Modeling Vehicle-Pedestrian Injury Collisions at Signalized Intersections

A Health Forecasting Approach to Informing Pro-active Pedestrian Safety Improvements

San Francisco Department of Public Health - Program on Health, Equity and Sustainability

Fall 2013

I. BACKGROUND

A substantial body of published research has found that the frequency of pedestrian-vehicle collisions is predicted by factors including pedestrian and traffic volumes, vehicle speeds, road and intersection design, land use, and socio-demographic population characteristics.[1] Vehicle and pedestrian volumes are the most important single determinants of pedestrian injury collision frequency, with non-linear but positive associations between pedestrian flows and traffic volumes with pedestrian injury collisions consistently found at the small-area level including intersections.[2] Studies also show heterogeneity in these associations, potentially due to differences in other environmental conditions including land use, transportation system, and socio-demographic factors. Roadway design and higher operating speeds are potent determinants of injury frequency as well as injury severity. The quality of pedestrian facilities is an important predictor of pedestrian injury in the research literature, based on analyses at the intersection, street segment, and small-areas levels.[3] Low-income neighborhoods are more heavily burdened with pedestrian injuries and fatalities due factors potentially including higher residential area traffic densities, greater use of active transport and public transit, and relatively poorer quality roadway facilities.[4]

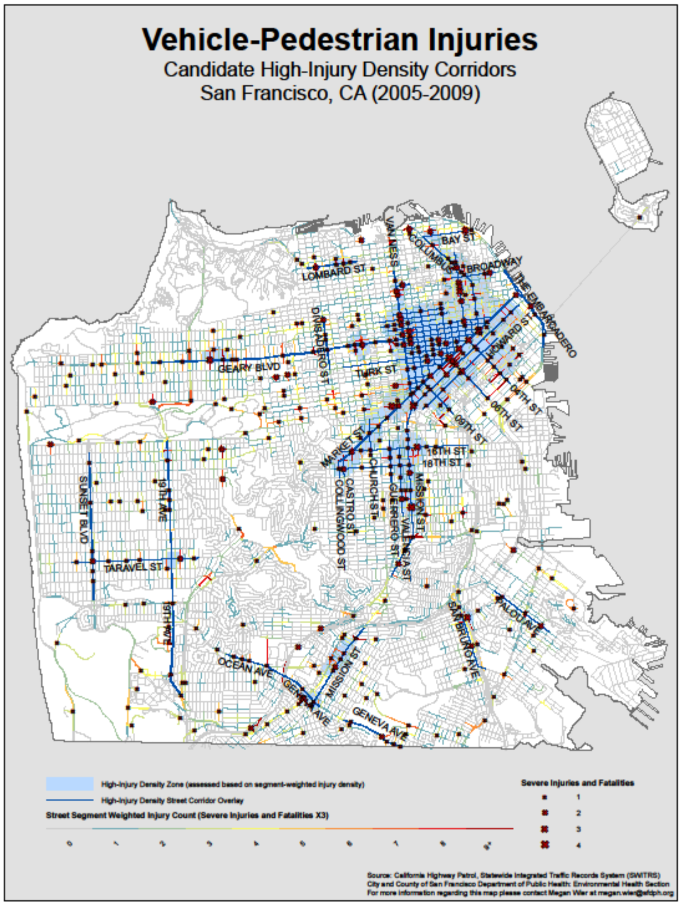
In 2009, the San Francisco Department of Public Health (SFDPH) published findings from its first pedestrian injury collision model in the peer-reviewed scientific journal *Accident Analysis & Prevention*.[5] The model predicts census-tract level changes in the number of collisions resulting in pedestrian injury or death associated with area-level changes in street, land use and population characteristics due to new development or transportation system factors. Significant predictors (census-tract level variables) in that model are as follows (+: positively correlated; -: negatively correlated):

- Traffic volume (+)
- Arterial streets (+; %, without MUNI transit)
- Neighborhood commercial areas (+; %, land area)
- Land area (-; square miles)
- Employee population (+)
- Resident population (+)
- Below poverty level (+; %, population)
- Age 65 and older (-; %, population)

Since initially developing that model, SFDPH has applied it in a number of health impact assessments, including the environmental impact analysis of the Eastern Neighborhoods Area Plans[6] and a health impact assessment of road pricing.[7] A primary criticism of the census-tract level modeling approach has been that census tracts are too large an area of estimation to inform targeted area improvements.

In 2011, SFDPH developed a methodology for identifying corridors with high absolute numbers of vehicle-pedestrian injury collisions in San Francisco (*see Figure 1*). The analysis demonstrated that 5% of the City's street length (the blue corridors) accounted for over 50% of serious and fatal injuries. The methodology was developed as a part of SFDPH's work as a co-lead of the San Francisco Citywide Pedestrian Safety Task Force. Since then, the high injury corridor network has been adopted by city transportation agencies to inform the prioritization of pedestrian safety improvements on the City's transportation network.

