

The effects of pedestrian and bicycle exposure on crash risk in Minneapolis

Tao Tao

University of Minnesota
taotao@umn.edu

Greg Lindsey (corresponding author)

University of Minnesota
linds301@umn.edu

Jason Cao

University of Minnesota
cao@umn.edu

Jueyu Wang

University of North Carolina, Chapel Hill
olivia.wang@unc.edu

Abstract: Exposure to risk is a theoretically important correlate of crash risk, but many safety performance functions (SPFs) for pedestrian and bicycle traffic have yet to include the mode-specific measures of exposure. When SPFs are used in the systematic approach to assess network-wide crash risk, the omission of the exposure potentially could affect the identification of high-risk locations. Using crash data from Minneapolis, this study constructs and compares two sets of SPFs, one with pedestrian and bicycle exposure variables and the other without, for network-wide intersection and mid-block crash models. Inclusion of mode-specific exposure variables improves model validity and measures of goodness-of-fit and increases accuracy of predictions of pedestrian and bicycle crash risk. Including these exposure variables in the SPFs changes the distribution of high-risk locations, including the proportion of high-risk locations in low-income and racially concentrated areas. These results confirm the importance of incorporating exposure measures within SPFs and the need for pedestrian and bicycle monitoring programs to generate exposure data.

Keywords: Pedestrian, bicycle, exposure, crash risk, safety performance function, equity

Article history:

Received: March 7, 2021

Received in revised form: July 13, 2021

Accepted: September 17, 2021

Available online: October 29, 2021

1 Introduction

Pedestrian and bicycle crashes are a major concern in contemporary transportation planning. During the past decades, federal, state, and local policies have encouraged multimodal transportation systems that integrate facilities for vehicular, transit, and non-motorized traffic. For example, many states and cities are participating in the nationwide Complete Streets initiative and implementing roadway improvements to accommodate cyclists, pedestrians, drivers, and transit riders of all ages and ability levels (Burden & Litman, 2011; La Plante & McCann, 2008). As non-motorized modes of transportation have been integrated with motorized modes, more and more pedestrians and bicyclists have been exposed to vehicular traffic, increasing concerns for their safety. In 2016, 5,987 pedestrians and 840

Copyright 2021 Tao Tao, Greg Lindsey, Jason Cao & Jueyu Wang

<http://dx.doi.org/10.5198/jtl.u.2021.1980>

ISSN: 1938-7849 | Licensed under the [Creative Commons Attribution – Noncommercial License 4.0](https://creativecommons.org/licenses/by-nc/4.0/)

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies.

bicyclists were killed in traffic crashes, accounting for 18.2% of all traffic fatalities in the year (NHTSA, 2018b, 2018a).

To identify and improve high-risk locations, many studies have reported safety performance functions (SPFs) to model the relationships between pedestrian/bicycle crashes and risk factors (FHWA, 2013). Exposure to risk, defined as the number of potential opportunities for a crash to occur, is one of the key risk factors (FHWA, 2018; Merlin et al., 2020). It is theoretically clear that pedestrian and bicycle exposure has significant relationships with pedestrian and bicycle crashes (Scheepers et al., 2011; Thomas et al., 2017). However, many, if not most studies have not incorporated measures of pedestrian and bicycle exposure, mainly due to lack of related datasets. The failure to include these measures in SPFs may lead to omitted variable bias and, in turn, misidentify high-risk locations when SPFs are used in systematic applications to assess relative crash risk. Potential consequences of misidentification of high-risk sites include erroneous prioritization of funds available for street improvements and misidentification of risk experienced by poor, marginalized, and vulnerable subpopulations.

Using crash data from Minneapolis between 2005 and 2017, this study develops SPFs to examine the effects of pedestrian and bicycle exposure on pedestrian and bicycle crashes, controlling for vehicular volume, built environment attributes, traffic facilities, and demographics. In particular, we develop two sets of SPFs: one set includes both pedestrian and bicycle exposure and the other set excludes both measures. Each set includes four models for pedestrian and bicycle crash risk at both intersections and mid-blocks (the street segments that connect two consecutive intersections), respectively. We also assess the consequences of excluding pedestrian and bicycle exposure from the SPFs. We address the following research questions:

(1) Are pedestrian and bicycle exposure variables statistically and practically significant in safety performance functions used to predict pedestrian and bicycle crashes?

(2) To what extent does inclusion of pedestrian and bicycle exposure variables change identification of high-risk locations in the entire city and in disadvantaged neighborhoods, respectively?

This study contributes threefold to the literature. First, in addition to vehicular traffic counts, it considers both pedestrian and bicycle exposure variables simultaneously in the SPFs and substantiates the statistically and practically significant effects of these measures on pedestrian and bicycle crashes. Second, this study shows that inclusion of pedestrian and bicycle exposure measures changes a substantial proportion of the high-risk locations, highlighting the important role of these exposure variables in planning practice. Finally, the inclusion of exposure also results in a different proportion of high-risk locations in low-income and racially concentrated areas.

This paper is organized as follows. We begin with a review of the current literature and then present our data and methods. Following presentation and discussion of our results, we conclude with a discussion of limitations, key findings, and policy implications.

2 Literature review

Researchers have conducted many insightful studies to estimate pedestrian and bicycle crash risk (Nordback et al., 2014; Schneider et al., 2010; Thomas et al., 2018) and to model the severity of crashes (Chen & Shen, 2016). Crash risk studies have been conducted at two scales: area-wide analyses and facility-specific investigations (Table 1). Area-wide analyses aggregate historical crashes to the area level using units of analysis such as census tracts (Loukaitou-Sideris et al., 2007), census block groups (CBGs) (Dumbaugh & Li, 2011), or traffic analysis zones (TAZs) (Yasmin & Eluru, 2016). Because pedestrian and bicycle crashes are rare events, area-wide studies help reduce the number of observations with zero crashes. However, these studies have low resolution and cannot be used to identify high-risk locations

or facilities. Facility-specific studies compare crash risk at the facility level. The units of analysis include intersections (Lee & Abdel-Aty, 2005) or mid-blocks (Williams et al., 2018). These studies help identify high-risk locations and provide guidance for pedestrian and bicycle facility improvement. For example, to prioritize street improvements, the city of Minneapolis (City of Minneapolis, 2016) established a ranking system, which incorporates safety (i.e., number of crashes), equity, and infrastructure conditions, to assess all street segments.

Facility-specific studies often employ the systemic approach to assess the traffic environment for pedestrians and bicyclists (Kim & Kim, 2015; Thomas et al., 2017). The core of this approach is to estimate a SPF (i.e., a model depicting the relationships between different risk factors and crash risk), predict crash risk, and examine high-risk locations. For example, using intersection crash data from Seattle between 2004 and 2017, Thomas et al. (2017) constructed negative binomial regression models to estimate the influences of pedestrian traffic count and other types of risk factors on total crashes for the study period. Using the estimated SPFs, they then predicted pedestrian crash risk for all 12,266 intersections in the city and identified high-risk locations for improvement. Our study follows a similar approach conceptually (i.e., predicting total crashes).

In systemic approaches, the choice of risk factors is critical to the validity of study outcomes. When specifying SPFs, previous studies often have considered four types of risk factors, including exposure to risk variables (Schneider et al., 2010; Thomas et al., 2017), built environment variables (Dumbaugh & Li, 2011; Guerra et al., 2019), traffic facility characteristics (Dumbaugh & Li, 2011; Zangenehpour et al., 2016), and socio-demographic characteristics (Loukaitou-Sideris et al., 2007; Yu et al., 2018). Among these factors, exposure to risk is theoretically important. Exposure measures include vehicular traffic volume such as annual average daily traffic (AADT) (Cottrill & Thakuriah, 2010; Loukaitou-Sideris et al., 2007) and comparable measures for pedestrian and bicycle traffic such as annual average daily bicyclists (AADB) and pedestrians (AADP) (Nordback et al., 2014; Thomas et al., 2017). Previous studies have shown that both pedestrian and bicycle exposure variables are positively associated with crash risk (Thomas et al., 2017; Zangenehpour et al., 2016). Given their contribution to crash risk, scholars argued that these variables should be included in the SPFs. In particular, Elvik (2013, p. 57) suggested including exposure to risk as “a count of the number of road users.” Merlin et al. (2020, p. 7) recommended including “mode-specific models of exposure” when studying pedestrian or bicycle crash risk. Elvik (2013) also has shown that it is important to include all relevant measures of exposure (e.g., AADT, AADB, AADP).

