CPBS
Center for Pedestrian and Bicyclist Safety

A Tier-1 University Transportation Center

The Role of Built Environment Factors in Enhancing Pedestrian Safety: A Comprehensive Analysis and Policy Implications

August 2025	A Report From the Center for Pedestrian and Bicyclist Safety

Asad Khattak
The University of Tennessee, Knoxville

Zeinab Bayati
The University of Tennessee, Knoxville

Bruce Appleyard
San Diego State University

About the Center for Pedestrian and Bicyclist Safety (CPBS)

The Center for Pedestrian and Bicyclist Safety (CPBS) is a consortium of universities committed to eliminating pedestrian and bicyclist fatalities and injuries through cutting-edge research, workforce development, technology transfer, and education. Consortium members include: The University of New Mexico, San Diego State University, The University of California, Berkeley, The University of Tennessee, Knoxville, and The University of Wisconsin, Milwaukee. More information can be found at: https://pedbikesafety.org

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

The U.S. Department of Transportation requires that all University Transportation Center reports be published publicly. To fulfill this requirement, the Center for Pedestrian and Bicyclist Safety provides reports publicly on its website, www.pedbikesafety.org The authors may copyright any books, publications, or other copyrightable materials developed in the course of, or under, or as a result of the funding grant; however, the U.S. Department of Transportation reserves a royalty-free, nonexclusive and irrevocable license to reproduce, publish, or otherwise use and to authorize others to use the work for government purposes.

Acknowledgments

This study was funded, partially or entirely, by a grant from the Center for Pedestrian and Bicyclist Safety (CPBS), supported by the U.S. Department of Transportation (USDOT) through the University Transportation Centers program. The authors would like to thank CPBS and the USDOT for their support of university-based research in transportation, particularly for the funding provided to support this project. We sincerely thank Dr. Latif Patwary and Dr. Iman Mahdinia for their input regarding databases and analysis.

- Patwary A., & A. Khattak, Exploring How Urban Form and Demographics are Linked with Pedestrian and Bicycle Safety, Presented at Transportation Research Board 103rd Annual Meeting, Transportation Research Board, TRBAM-24-04785, 2024 (forthcoming in *Journal of Transportation Safety and Security*).

TECHNICAL DOCUMENTATION

1. Project No.	2. Government Accession No.	3. Recipient's Catalog No.
24UTK02		

4. Title and Subtitle The Role of Built Environment Factors in Enhancing Pedestrian Safety: A Comprehensive Analysis and Policy Implications	**5. Report Date** August 2025 **6. Performing Organization Code** N/A
7. Author(s) Asad Khattak https://orcid.org/0000-0002-0790-7794 Zeinab Bayati https://orcid.org/0000-0002-5815-4657 Bruce Appleyard https://orcid.org/0000-0003-2105-8079	**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address Center for Pedestrian and Bicyclist Safety Centennial Engineering Center 3020 The University of New Mexico, Albuquerque, NM 87131	**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** 69A3552348336
12. Sponsoring Agency Name and Address United States Department of Transportation Office of Research, Development, and Technology (RD&T)	**13. Type of Report and Period Covered** Final Research Report **14. Sponsoring Agency Code** USDOT OST-R

15. Supplementary Notes

Report available at the CPBS website [https://pedbikesafety.org] and DOI https://doi.org/10.21949/te0a-jb16

16. Abstract

Pedestrian fatalities and injuries are a growing concern across the US. Understanding how pedestrians' and drivers' behaviors, roadway environments, and neighborhood-level factors interact to shape pedestrian crash outcomes is essential for developing targeted safety interventions. This study presents a comprehensive, multi-level analysis of pedestrian crash severity and frequency in North Carolina, utilizing a combination of crash, environmental, and demographic data. Crash-level data from the Pedestrian and Bicycle Crash Analysis Tool (PBCAT) were integrated with census tract-level demographics datasets to create a unique database. A Bayesian cumulative ordinal logit model that accounts for uncertainty was estimated for injury severity (2019 to 2022). A Zero-Hurdle Negative Binomial model was also estimated to investigate pedestrian crash frequency across 2,660 census tracts from 2019 to 2022, focusing on burden indicators. Findings reveal that built environment features such as lighting, pedestrian infrastructure, and roadway configuration significantly affect injury severity, given a crash. Behavioral crash types (e.g., working or playing on the roadway) were consistently associated with more severe outcomes. Burdens such as transportation costs were associated with higher crash frequencies and severities. In contrast, access burden was associated with fewer crashes but greater injury severity when crashes occurred. This study contributes uniquely to pedestrian safety by (1) revealing how different types of burdens are associated with crash outcomes, (2) demonstrating the value of integrating diverse datasets for a deeper understanding of crashes, and (3) applying advanced models to deal with uncertainty and inform place-based safety strategies.