Figure 1. High Vehicle-Pedestrian Injury Corridors in San Francisco, CA

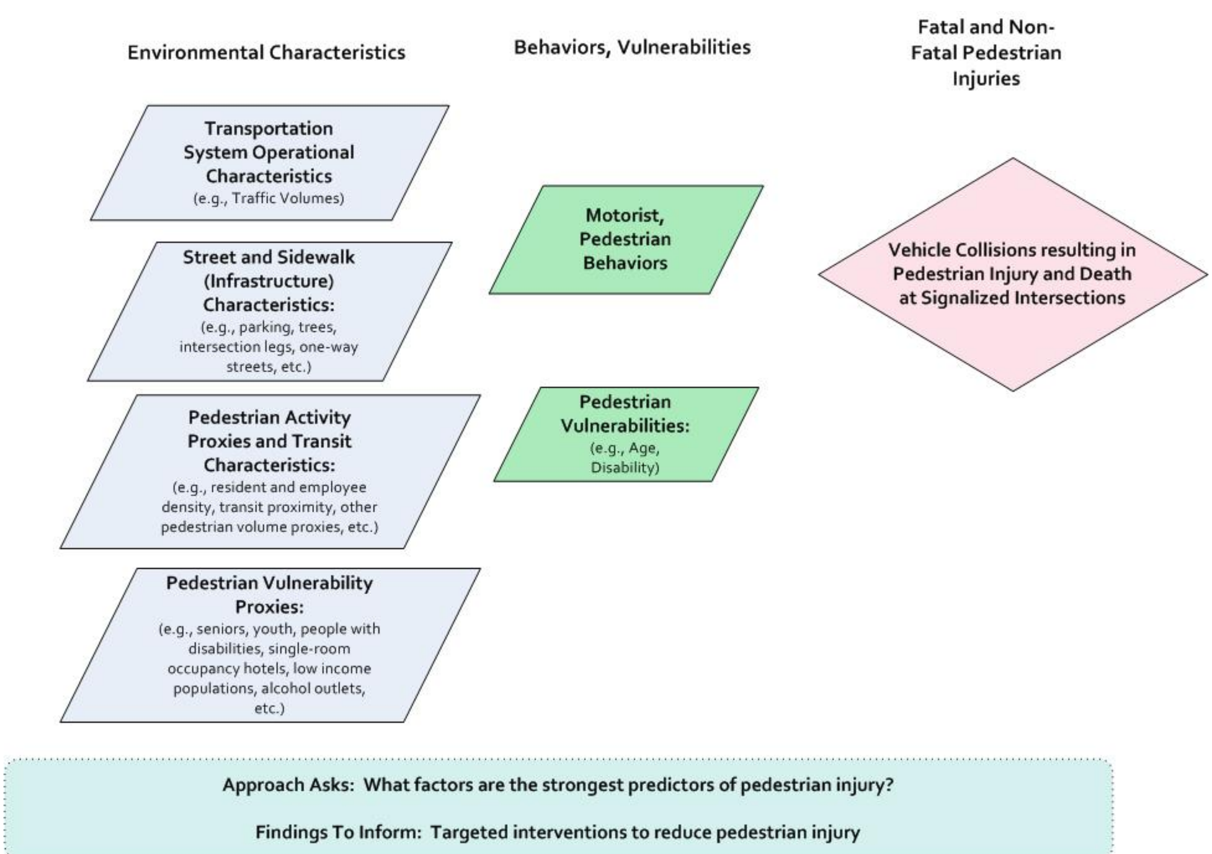


While the institutionalization of the high injury corridor approach has been a significant step forward with respect to data-drive pedestrian safety investments, a valid critique is that is does not consider future changes – particularly with respect to land use and transportation factors – that would impact on pedestrian safety conditions. Given the above critiques regarding the scale of the initial model, SFDPH thus embarked on the work summarized in this report to develop a more refined model of vehicle-pedestrian injury collisions at signalized intersections in San Francisco. Similar to the high-injury corridor approach – these intersections account for only 16% of San Francisco's intersections but almost 50% of vehicle-pedestrian injury collisions. As with the census-tract level approach, model inputs are conditions data that are typically generated as a part of transportation and land use planning. In addition to SFDPH's previous work, the model's development is informed by the Pedestrian Safety Prediction Methodology research conducted by the National Cooperative Highway Research Program (NCHRP). The NCHRP developed a "…methodology for quantifying the pedestrian safety effects related to existing site characteristics and/or proposed improvements on urban and suburban arterials" at signalized intersections using datasets from Toronto, Ontario and Charlotte, North Carolina. [8] Our work utilizes a similar approach to develop a methodology using data from San Francisco.

SFDPH began this work by consulting with the Citywide Pedestrian Safety Task Force's Data Subcommittee who provided feedback on the following model framework (Figure 2) and model inputs detailed in Table 1. The Data Subcommittee included representatives from City agencies including the San Francisco Police Department, the San Francisco Planning Department, the San Francisco County Transportation Agency, the San Francisco Injury Center, the San Francisco Municipal Transportation Agency, the San Francisco Mayor's Office on Disability, the San Francisco General Hospital Trauma Center and the San Francisco Fire Department – Emergency Medical Services, as well as community organizations Walk SF, Senior Action Network/California WALKS and the Pedestrian Safety Advisory Committee to the Board of Supervisors.

Figure 2 describes the conceptual framework that informed our model development. Specifically, that characteristics of the environment around signalized intersections – including Transportation System Operational characteristics (e.g., traffic volumes), Street and Sidewalk characteristics (e.g., presence of parking, trees), and characteristics related to Pedestrian Activity (or its proxies), Transit, and Vulnerable Pedestrians all predict pedestrian injuries at signalized intersections. These factors all shape behaviors and vulnerabilities that occur or exist at the level of pedestrian/motorist behavior and the vulnerability of individual pedestrians – however those specific behaviors and/or individual pedestrian characteristics while on the pathway to injury are not part of the model.

Figure 2. Vehicle-Pedestrian Injury Collisions at Signalized Intersections: Conceptual Framework