Few studies, however, have included pedestrian and bicycle exposure, mainly because comprehensive datasets for non-motorized modes historically have not been available. While the Federal Highway Administration (FHWA), state transportation agencies, and regional and local agencies cooperate in traffic monitoring programs to produce estimates of vehicular AADT for virtually all major roads in the U.S., no similar and comprehensive monitoring programs exist for non-motorized modes, partly because most bicycle and pedestrian traffic operates on local roads. Although an increasing number of state departments of transportation, metropolitan planning organizations (MPOs) and local public works agencies are monitoring non-motorized traffic, most do not manage comprehensive programs. As a result, most studies have yet to account for risks associated with bicycling and walking. As shown in Table 1, nine of the 20 studies incorporated vehicular traffic counts, but only three used pedestrian counts, three used bicycle counts, and none used both. To confirm that the absence of pedestrian and bicycle measures was mainly due to the unavailability of data, we checked the 20 papers for explicit statements and/or contacted authors about the availability of exposure data. For the papers in Table 1, both pedestrian and bicycle counts were available for only two studies (Nordback et al., 2014; Thomas et al., 2017). Thomas et al. (2017) noted that bicycle exposure data were not included in final models; Nordback et al. (2014) did not include pedestrian exposure in bicycle crash models because continuous

count data were not available for estimating AADP from available short-duration counts. The failure to include pedestrian and bicycle exposure variables from SPFs will lead to omitted variable bias (Elvik, 2013; Merlin et al., 2020).

In lieu of pedestrian and bicycle exposure variables, some studies have used proxy measures, such as number of commuters (Yasmin & Eluru, 2016) or population density and job density (Loukaitou-Sideris et al., 2007). Although these proxies are correlated with pedestrian and bicycle exposure, the exposure is also associated with other built environment variables and socio-demographics in the vicinity of the facility (Griswold et al., 2011; Hankey & Lindsey, 2016). Because these proxies explain only a limited proportion of the variation in pedestrian and bicycle exposure, they result in only marginal improvements in the SPFs. As the predicted crash risk is used to determine high-risk locations (Thomas et al., 2018), both the absence of exposure variables and poor proxies may lead to misidentification of high-risk locations and thence misallocation of scarce resources.

Table 1. Selected studies on pedestrian and bicycle crashes

Studies	Crash type	Analysis unit type	Analysis unit	Exposure to risk considered		
				Pedestrian	Bicycle	Vehicle
Lee and Abdel-Aty (2005)	Pedestrian	Facility-specific	Intersection			
Loukaitou-Sideris et al. (2007)	Pedestrian	Area-wide	CBG	Population density and employment density ^a		AADT
Daniels et al. (2009)	Bicycle	Facility-specific	Intersection			
Cottrill and Thakuriah (2010)	Pedestrian	Area-wide	Census tract	Population density ^a		AADT
Schneider et al. (2010)	Pedestrian	Facility-specific	Intersection	Estimated pedestrian count in ten years		Estimated vehicle traffic in ten years
Dumbaugh and Li (2011)	Pedestrian and bicycle	Area-wide	CBG	Population density ^a		Vehicle Mile Travelled (VMT) ^a
Turner et al. (2011)	Bicycle	Facility-specific	Intersections and mid-blocks		Daily bicycle traffic count	Daily vehicle traffic count
Schepers et al. (2011)	Bicycle	Facility-specific	Intersection		Estimated daily bicycle traffic count	Estimated daily vehicle traffic count
Siddiqui et al. (2012)	Pedestrian and bicycle	Area-wide	TAZ	Population density and number of jobs ^a		
Ukkusuri et al. (2012)	Pedestrian	Area-wide	TAZ	Population and population under 17 ^a		
Gladhill and Monsere (2012)	Pedestrian and bicycle	Area-wide	0.1-mile grid	Population and number of jobs ^a		VMT ^a
Nordback et al. (2014)	Bicycle	Facility-specific	Intersection		Estimated AADB	AADT
Park et al. (2015)	Bicycle	Facility-specific	Mid-block	Proportion of commuters by walk in total commuters ^a	Proportion of commuters by bicycle in total commuters ^a	AADT
Chen (2015)	Bicycle	Area-wide	TAZ		Number of bicycle trips ^a	Number of trips ^a
Yasmin and Eluru (2016)	Bicycle	Area-wide	TAZ	Number of walk commuters ^a	Number of bike commuters ^a	Number of vehicles ^a
Zegeer et al. (2017)	Pedestrian	Facility-specific	Intersection	Estimated daily pedestrian traffic count		AADT
Thomas et al. (2017)	Pedestrian	Facility-specific	Intersection	Estimated AADP		Arterial classification ^a
Williams et al. (2018)	Bicycle	Facility-specific	Mid-block			
Yu et al. (2018)	Pedestrian	Area-wide	CBG	Population density and employment density ^a		VMT ^a
Guerra et al. (2019)	Pedestrian	Facility-specific	Intersection and mid-block	Number of trips, population density, and employment density ^a		AADT

Note. ^a Proxy variables for exposure to risk.

Furthermore, the omission of pedestrian and bicycle exposure in SPFs may result in failures to prioritize improvements in non-motorized infrastructure in disadvantaged neighborhoods. Low-income people and minority groups use non-motorized modes of transportation frequently (Renne & Bennett, 2014). Accordingly, low-income and minority-concentrated neighborhoods have more pedestrian and bicycle activities (Cottrill & Thakuriah, 2010; Yu et al., 2018) and a higher proportion of crashes than other neighborhoods (Loukaitou-Sideris et al., 2007; Siddiqui et al., 2014). The disproportionate safety burden of these disadvantaged neighborhoods raises important equity issues and concerns (Loukaitou-Sideris et al., 2007; Yu et al., 2018). Furthermore, omitting pedestrian and bicycle exposure variables or using poor proxies will systematically understate the crash risk of the facilities located in these neighborhoods. In turn, high-risk locations in these neighborhoods will be under-identified. Potential misallocation of budgetary resources for capital improvements from disadvantaged neighborhoods to affluent neighborhoods will maintain transportation disparities for disadvantaged people. However, few studies examine the extent to which high-risk locations in disadvantaged neighborhoods are misidentified and the implications for equity planning.

In summary, most SPFs in published studies of pedestrian and bicycle crash risk do not include both pedestrian and bicycle exposure measures. Few studies examine the impacts of omitting these measures on pedestrian and bicycle planning practice. To illustrate the importance of exposure measures, we incorporate both pedestrian and bicycle exposure variables in the SPFs and compare the performance of the SPFs with and without these variables. We further assess the implications for identification of high-risk locations and social equity.

3 Data and method

3.1 Data

We acquired a dataset of pedestrian and bicycle crashes occurring in the city of Minneapolis reported by police between 2005 and 2017 from the Minnesota Department of Public Safety (DPS, 2018). The data include crash location, road user type, and other detailed information regarding the circumstances of crashes.

Figure 1 illustrates the distribution of crashes in each CBG and two specific areas: the central business district (CBD) and Areas of Concentrated Poverty (ACP50). The ACP50 are designated by the Metropolitan Council, the metropolitan planning organization. These areas are CBGs where “50% or more of residents are people of color, and 40% or more of the residents have family or individual incomes that are less than 185% of the federal poverty threshold” (City of Minneapolis, 2017, pp. 5–23). The Metropolitan Council and cities in the region, including Minneapolis, often consider these areas when spatially targeting equity initiatives. Between 2005 and 2017, 3,812 pedestrian crashes and 3,490 bicycle crashes occurred in Minneapolis. The number of crashes was highest in the CBD, and the number of crashes in the ACP50 was higher than in the areas of the city that were more affluent and majority white. There were more pedestrian crashes than bicycle crashes in the CBD but not in other areas of the city.

