

2.1. Outcome variable

We used data on vehicle-pedestrian injury collisions in San Francisco, 2001–2005, from the Statewide Integrated Traffic Records System (SWITRS) which contains data on reported vehicle collisions on public roadways (CHP, 2008). SWITRS vehicle-pedestrian injury collision data were imported into ArcGIS (version 9.2; ESRI Inc., Redlands, CA, USA) and geocoded to the intersection of the reported primary and secondary streets (exact street address is not collected). We used a spatial join to assign vehicle-pedestrian injury collisions to one of the 176 census tracts in San Francisco (Geolytics Inc., 2004). We excluded non-injury collisions which are reported as "Property Damage Only". We included collisions resulting in pedestrian injuries and/or fatalities, hereafter referred to as "vehicle-pedestrian injury collisions."

2.2. Independent variables

We obtained street segment traffic counts and street length and type data from researchers at the San Francisco Department of Public Health and the University of California - Berkeley. This dataset

Fig. 3. Vehicle-pedestrian injury collisions: San Francisco, California census tracts (2001-2005). Source: California Highway Patrol, Statewide Integrated Traffic Records System (SWITRS).

Table 1

Descriptive statistics for San Francisco, California census tracts (*n* = 176).

 $^{\rm a}$ Includes collisions resulting in pedestrian injuries and/or fatalities.

b Excludes grade-separated street segments inaccessible to pedestrians.

 $c \geq 16$ years old, in the civilian labor force, unemployed.

 $\frac{1}{d}$ Workers 16 years and older.

was created from existing street segment daily traffic counts collected principally from 1997 to 2002 and maintained by the San Francisco Municipal Transportation Agency and the County Transportation Authority. Researchers consolidated the traffic counts, adjusted them by street direction (one-way, two way, and divided streets), and applied them to their respective street centerline network number for geocoding purposes. Traffic counts were imputed for street segments with unknown counts based on the average traffic count for the street segment type (highway/freeway, arterial with public transit (e.g., streetcars, light rail vehicles, buses, cable cars), arterial without public transit, residential) within each planning neighborhood as detailed in Seto et al. (2007). We excluded grade-separated street segment counts (i.e., streets without pedestrian access) and then aggregated the remaining street segment traffic counts by census tract to obtain the study traffic estimates—hereafter referred to as *traffic volume*.We obtained 2005 zoning district area data from the San Francisco Planning Department, and aggregated zoning use districts into census-tract level land use characteristic categories (Table 1).

We obtained the following census-tract level aggregate variables from the U.S. Census Bureau (2000), Summary Files 1 and 3: land area (square miles); population age 65 and older and under age 18; unemployed population; workers traveling to work by walking or by public transit; population living below the poverty level last year. We used the census-tract level variables as numerators and the total population from the corresponding Summary File as the denominator to create census-tract level proportions (Table 1). The number of workers-at-work in the census tract was obtained from the Census Transportation Planning Package (2000). We determined the number of intersections in each census tract using a spatial join with intersection, node, and street CNN data provided by the City and County of San Francisco (SFDTIS, 2006).

2.3. Data analysis

We first assessed the distribution of vehicle-pedestrian injury collisions, land use, street characteristics, pedestrian exposure proxies, and demographic characteristics in San Francisco census tracts. We then used ordinary least squares regression (OLS) to model the natural log of the number of vehicle-pedestrian injury collisions over a 5-year period. We added one collision to the three census tracts with zero collisions reported so they would not be dropped from the analysis. The model form used for our analyses is

$$
ln(PLC) = \beta_0 + \sum \beta_i X_i
$$
\n(1.2)

where ln(PIC) is the predicted natural log of vehicle-pedestrian injury collisions per census tract; β_0 , the intercept; β_i , the model coefficient for 1-unit change in predictor variable *i*; and, *Xi*, the census-tract level data for predictor variable *i*. All variables are continuous and at the census-tract level. To better approximate the normality assumptions of the linear model, we applied a natural log transformation to both the traffic volume and employee variables, as in previous research (Hadayeghi et al., 2003; Jacobsen, 2003; Lovegrove and Sayed, 2006).