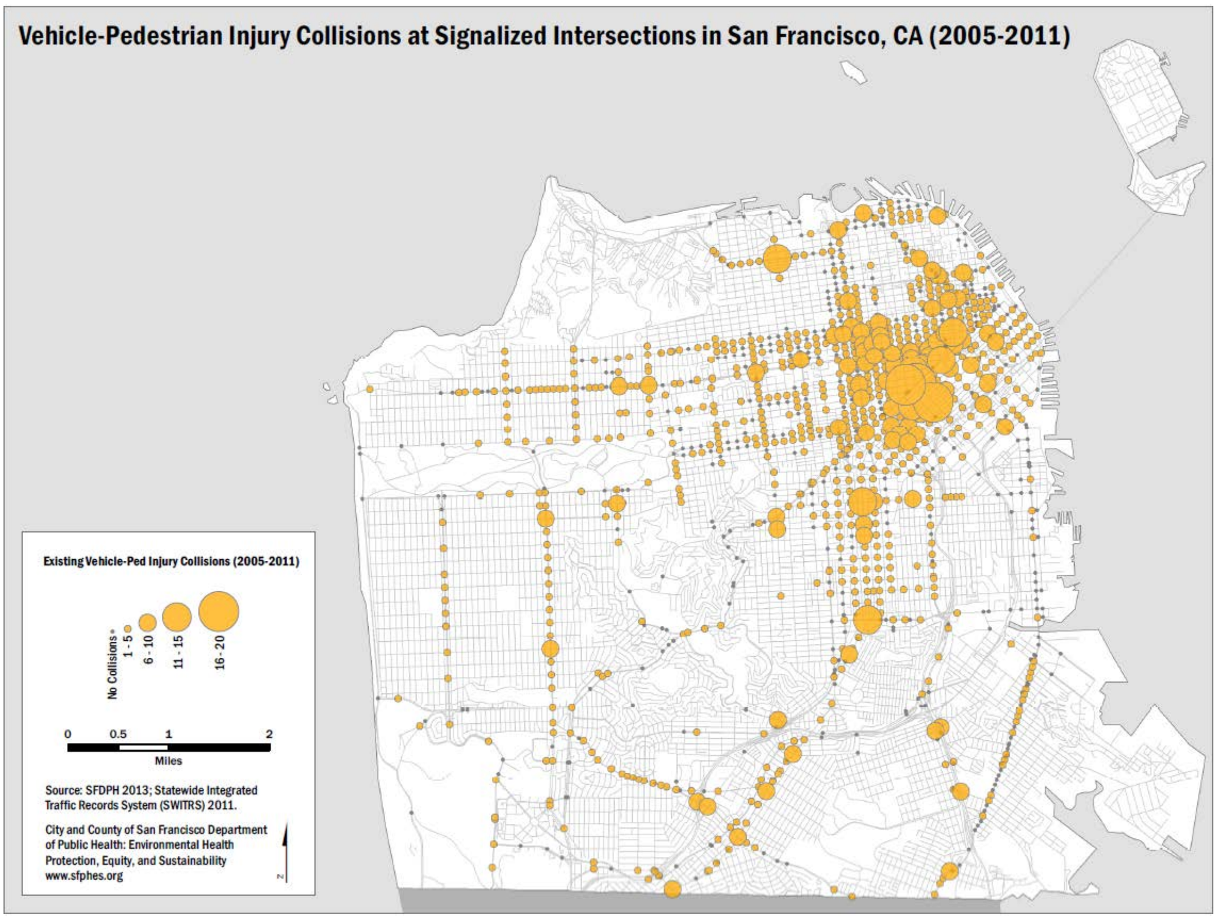


The next sections of this paper detail model development and application informed by this conceptual framework.

II. METHODS

The signalized intersection level model is based on cross-sectional data for the n=1,230 signalized intersections in the City and County of San Francisco, California, depicted in Figure 3, below.

Figure 3. Vehicle-Pedestrian Injury Collisions at Signalized Intersections in the City and County of San Francisco, California (n=2,441 collisions at n=1,230 intersections, 2005-2011)



We used data on vehicle-pedestrian injury collisions in San Francisco, 2005–2011, from the Statewide Integrated Traffic Records System (SWITRS) which contains data on reported vehicle collisions on public roadways.[9] SWITRS vehicle-pedestrian injury collision data were imported into ArcGIS (version 10.1; ESRI Inc., Redlands, CA, USA) and geocoded to the intersection of the reported primary and secondary streets (exact street address is not collected). 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." We restricted our analysis to collisions occurring at or within 20 feet of a signalized intersection, consistent with the definition of intersection collisions used by San Francisco Municipal Transportation Agency traffic engineers. This resulted in a total of n=2,441 collisions at the n=1,230 intersections in the seven-year analysis timeframe.

Table 1 describes the independent variables considered in our analysis and data source and year. We selected our analytic variables based on the previous literature and our interest in environmental predictors of vehicle-pedestrian injury collisions at signalized intersections as detailed in Figure 2.

We used negative binomial regression (NBR) to model the count of vehicle-pedestrian injury collisions over a 7-year period, a model commonly employed in the traffic safety literature.[10] The model form used is the Poisson model with an additional variance parameter (alpha) that accounts for over-dispersion (when the variance is much greater than the mean), which was evident in the collision data. While the Poisson modeling approach assumes that locations with the same covariates will have the same underlying rate, the NBR approach allows this rate to vary (an approach used when there are known important unmeasured covariates, also true in our data with respect to pedestrian volume).

The model form used for our analyses is:

$E(PI) = \exp(b_0 + b_1X_1 + \ldots b_nX_n)$

Where:

E(PI)=predicted vehicle-pedestrian injury collisions per signalized intersection

b_0 = intercept

b_1 = model coefficient for unit change in predictor variable i (signalized intersection level)

X_1 = signalized intersection-level data for predictor variable i

b_x = model coefficient for unit change in predictor variable x (signalized intersection level)

X_x = signalized intersection-level data for predictor variable x

We applied a natural log transformation to the traffic volume, bus volume, employee, and resident population variables, similar to previous research.[11]

We used the conceptual framework for our model building approach (Figure 2). We started with a base model including the Transportation System Operational Characteristics in Table 1, along with a ratio of minimum / maximum traffic volume (consistent with the NCHRP recommended approach).[12] We then added Street Characteristics followed by Pedestrian Activity Proxies and then Vulnerable Population Proxies. In each step, variables were dropped from the model based on coefficient p-value.

We assessed model fit based on goodness-of-fit statistics recommended in the literature: the likelihood ratio R^2, and the Mean Pearson X^2 and Mean Deviance statistics.[13] The likelihood ratio R^2 "represents the extent to which the model explains more of the variation of the dependent variable than an intercept-only model."[14] Notably there is not an equivalent goodness-of-fit test for negative binomial regression as the R^2 for an ordinary least squared regression model and we thus relied on the best estimates based on the literature. Statistical analyses were conducted using STATA software (version 9.2; StataCorp, College Station, TX, USA).

III. RESULTS

There were at total of 5,225 vehicle-pedestrian injury collisions recorded by the police and geocodable in San Francisco from 2005 to 2011. Of those, 47% (n=2,441) occurred at signalized intersections (Figure 4); 31% occurred mid-block (>20 feet from an intersection). Signalized intersections account for 16% of intersections in San Francisco (1230/7723 total intersections) – and almost half of all vehicle-pedestrian injury collisions. Signalized intersections in San Francisco are generally located on busier arterials with heavier traffic volumes.

Figure 4. Vehicle-Pedestrian Injury Collisions by Location Type (2005-2011), n=5225

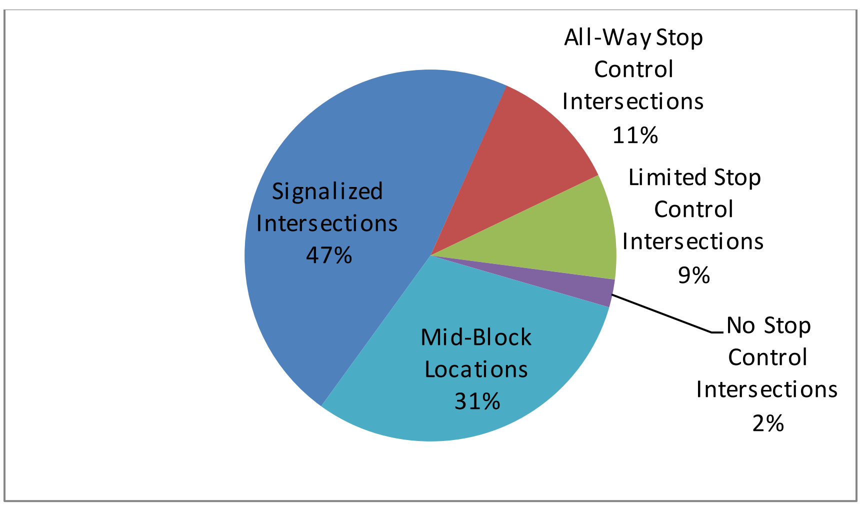


Figure 5 illustrates the frequency distribution of vehicle-pedestrian injury collisions at signalized intersections in San Francisco; Figure 3 depicts the geographic distribution of those collision.