We used ArcGIS to categorize crashes into intersection crashes and mid-block crashes. Our rationale for this distinction is that different factors may be associated with crashes at these two types of locations. We defined an intersection crash as a crash that occurred within a 35-meter buffer of the center of an intersection on main and secondary roads (e.g., minor arterials and collectors) and within a 15-meter buffer of the center of an intersection on local roads. We developed this definition based on conversations with practitioners and through empirical tests of different buffer measures. Practitioners

noted that the influence zone of an intersection with respect to crashes varies with road geometry. We tested several combinations of radii and reviewed the number of intersection and mid-block crashes with each scenario, ultimately choosing the combination of 35-meter and 15-meter buffers to identify intersection crashes for main/secondary and local roadways, respectively. All crashes that were not classified as intersection crashes were categorized as mid-block crashes.

We obtained three measures of exposure to risk. The Minneapolis Department of Public Works (DPW) is distinctive among transportation agencies of the major cities in the U.S. in that it has recorded peak-hour pedestrian and bicycle traffic volumes at multiple locations in the city during summer and fall for years (2007-2014). We assigned DPW's pedestrian and bicycle counts to intersections and mid-blocks for further analysis (see Lindsey et al., 2019, p. 12). In total, our intersection and mid-block crash models have sample sizes of 173 and 437, respectively.

Because we aim to predict crash risk for all intersections ($N = 6,639$) and mid-blocks ($N = 12,589$) in the city, we need to produce pedestrian and bicycle exposure for all the locations where actual counts were not available. To achieve this objective, we adapted the estimates of pedestrian and bicycle peak-hour, mid-block traffic produced by the pedestrian and bicycle demand models developed by Hankey and Lindsey (2016). These models estimate mid-block traffic as a function of adjacent land use, street functional class, and other variables (see Appendix A). After obtaining the estimated pedestrian and bicycle traffic counts for the 12,152 mid-blocks where actual counts were not available, we aggregated them for the 6,639 intersections. More information about the data processing could be found in this report (Lindsey et al., 2019).

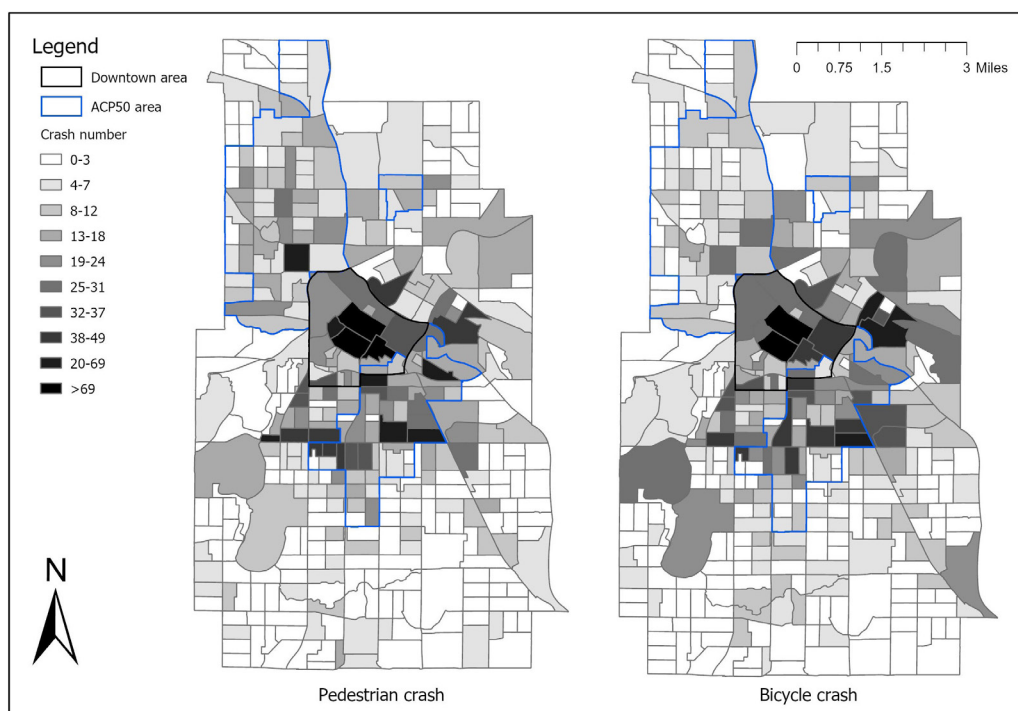


Figure 1. Distribution of pedestrian and bicycle crashes by each CBG in Minneapolis from 2005 to 2017 (ACP50 boundaries were extracted from the Minnesota Geospatial Commons, <http://gisdata.mn.gov>).

We used vehicular AADT of mid-blocks from the DPW to measure vehicular exposure and, following DPW practice, imputed 500 cars per day for missing vehicular AADT on local streets. Although

methods for imputing traffic volumes for local streets have been proposed (Chen et al., 2019), we used the DPW estimate to maintain consistency with local engineering practice. We then used the same method as aggregating pedestrian and bicycle traffic counts to calculate the vehicular AADT for each intersection in Minneapolis.

Besides these three measures of exposure, we assembled data for 24 potential correlates of crash risk (Table 2), including ten built environment variables, eleven traffic facility variables, and six socio-demographic variables. Specifically, traffic facility variables were selected within categories associated with risk: including pedestrian and bicycle facilities, visibility, traffic control, and geometric characteristics (Schneider et al., 2021). We checked for multicollinearity among the variables by assessing the variance inflation factors of all independent variables. The value of each was smaller than 5, indicating multicollinearity is not a threat to validity.

3.2 Method

We used negative binomial regression to model the SPFs for four dependent variables because the dependent variables (i.e., total crashes) are count data, and their variance is larger than their mean (Table 2). These SPFs estimated the number of pedestrian crashes at intersections (called the pedestrian intersection model for simplicity), the number of pedestrian crashes at mid-blocks (called the pedestrian mid-block model), the number of bicycle crashes at intersections (called the bicycle intersection model), and the number of bicycle crashes at mid-blocks (called the bicycle mid-block model). We chose to use total crashes during a long period (2005-2017) for three reasons. First, because pedestrian and bicycle crashes are rare events, using total crashes over a multi-year period years increases the chances of obtaining sufficient crash events for developing SPFs (Thomas et al., 2017). Second, we do not have the annual measures for all the independent variables (e.g., road geometry) in our study period, so we cannot correlate annual changes with annual crashes. Third, using total crashes is more consistent with the practices used by Minneapolis DPW, which incorporates total crashes over time to prioritize street improvements (City of Minneapolis, 2016).

We developed cumulative residual (CURE) plots to determine the appropriate forms of exposure variables in the SPFs. A CURE plot is a line graph showing the relationship between the cumulative residuals of the SPF and an independent variable (Hauer, 2015). CURE plots can help examine whether a selected transformation of an independent variable provides the best fit (Srinivasan & Bauer, 2013; Srinivasan et al., 2013). Several studies have applied transformations of exposure variables in SPFs, including the original form (Thomas et al., 2017), quadratic form (Carlson et al., 2019; Cottrill & Thakuriah, 2010), and logarithmic form (Schepers et al., 2011; Turner et al., 2011). After reviewing the CURE plots for the three exposure variables, we included the logarithmic form of exposure variables in our SPFs (see the CURE plots in Appendix B). In addition to exposure variables, we included built environment measures, traffic facilities, and socio-demographic variables as independent variables in the SPFs. The general form of the SPF is:

$$y = e^{\beta_0} \times B^{\beta_1} \times P^{\beta_2} \times AADT^{\beta_3} \times e^{\beta X} \quad (1)$$

where y is the total crash number for 13 years, B is actual bicycle count, P is actual pedestrian count, $AADT$ is vehicular volume, β_1 to β_3 are the coefficients of the corresponding exposure variables, X is the matrix of other independent variables, β is the corresponding vector of coefficients, and β_0 is the intercept.

We constructed two sets of SPFs: one with pedestrian and bicycle exposure variables and the other without. We included theoretically relevant independent variables in both sets of SPFs. This helps pres-

ent results of all the variables we have tested and reduce the possibility of publication bias. We included the deviance R^2 , Akaike information criterion (AIC), and Bayesian information criterion (BIC) to compare the model fit to the dataset. We also constructed two sets of parsimonious SPFs and presented the results in Appendix C. The rest of this paper focuses on the results of the full SPFs.