We used the conceptual framework for our model building approach (Fig. 2). We started with a base model including the street and land use characteristics in Table 1, then added population characteristics followed by commute behavior variables. In each step, variables were dropped from the model based on coefficient *p*value. We assessed model fit based on the values, distribution, and tests of spatial autocorrelation (Moran's I) of the residuals. One census tract was an evident outlier based on our assessment of model fit based on residual plots. This census tract was one of three census tracts to which we had added one collision because it had zero collisions reported. We reviewed the census tract's predictor variable values and found no extreme or erroneous values. We then visually assessed the tract's geographic characteristics using ArcGIS and GoogleTM maps, which provided evidence that it is likely a true outlier. This tract's street network has numerous dead ends, lacks connectivity, and has large portions of area densely forested and without streets, in contrast with the grid street network and sparser street trees in most of the city. Additionally, the large medical center that employs most of the tract's >5000 employees is on the tract border, its campus split by the boundary and the area in the census-tract largely surrounded by forested land. Based on these anomalous environmental conditions, we dropped the tract from our analysis, which improved model fit. Our final model is therefore based on 175 census tracts. All analyses were conducted using STATA software (version 9; StataCorp, College Station, TX, USA).

3. Results

There were 4039 recorded vehicle-pedestrian injury collisions in San Francisco's 176 census tracts from 2001 to 2005, with a median 14 and mean 23, ranging from 0 to 191 vehicle-pedestrian injury collisions in a tract (Table 1). As illustrated in Fig. 3, vehiclepedestrian injury collisions were dispersed throughout the city, with evident concentrations in areas near freeways and highways that carry high traffic volumes from bridges and highways, as noted in previous literature (UCSF SFDPH, 2004).

San Francisco census tracts exhibit a wide range in aggregate traffic volume estimates, with a median of 703,145 and a mean of 925,544 aggregated vehicles, largely influenced by the presence of arterial streets and freeway ramps (data not shown). A scatter plot of traffic volume by vehicle-pedestrian injury collisions shows a positive linear association (unadjusted Pearson *R*² = .359, natural logs, data not shown).

Table 1 shows the mean, median, standard deviation, and range of street and land use characteristics, population characteristics, and commuting behaviors in San Francisco, revealing the diversity in compositional and contextual characteristics of the city's census tracts. A median of 64% of census-tract street length was residential (range, 23–100%), while the median percentages of census-tract streets that were arterial with and without public transit were similar (16% and 17%, respectively), the range in values across census tracts was large (0–67% and 0–48%, respectively). The census-tract median population was close to 4000—consistent with the average tract size as defined by the U.S. Census Bureau (2001). The median number of workers in a census tract was 1063, with a higher mean of 3337 reflecting the skewed range of less than 100 to almost 95,000 workers. Median proportions of youth and seniors were similar—approximately 13–14%, though both subgroups had wide ranges across census tracts of approximately 1 to >40%. The median proportion of residents living in poverty was 9%, and ranged from 0 to >50% across the city. A median of 6% of residents walk to work (range, 0–42%) while an average of 16% take public transit (range, 5–49%).

With the exception of land area and proportion of residents who are seniors, all final model variables had a positive association with vehicle-pedestrian injury collisions (Table 2). Increases in traffic volume, proportion of arterial streets without transit, proportion of land area zoned for neighborhood commercial and mixed residential/neighborhood commercial use, employee and resident populations, and proportion of people living in poverty predicted increased vehicle-pedestrian injury collisions.

The final model explains approximately 72% of the systematic variation in census-tract vehicle-pedestrian injury collisions.

Final ordinary least squares regression model of census-tract level vehicle-pedestrian injury collisions: San Francisco, California, 2001–2005 (*n* = 175).

^a Excludes grade-separated street segments inaccessible to pedestrians.

Traffic volume had the highest adjusted partial correlation with vehicle-pedestrian injury collisions (*r* = .454), followed by the number of employees (*r* = .358), proportion of land zoned neighborhood commercial (*r* = .323), proportion of arterial streets without transit (*r* = .314), and resident population (*r* = .303).