Figure 5. Histogram of Vehicle-Pedestrian Injury Collision Counts at Signalized Intersections in San Francisco, California (n=2,441, 2005-2011)

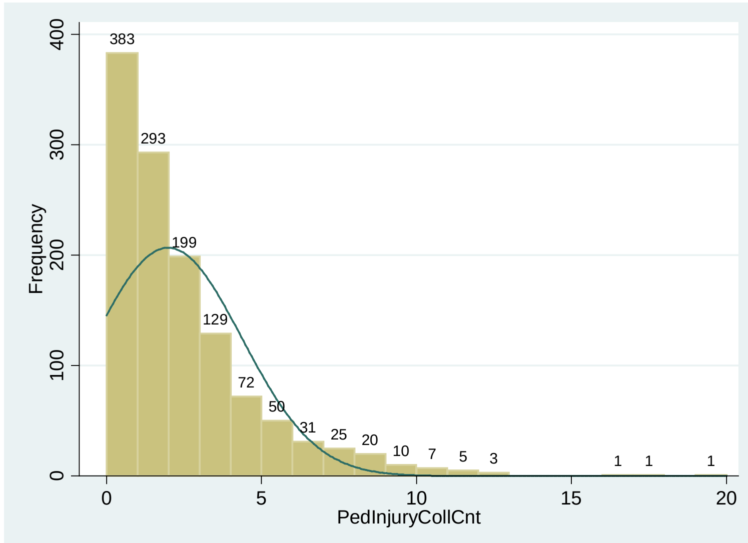


Table 1 shows the mean, median, standard deviation, and range for Transportation System Operational Characteristics, Street Characteristics, Pedestrian Activity Proxies and Vulnerable Population Proxies at signalized intersections in San Francisco. There is notable diversity in conditions across signalized intersections. 58% had a maximum speed limit of 25 mph (miles per hour), while 26% had 30 mph, and 14% had 35 mph with 3% at 40 mph or higher. 20% of signalized intersections have only three legs, while 77% have four legs and just 2% have 5 legs or more. There was an average of 3.3 trees at each intersection, ranging from 0-14 trees. Residential and employee density also varied notably – with the residential average of 5,688 residents within ¼ mile, ranging from 0 – 25,238, and 7,344 employees per ¼ mile, ranging from 0-56,316. Average household income within ¼ mile ranged from $28,835-246,475, with an average of $108,315. There was also notable diversity in the proportion of youth and seniors living within ¼ mile, with the youth average of 11% (ranging from 0-26%) and the senior average of 14% (ranging from 0-47%).

Table 2 details the results of the best fitting model. Increases in traffic volume (log), traffic ratio (minimum/maximum leg), residents (log, number within ¼ mile), employees (log, number within ¼ mile), muni stops within 100 feet, bus volume (log at intersection), and number of single room occupancy hotels within ¼ mile of the intersection all predict increased vehicle-pedestrian injury collisions at signalized intersections. Three-legged intersections, number of trees within 100 feet of intersection, slope (maximum at intersection), and average household income within ¼ mile of intersection are all inversely associated with vehicle-pedestrian injury collisions at signalized intersections. Goodness-of-fit tests indicated that the negative binomial model was preferred to Poisson, with the dispersion parameter alpha significantly contributing to the results at $p<0.001$. The likelihood ratio R^2 statistic was relatively high at 0.67 indicating the model explains substantially more of the variance in the data than an intercept only model, higher than the statistics report for models in the NCHRP report.[15] The mean Pearson chi-square and mean Deviance are also within the recommended acceptable range of 0.8-1.2.[16]

IV. APPLICATION

We then applied the model in a case study application to estimate future changes in vehicle-pedestrian injury collisions in the Central Corridor project area, an area consisting of 12 large city blocks in San Francisco's South of Market neighborhood currently undergoing substantial rezoning that would impact on residential and employee density as well as transportation conditions. This area is also divided by the I-80 freeway, with its freeway ramps bringing thousands of vehicles into and out of the community each day including commuters coming across the Bay Bridge. For this initial application, we used estimates from existing reports of estimated increases in residential and employee density[17] and traffic volumes.[18] These sources provide the following estimates:

- 15-35% increase in area traffic volumes: we chose the mid-point, 25%, for our application
- 1.94% increase in area employees
- 2.35% increase in area residents

We applied these estimates to the final model equation (below), assuming all other model covariates were held constant:

Vehicle-Pedestrian Injury Collisions (n) **=** $\mathbf{exp(-5.428 + b_{tvolln}(.2561) + b_{mintv/maxtv}(.3447) + b_{3legs}(-.6893) + b_{trees}(-.0295) + b_{empln}(.0665) + b_{resln}(.2762) + b_{muni}(.1021) + b_{busln}(.1749) + b_{slope}(-.0260) + b_{hhinck}(-.0028) + b_{sro}(.0076))}$

Where:

tvolln =	Traffic Volume at intersection (aggregate, log N)
mintv/maxtv =	Minimum Traffic Volume / Maximum Traffic Volume ratio
3legs =	3-Leg Intersection
trees =	Trees within 100 feet of intersection (N)
empln =	Employees within 1/4 mile of intersection (log N)
resln =	Residents within 1/4 mile of intersection (log N)
muni =	Muni stops within 100 feet of intersection (N)
busln =	Bus volume at intersection (log N)
slope =	Slope at intersection (maximum)
hhinck =	Household income, average within 1/4 mile of intersection ($1,000)
sro =	Single resident occupancy hotels within 1/4 mile of intersection (N)

We then subtracted the future estimated conditions results for each signalized intersection from the estimated existing conditions count of vehicle-pedestrian injury collisions and divided that amount by the estimated existing conditions count of vehicle-pedestrian injury collisions to obtain an estimated percent change in future conditions. For this example, given that the same predicted future changes were applied to all area intersections and that all estimates that change in future conditions (traffic volume, residents, employees) were natural log transformations, all signalized intersections in the project area had the same predicted increase in vehicle-pedestrian injury collision of 6.7% in the future scenario. Using the above formula, this means that a 25% increase in traffic volume is associated with a 6% increase in vehicle-pedestrian injury collisions (simplified as a power function, $((1.25^{.2561}) - 1)$), a

1.94% increase in area employees is associated with a 0.1% increase in in vehicle-pedestrian injury collisions (or $((1.0194^{0.0665})-1)$), and a 2.35% increase in area residents is associated with a 0.6% increase in vehicle-pedestrian injury collisions (or $(1.0235^{0.2762})-1)$). Traffic volume is clearly the largest contributor to increased vehicle-pedestrian injury collisions in this case study.