We then used the SPFs with and without exposure to estimate pedestrian and bicycle crashes for all intersections and mid-blocks in the city. Following the guidance by FHWA (Srinivasan & Bauer, 2013; Srinivasan et al., 2013), we applied the empirical Bayes (EB) method (Elvik, 2008; Hauer et al., 2002) to calculate the final estimation of the pedestrian and bicycle crashes for all intersections and mid-blocks in the city. The EB method helps to address the issues of regression to the mean and imperfect data quality (Thomas et al., 2017). Specifically, we used the equation below estimate crashes:

$$EB = wP + (1-w)x, \quad (2)$$

where EB is the empirical Bayes estimate of the crash number, P is the estimated crash number calculated by the SPFs, x is the historical crash number, and w is the weight. We can calculate w with the equation below:

$$w = \frac{1}{1 + \frac{P}{k}}, \quad (3)$$

where k is the inverse of the dispersion factor estimated by the SPFs (Elvik, 2008). Sengupta et al. (2021) recommended using the same time period for the dependent variable and the EB estimation to minimize biased estimation. Following this suggestion, our EB estimation is the total crashes for 13 years, which is consistent with the dependent variable in the SPFs.

To evaluate the performance of the EB estimation by the two sets of SPFs, we validated the results using the crash history for 2018-2019, the two-year period following our study period (2005-2017). Specifically, we applied the EB estimates for the SPFs with and without exposure to predict the total crashes for 13 years. We then transferred these estimates to 2-year estimates by dividing them by 6.5. We measured the accuracy of the predictions and compared the SPFs using root mean squared error (RMSE) for performance assessment.

The other indicator is the discrepancy in high-risk locations identified by the two sets of SPFs. For each set of SPFs, we ranked all the intersections and mid-blocks based on their EB estimation of crash risk, respectively, and selected the locations within the top 1%, 5%, and 10% of the rankings as high-risk locations. We then compared the high-risk locations identified by the two sets of SPFs in the entire city and in the ACP50, respectively. We calculated the proportion of high-risk locations identified by both the models with exposure and the models without exposure. A smaller proportion indicates a larger discrepancy.

Table 2. Variable definitions, sources, and descriptive statistics

Variables	Definition	Data source	Year	Intersection (N = 173)		Mid-block (N = 437)	
				Mean	Standard deviation	Mean	Standard deviation
Number of pedestrian crashes	Total number of pedestrian crashes occurred at the intersection or mid-block segment	DPS	2005-2017	3.13	5.16	0.24	0.83
Number of bicycle crashes	Total number of total bicycle crashes occurred at the intersection or mid-block segment	DPS	2005-2017	2.21	3.03	0.24	0.75
Exposure to risk							
Ln(Actual pedestrian count)	Natural logarithm of average 2-hour (4-6 pm) pedestrian count	DPW	2007-2014	6.72	1.20	5.91	1.27
Ln(Actual bicycle count)	Natural logarithm of average 2-hour (4-6 pm) bicycle count	DPW	2007-2014	6.10	0.90	5.41	1.11
Ln(Vehicular AADT)	Natural logarithm of vehicular AADT collected from the most recent years	DPW	2005-2015	9.49	0.63	8.71	0.79
Built environment							
Population density	Number of people per hectare of the census block where the intersection or the mid-block centroid is located	USCB ^a	2010	34.93	59.83	33.92	48.84
Job density	Number of jobs per hectare of the census block where the intersection or the mid-block centroid is located	USCB	2015	165.61	940.63	155.34	939.59
Number of intersections	Number of intersections within the 400-meter buffer of the intersection or the mid-block centroid	DPW	-	29.13	7.83	28.92	8.50
Presence of transit stop	A dummy variable indicating whether there is a transit stop within the 30-meter buffer of the intersection or mid-block segment	MGC	2018	0.60	0.49	0.12	0.32
Share of commercial area	Proportion of commercial land use within the 100-meter buffer of the intersection or the mid-block centroid	MGC	2016	0.21	0.25	0.17	0.24
Share of office area	Proportion of office land use within the 100-meter buffer of the intersection or the mid-block centroid	MGC	2016	0.14	0.18	0.12	0.19
Share of industrial area	Proportion of industrial land use within the 100-meter buffer of the intersection or the mid-block centroid	MGC	2016	0.04	0.13	0.05	0.17
Share of open space	Proportion of open space within the 100-meter buffer of the intersection or the mid-block centroid	MGC	2016	0.06	0.15	0.09	0.20
Land use entropy^b	Land use entropy index within the 100-meter buffer of the intersection or the mid-block centroid	MGC	2016	0.63	0.30	0.58	0.32
Downtown	A dummy variable indicating whether the intersection or mid-block is in the downtown area	DPW	2018	0.17	0.37	0.16	0.37
Pedestrian and bicycle facilities							
Presence of sidewalk	A dummy variable indicating whether there is a sidewalk linked to the intersection or present in any side of the mid-block	DPW	-	0.97	0.18	0.81	0.40

Variables	Definition	Data source	Year	Intersection (N = 173)		Mid-block (N = 437)	
				Mean	Standard deviation	Mean	Standard deviation
Presence of bicycle lane	A dummy variable indicating whether there is a bicycle lane linked to the intersection or present in any side of the mid-block	DPW	2011	0.64	0.48	0.21	0.41
Presence of trail	A dummy variable indicating whether there is a trail linked to the intersection	DPW	2014	0.06	0.23	-	-
Visibility							
Presence of streetlight	A dummy variable indicating whether there is a streetlight in the 30-meter buffer of the intersection	DPW	2018	0.77	0.42	-	-
Traffic control							
Presence of traffic signal	A dummy variable indicating whether there is a traffic signal in the 30-meter buffer of the intersection	DPW	2018	0.66	0.47	-	-
Geometric characteristics							
Travel width of road	Travel width in feet (from shoulder to shoulder) of the mid-block, or maximum travel width of the legs linked to the intersection	Mn-DOT	-	30.47	12.92	24.22	11.49
Number of legs	Number of legs linked to the intersection	DPW	-	3.94	0.42	-	-
Number of main roads	Number of main roads (e.g., minor arterials) linked to the intersection ^c	DPW	-	0.08	0.44	-	-
Number of secondary roads	Number of secondary roads (e.g., collectors) linked to the intersection	DPW	-	1.61	1.49	-	-
Presence of main road	A dummy variable indicating whether the mid-block is a main road	DPW	-	-	-	0.04	0.19
Presence of secondary road	A dummy variable indicating whether the mid-block is a secondary road	DPW	-	-	-	0.41	0.49
Socio-demographics							
Share of children	Percentage of children (age 14 -) in total population of the CBG where the intersection or the mid-block centroid is located	USCB	2016	0.15	0.10	0.15	0.11
Share of seniors	Percentage of old people (age 65 +) in total population of the CBG where the intersection or the mid-block centroid is located	USCB	2016	0.09	0.07	0.09	0.07
Share of men	Percentage of male people in total population of the CBG where the intersection or the mid-block centroid is located	USCB	2016	0.52	0.08	0.52	0.07
Average household size	Average household size of the CBG where the intersection or the mid-block centroid is located	USCB	2016	2.23	0.67	2.26	0.69
Share of white population	Percentage of white population in the CBG where the intersection or the mid-block centroid is located	USCB	2016	0.60	0.27	0.62	0.26
Share of poverty population	Percentage of poverty population in the CBG where the intersection or the mid-block centroid is located, in which poverty population is people whose income less than 185% of poverty threshold	USCB	2016	0.39	0.25	0.39	0.25

Note: ^a USCB = United States Census Bureau.

^b Calculation of land use entropy follows the equation below (Hu et al., 2018),

$$Land\ use\ entropy = - \left(\frac{\sum_{i=1}^N p_i \ln(p_i)}{\ln(N)} \right),$$

where N is the number of land-use types in the buffer and p_i is the percentage of the i th land-use type in terms of area in the buffer.