Since we used a natural log transformation for both the traffic volume and employee population variables, the interpretation of their coefficients is equivalent to the power function described in Formula (1.1) (assuming all other final model covariates are held constant). Fig. 1 illustrates the power function's (PF) predictions of percent change in vehicle-pedestrian injury collisions based on percent change in traffic volume at varying β . Adjusting for the other covariates, our final model (Fig. 1, FM) is equivalent to a power function with β =0.753 (Table 2, coefficient on log traffic volume). Therefore, a 15% increase in census-tract traffic volume is associated with an approximate 11% increase in vehiclepedestrian injury collisions ($(1.15)^{0.753} - 1 = 11.1\%$). Similarly, a 15% increase in area employees (e.g., from 10% to 11.5%) predicts an approximately 3% increase in vehicle-pedestrian injury collisions $((1.15)^{0.228} - 1 = 3.2\%).$

Model coefficients for independent variables that were not log transformed estimate the change in the log count of vehiclepedestrian injury collisions per unit increase in the predictor. For example, a 5-unit increase in the proportion of censustract street length that is arterial (without transit) is associated with an approximately 8% change in vehicle-pedestrian injury collisions ($exp(5^*0.017) = 1.08$). Similarly, an increase in resident population of 500 people would predict an approximately 5% $(exp(500*0.0001) = 1.05)$ increase in vehicle-pedestrian injury collisions.

4. Discussion

In San Francisco, California, in a multivariate regression model at the census-tract level, statistically significant predictors of vehiclepedestrian injury collisions include traffic volume, arterial streets without public transit, proportions of land area zoned for neighborhood commercial use and residential-neighborhood commercial use, land area, employee population, resident population, proportion of people living in poverty, and proportion of people aged 65 and over. All model variables had a positive association with vehicle-pedestrian injury collisions, with the exception of land area (its increase potentially capturing decreasing population density) and proportion of the population that are seniors.

Comparing predicted percent change in vehicle-pedestrian injury collisions based on our final model with those based on the simpler power function illustrates that variables in addition to traffic volume – including built environment characteristics that are potential pedestrian attractors (neighborhood commercial districts) and area-level population characteristics that are potential proxies for pedestrian activity (resident and employee populations) – contribute significant explanatory power to the model. While our results are specific to San Francisco, California, our conceptual model (Fig. 2) and the findings of this novel approach to estimating the impact of area-level changes on vehicle-pedestrians collisions may inform models in other urban areas. We next discuss potential strengths and limitations of this approach.

We used census tracts as our unit of analysis, ideal for small arealevel analysis as they are created to be relatively homogeneous with respect to demographic characteristics (U.S. Census Bureau, 2001) and census data is publicly available. Additionally, U.S. Census data for number of employees, resident socioeconomic data, and commute behaviors is available at the census-tract level, but not at a smaller area level. Other potential data sources include the American Community Survey. Area-specific data on number of residents, employees, traffic, streets, land use and other environmental factors are routinely collected, analyzed and reported in local planning processes—which would enable the model to be applied to predict the impacts of large development projects on vehicle-pedestrian injury collisions.

Our final model of vehicle-pedestrian injury collisions is an ecological analysis (i.e., all units of analysis are at the census-tract level) using both aggregate (summaries of observations derived from individuals) and environmental (physical characteristics) variables. An*ecological fallacy* occurs when onemakes (incorrect) causal inferences about associations between individual-level variables based on observed associations in ecological analyses. In applying and interpreting our area-level pedestrian injury collision model, we intend to make inferences to areas; no causal inferences are made at the level of the individual.

Underreporting of collisions could affect model results. Based on a comparison of SWITRS and hospital records in 2000–2001, Sciortino et al. (2005b) found that SWITRS under-reported San Francisco pedestrian injuries by 21% (using San Francisco General Hospital medical records as a gold standard), with African Americans and males less likely to have a SWITRS-reported injury. This ascertainment bias could have caused our model to underestimate area-level pedestrian injuries, particularly in predominantly African American neighborhoods. Because area-level racial/ethnic composition is highly correlated with poverty, this bias may have resulted in an underestimate of the effect of poverty, which may partially capture disparities in built environmental conditions or increased pedestrian activity among less auto-dependent populations.