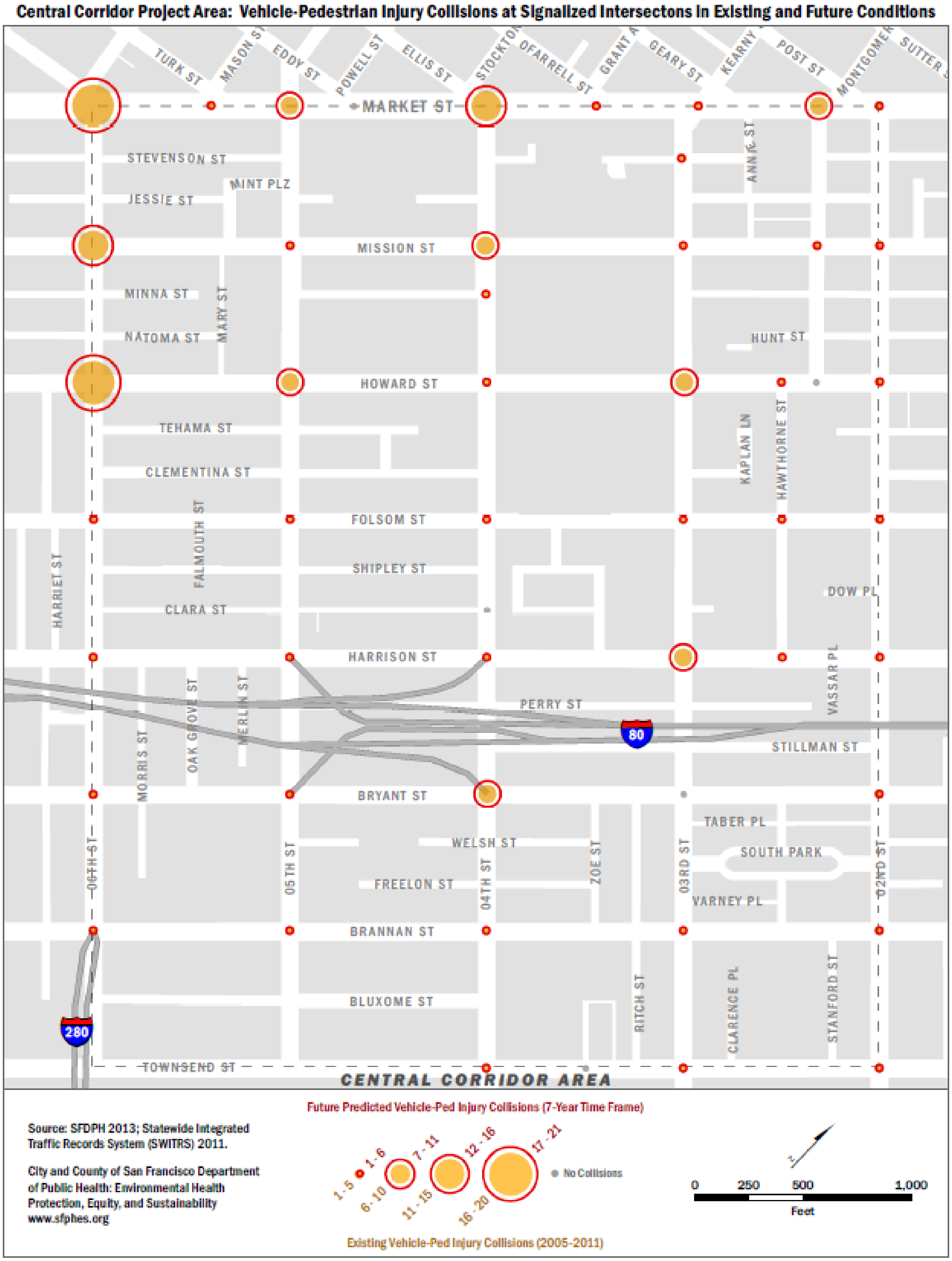
In existing conditions there are on average 210 vehicle-pedestrian injury collisions at the 50 project area intersections in a 7-year period. Using the above estimated changes in traffic volume, residents, and employees and assuming no other changes in other conditions in the project area, the model predicts an increase of 14 vehicle-pedestrian injury collisions in that same area in a seven-year period – or two per year in the 12 block area. Intersections with existing high injury collisions are estimated to have the highest increases – the highest including the corridor from 6th/Market, 6th/Mission and 6th/Howard as well as 4th/Market. Figure 6 depicts the project area in existing conditions, as well as predicted increases in future conditions for a seven-year period.

Different results would be estimated in a future scenario that included changes in other covariates. The model coefficients for the independent variables that are not log-transformed estimate the change in the log count of vehicle-pedestrian injury collisions per unit increase in that predictor. For example, an increase in 5 trees within 100 feet of an intersection is associated with a 14% *decrease* in vehicle-pedestrian injury collisions (exp(5*-0.0295069) = 1-0.86 = 14%), whereas reducing the number of intersection legs to three legs (exp(1*-0.6893) = 1-0.50 = 50%) is associated with a 50% decrease in vehicle-pedestrian injury collisions.

An important caveat regarding the model and this application is that only collisions at signalized intersections are estimated. An additional 210 vehicle-pedestrian injury collisions (147 mid-block, 63 at unsignalized intersections) occurred in the area from 2005-2011, that are not reflected in Figure 6 or modeled in future conditions. Model findings are thus informative regarding the need for additional attention to pedestrian safety in the context of this project, and focus particularly at signalized intersections.

Notably, there are a number of pedestrian safety improvements slated for the Central Corridor project area, including lane reductions, sidewalk widening, signalized crosswalk installations, corner bulb-outs, street trees and furnishings (though not yet site specific), and restricting curb cuts – many of which SFDPH has already supported through work with Central Corridor project staff. Using the above findings, SFDPH will continue to work with transportation and planning staff to better understand the slated changes for the area, to obtain more site-specific data as it becomes available, and to better understand how model results can better demonstrate the need for targeted pedestrian safety improvements.