17. Key Words	18. Distribution Statement
Pedestrian Safety, Built Environment, Burdened Areas, Bayesian Models,	No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	70	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized.**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

The Role of Built Environment Factors in Enhancing Pedestrian Safety: A Comprehensive Analysis and Policy Implications

A Center for Pedestrian and Bicyclist Safety Research Report

August 2025

Asad Khattak
The University of Tennessee, Knoxville

Zeinab Bayati
The University of Tennessee, Knoxville

Bruce Appleyard
San Diego State University

TABLE OF CONTENTS

List of Figures

List of Tables

Acronyms, Abbreviations, and Symbols

EXP	Example
NHTSA	National Highway Traffic Safety Administration
UCR	Uniform Crash Report
DUI	Driving Under the Influence
WHO	World Health Organization
PBCAT	Pedestrian and Bicycle Crash Analysis Tool
NaNDA	National Neighborhood Data Archive
NC	North Carolina
FARS	Fatality Analysis Reporting System
Built-Env	Built Environmental
Envir.	Environmental
SD	Standard Deviation
ZHNB	Zero-Hurdle Negative binomial
OR	Odds Ratio
LOOIC	Leave-One-out information criterion
ELPD_LOO	Expected Log Predictive Density
BIC	Bayesian Information Criterion
AIC	Akaike Information Criterion
Ped Position	Pedestrian Position
CI	Confidence Interval

Abstract

Pedestrian fatalities and injuries are a growing concern across the US. Understanding how pedestrians' and drivers' behaviors, roadway environments, and neighborhood-level factors interact to shape pedestrian crash outcomes is essential for developing targeted safety interventions. This study presents a comprehensive, multi-level analysis of pedestrian crash severity and frequency in North Carolina, utilizing a combination of crash, environmental, and demographic data. Crash-level data from the Pedestrian and Bicycle Crash Analysis Tool (PBCAT) were integrated with census tract-level demographics datasets to create a unique database. A Bayesian cumulative ordinal logit model that accounts for uncertainty was estimated for injury severity (2019 to 2022). A Zero-Hurdle Negative Binomial model was also estimated to investigate pedestrian crash frequency across 2,660 census tracts from 2019 to 2022, focusing on burden indicators. Findings reveal that built environment features such as lighting, pedestrian infrastructure, and roadway configuration significantly affect injury severity, given a crash. Behavioral crash types (e.g., working or playing on the roadway) were consistently associated with more severe outcomes. Burdens such as transportation costs were associated with higher crash frequencies and severities. In contrast, access burden was associated with fewer crashes but greater injury severity when crashes occurred. This study contributes uniquely to pedestrian safety by (1) revealing how different types of burdens are associated with crash outcomes, (2) demonstrating the value of integrating diverse datasets for a deeper understanding of crashes, and (3) applying advanced models to deal with uncertainty and inform place-based safety strategies.

Executive Summary

Pedestrian safety remains a pressing concern in transportation planning, particularly as urban areas evolve, and transportation challenges persist or intensify in specific contexts. Despite national efforts to reduce such crashes, the rates remain elevated, particularly in areas with inadequate built environment conditions, where infrastructure limitations contribute to a greater risk. This project investigates pedestrian safety in two important contexts: one that focuses on individual crash events and injury severity, and another that examines broader neighborhood characteristics associated with crash frequency. By addressing both scales, the study aims to understand how planning-level built environment features associate with safety outcomes and guide more effective roadway design across contexts.

To explore how built environment conditions contribute to crash frequency and severity, this study integrates diverse data sources, including police-reported crash data from PBCAT (a unique database of accurately coded pedestrian crashes), U.S. Census indicators, street connectivity measures, and community-level burden indicators. At the crash level, a Bayesian cumulative ordinal logit model that accounts for uncertainty was estimated to assess injury severity using crash data from 2019 to 2022, informed by priors derived from a 2018 crash dataset. This analysis enables a nuanced understanding of correlates than traditional methods. At the census tract level, a ZHNB model was estimated for crash frequency across 2,660 census tracts, with pedestrian exposure controlled through population offsets. The study framework is illustrated in Figure 1.

The results reveal clear and sometimes complex patterns. At the neighborhood level, street network length and community burdens, particularly transportation costs, environmental, and weather-related burdens, were associated with increased crash frequency and severity. Interestingly, the access burden was associated with fewer crashes but greater injury severity when they did occur. At the crash level, pedestrian behaviors, such as dash or dart movements and working or playing in the roadway, emerged as a dominant factor in injury severity, often more so than other driver actions. However, these behaviors interacted differently depending on the roadway context. For instance, risky behavior like playing in the roadway showed lower severity on curved roads, possibly due to lower speeds or heightened driver caution.

Lighting conditions significantly predicted crash outcomes, consistent with other studies. Crashes during daylight or dusk/dawn showed consistently lower injury severity compared to nighttime incidents. These findings support the need for targeted nighttime infrastructure investments, such as improved street lighting, especially in high-risk communities. The analysis further showed that pedestrian-designated spaces, such as sidewalks and shared-use paths, significantly reduce crash severity, reinforcing the importance of physically separated pedestrian infrastructure.

Overall, by identifying specific pedestrian and driver behaviors, as well as built environmental features, the study aims to inform targeted interventions that address both behavioral and environmental factors. Ultimately, the findings provide a data-driven foundation for enhancing planning-level design strategies and promoting transportation safety across diverse urban, suburban, and rural settings.