We aggregated vehicle-pedestrian injury collision data by census tract, after geocoding collisions to the nearest intersection. This could result in erroneous census-tract assignment for some collisions that fall on census-tract boundaries. However, we do not have reason to believe that there would be systematic bias in this error.

Actual pedestrian volume data was not available. Significant predictors in our final model were number of residents, employees, and proportion of land zoned neighborhood commercial/residentialneighborhood commercial—potential partial proxies for pedestrian activity and pedestrian attractors. As previously mentioned, researchers have found that pedestrian volumes or proxy variables are associated with increased pedestrian injury risk that is attenuated as pedestrian volumes increase (Geyer et al., 2005; Jacobsen, 2003). While the commuting via walking variable was not a significant predictor in our final model, the log-transformed employee population variable was a significant strong predictor. An attenuated relationship was not found for resident population, potentially because census-tract boundaries are informed by resident population size and therefore have less variation across the city. A spatial analysis of pedestrian collisions in Hawaii also found both resident population and commercial areas were positively associated with pedestrian collisions; however, total jobs was not a statistically significant predictor for this analysis which focused on land use, population, employment and economic variables—potentially due to regional differences (Kim et al., 2006).

We were not able to include a reliable vehicle speed assessment variable. Vehicle speed strongly predicts injury severity—the chance of a fatal vehicle-pedestrian collision increasing from 5% at 20 mph to 85% at 40 mph (UK Department of Transportation, 1987). Our model did not distinguish collisions by severity, a question for which vehicle speed data would be more salient. The street type variables we did include were associated with both vehicle-pedestrian injury collisions and traffic volume, and were potentially proxies for vehicle speeds or other street characteristics for which we did not have citywide data (e.g., number of lanes, street width).

We repeated these analyses using a negative binomial regression model and obtained very similar coefficients and standard errors. We used the OLS model for our final analysis based on our stated interest in developing a model for practical application that can be readily applied and interpreted, its transparency preferable for establishing and understanding the causal relationship between traffic volume and vehicle-pedestrian injury collisions. An assumption when using a simple OLS approach is that the dependent variable values are linearly distributed, with a 1-unit change in an independent variable *x* predicting the same corresponding change in the dependent variable across all values of *x*. We adjusted for the non-linear relationship between collisions and the independent variables traffic volume and employee population using a natural log transformation of those variables. Once a causal relationship between traffic volume and collisions is established, it is likely that advanced statistical techniques incorporating both linear and nonlinear approaches, such as neural networks, may improve model prediction (Tu, 1996).

Our results are partly consistent with those reported in a previous study of 1990 vehicle-pedestrian injury collisions in San Francisco census tracts. LaScala et al. (2000) reported a significant, positive association with traffic flow, resident population, and proportion unemployed, and a significant, negative association with proportion with a high school diploma or higher—similar to our findings regarding increased risk with a higher proportion of poverty, higher traffic volumes and more residents. Proportion of the resident population that was male also had a significant positive association in their model and proportion aged 0–15 was inversely associated with vehicle-pedestrian injury collisions; we did not include proportion male population as a potential predictor, and proportion of the population aged 0–15 was not a significant predictor in our final model. This difference could be explained by the correlation of land use and street characteristics (only included in our model) with population characteristics (such as age distribution) in San Francisco census tracts. Similar to our results, proportion of seniors age 55 and older was inversely associated with pedestrian injury collisions. This 1990 study did not explore employee population, street type or land use variables (other than bars, restaurants, alcohol outlets per kilometer roadway, which were not significantly associated with overall vehicle-pedestrian injury collisions). La Scala et al. standardized their log-transformed pedestrian injury outcome by roadway length, which limits comparisons.

Similar to our findings, a recent Los Angeles study found population and employment density, traffic density, and land use variables – as well as proportion of population that was Hispanic (described as a socioeconomic variable) – predicted pedestrian collision density (Loukaitou-Sideris et al., 2007). However, the researchers did not explore street type variables—and proportion of population over age 65 was not significant in their final model. Their findings also differed from ours in that – based on ranking of independent variable beta weights – population density was the most predictive variable, followed by traffic density and employee density—whereas traffic volume was the most predictive variable in our model, followed by employee population and neighborhood commercial land use proportion (data not shown). Notably, traffic volume and employee populations were strong predictors in both models.