Figure 6. Central Corridor Project Area: Vehicle-Pedestrian Injury Collisions at Signalized Intersections in Existing and Estimated Future Conditions (Seven-Year Time Frame)



V. DISCUSSION AND NEXT STEPS

In San Francisco, California, a multivariate, negative binomial regression model of vehicle-pedestrian injury collisions at signalized intersections found statistically significant predictors of collisions to include traffic and bus volumes, intersection leg count, trees, employee and resident population, bus stops, street slope, and average household income and single resident occupancy hotels. Model variables had a positive association with vehicle-pedestrian injury collisions, with the exception of three-leg intersections, slope, trees, and average household income within ¼ mile, which had an inverse association with vehicle-pedestrian injury collisions.

Predictive modeling and the resulting findings should be applied and interpreted with caution. Consistent with previously developed models, this model was developed to calculate expected numbers of injuries absent preventive measures. Once those future collision conditions are estimated, knowledge about the effectiveness of specific prevention efforts (e.g., crash reductions factors for engineering countermeasures) can be used to estimate the effectiveness of particular interventions on collision reduction**.** With respect to HIA, this type of modeling can be helpful to estimate potential changes in future conditions to inform the need for interventions to protect or promote health. In the application example, estimated increases in vehicle-pedestrian injury collisions provide information regarding the need for pedestrian safety measures to be considered in the context of a project. While it is important to be cautious, model findings are informative regarding associations between vehicle-pedestrian injury collisions and transportation system, street and sidewalk, and other factors that are potential proxies for pedestrian activity and vulnerable populations.

The model was created based on police-reported injury data. Observed counts of vehicle-pedestrian injuries based upon the collisions reported to and recorded by police generally underestimate the true burden of vehicle-pedestrian injuries. An analysis in San Francisco comparing data from 2000-2001 police records with hospital data from San Francisco General Hospital (SFGH, the City's Level-I Trauma Center which sees the majority of more severely injured pedestrians) found that 22% of pedestrians injured and seen at SFGH were not reported in police records.[19] Undercounts of baseline injuries based on police record data thus will translate into undercounts of predicted future injuries.

SFDPH's previous census-tract level model of vehicle-pedestrian injury collisions[20] as well as a number of other published studies have found traffic volumes to be predictive of vehicle-pedestrian injury collisions in a variety of contexts and geographic levels.[21] The coefficient for traffic volume (0.256) is notably lower than in SFDPH's census tract level model (0.753), however this signalized intersection model has a separate coefficient for bus volume (0.175), and includes an additional variable for the minimum / maximum traffic volume (0.345) which likely explains some of these differences. Our model coefficients for traffic volume and the minimum/maximum volume ratio were similar to the coefficients in the final base model recommended for four-leg intersections (0.40 and 0.26, respectively) developed by the National Cooperative Highway Research Program.[22] Distinct from that approach, we included intersection legs as a potential covariate in our model so that we could directly assess its contribution to vehicle-pedestrian injury collisions.

Both bus volumes and the number of bus stops within 100 feet of an intersection were positively associated with increases in vehicle-pedestrian injury collisions in our model. This makes intuitive sense, in that public transit is strong attractor for pedestrians and that corridors with high bus volumes may also have relatively higher numbers of both traffic and pedestrian volumes. These findings suggest additional analyses could be informative to better understand the specific factors contributing to increases in vehicle-pedestrian injury collisions around transit, and potential transit-specific pedestrian safety improvements that could be implemented.

Increases in vehicle-pedestrian injury collisions were also associated with decreasing average income (within one-quarter mile) and increases in single resident occupancy hotels (SROs) in San Francisco – independent of traffic volume. Multiple factors could potentially explain this finding, including increased reliance on walking among lower income populations, concentrations of populations particularly vulnerable to pedestrian injury living in SROs including seniors and people with disabilities, or potential unadjusted confounding factors that increase risk for pedestrian injury (e.g., vehicle speeds) also being concentrated in these communities. Though we tested the contribution of maximum speed limit at the intersection, we were not able to include a reliable citywide vehicle speed assessment variable in our model, which strongly predicts injury severity. The finding of increased vehicle-pedestrian injuries in lower income communities is consistent with previous research in San Francisco and other cities.[23,24] Recent research stresses the importance of reducing traffic volume and improving roadway design, particularly in these communities, to help addressing existing injury disparities.[25]

A number of factors significant in the model are potential proxies for pedestrian activity – including resident and employee populations, the presence of bus stops, and the number of SROs. The application of the model and its results are not intended to be interpreted to inform decisions or measures that discourage walking – but rather provide additional information regarding where pedestrian safety measures can be targeted to prevent injury and death in the context of ongoing land use and transportation system decisions.

As noted in the application example, an important caveat in the application of this model to a project area is that it is only applicable to signalized intersections. As a next step, SFDPH plans to continue model development for non-signalized intersections and mid-block collisions. In the absence of models for those locations, a signalized intersection model is informative to estimate the potential impacts of future conditions associated with transportation and land use changes at signalized intersections where almost half of vehicle-pedestrian injury collisions in San Francisco occur.

We did not include some variables in the modeling process due to their high correlation with multiple other variables of interest. These variables include: number of lanes at the intersection, intersection density within ¼ mile, off-street parking spaces, residential and commercial zoning, schools or senior centers within ¼ mile, and proportion of residents with disabilities and non-English speaking.

Intersection-level engineering factors were not included in this model of signalized intersections for a number of reasons. The presence of such countermeasures are confounded by other model covariates (e.g., traffic volumes) which are used to determine installation. This model also used a cross-sectional approach that would not be appropriate to assess the effectiveness of such measures given the lack of data regarding the timing of their implementation. Clear guidance regarding the use of multivariate models to conduct such research exists but is not within the scope of this modeling effort.[26]

VI. CONCLUSION

The results of this model reveal the significant contribution of environmental factors – specifically traffic volumes as well as proxies for pedestrian volume and transportation system features including transit – to the occurrence and location of pedestrian injuries in San Francisco. SFDPH's findings illustrate that identifying where land use and transportation factors are predicted to change can help predict where pedestrian injury would be likely to increase. Understanding these relationships provides a strategic opportunity to prevent pedestrian injury by proactively incorporating pedestrian safety improvements in these plans, which can be advanced through SFDPH's ongoing engagement in inter-agency efforts to reduce pedestrian injury in San Francisco.