^c Road classifications applied in this study have major roads (e.g., minor arterials), secondary roads (e.g., collectors), and local roads. Highways (e.g., principal arterials, serve only vehicles) were excluded from this study.

4 Results

4.1 Model comparison

Table 3 and Table 4 present the comparison between two sets of SPFs: models with and without exposure. The models with exposure have better goodness of fit than the models without exposure in terms of deviance R^2 , AIC, BIC, and RMSE. Taking the pedestrian intersection models as an example, the model with exposure has a deviance R^2 of 0.81, whereas the fit for the model without exposure is 0.74. In addition, the model with exposure has a lower AIC (577) and BIC (668) than those of the model without exposure, which are 607 and 692, respectively. Thus, pedestrian and bicycle exposure variables substantially improve model performance. Compared with intersection models, the mid-block models have relatively lower deviance R^2 . During the study period, 87% of the observations in the mid-block pedestrian crash model and 84% of the observations in the mid-block bicycle crash model had zero historical crashes; the corresponding statistics for intersection models are 38% and 34%, respectively. The larger proportion of the observations with zero crashes in the mid-block models leads to a lower explanatory power of the independent variables.

The pedestrian exposure variable is significant in two of the four crash models, and the bicycle exposure variable is significant in three models. In general, these exposure variables have positive associations with the number of crashes, consistent with the literature (Schepers et al., 2011; Thomas et al., 2017). However, bicycle exposure is negatively associated with pedestrian crash risk at intersections. This may be because when there are more bicycles on roads, drivers may become more cautious when driving, also creating a safer traffic environment for pedestrians.

Our analyses of elasticities also suggest the important role of pedestrian and bicycle exposure in predicting pedestrian and bicycle crash risk. In the models with exposure, the elasticities of pedestrian and bicycle exposure variables are among the largest for all variables. Specifically, in the pedestrian intersection model, pedestrian exposure has an elasticity of 0.66, the fourth largest elasticity, following number of legs (-1.51), vehicular AADT (1.19), and average household size (-1.02). In the pedestrian mid-block model, pedestrian exposure has the sixth largest elasticity of 0.47, smaller than average household size (-1.40), vehicular AADT (1.00), share of poverty population (0.79), travel width of road (-0.59), and share of men (0.51). In the bicycle intersection model, bicycle exposure has an elasticity of 0.58, the second largest one after average household size (-0.82). In the bicycle mid-block model, the elasticity of bicycle exposure is 0.47, the fourth largest elasticity after travel width of lane (-0.75), vehicular AADT (0.57), and average household size (0.49).

Among built environment variables, population density is positively correlated with both pedestrian and bicycle crash risk at intersections. Even after controlling for exposure, density of development (which has been shown in other studies to be correlated with pedestrian and bicycle volumes), is positively associated with both pedestrian and bicycle crash risk (Dumbaugh & Li, 2011; Siddiqui et al., 2012). The presence of a transit stop is positively correlated with pedestrian crash risk in the intersection model, likely because transit stops attract pedestrians and cross-walkers (Thomas et al., 2017). Consistent with studies that have found higher volumes of road users such as bicyclists and vehicles at commercial areas (Griswold et al., 2011), the share of commercial area has a positive relationship with bicycle crash risk in the mid-block model. Share of industrial area is positively correlated with bicycle crash in the mid-block model. Share of open space is positively correlated with bicycle crash risk at mid-block. This is plausible because many bicyclers ride there for purposes of workout or entertainment. Mid-blocks in the downtown area have higher bicycle crash risk than those which are located outside. Some results are counter-intuitive and need more investigation in future research (e.g., include more

Table 3. Pedestrian model results

	Pedestrian intersection models				Pedestrian mid-block models			
	With exposure		Without exposure		With exposure		Without exposure	
	Coefficient	Elasticity	Coefficient	Elasticity	Coefficient	Elasticity	Coefficient	Elasticity
	Exposure to risk							
Ln(Actual pedestrian count)	0.66	0.66			0.47	0.47		
Ln(Actual bicycle count)	-0.34	-0.34			-0.12	-0.12		
Ln(Actual AADT)	1.19	1.19	1.24	1.24	1.00	1.00	1.19	1.19
	Built environment							
Population density	1.82×10^{-3}	0.06	1.60×10^{-3}	0.06	0.01	0.17	0.01	0.18
Job density	-1.55×10^{-4}	-0.03	-1.18×10^{-4}	-0.02	8.35×10^{-5}	0.01	1.63×10^{-4}	0.03
Number of intersections	0.01	0.33	0.01	0.21	-1.92×10^{-4}	-0.01	-4.31×10^{-3}	-0.12
Presence of transit stop	0.50	0.30	0.56	0.34	-0.92	-0.11	-0.95	-0.11
Share of commercial area	-0.52	-0.11	0.40	0.09	-0.10	-0.02	0.59	0.10
Share of office area	-0.91	-0.12	-0.65	-0.09	-0.50	-0.06	-0.22	-0.03
Share of industrial area	-0.90	-0.04	-2.37	-0.11	0.53	0.03	-0.28	-0.02
Share of open space	-2.65	-0.16	-3.16	-0.20	-0.25	-0.02	-0.04	-3.35×10^{-3}
Land use entropy	0.03	0.02	1.88×10^{-3}	1.19×10^{-3}	-0.32	-0.19	-0.17	-0.10
Downtown	-0.37	-0.06	-0.25	-0.04	0.50	0.08	0.60	0.10
	Traffic facilities							
Presence of sidewalk	-0.31	-0.30	0.18	0.18	0.58	0.47	0.59	0.48
Presence of trail	0.56	0.03	0.51	0.03				
Presence of street light	0.08	0.06	0.34	0.26				
Presence of traffic signal	0.71	0.47	0.76	0.50				
Travel width of lane	0.01	0.20	0.01	0.29	-0.02	-0.59	-0.03	-0.69
Number of legs	-0.38	-1.51	-0.28	-1.08				
Number of main roads	0.43	0.04	-0.08	-0.01				
Number of secondary roads	0.14	0.23	0.11	0.18				
Presence of main road					0.56	0.02	0.04	0.00
Presence of secondary road					0.27	0.11	0.20	0.08
	Socio-demographics							
Share of children	0.94	0.14	0.94	0.14	0.64	0.09	0.01	1.80×10^{-3}
Share of seniors	-4.97	-0.44	-4.90	-0.43	0.88	0.08	0.42	0.04
Share of men	-0.97	-0.51	-0.04	-0.02	0.99	0.51	1.19	0.61
Average household size	-0.46	-1.02	-0.40	-0.90	-0.62	-1.40	-0.62	-1.40

	Pedestrian intersection models				Pedestrian mid-block models			
	With exposure		Without exposure		With exposure		Without exposure	
	Coefficient	Elasticity	Coefficient	Elasticity	Coefficient	Elasticity	Coefficient	Elasticity
Share of white population	-0.59	-0.36	-0.37	-0.22	0.40	0.24	0.69	0.43
Share of poverty population	-0.04	-0.02	0.37	0.15	2.04	0.79	2.59	1.00
Constant and model fitness								
Constant	-10.94		-11.01		-13.22		-13.07	
Dispersion factor	0.04		0.17		2.22		2.57	
Deviance R ²	0.81		0.74		0.33		0.31	
AIC	577		607		472		473	
BIC	668		692		574		567	

Note: The variable with coefficient in the shading cell is significant at level 95% in the model.