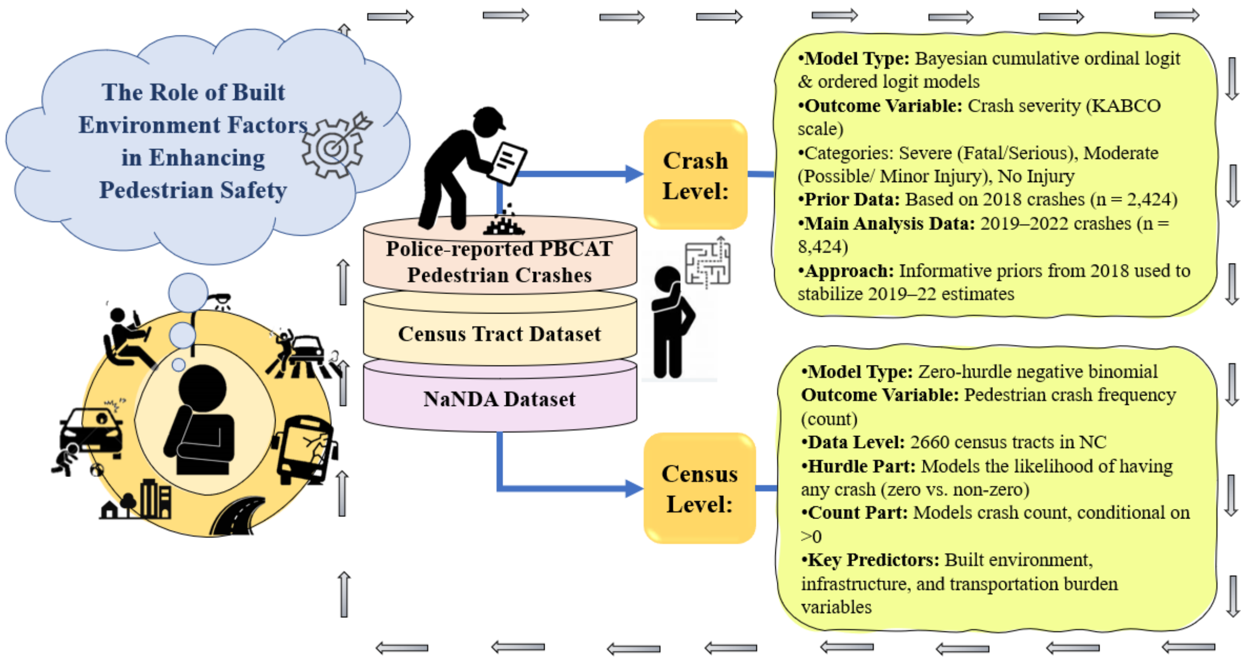

Figure 1. Study Framework

Introduction

Pedestrian crashes in the United States have continued to rise despite extensive interventions aimed at reducing them (WHO, 2023). This concerning trend highlights the complexity of pedestrian safety and underscores the need for a comprehensive approach to understand and address the issue effectively (NHTSA, 2023). Numerous factors contribute to pedestrian safety risks. These factors include road characteristics, driver behavior, pedestrian actions, vehicle features, and unpredictable environmental hazards. Understanding the interplay between these elements is crucial, as addressing pedestrian safety involves coordinated improvements in infrastructure design, increased public awareness, effective policy implementation, and advancements in vehicle safety technology.

Among these factors, the built environment is crucial in pedestrian safety. The built environment encompasses a wide range of artificial and natural elements, each potentially influential in its own right (Clifton et al., 2009; Mansfield et al., 2018). Artificial elements typically refer to human-made infrastructure such as roadway design, traffic signals, pedestrian crossings, signage, sidewalks, street lighting, crosswalk markings, and street furniture (Prato et al., 2018; Soto et al., 2022; Ukkusuri et al., 2012; Yu, 2015). Natural elements relevant to pedestrian safety include weather conditions such as rain, fog, snow, and sunlight glare, all of which can significantly impact visibility, surface conditions, and the behaviors of drivers and pedestrians (Qiu & Nixon, 2008; Wang et al., 2022; Wu et al., 2018). Together, these artificial and natural elements significantly shape pedestrian safety and crash outcomes by influencing road-user interactions, visibility conditions, and overall roadway safety.

The road context and land use patterns are associated with the likelihood and severity of pedestrian crashes. The presence of urban or rural characteristics, the structure of the roadway, posted speed limits, and whether the location is at an intersection can all influence pedestrian safety outcomes (Chen, 2015; Clifton et al., 2009). These contextual factors shape exposure and risk levels, highlighting the importance of integrating land use and roadway characteristics into safety assessments and interventions.

Moreover, the nature of the crash can result in varying degrees of injury severity, shaped by a combination of factors, including lighting, land use, and the behaviors of those involved. In some instances, pedestrian behavior is more closely linked to the occurrence of crashes than driver actions. Understanding the relationship between crash types and built environment features is essential for developing targeted safety measures (Salon & McIntyre, 2018; Yue et al., 2020). For instance, identifying whether certain crash types are more prevalent in poorly lit urban areas versus well-lit rural locations, or examining how pedestrian positions influence crash outcomes, can inform the development of more effective and context-specific pedestrian safety interventions (Ackaah et al., 2020; Hu & Cicchino, 2022; Kim et al., 2006).