REFERENCES

[1] Ewing R. 2006. Fatal and non-fatal injuries. In: understanding the relationship between public health and the built environment: A report prepared for the LEED-ND Core Committee. Available at: www.activeliving.org/files/LEED_ND_report.pdf.

Harwood D, Zegeer C, Lyon C, et al. 2008. NCHRP Web-Only Document 129: Pedestrian Safety Prediction Methodology. National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C. Available at: onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w129p3.pdf.

NHTSA (National Highway Traffic Safety Administration). 1999. Literature Review on Vehicle Travel Speeds and Pedestrian Injuries. Washington, DC: DOT HS 809 021. Available at: http://www.nhtsa.gov/people/injury/research/pub/hs809012.html.

Retting RA, Ferguson SA, McCartt AT. 2003. A review of evidence-based traffic engineering measures designed to reduce pedestrian-motor vehicle crashes. American Journal of Public Health 93(9): 1456-63.

[2] Bhatia R, Wier M. 2011. "Safety In Numbers" Re-examined: Can we make valid or practical inferences from available evidence? Accident Analysis & Prevention 43(1): 235-240.

[3] Brüde U, Larsson J. 1993. Models for predicting accidents at junctions where pedestrians and cyclists are involved. How well do they fit? Accident Analysis and Prevention 25(5): 499-509.

Geyer J, Raford N, Ragland D, Pham T. 2006. The continuing debate about Safety in Numbers – Data from Oakland, CA. TRB Annual Meeting CD-ROM.

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

Lee C, Abdel-Aty M. 2005. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. Accident Analysis and Prevention 37: 775-786.

Loukaitou-Sideris A, Ligget R, Sung HG. 2007. Death on the crosswalk: A study of pedestrian-automobile collisions in Los Angeles. Journal of Planning Education and Research 26: 338-351.

Lyon C, Persaud B. 2002. Pedestrian collision prediction models for urban intersections. Transportation Research Record 1818: 102-107.

Wier M, Weintraub J, Humphreys EH, Seto E, Bhatia R. 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention 41(1): 137-145.

[4] Cottrill CD, Thakuriah PV. 2010. Evaluating pedestrian crashes in areas with high low-income or minority populations. Accident Analysis and Prevention 42(6): 1718-1728.

[5] Wier M, Weintraub J, Humphreys EH, Seto E, Bhatia R. 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention 41(1): 137-145.

[6] Bhatia R, Wier M, Weintraub J. 2007. Impacts of Urban Land Use Development on Pedestrian-Motor Vehicle Collisions: An Application of the San Francisco Pedestrian Injury Model to Five Neighborhood Plans. San Francisco Department of Public Health, Environmental Health Section. San Francisco, CA.

[7] Wier M, Bhatia R, McLaughlin J, Morris D, Comerford Scully C, Harris M, Bedoya J, Cowles S, Rivard T.

Health Effects of Road Pricing in San Francisco, California: Findings from a Health Impact Assessment.

San Francisco, CA: San Francisco Department of Public Health, September 2011.

[8] National Cooperative Highway Research Program. 2008. Pedestrian Safety Prediction Methodology. Web-Only Document 129: Phase III. Transportation Research Board of the National Academies.

[9] California Highway Patrol. 2013. Available at: http://iswitrs.chp.ca.gov/Reports/jsp/userLogin.jsp,

[10] National Cooperative Highway Research Program. 2008. Pedestrian Safety Prediction Methodology. Web-Only Document 129: Phase III. Transportation Research Board of the National Academies.

Pulugurtha S, Sambhara V. 2011. Pedestrian crash estimation models for signalized intersections. Accident Analysis and Prevention 43:439-446.

[11] Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. Accident Analysis and Prevention 37, 775–786.

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.

Wier M, Weintraub J, Humphreys EH, Seto E, Bhatia R. 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention 41(1): 137-145.

[12] National Cooperative Highway Research Program. 2008. Pedestrian Safety Prediction Methodology. Web-Only Document 129: Phase III. Transportation Research Board of the National Academies.

[13] Hadayeghi A, Shalaby AS, Persaud BN. Macrolevel accidente prediciton models for evaluating safety of urban transportation systems.Transportation Research Record No. 1840, Statistical Methods and Modeling and Safety Data, Analysis, and Evaluation, 2003, 87-95.

National Cooperative Highway Research Program. 2008. Pedestrian Safety Prediction Methodology. Web-Only Document 129: Phase III. Transportation Research Board of the National Academies.

[14] National Cooperative Highway Research Program. 2008. Pedestrian Safety Prediction Methodology. Web-Only Document 129: Phase III. Transportation Research Board of the National Academies.

[15] National Cooperative Highway Research Program. 2008. Pedestrian Safety Prediction Methodology. Web-Only Document 129: Phase III. Transportation Research Board of the National Academies.

[16] Hadayeghi A, Shalaby AS, Persaud BN. Macrolevel accident prediction models for evaluating safety of urban transportation systems. Transportation Research Record No. 1840, Statistical Methods and Modeling and Safety Data, Analysis, and Evaluation, 2003, 87-95.

[17] San Francisco Planning Department. 2013. Central Corridor Plan – Draft. San Francisco, CA.

[18] San Francisco Municipal Transportation Agency. 2011. Eastern Neighborhoods Transportation Implementation Study Final Report. San Francisco, CA.

[19] Sciortino S, Vassar M, Radetsky M, Knudson MM. 2005. San Francisco pedestrian injury surveillance: mapping, under-reporting, and injury severity in police and hospital records. Accident Analysis and Prevention 37(6): 1102-1113.

[20] Wier M, Weintraub J, Humphreys EH, Seto E, Bhatia R. 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention 41(1): 137-145.

[21] Elvik R. 2009. The non-linearity of risk and the promotion of environmentally sustainable transport. Accident Analysis and Prevention 41:849-855.

[22] National Cooperative Highway Research Program. 2008. Pedestrian Safety Prediction Methodology. Web-Only Document 129: Phase III. Transportation Research Board of the National Academies.

[23] Wier M, Weintraub J, Humphreys EH, Seto E, Bhatia R. 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention 41(1): 137-145.

[24] Chakravarthy B, Anderson CL, Ludlow J, Lotfipour S, Vaca FE. 2010. The relationship of pedestrian injuries to socioeconomic characteristics in a large Southern California County. Traffic Injury Prevention. 11(5):508-13.

[25] Morency P. et al. 2012. Neighborhood social inequalities in road traffic injuries: the influence of traffic volume and road design. American Journal of Public Health. 102(6):1112-9.

[26] Elvik R. 2011. Assessing causality in multivariate accident models. Accident Analysis and Prevention. 43: 253-264.