The coefficient for (log)traffic volume in our model, 0.753, was notably higher than the 0.5 reported for the simpler road safety function (Lee and Abdel-Aty, 2005) as well as the 0.221 from the Los Angeles study (Loukaitou-Sideris et al., 2007). A primary reason for these differing estimates may be regional and/or geographic differences in land use, transportation, population or other characteristics (e.g., weather) that result in differences in the predictive value of traffic volume. Understanding these differences is another research question, requiring multi-level models and regional data.

5. Conclusion

Consistent with previous national and international findings (Roberts et al., 1995; Lee and Abdel-Aty, 2005; Brugge et al., 2002; LaScala et al., 2000), our study provides additional evidence that traffic volume is a primary environmental cause of vehiclepedestrian injury collisions at the area level. In addition to traffic volume, employee and resident populations, arterial streets without public transit, proportions of land area zoned for neighborhood commercial use and residential-neighborhood commercial use, land area, proportion of people living in poverty, and proportion of people aged 65 and over are statistically significant predictors of vehicle-pedestrian injury collisions in a multivariate model at the census-tract level in San Francisco, California.

We developed this model to predict vehicle-pedestrian injury collisions resulting from land use and transportation planning decisions—specifically, in the context of environmental impact assessment and as required by the National Environmental Policy Act and related state laws (Bhatia and Wernham, 2008). A bivariate power function may be used as a simple predictive tool to forecast the impact of increased traffic on vehicle-pedestrian injury collisions; however, a multivariate approach may provide more defensible estimates in planning or development scenarios which have broad impacts on an area's land use, transportation and population characteristics. We have used this multivariate model to analyze the impacts of San Francisco neighborhood rezoning plans on vehicle-pedestrian injury collisions, and our findings

were incorporated in the plans' environmental impact assessment (SFPD, 2007). San Francisco also intends to use this model to predict collision impacts associated with area-level congestion pricing proposals. Subsequent reports and publications will describe these practical applications.

Micro-level (e.g., intersection) approaches that identify specific locations with existing high numbers of vehicle-pedestrian injury collisions support targeted pedestrian safety countermeasures. This area-level model can similarly support pedestrian injury prevention by justifying area-level interventions in development and planning processes. Examples of these interventions include: transit-oriented development that coordinates high-density land use with public transit locations and includes street amenities and design features that slow traffic and support safe walking; employer-based transportation demand management programs to incentivize commuting to work via walking, biking and public transit and decrease driving; and street design that slows traffic and improves the quality and safety of the pedestrian environment near land uses including residences, schools, or senior centers (VTPI, 2008).

Acknowledgements

The authors acknowledge the helpful contributions of Cynthia Comerford Scully, MA, Environmental Planner, San Francisco Department of Public Health, for ArcGIS data management; and Tom Rivard, Senior Environmental Health Specialist, San Francisco Department of Public Health, for guidance and support in utilizing the traffic count database.

Appendix A. Application of the bivariate power function to a local development project

The following application was conducted by one of the authors (R. Bhatia) in the context of a health impact assessment of the Oak to Ninth Development Project proposed in Oakland, California (UCBHIG, 2006).

Traffic analysis in the proposed project's environmental impact report provided data on changes in traffic volume on area roadways. Estimates projected that the development, which includes 3100 residential units and 3500 parking spaces, would result in an additional 27,110 daily vehicle trips external to the project. An intersection-level traffic analysis for 51 intersections demonstrated that those trips would increase traffic volume on surrounding local streets, with 5% or greater cumulative increases at several intersections. Overall, the increase in intersection vehicle volumes varied considerably, ranging from 2% to 127%. The average weighted project-related increase in vehicle volume at studied intersections was approximately 11% after project completion; the average cumulative increase in vehicle volume by 2025 was 45%, including other proposed area development projects at these intersections.

The Statewide Integrated Traffic Records System (SWITRS) provided data on reported pedestrian injuries occurring in Oakland from 2000 to 2005. Pedestrian injuries were mapped to intersections using ArcGIS (>90% successfully geocoded). 545 pedestrian-vehicle collisions occurred at the 51 study intersections during 2000–2005. Since approximately 10% of collisions could not be geocoded, the current annual average number of pedestrian injuries in areas affected by project-traffic was assumed, approximately 100 per year. Because some pedestrian injuries may not be reported, this may underestimate the actual number of pedestrian injuries.