This study makes a unique contribution to the existing literature by integrating three distinct and comprehensive datasets that collectively encompass a wide range of relevant features, including built environment characteristics, detailed crash information, and socioeconomic burden indicators; the study primarily leverages the Pedestrian and Bicycle Crash Analysis Tool (PBCAT) dataset focused on North Carolina police-reported pedestrian crashes from 2018 to 2022. Building on these insights, this study aims to address several critical research questions:

1. What neighborhood-level built environment features, such as street lighting, sidewalk availability, road design, land use type/mix and density, and traffic/pedestrian volumes, are key contributors to pedestrian crashes and injury severity?
2. How do built environments correlate differently in nighttime versus daytime crashes?
3. What are the pathways between neighborhood-built environment measures and pedestrian safety outcomes?
4. What behaviors of pedestrians and drivers in different built-environment contexts can be targeted for interventions to reduce the risk of pedestrian crashes?

To answer these questions, a particular emphasis is placed on examining interaction effects among these diverse variables, offering deeper insights into how combined environmental and situational factors contribute to pedestrian crash severity in crash-level data and crash frequency at census tract-level data, applying a multi-scale analytical approach. This dual-level examination enables the capture of both micro-level crash dynamics and broader area-level patterns, facilitating a nuanced understanding of pedestrian safety issues. The findings from this research are expected to inform targeted interventions and guide policy-makers and urban planners toward creating safer, more pedestrian-friendly environments.

Literature Review

Vision Zero, a global traffic safety initiative aimed at eliminating all traffic-related fatalities and serious injuries (Evenson et al., 2018), has become a major focus in research due to the rising risks faced by Vulnerable Road Users (VRUs), particularly pedestrians (NHTSA, 2023). Studies have examined factors such as human error, vehicle design, pedestrian behavior, and environmental conditions (Ferenchak & Abadi, 2021; M Rahman et al., 2022; Mahmoud et al., 2021). Among these, environmental and built environment factors, such as visibility, weather, road quality, sidewalks, and crossings, are particularly important, as they shape both driver and pedestrian outcomes (Ferenchak & Abadi, 2021; Fonsecaa et al., 2022; M Rahman et al., 2022; Mahmoud et al., 2021).

Many national-level studies have explored built environment associations with pedestrian safety, often using fatal crash datasets, such as the Fatality Analysis Reporting System (FARS) (Ewing et al., 2014; Mansfield et al., 2018; Wali & D. Frank, 2024). These analyses typically use macro-level variables and composite indices, such as walkability, revealing paradoxes. For instance, while walkable areas may have more pedestrian activity and crashes, they may show lower per-capita risk. However, these studies often overemphasize relatively rare fatalities, potentially skewing conclusions about overall safety. Moreover, national datasets often lack detailed roadway information (e.g., sidewalk presence or crossing aids), which limits their usefulness for evaluating local safety interventions (Chen, 2015; Mansfield et al., 2018).

To address these gaps, several urban-level studies have examined built environment features using more detailed data. In New York City, for example, crash frequencies were linked to industrial land use, road width, and transit density (Ukkusuri et al., 2012). Similarly, in Montreal, pedestrian activity was linked to environmental features, but collision frequency showed only weak associations (Miranda-Morenoa et al., 2011). These studies reveal the complexity of the built environment–crash relationship and the importance of micro-level data, such as lighting and intersection design.

Crash severity studies further suggest that while macro-level features, such as land use density, increase crash frequency, micro-level factors, including sidewalk coverage, lighting, and transit access, are associated with less severe injuries (Clifton et al., 2009; Congiu et al., 2019; Yu, 2015). Proximity to schools and parks may reduce crash severity due to calmer traffic, while features like parking spots near bus stops may increase crash risk (Yu, 2024; Zahabi et al., 2011).

Sociodemographic context adds another layer of complexity. Lower-income and burdened neighborhoods tend to face a higher risk of pedestrian crashes, often due to inferior infrastructure, increased exposure, and limited access to safety interventions (Clifton et al., 2009). Nevertheless, most studies suffer from common limitations, including the approximation of crash locations, the omission of infrastructure maintenance data, and the failure to account for spatial autocorrelation—issues that can compromise the precision of crash modeling.

In response to these challenges, many studies have used Bayesian methods to better model spatial dependencies, address uncertainty, and integrate multiple data sources (Table 1). For example, (Munira et al., 2020) used a Bayesian spatial Poisson–lognormal model to analyze intersection-level crashes in Austin; (Kitali et al., 2017) used Bayesian clog-log models for older pedestrian crashes in Florida; and (Siddiqui et al., 2012) modeled crash counts using spatial CAR priors at the TAZ level. Other studies have applied Bayesian Additive Regression Trees (BART) to identify highway hotspots (Krueger et al., 2020) and hierarchical Bayesian models to evaluate policy impacts in Vision Zero cities (Auerbach et al., 2021). Despite these advances, few studies apply dual-scale modeling, and most utilize flat or weakly informative priors, which limits their ability to stabilize estimates for low-frequency events or rare settings.