Table 1. Bivariate Statistics for Independent Variables:
Vehicle-Pedestrian Injury Collisions at Signalized Intersections in San Francisco, California
(SWITRS 2005-2011, N=2,441 Collisions at N=1,230 intersections)

Variables	Mean	Median	Minimum	Maximum	Standard Deviation	Data Source (Year)
Transportation System Operational Characteristics						
Traffic volume, intersection aggregated (Avg. Annual Daily)	57,329	50,416	2	645,645	44,015	SFCTA (2005)
Traffic volume: Minimum Leg Volume / Maximum Leg Volume (N)	0.4	0.3	0.0	1.0	0.3	SFCTA (2005)
Speed limit, maximum at intersection (%)						SFMTA (2013)
25 mph	58%	100%	0%	100%	49%	
30 mph	26%	0%	0%	100%	44%	
35 mph	14%	0%	0%	100%	35%	
40 mph	1%	0%	0%	100%	8%	
45 mph	2%	0%	0%	100%	13%	
Street Characteristics						
Legs at Intersection (%)						SFDPW (2010)
3	20%	0%	0%	100%	40%	
4	77%	100%	0%	100%	42%	
5 or 6	3%	0%	0%	100%	17%	
Freeway ramps within 1/4 mile (N)	0.5	0.0	0.0	8.0	1.3	SFMTA (2010)
Truck route at intersection (%)	81%	100%	0%	100%	40%	SFMTA (2009)
One way street at intersection (%)	42%	0%	0%	100%	49%	SFCTA (2010)
Bike lane at intersection (%)	42%	0%	0%	100%	49%	SFMTA (2010)
Street trees within 100 feet (N)	3.3	3.0	0.0	14.0	3.0	SAIC (2007)
Metered parking at intersection (%)	8%	0%	0%	100%	28%	SFMTA (2010)
Pedestrian Activity Proxies						
Residential population density within 1/4 mile (N)	5,688	4,876	0	25,238	4,078	Census (2010)
Employee population density within 1/4 mile (N)	7,344	2,448	33	56,316	10,835	LEHD (2009)
Bus stops within 100 feet (N)	0.9	0.0	0.0	8.0	1.1	SFMTA (2010)
Bus volume, intersection aggregated (Avg. Annual Daily)	128	100	0	1,092	130	SFCTA (2005)
Regional rail transit station within 1/4 mile (n)	0.2	0	0	6.0	0.7	SFMTA (2010)
University within 1/4 mile (%)	6%	0%	0%	100%	24%	SF Planning (2011)
Parks within 1/4 mile (N)	1.8	2.0	0.0	7.0	1.3	SFRP (2008)
Zoning within 1/4 mile (%, land area)						SF Planning (2011)
Industrial/PDR	4%	0%	0%	85%	12%	
Neighborhood Commercial	7%	5%	0%	32%	7%	
Mixed Use	1%	0%	0%	26%	4%	
Public Use	10%	6%	0%	89%	12%	
Slope, major street (N)	5.4	4.0	0.0	29.0	3.7	SFDPH (2007)
Vulnerable Population Proxies						
Alcohol outlets within 1/4 mile (N)	9	7	0	43	8	CDABC (2011)
Single resident occupancy hotels within 1/4 mile (N)	4	0	0	78	11	SFDPH (2012)
Public health centers within 1/4 mile (N)	2	1	0	11	2	SFDPH (2010)
Average household income within 1/4 mile ($)	$ 108,315	$ 101,855	$ 28,835	$ 246,475	$ 42,881	
People aged 17 and under within 1/4 mile (%)	11%	10%	0%	26%	5%	Census (2010)
People aged 65 and older within 1/4 mile (%)	14%	13%	0%	47%	7%	Census (2010)

CDABC: California Department of Alcoholic Beverage Control
Census: U.S. Census
LEHD: Longitudinal Employer andHousehold Dynamics Program (Census)
SAIC: Science Applications International Corporation
SFCTA: San Francisco County Transportation Authority's Travel Forecasting Model, SF-CHAMP 4
SFDPH: San Francisco Department of Public Health
SFDPW: San Francisco Department of Public Works
SFDTIS: San Francisco Department of Telecommunications and Information Services
SFMTA: San Francisco Municipal Transportation Agency
SF Planning: San Francisco Planning Department
SFRP: San Francisco Department of Recreation and Parks
SWITRS: Statewide Integrated Traffic Records System

Table 2. Final Model of Vehicle-Pedestrian Injury Collisions at Signalized Intersections in San Francisco, California (SWITRS 2005-2011, N=2,441 Collisions at N=1,230 intersections)

Negative Binomial Regression

Parameter	Estimate	Standard error (SE)	p-value	95% CI, Lower	95% CI, Upper
Traffic Volume at intersection (log N)	0.256	0.037	0.000	0.184	0.329
Minimum Traffic Volume / Maximum Traffic Volume	0.345	0.101	0.001	0.148	0.542
3-Leg Intersection	-0.689	0.091	0.000	-0.867	-0.511
Trees within 100 feet of intersection (N)	-0.030	0.010	0.002	-0.048	-0.011
Employees within 1/4 mile of intersection (log N)	0.067	0.020	0.001	0.028	0.105
Residents within 1/4 mile of intersection (log N)	0.276	0.049	0.000	0.180	0.373
Bus stops within 100 feet of intersection (N)	0.102	0.024	0.000	0.055	0.149
Bus volume at intersection (log N)	0.175	0.023	0.000	0.130	0.219
Slope at intersection (maximum)	-0.026	0.008	0.001	-0.042	-0.010
Household income, average within 1/4 mile of intersection ($1,000)	-0.003	0.001	0.001	-0.004	-0.001
Single resident occupancy hotels within 1/4 mile of intersection (N)	0.008	0.002	0.002	0.003	0.012
intercept	-5.428	0.621	0.000	-6.644	-4.212
Dispersion parameter (α), model	*0.301*	*0.036*		*0.238*	*0.381*

Model Characteristics	Estimate or Measure
Number of observations (n)	1230
Number of parameters (p)	11
Log-likelihood	-2074.337
Dispersion parameter (α), model with only the *constant*	0.8783203

Goodness of Fit	**Estimate or Measure**
Likelihood ratio test of α=0	LR=166.96 (p< 0.001)
$R_{\alpha}^2 = 1 - \alpha_{final} / \alpha_{constant\text{-}only}$	0.65716755
Mean/Scaled Pearson X^2/(df)	1.08
Mean/Scaled Deviance/(df)	1.10