Table 4. Bicycle model results

	Bicycle intersection models				Bicycle mid-block models			
	With exposure		Without exposure		With exposure		Without exposure	
	Coefficient	Elasticity	Coefficient	Elasticity	Coefficient	Elasticity	Coefficient	Elasticity
Exposure to risk								
Ln(Actual pedestrian count)	0.16	0.16			0.14	0.14		
Ln(Actual bicycle count)	0.58	0.58			0.47	0.47		
Ln(Actual AADT)	0.40	0.40	0.55	0.55	0.57	0.57	0.68	0.68
Built environment								
Population density	2.07×10^{-3}	0.07	2.75×10^{-3}	0.10	2.70×10^{-3}	0.09	3.11×10^{-3}	0.11
Job density	-4.30×10^{-5}	-0.01	-6.88×10^{-5}	-0.01	-3.05×10^{-4}	-0.05	-2.74×10^{-4}	-0.04
Number of intersections	-0.02	-0.57	-0.02	-0.60	-2.65×10^{-5}	-7.67×10^{-4}	-0.01	-0.32
Presence of transit stop	-0.11	-0.07	-0.11	-0.07	-0.33	-0.04	-0.26	-0.03
Share of commercial area	-0.10	-0.02	0.18	0.04	2.04	0.34	2.19	0.36
Share of office area	0.23	0.03	-0.11	-0.02	0.13	0.02	-0.08	-0.01
Share of industrial area	0.64	0.03	-0.06	-2.85×10^{-3}	1.67	0.09	1.73	0.09
Share of open space	-1.11	-0.07	-0.46	-0.03	2.16	0.20	2.88	0.26
Land use entropy	0.28	0.18	0.68	0.43	-0.70	-0.41	-0.46	-0.27
Downtown	-0.40	-0.07	-0.40	-0.07	1.47	0.24	1.64	0.27
Traffic facilities								

Presence of bike lane	-0.12	-0.07	0.05	0.03	-0.06	-0.01	0.09	0.02
Presence of trail	-0.43	-0.02	-0.04	0.00				
Presence of street light	0.18	0.14	0.39	0.30				
Presence of traffic signal	0.34	0.22	0.36	0.24				
Travel width of lane	0.01	0.21	0.01	0.35	-0.03	-0.75	-0.03	-0.72
Number of legs	-0.03	-0.12	0.03	0.10				
Number of main roads	-0.10	-0.01	-0.10	-0.01				
Number of secondary roads	0.26	0.42	0.18	0.30				
Presence of main road					-1.15	-0.04	-1.16	-0.04
Presence of secondary road					0.35	0.14	0.30	0.12
Socio-demographics								
Share of children	-0.48	-0.07	-2.25	-0.33	1.15	0.17	-0.72	-0.11
Share of seniors	-1.95	-0.17	-1.46	-0.13	3.45	0.31	2.75	0.25
Share of men	-1.06	-0.55	-0.68	-0.35	0.33	0.17	0.09	0.04
Average household size	-0.37	-0.82	-0.35	-0.78	0.22	0.49	0.14	0.32
Share of white population	-0.54	-0.32	-0.06	-0.04	0.36	0.22	0.67	0.41
Share of poverty population	-0.07	-0.03	1.10	0.43	0.59	0.23	1.63	0.63
Constant and model fitness								
Constant	-6.34		-4.78		-11.81		-9.18	
Dispersion factor	0.16		0.29		1.54		1.92	
Deviance R ²	0.56		0.47		0.30		0.25	
AIC	611		635		493		503	
BIC	703		720		595		597	

Note: The variable with coefficient in the shading cell is significant at level 95% in the model.

other influencers in the models to check the potential issue of endogeneity). Job density is negatively correlated with pedestrian crash risk in intersections, and open space is negatively correlated with pedestrian crash risk at intersections. These findings may reflect features not included in our models such as specific traffic controls or signs.

4.2 Validation of the EB estimation

We compared the empirical Bayes estimation of the crash risk by the two sets of SPFs with crash history from 2018 to 2019 (Table 5). While the RMSEs are very close, three of the four SPFs with exposure have slightly lower RMSEs than those of the SPFs without exposure (specifically, the bicycle intersection, bicycle mid-block, and pedestrian mid-block). The difference is greatest for the bicycle intersection model. This result confirms that, in general, incorporation of pedestrian and bicycle exposure variables in SPFs can improve their prediction of crash risk in the future.

Table 5. Comparison between Empirical Bayes estimation and two-year historical crash numbers (2018-2019) in terms of RMSE

Pedestrian	Intersection models		Mid-block models	
	With exposure	Without exposure	With exposure	Without exposure
Average historical crash number	0.0738		0.0127	
Average estimated crash number	0.0613	0.0924	0.0080	0.0122
RMSE	0.2998	0.2917	0.1203	0.1205
Bicycle	Intersection models		Mid-block models	
	With exposure	Without exposure	With exposure	Without exposure
Average historical crash number	0.0452		0.0084	
Average estimated crash number	0.0388	0.0977	0.0068	0.0149
RMSE	0.2199	0.2310	0.0941	0.0953

4.3 Differences in high-risk locations

To illustrate the changes in high-risk locations identified by the two sets of SPFs, we ranked all 6,639 intersections and 12,589 mid-blocks by the EB estimation of crash risk and then identified the locations comprising the top 1%, top 5%, and top 10% of the rankings. Table 6 presents the share of the high-risk locations identified by both the models with and without exposure. When high-risk locations are defined as top 1%, approximately 60% of the intersections and 80% of the mid-blocks identified by the two sets of SPFs are the same. For example, 60% of the high-risk intersections identified by the bicycle model with exposure measures are the same as those identified by the bicycle model without exposure. As the proportion used to define high-risk locations increases, the share of the same high-risk locations identified by the two sets of models change slightly. However, the discrepancy is still large: about 20% of the top 10% high-risk locations differ.

Table 6. Share of high-risk locations in the city identified by both the SPFs with and without exposure

	Pedestrian		Bicycle	
	Intersection (N = 6,639)	Mid-block (N = 12,589)	Intersection (N = 6,639)	Mid-block (N = 12,589)
Top 1%	54%	88%	60%	75%
Top 5%	72%	81%	71%	75%
Top 10%	87%	86%	79%	77%

Theoretically, the models with exposure predict high-risk locations better than those without exposure. Empirically, the former models fit the data better than the latter ones in terms of model performance (i.e., deviance R^2 , AIC, and BIC) and variables' significance and elasticities (Section 4.1). If we assume for sake of illustration that the high-risk locations identified by the models with exposure have greater validity, the models without exposure potentially misclassify a substantial number of locations as high-risk locations (Table 6). Figures D1 and D2 in the Appendix D visually compare the top 1% high-risk locations predicted by the two sets of SPFs.

Overall, the models with exposure produce different high-risk locations, especially the top 1% high-risk locations, from the models without exposure. In Minneapolis, 1% of high-risk locations includes 66 intersections and 126 mid-blocks. Given the objective of increasing safety, and assuming that funds are prioritized for the intersections and mid-blocks with highest relative risk, the choice of models potentially affects the allocation of millions of dollars of scarce public resources.

We next compared the overlap of high-risk locations in the ACP50s (i.e., low income, majority-minority areas) identified by the two sets of SPFs (Table 7). While there was overlap in the sets of intersections and mid-blocks identified as higher risk, there were also substantial differences in the distributions. For example, only 42% and 68% of the intersections with the highest pedestrian and bicycle crash risk, respectively, were the same.

Table 7. Share of high-risk locations in the ACP50s identified by both the SPFs with and without exposure

	Pedestrian		Bicycle	
	Intersection (N = 6,639)	Mid-block (N = 12,589)	Intersection (N = 6,639)	Mid-block (N = 12,589)
Top 1%	42%	91%	68%	67%
Top 5%	73%	83%	79%	64%
Top 10%	92%	87%	82%	67%

We also compared the number of high-risk locations identified by the sets of SPFs with and without exposure in the ACP50s (Table 8). As expected, the numbers of high-risk locations change. For example, the pedestrian mid-block model with exposure identified 19 mid-blocks as top 1% high-risk locations, while the corresponding model without exposure identified 16 high-risk mid-blocks, a relative difference of 16%. Conversely, four cases show that SPFs with exposure predicted fewer high-risk locations than those without in the APC50 areas. These results confirm that inclusion of pedestrian and bicycle exposure variables has an influence on the number of identified high-risk locations in ACP50s. However, all mechanisms that affect whether differential crash risk exists in these racial and low-income neighborhoods are not clear and additional studies are needed.