Recent research utilizing the PBCAT has estimated diverse statistical models to examine crash dynamics in depth. A 2024 study analyzing North Carolina data (2009–2019, N≈24,886) used finite mixture ordered probit models to uncover latent crash classes, revealing that 50.6% of crashes involved pedestrian behaviors, while driver recognition errors and violations contributed roughly 22% and 29%, respectively (Ahmad & Khattak, 2024). Meanwhile, traffic risk models using North Carolina PBCAT and road segment data applied binary logistic regression to estimate the likelihood of any crash occurring, and negative binomial regression (NBR) for crash frequency, stratified by roadway functional classes (Gayah et al., 2022). Additional studies have utilized PBCAT-coded crash narratives to develop machine-learning classification algorithms and CHAID-based decision trees, thereby enhancing the granularity and consistency of crash typing across states (Das et al., 2020; Lopez et al., 2022). While prior PBCAT studies have effectively identified behavioral and roadway-related risk patterns, most rely on either cross-sectional analysis, traditional regression techniques, or single-scale modeling.

This study advances the literature by applying a dual-scale framework, which applies a Bayesian ordered logistic model for crash severity and a zero-hurdle negative binomial model for census tract-level crash frequency. It integrates three complementary datasets to create a unique database: detailed crash reports from PBCAT, socioeconomic indicators at the census tract level, and built environment and connectivity features from the NaNDA. At the crash level, a Bayesian ordered logistic model is used to estimate injury severity, utilizing informative priors derived from 2018 crash data to enhance temporal transferability and model stability across the 10918 crashes from the 2019–2022 period. At the macro scale, a zero-hurdle negative binomial model is used at the 2660 census tract levels to model crash frequency, explicitly accounting for overdispersion and excess zeros often found in areal crash data. The model also incorporates interaction terms between crash type and built environment variables to capture contextual effects.

This dual-scale approach, enhanced by temporally informed priors, enables more stable and generalizable estimates of pedestrian risk—especially in lower-frequency or underserved settings—and provides a robust foundation for informed policy and infrastructure decisions.

Table 1. Studies on Built-Env with Bayesian Approach

Study	Bayesian Method	Data Years & Area	Built-Env Focus	Priors Used
This study	Bayesian ordered logistic (crash-level) & ZHNB (tract-level)	2019–2022, North Carolina (PBCAT)	Built environment, crash characteristics, socioeconomic burden	Informative priors from 2018 posteriors (temporally transferred)
Munira et al. (2020)	Multivariate spatial Poisson–lognormal	2010–2019, Austin, TX	Intersection geometry, bus stops, pedestrian volume	Spatial (CAR*); flat priors for fixed effects
Alluri et al. (2023)	Bayesian complementary log–log (hotspot)	2011–2014, Miami-Dade, FL	Sidewalks, lighting, transit access	Weakly informative Normal priors
Kitali et al. (2024)	Bayesian complementary log–log (severity)	2009–2013, Florida (older pedestrians)	Lighting, roadway geometry	Vague priors for link function parameters; flat
Siddiqui et al. (2012)	Spatial Poisson–lognormal (TAZ level)	Multi-year, multiple U.S. TAZs	Road length, intersection density, land use	Spatial (CAR); flat priors
Davis (2013)	Bayesian network	Crash reconstruction data	Speeding, environmental conditions	Uniform (non-informative) priors
Krueger et al. (2020)	Bayesian additive regression trees (BART)	Highway crash hotspots	Spatial network & built environment	Informative (nonparametric BART priors)
Bansal et al. (2020)	Bayesian zero-inflated Poisson/negative binomial	2008-2012 NYC census tracts	Pedestrian injury frequency	Vague Normal priors; Gamma on dispersion
Auerbach et al. (2017)	Hierarchical Bayesian model	2009-2015 NYC (Vision Zero policy analysis)	Built environment covariates	Informative hierarchical priors

*CAR: Conditional Autoregressive

Data and Methodology

Data

This study used a unique dataset of police-reported pedestrian crashes obtained from PBCAT for North Carolina, spanning 2018 to 2022. Two additional data sources were integrated to obtain a more complete picture of crashes: (1) census tract-level demographic data and (2) built environment measures, including street connectivity and network density, NaNDA (Ailshire et al., 2023). These three datasets were merged based on crash location using ArcGIS and R, enabling two levels of analysis:

1- **Crash-Level Analysis**: Individual crash records were spatially linked to corresponding built environment and census tract attributes. After merging, 30 crashes were excluded due to missing values in key variables, resulting in a final sample of 10,918 pedestrian crashes. This portion of the study focuses on modeling crash severity, which is classified using the KABCO scale. The KABCO scale is a standardized injury severity classification used in police crash reports: K: Fatal injury, A: Suspected serious injury, B: Suspected minor injury, C: Possible injury, O: No apparent injury; this ordinal scale serves as the dependent variable for the crash-level analysis.