Based on a power function of vehicle volume described in Formula (1.1), an 11% increase in vehicle volume on all roadways in an area with a baseline of 100 pedestrian injuries per year predicts an increase in 5.4 injuries per year, or 268 injuries between 2025 and 2075. Based on a cumulative increase in average daily trips of 45% in 2025, the impact is 20 injuries per year or 1000 injuries between 2025 and 2075.

References

- Bhatia, R., Wernham, A., 2008. Integrating human health into environmental impact assessment: an unrealized opportunity for environmental health and justice. Environmental Health Perspectives 116, 991–1000.
- Brugge, D., Lai, Z., Hill, C., Rand, W., 2002. Traffic injury data, policy, and public health: lessons from Boston Chinatown. Journal of Urban Health 79 (1), 87–103.
- Campbell, B.J., Zegeer, C.V., Huang, H.F., Cynecki, M.J., McLean, V.A., 2004. A Review of Pedestrian Safety Research in the United States and Abroad. Federal Highway Administration, Publication No. FHWA RD-03-042. http://www. walkinginfo.org/pdf/PedSynth/Ped Synthesis Report.pdf, accessed on 15 February 2007.
- Cavill, N., 2001. Walking and health: making the links. World Transport Policy and Practice 7 (4), 33–38.
- CCSF MTA (City and County of San Francisco Municipal Transportation Agency Traffic Engineering Division), 2007. San Francisco 2006 Collision Report. San Francisco, CA. http://www.sfmta.com/cms/rtraffic/documents/Collision report 2006.pdf, accessed on 8 April 2008.
- CCSF MTA (City and County of San Francisco Municipal Transportation Agency Traffic Engineering Division), 2006. San Francisco 2005 Collision Report. San Francisco, Γ A
- Census Transportation Planning Package, 2000. http://www.mtc.ca.gov/maps and data/datamart/census/ctpp2000, accessed on 13 November 2006.
- CHP (California Highway Patrol), 2008. Accident Investigation Unit. Statewide Integrated Traffic Records System. http://www.chp.ca.gov/switrs/index.html, accessed on 7 April 2008.
- Christoffel, T., Gallagher, S.S., 1999. Injury Prevention and Public Health: Practical Knowledge, Skills, and Strategies. Aspen Publishers Inc., Gaithersburg, MD.
- Crossroads Software Inc., 2007. http://www.crossroadssoftware.com/. Accessed on 1 March 2007.
- Demetriades, D., Murray, J., Martin, M., Velmahos, G., Salim, A., Alo, K., Rhee, P., 2004. Pedestrians injured by automobiles: relationship of age to injury type and severity. Journal of the American College of Surgeons 199 (3), 382–387.
- Ewing, R., 2006. Fatal and Non-fatal Injuries. Understanding the Relationship Between Public Health and the Built Environment: A Report Prepared for the LEED-ND Core Committee. http://www.activeliving.org/files/LEED ND report.pdf, accessed on 7 April 2008.
- Frumkin, H., Frank, L., Jackson, R., 2004. Urban Sprawl and Public Health: Designing, Planning, and Building for Healthy Communities. Island Press, Washington, DC. Geolytics Inc., 2004. Neighborhood Change Database (NCDB) 1970–2000.
- Geyer, J., Raford, N., Ragland, D., Pham, T., 2005. The Continuing Debate about Safety in Numbers—Data from Oakland, CA. TRB Annual Meeting CD-ROM, TRB Paper #06-2616. http://repositories.cdlib.org/cgi/viewcontent.cgi? article=1029&context=its/tsc.
- Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2003. Macrolevel accident prediction models for evaluating safety of urban transportation systems. Transportation Research Record 1840, 87–95.
- Jacobsen, P.L., 2003. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. Injury Prevention 9 (3), 205–209.
- Johnson, E., Geyer, J., Rai, N., Ragland, D.R., 2004. Low Income Childhood Pedestrian Injury: Understanding the Disparate Risk. U.C. Berkeley Traffic Safety Center, Paper UCB-TSC-RR-2004-20. http://repositories.cdlib.org/its/tsc/UCB-TSC-RR-2004-20.
- Kim, K., Brunner, I.M., Yamashita, E.Y., 2006. Influence of land use, population, employment, and economic activity on accidents. Transportation Research Record 1953, 56–64.
- Ladron de Guevara, F., Washington, S.P., Oh, J., 2004. Forecasting crashes at the planning level: simultaneous negative binomial crash model applied in Tucson, Arizona. Transportation Research Record 1897, 191–199.
- Laflamme, L., Diderichsen, F., 2000. Social differences in traffic injury risks in childhood and youth—a literature review and a research agenda. Injury Prevention 6 (4) , 293–298.
- LaScala, E.A., Gruenewald, P.J., Johnson, F.W., 2004. An ecological study of the locations of schools and child pedestrian injury collisions. Accident Analysis and Prevention 36 (4), 569–576.
- LaScala, E., Gerber, D., Gruenwald, P.J., 2000. Demographic and environmental correlates of pedestrian injury collisions: a spatial analysis. Accident Analysis and Prevention 32 (5), 651–658.
- Lavizzo-Mourney, R., McGinnis, J.M., 2003. Making the case for active living communities. American Journal of Public Health 93 (9), 1386–1388.
- Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. Accident Analysis and Prevention 37, 775–786. Leyden, K.M., 2003. Social capital and the built environment: the importance of
- walkable neighborhoods. American Journal of Public Health 93 (9), 1546–1551.
- Loukaitou-Sideris, A., Ligget, R., Sung, H.G., 2007. Death on the crosswalk: a study of pedestrian-automobile collisions in Los Angeles. Journal of Planning Education and Research 26, 338–351.