Table 8. Comparison of the number of high-risk locations in the ACP50s identified by both the SPFs with and without exposure

Pedestrian						
	Intersection (N = 6,639)			Mid-block (N = 12,589)		
	SPFs with exposure	SPFs without exposure	Difference	SPFs with exposure	SPFs without exposure	Difference
Top 1%	19	16	-16%	32	32	0%
Top 5%	93	115	24%	185	181	-2%
Top 10%	207	221	7%	359	350	-3%
Bicycle						
	Intersection (N = 6,639)			Mid-block (N = 12,589)		
	SPFs with exposure	SPFs without exposure	Difference	SPFs with exposure	SPFs without exposure	Difference
Top 1%	22	23	5%	18	19	6%
Top 5%	118	107	-9%	121	101	-17%
Top 10%	192	184	-4%	285	229	-20%

5 Conclusions

In this study we explored the influences of pedestrian and bicycle exposure on pedestrian and bicycle crashes at intersections and mid-blocks in the city of Minneapolis. We also assessed the implications of including or excluding estimates of exposure by constructing two sets of SPFs, one with, and one without, both exposure variables, for both the intersection and mid-block models. The results show that including pedestrian and bicycle exposure variables in the SPFs improves model goodness of fit. Furthermore, the elasticities of the exposure measures and their significant relationships with pedestrian and bicycle crash risk suggest that they are essential to increasing the validity of predictions of pedestrian and bicycle crash risk.

Transportation agencies have scarce resources for street improvements, so they must prioritize hot spots with the highest pedestrian and bicycle crash risk. They also must redress historic inequities in investment in poor, marginalized neighborhoods. This study shows that many of the intersections and mid-blocks identified as high-risk by the models with exposure and those without are different. Among the top 1% high-risk intersections in Minneapolis, about 40% of the locations differ; among the top 1% of mid-blocks, about 20% of the locations differ. Because an improvement project may cost hundreds of thousands of dollars, the choice of models matters, and it is important to use tools that account for as many variables associated with crash risk as possible. Using Minneapolis as an example, the annual budget for street facility improvement is \$21.2 million (City of Minneapolis, 2016). If, as estimated, 20% to 40% of the high-risk locations were misidentified, millions of dollars potentially could be misallocated to lower priority sites.

The high-risk locations identified by the two sets of SPFs are also different in ACP50 neighborhoods that are the focus of local efforts to address historical inequities. Among the top 1% high-risk intersections, roughly 30% to 60% locations differ; among the top 1% high-risk mid-blocks, 10% to 30% differ. Given these inclusion of exposure variables clearly matters, it is important to include them to ensure validity of prioritization of high-risk locations in ACP50 areas. Additional studies to confirm these findings and address equity-related issues are needed.

Overall, these findings add to the growing evidence of the importance of including pedestrian and bicycle exposure variables in the SPFs. Models that include measures of pedestrian and bicycle crash risk simultaneously have the greatest theoretical validity. In addition to producing more valid estimates of crash risk, they also can facilitate more robust assessments of other measures, like safety in numbers (Elvik & Bjørnskau, 2017). To increase the availability of exposure measures, local transportation agencies will need pedestrian and bicycle monitoring programs that are the foundation of developing pedestrian and bicycle exposure data.

Our research has limitations that can be addressed over time through additional analyses and as more data are collected. First, our crash dataset extends over a period of 13 years. During this period, exposure measures and other independent variables such as geometric characteristics, traffic controls, and transportation policies have changed. This limitation, which is common in the literature, reduces the validity of our estimates, but can be addressed in the future as better databases become available. Second, the time periods for our crash data, our estimates of exposure to risk, and our other independent variables are different. For example, we use two-hour peak hour counts as measures of pedestrian and bicycle exposure, while our crash dataset includes crashes at all times of day. This limitation could be addressed with better measures of exposure such as AADP or AADB that will become available as more communities implement monitoring programs. A related limitation is that a small share of our measures of bicycle and pedestrian exposure are actual counts; the vast majority are estimates from a validated demand model. A specific limitation of our pedestrian mid-block model is that our estimate of pedestrian exposure is pedestrians on the sidewalk, not pedestrians in the street or crossing at mid-block. We believe that these two measures are to be highly correlated, but field studies would be required to confirm this hypothesis. In general, new data collection initiatives can help address both limitations. Another limitation not accounted for in our analyses is that some behavioral or cultural factors may be associated with crashes. For example, certain driver behaviors, such as failure to yield for pedestrians, or certain activities by vulnerable users (e.g., pedestrians jaywalking or cyclists running red lights) likely are important. Studies of specific factors associated with crashes may provide insight into these issues. Although we considered several socio-demographic variables as proxies, but most of them are not significant, thus precluding any generalizations about associations with crashes.

References

- Burden, D., & Litman, T. (2011). America needs complete streets. *ITE Journal (Institute of Transportation Engineers)*, 81(4), 36–43.
- Carlson, K., Ermagun, A., Murphy, B., Owen, A., & Levinson, D. (2019). Safety in numbers for bicyclists at urban intersections. *Transportation Research Record*, 2673(6), 677–684. <https://doi.org/10.1177/0361198119846480>
- Chen, P. (2015). Built environment factors in explaining the automobile-involved bicycle crash frequencies: A spatial statistic approach. *Safety Science*, 79, 336–343. <https://doi.org/https://doi.org/10.1016/j.ssci.2015.06.016>
- Chen, P., Hu, S., Shen, Q., Lin, H., & Xie, C. (2019). Estimating traffic volume for local streets with imbalanced data. *Transportation Research Record*, 2673(3), 598–610. <https://doi.org/10.1177/0361198119833347>
- Chen, P., & Shen, Q. (2016). Built environment effects on cyclist injury severity in automobile-involved bicycle crashes. *Accident Analysis & Prevention*, 86, 239–246. <https://doi.org/10.1016/j.aap.2015.11.002>
- City of Minneapolis. (2016). 20 year streets funding plan. Minneapolis, MN: City of Minneapolis. Retrieved from <https://www2.minneapolismn.gov/government/departments/public-works/tpp/20-year-plan/>
- City of Minneapolis. (2017). 2017 City of Minneapolis pedestrian crash study. Minneapolis: City of Minneapolis.
- Cottrill, C. D., & Thakuriah, P. V. (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis & Prevention*, 42(6), 1718–1728. <https://doi.org/10.1016/j.aap.2010.04.012>
- Daniels, S., Brijs, T., Nuyts, E., & Wets, G. (2009). Injury crashes with bicyclists at roundabouts: Influence of some location characteristics and the design of cycle facilities. *Journal of Safety Research*, 40(2), 141–148. <https://doi.org/10.1016/j.jsr.2009.02.004>
- DPS. (2018). Crashes in Minnesota from 2005 to 2017. St. Paul, MN: Department of Public Safety.
- Dumbaugh, E., & Li, W. (2011). Designing for the safety of pedestrians, cyclists, and motorists in urban environments. *Journal of the American Planning Association*, 77(1), 69–88. <https://doi.org/10.1080/01944363.2011.536101>
- Elvik, R. (2008). The predictive validity of empirical Bayes estimates of road safety. *Accident Analysis & Prevention*, 40(6), 1964–1969. <https://doi.org/10.1016/j.aap.2008.07.007>
- Elvik, R. (2013). Can a safety-in-numbers effect and a hazard-in-numbers effect co-exist in the same data? *Accident Analysis & Prevention*, 60, 57–63. <https://doi.org/10.1016/j.aap.2013.08.010>
- Elvik, R., & Bjørnskau, T. (2017). Safety-in-numbers: A systematic review and meta-analysis of evidence. *Safety Science*, 92, 274–282. <https://doi.org/10.1016/j.ssci.2015.07.017>
- FHWA. (2013). Systemic safety project selection tool. Washington, DC: FHWA. Retrieved from <https://safety.fhwa.dot.gov/systemic/fhwasa13019/>
- FHWA. (2018). *Guide for scalable risk assessment methods for pedestrian and bicyclists*. Washington, DC: FHWA.
- Gladhill, K., & Monsere, C. (2012). Exploring traffic safety and urban form in Portland, Oregon. *Transportation Research Record*, 2318, 63–74. <https://doi.org/10.3141/2318-08>
- Griswold, J., Medury, A., & Schneider, R. (2011). Pilot models for estimating bicycle intersection volumes. *Transportation Research Record*, 2247, 1–7. <https://doi.org/10.3141/2247-01>
- Guerra, E., Dong, X., & Kondo, M. (2019). Do denser neighborhoods have safer streets? Population density and traffic safety in the Philadelphia region. *Journal of Planning Education and Research*,