2- **Census Tract-Level Analysis**: Crashes were aggregated by census tract to explore spatial patterns across North Carolina. The outcome variable in this analysis is the frequency of pedestrian crashes per tract. The final dataset includes 2,660 census tracts with crash data. The main variables used in this section focused on census tract characteristics and burden indicators, as mentioned further in Table 3.

Together, these complete datasets support both micro-level (individual crash) and macro-level (area-wide) investigations into the relationships between crash outcomes and built environment features.

Crash-Level Descriptive Analysis

Descriptive statistics for the crash-level data are presented in Table 2. This study captures various characteristics related to pedestrian crashes, including crash types, roadway features, and individual-level factors, to better understand patterns and potential risk factors associated with crash occurrences. Figure 2 illustrates the distribution of crash outcomes based on the KABCO scale. Among the 10,918 pedestrian-involved crashes recorded from 2018 to 2022, the majority resulted in non-fatal injuries. Specifically, suspected minor injuries (B) were the most frequent outcome, accounting for 4,186 crashes (38.3%), followed by possible injuries (C) with 3,210 cases

(29.4%). More severe outcomes included 1,528 suspected serious injuries (A) (14.0%) and 1,229 fatal crashes (K) (11.3%). Only 765 crashes (7.0%) resulted in no apparent injury (O). These distributions highlight the high likelihood of injury in pedestrian-involved crashes, with relatively few resulting in no harm.

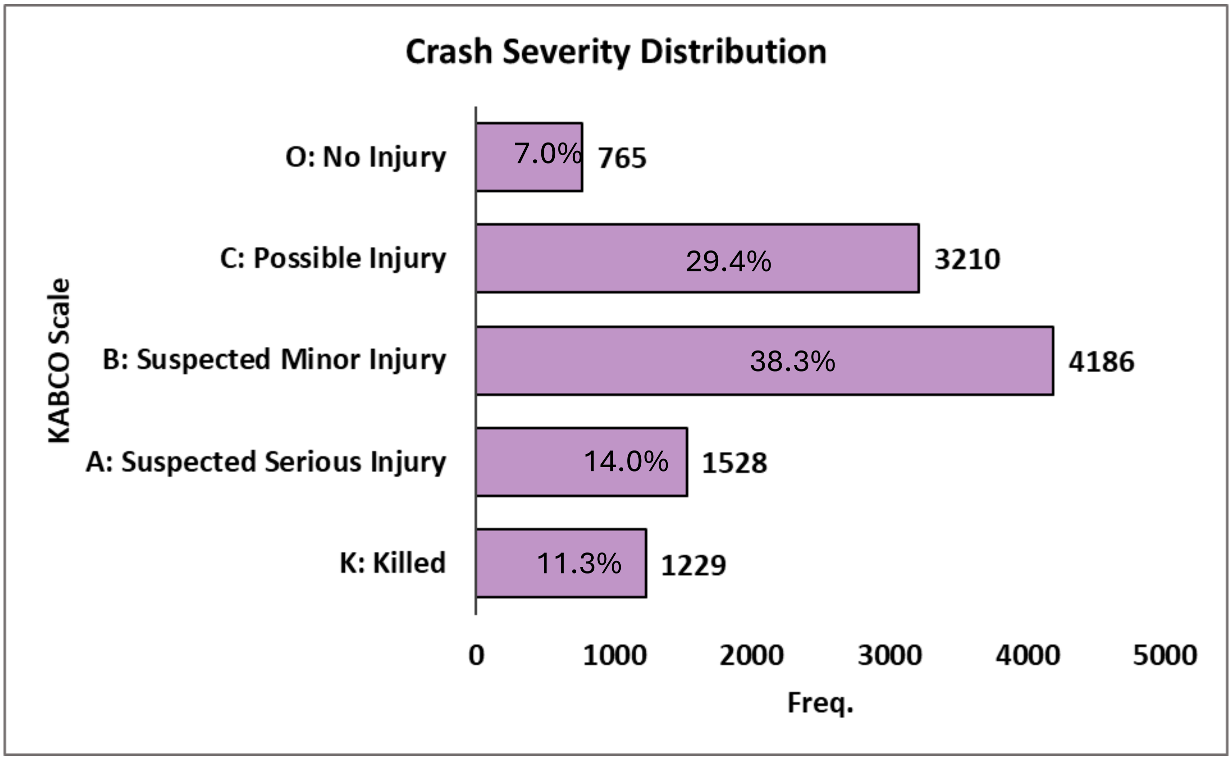

Figure 2. Distribution of Crash Severity Across 10,918 Crashes

Figure 3 illustrates the distribution of pedestrian crashes across various crash types, each representing different pedestrian actions or scenarios at the time of the incident. Among the 10,918 crashes analyzed, the most common situations involved pedestrians crossing the roadway, either with or without vehicle turning movements. Specifically, crashes involving non-turning vehicles accounted for the highest frequency—2,327 crashes (21.3%)—followed by turning vehicles with 1,821 crashes (16.7%), and pedestrians walking along the roadway with 1,784 crashes (16.3%). Other notable categories included unusual circumstances (1,486 crashes, 13.6%) and pedestrians in the roadway under unknown circumstances (1,047 crashes, 9.6%). Less common scenarios involved bus-related incidents (80 crashes, 0.7%) and waiting to cross (38 crashes, 0.3%). Overall, the figure highlights that crossing-related and walking-along-roadway scenarios are the most prevalent, pointing to key targets for interventions aimed at improving pedestrian safety.