- Lovegrove, G.R., Sayed, T., 2006. Macrolevel collision prediction models for evaluating neighborhood traffic safety. Canadian Journal of Civil Engineering 33 (5), 609–621.
- Mokdad, A.H., Marks, J.S., Stroup, D.F., Gerberding, J.L., 2004. Actual causes of death in the United States, 2000. Journal of the American Medical Association 291 (10), 1238–1245.
- Morency, P., Cloutier, M.S., 2006. From targeted "black spots" to area-wide pedestrian safety. Injury Prevention 12, 360–364.
- NHTSA (National Highway Traffic Safety Administration), 1999. Literature Review on Vehicle Travel Speeds and Pedestrian Injuries. USDOT, Washington, DC.
- NHTSA (National Highway Traffic Safety Administration), 2006a. Fatality Analysis Reporting System, National Statistics. http://www-fars.nhtsa.dot.gov/, accessed on 12 February 2008.
- NHTSA (National Highway Traffic Safety Administration), 2006b. Traffic Safety Facts. http://www.nhtsa.dot.gov/portal/site/nhtsa/menuitem.dfedd570f698cabbbf-30811060008a0c/, accessed on 12 February 2008.
- PBCAT (Pedestrian and Bicycle Crash Analysis Tool), 2007. Pedestrian and Bicycle Information Center. http://www.walkinginfo.org/pc/pbcat.cfm, accessed on 14 March 2007.
- Retting, R.A., Ferguson, S.A., McCartt, A.T., 2003. A review of evidence-based traffic engineering measures designed to reduce pedestrian-motor vehicle crashes. American Journal of Public Health 93 (9), 1456–1463.
- Roberts, I., Norton, R., Jackson, R., Dunn, R., Hassall, I., 1995. Effect of environmental factors on risk of injury of child pedestrians by motor vehicles: a case-control study. British Medical Journal 310 (6972), 91–94.
- Ryb, G.E., Dischinger, P.C., Kufera, J.A., Soderstrom, C.A., 2007. Social, behavioral and driving characteristics of injured pedestrians: a comparison with other unintentional trauma patients. Accident Analysis and Prevention 39 (2), 313–318.
- Sciortino, S., Chiapello, E., 2005a. Environmental, Behavioral, and Demographic Factors that Affect the Severity of Pedestrian Injuries. http://www.sfdph.org/ dph/files/trafficsafety/EnviroFactorsSeverity4.pdf, accessed on 19 September 2006.
- Sciortino, S., Vassar, M., Radetsky, M., Knudson, M.M., 2005b. San Francisco pedestrian injury surveillance: mapping, under-reporting, and injury severity in police and hospital records. Accident Analysis and Prevention 37, 1102–1113.
- Seto, E.Y., Holt, A., Rivard, T., Bhatia, R., 2007. Spatial distribution of traffic induced noise exposures in a US city: an analytic tool for assessing the health impacts of urban planning decisions. International Journal of Health Geography 6, 24.
- SFCTA (San Francisco County Transportation Authority), 2004. Countywide Transportation Plan. San Francisco, CA. http://www.sfcta.org/content/view/13/30/, accessed on 18 January 2008.
- SFDTIS (San Francisco Department of Telecommunications Information Services GIS Server), 2006. San Francisco Department of Parking and Traffic, San Francisco Street Basemap San Francisco, CA. SFDTIS.