- 0739456X1984504. <https://doi.org/10.1177/0739456X19845043>
- Hankey, S., & Lindsey, G. (2016). Facility-demand models of peak period pedestrian and bicycle traffic. *Transportation Research Record*, 2586, 48–58. <https://doi.org/10.3141/2586-06>
- Hauer, E. (2015). *The art of regression modeling in road safety*. New York: Springer. <https://doi.org/10.1007/978-3-319-12529-9>
- Hauer, E., Harwood, D. W., Council, F. M., & Griffith, M. S. (2002). Estimating safety by the empirical Bayes method: A tutorial. *Transportation Research Record*, 1784, 126–131. <https://doi.org/10.3141/1784-16>
- Hu, S., Chen, P., Lin, H., Xie, C., & Chen, X. (2018). Promoting carsharing attractiveness and efficiency: An exploratory analysis. *Transportation Research Part D: Transport and Environment*, 65, 229–243. <https://doi.org/10.1016/j.trd.2018.08.015>
- Kim, D., & Kim, K. (2015). The influence of bicycle-oriented facilities on bicycle crashes within crash-concentrated areas. *Traffic Injury Prevention*, 16(1), 70–75. <https://doi.org/10.1080/15389588.2014.895924>
- La Plante, J., & McCann, B. (2008). Complete streets: We can get there from here. *ITE Journal (Institute of Transportation Engineers)*, 78(5), 24–28.
- Lee, C., & Abdel-Aty, M. (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. *Accident Analysis & Prevention*, 37(4), 775–786. <https://doi.org/10.1016/j.aap.2005.03.019>
- Lindsey, G., Tao, T., Wang, J., & Cao, J. (2019). *Pedestrian and bicycle crash risk and equity: Implications for street improvement*. Minneapolis, MN: Roadway Safety Institute. Retrieved from <http://conservancy.umn.edu/handle/11299/203635>
- Loukaitou-Sideris, A., Liggett, 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(3), 338–351. <https://doi.org/10.1177/0739456X06297008>
- Merlin, L. A., Guerra, E., & Dumbaugh, E. (2020). Crash risk, crash exposure, and the built environment: A conceptual review. *Accident Analysis & Prevention*, 134(May), 105244. <https://doi.org/10.1016/j.aap.2019.07.020>
- NHTSA. (2018a). *2016 Bicyclists and other cyclists traffic safety fact sheet*. Washington, DC: National Highway Traffic Safety Administration.
- NHTSA. (2018b). *2016 Pedestrians traffic safety fact sheet*. Washington, DC: National Highway Traffic Safety Administration.
- Nordback, K., Marshall, W. E., & Janson, B. N. (2014). Bicyclist safety performance functions for a U.S. city. *Accident Analysis & Prevention*, 65, 114–122. <https://doi.org/10.1016/j.aap.2013.12.016>
- Park, J., Abdel-Aty, M., Lee, J., & Lee, C. (2015). Developing crash modification functions to assess safety effects of adding bike lanes for urban arterials with different roadway and socio-economic characteristics. *Accident Analysis & Prevention*, 74, 179–191. <https://doi.org/10.1016/j.aap.2014.10.024>
- Renne, J. L., & Bennett, P. (2014). Socioeconomics of urban travel: Evidence from the 2009 National Household Travel Survey with implications for sustainability. *World Transport Policy & Practice*, 20(4).
- Schepers, J. P., Kroeze, P. A., Sweers, W., & Wüst, J. C. (2011). Road factors and bicycle–motor vehicle crashes at unsignalized priority intersections. *Accident Analysis & Prevention*, 43(3), 853–861. <https://doi.org/10.1016/j.aap.2010.11.005>
- Schneider, R., Diogenes, M., Arnold, L., Attaset, V., Griswold, J., & Ragland, D. (2010). Association between roadway intersection characteristics and pedestrian crash risk in Alameda County, California. *Transportation Research Record*, 2198, 41–51. <https://doi.org/10.3141/2198-06>
- Schneider, R. J., Schmitz, A., Lindsey, G., & Qin, X. (2021). Exposure-based models of trail user

- crashes at roadway crossings. *Transportation Research Record*, 2675(8), 468–480. <https://doi.org/10.1177/0361198121998692>
- Sengupta, A., Gayah, V. V., & Donnell, E. T. (2021). Examining the impacts of crash data aggregation on SPF estimation. *Accident Analysis & Prevention*, 160(January), 106313. <https://doi.org/10.1016/j.aap.2021.106313>
- Siddiqui, C., Abdel-Aty, M., & Choi, K. (2012). Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accident Analysis & Prevention*, 45, 382–391. <https://doi.org/10.1016/j.aap.2011.08.003>
- Siddiqui, C., Abdel-Aty, M., & Choi, K. (2014). Implications of pedestrian safety planning factors in areas with minority and low-income populations. *International Journal of Sustainable Transportation*, 8(5), 360–381. <https://doi.org/10.1080/15568318.2012.702853>
- Srinivasan, R., & Bauer, K. (2013). *Safety performance function development guide: Developing jurisdiction-specific SPFs*. Washington, DC: Federal Highway Administration, Office of Safety.
- Srinivasan, R., Carter, D., & Bauer, K. (2013). *Safety performance function decision guide: SPF calibration vs SPF development*. Washington, DC: Federal Highway Administration, Office of Safety.
- Thomas, L., Lan, B., Sanders, R. L., Frackelton, A., Gardner, S., & Hintze, M. (2017). Changing the future? Development and application of pedestrian safety performance functions to prioritize locations in Seattle, Washington. *Transportation Research Record*, 2659(1), 212–223. <https://doi.org/10.3141/2659-23>
- Thomas, L., Sandt, L., Zegeer, C., Kumfer, W., Lang, K., Lan, B., ... Schneider, R. J. (2018). *Systemic pedestrian safety analysis* (Research report 893). Washington, DC: National Cooperative Highway Research Program. <https://doi.org/10.17226/25255>
- Turner, S., Wood, G., Hughes, T., & Singh, R. (2011). Safety performance functions for bicycle crashes in New Zealand and Australia. *Transportation Research Record*, 2236(1), 66–73. <https://doi.org/10.3141/2236-08>
- Ukkusuri, S., Miranda-Moreno, L. F., Ramadurai, G., & Isa-Tavarez, J. (2012). The role of built environment on pedestrian crash frequency. *Safety Science*, 50(4), 1141–1151. <https://doi.org/10.1016/j.ssci.2011.09.012>
- Williams, T., Page, S., & Doscher, C. (2018). Spatial characteristics of bicycle–motor vehicle crashes in Christchurch, New Zealand: A case-control approach. *Journal of Transport and Land Use*, 11(1), 849–864. <https://doi.org/10.5198/jtlu.2018.1147>
- Yasmin, S., & Eluru, N. (2016). Latent segmentation based count models: Analysis of bicycle safety in Montreal and Toronto. *Accident Analysis & Prevention*, 95, 157–171. <https://doi.org/10.1016/j.aap.2016.07.015>
- Yu, C.-Y., Zhu, X., & Lee, C. (2018). Income and racial disparity and the role of the built environment in pedestrian injuries. *Journal of Planning Education and Research*, 0739456X1880775. <https://doi.org/10.1177/0739456X18807759>
- Zangenehpour, S., Strauss, J., Miranda-Moreno, L. F., & Saunier, N. (2016). Are signalized intersections with cycle tracks safer? A case-control study based on automated surrogate safety analysis using video data. *Accident Analysis & Prevention*, 86, 161–172. <https://doi.org/10.1016/j.aap.2015.10.025>
- Zegeer, C., Lyon, C., Srinivasan, R., Persaud, B., Lan, B., Smith, S., ... Sundstrom, C. (2017). Development of crash modification factors for uncontrolled pedestrian crossing treatments. *Transportation Research Record*, 2636(1), 1–8. <https://doi.org/10.3141/2636-01>

Appendices

Appendices A–D available at <https://www.jtlu.org/index.php/jtlu/article/view/1980>.