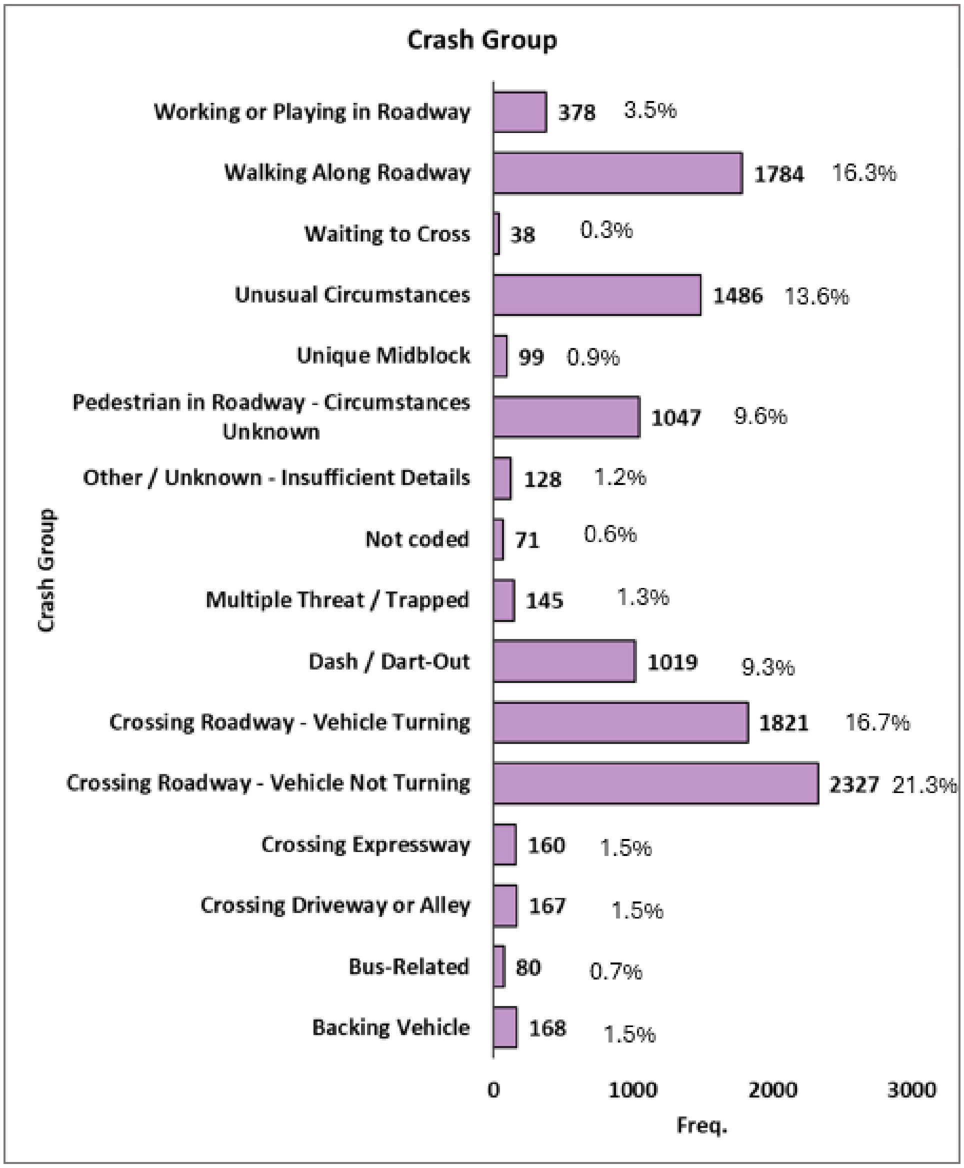

Figure 3. Crash Injury Severity Distribution Based on the KABCO Scale

Tables 2 and 3 present descriptive statistics for the variables used in the analysis, grouped into three main categories: built environment factors, pedestrian characteristics, and driver characteristics. The dataset included 10,918 pedestrian-involved crashes across North Carolina from 2018 to 2022. The distribution of crashes across years was fairly balanced, with the highest number in 2018 (22.2%) and the lowest in 2020 (18.2%), likely reflecting changes in travel patterns during the COVID-19 pandemic.

Built environment factors indicated that most crashes occurred on roads with speed limits between 30–35 mph (35.0%) and 40–45 mph (23.3%), while a smaller share happened on lower (2.3%) or higher-speed roads (3.6%). Lighting conditions revealed that nearly half of the crashes occurred in daylight (45.5%), but nighttime conditions were also significant, with 25.1% occurring under dark but lighted roadways and 24.4% on unlit roads. Regarding road geometry, most crashes occurred on straight road segments (92.0%) and under dry surface conditions (84.7%), with only a small share happening on curves or in adverse surface conditions. Most crashes occurred during clear weather (78%), followed by cloudy conditions (11%). Rain and similar weather types were less common, while snow and fog were rare.