- SFPD (San Francisco Planning Department), 2007. Eastern Neighborhoods Rezoning and Area Plans, Draft Environmental Impact Report: Transportation, Circulation, and Parking. http://www.sfgov.org/site/uploadedfiles/planning/EN DEIR Part-7 Trans-Noise-AQ.pdf, accessed on 5 September 2008.
- STPP (Surface Transportation Policy Project) and California Walks, 2002. Pedes-trian Safety in California: Five Years of Progress & Pitfalls. STPP, California. http://www.transact.org/Ca/ped safety 2002 finalreport.htm, accessed on 18 April 2008.
- Subramanian, R., 2006. Motor Vehicle Traffic Crashes as a Leading Cause of Death in the United States, 2003. NHTSA Traffic Safety Facts Research Note. http://wwwnrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/RNotes/2006/810568.pdf, accessed on 9 March 2007.
- Tu, J., 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. Journal of Clinical Epidemiology 49, 1225–1231.
- UCBHIG (University of California-Berkeley Health Impact Group), 2006. Oak to Ninth Avenue Health Impact Assessment. http://ehs.sph.berkeley.edu/hia/ projects/projects.htm, accessed on 18 January 2008.
- UCSF SFDPH (University of California-Injury Center and the San Francisco Department of Public Health), 2004. Profile of Injury in San Francisco. http://www. surgery.ucsf.edu/sfic/profile05.pdf, accessed on 16 January 2007.
- UK Department of Transportation, 1987. Killing Speeds and Saving Lives. UK Dept of Transportation, London, England.
- U.S. Census Bureau, 2001. Introduction to Census 2000 Data Products. U.S. Department of Commerce. http://www.census.gov/prod/2001pubs/mso-01icdp.pdf, accessed on 15 September 2006.
- U.S. Census Bureau, 2000. American Factfinder, Census 2000. http://factfinder. census.gov/, accessed on 22 September 2006.
- USDHHS (U.S. Department of Health Human Services), 2000. Healthy People 2010: Understanding and Improving Health, 2nd ed. U.S. Government Printing Office, Washington, DC.
- USDOT (U.S. Department of Transportation), 1998. Zone Guide for Pedestrian Safety. DOT HS 808 742.
- VTPI (Victoria Transport Policy Institute), 2008. Transportation Demand Management Encyclopedia. Victoria Transport Institute, Victoria, BC. http://www. vtpi.org/nmtguide.doc, accessed on 29 August 2008. Walton-Haynes, L., 2002. Pedestrian injuries to young children. EPICGram, No.
- 5. California Department of Health Services, Epidemiology and Prevention for Injury Control (EPIC) Branch. http://www.dhs.ca.gov/epic/publications/ EPICgrams/EPICgram5 Pedestrian.pdf, accessed on 12 Feb 2008.
- Wazana, A., Krueger, P., Raina, P., Chambers, L., 1997. A review of risk factors for child pedestrian injuries: are they modifiable? Injury Prevention 3 (4), 295–304.
- Wedagama, D.M.P., Bird, R., Metcalfe, A.V., 2006. The influence of urban land-use on non-motorised transport casualties. Accident Analysis and Prevention 38, 1049–1057.