Most crashes occurred on two-way, undivided roads (65%), followed by divided roads with unprotected medians (19%). Roads with positive median barriers and other configurations were less common. The majority of crashes occurred in urban areas (67.8%), while rural and mixed-use areas accounted for 16.8% and 14.8%, respectively. Around 58.2% of crashes took place outside intersections, and traffic control was absent in 52.2% of cases. Furthermore, 75.7% of crashes occurred on weekdays, and 57.6% of the sample came from census tracts classified as transportation-cost burdened. Street density values varied widely across locations, with an average of 20,347 links per square mile and a standard deviation of 11,585, reflecting a mix of high- and low-density urban environments.

According to pedestrian-related factors, 11.9% of pedestrians involved in the crashes were flagged for alcohol involvement, while 83.7% were not. Males made up the majority of pedestrians (63.7%), followed by females (33.7%). In terms of pedestrian position, two-thirds (66.7%) were coded as category 1 (likely roadway or crosswalk), while 23.7% were in position 2, and the remaining were distributed across other positions. The average pedestrian age was 38 years (ranging from 1 to 121)

Across all crashes, a majority occurred in tracts under high transportation-cost stress (57.6%) and in those flagged for social burdens (53.4%). Health-burdened areas accounted for 42.5% of crashes, while environmental-burdened tracts saw 39.0%, and weather-burdened ones 31.9%. Only about one-quarter of crashes (24.5%) happened in areas facing transportation-access challenges, highlighting how economic and social strains more strongly coincide with higher crash counts.

Driver-related characteristics showed that alcohol involvement was relatively low (3.1%), though 9.2% of records lacked reliable information. Males represented 48.3% of drivers, females 32.0%, and 19.7% were unknown or not coded. The majority of crashes involved passenger vehicles (85.5%), followed by trucks and commercial vehicles (2.1%), buses and emergency vehicles (0.9%), and motorcycles or small vehicles (0.7%). The remaining cases involved unknown vehicle types. Additionally, the average driver age was 42 years (ranging from 11 to 98).

Table 2. Descriptive Statistics (N = 10,918 Pedestrian Crashes)

Variable	Categories	Freq.	Percent
Built Environmental Factors			
Crash Year	**2018**	2,424	22.2%
	2019	2,267	20.8%
	2020	1,988	18.2%
	2021	2,029	18.6%
	2022	2,210	20.2%
Speed Limit (mph)	**5 - 15**	252	2.3%
	20 - 25	1,586	14.5%
	30 - 35	3,814	35.0%
	40 - 45	2,542	23.3%
	50 - 55	1,298	11.9%
	60 - 75	395	3.6%
	Unknown	1031	9.4%
Lighting	**Dark – Lighted Roadway**	2,741	25.1%
	Dark – Roadway Not Lighted	2,664	24.4%
	Dark – Unknown Lighting	47	0.4%
	Dusk	308	2.8%
	Dawn	166	1.5%
	Daylight	4,972	45.5%
	Unknown	20	0.2%

Road Characteristics	**Straight**	10,040	92.0%
	Curved	565	5.2%
	Unknown	313	2.9%
Land Use	**Urban (>70% Developed)**	7,401	67.8%
	Rural (<30% Developed)	1,829	16.8%
	Mixed (30% to 70% Developed)	1,617	14.8%
	Not coded	71	0.7%
Traffic Control Present	**No Control**	5,701	52.2%
	Yes	5,217	47.8%
Road Condition	**Dry**	9,252	84.7%
	Wet	1,627	14.9%
	Ice	35	0.3%
	Sand, Mud, Dirt, Gravel	4	0.0%
Weather	**Clear**	8,559	78.3%
	Cloudy	1,221	11.2%
	Rain, Sleet, Hail, Freezing Rain/Drizzle	1,046	9.6%
	Snow	35	0.3%
	Fog, Smog, Smoke	57	0.5%
Road Configuration	**One-Way, Not Divided**	422	3.9%
	Two-Way, Divided, Positive Median Barrier	1,094	10.0%
	Two-Way, Divided, Unprotected Median	2,036	18.6%
	Two-Way, Not Divided	7,061	64.7%
	Unknown	305	2.8%
Burden Indicators	**Weather Burden**	3,488	31.9%
	Environmental Burden	4,256	39.0%
	Social Burden	5,830	53.4%
	Health Burden	4641	42.5%
	Transportation Access Burden	2,674	24.5%
	Transportation Cost Burden	6,293	57.6%

Non-Intersection		6,354	58.2%
Pedestrian characteristics			
Pedestrian Alcohol Flag	**No**	9,134	83.7%
	Yes	1,303	11.9%
	Unknown	481	4.4%
Gender	**Male**	6,961	63.7%
	Female	3,681	33.7%
	Unknown	276	2.5%
Pedestrian Position	**Travel way**	7,287	66.7%
	Crosswalk, Sidewalk	2,592	23.7%
	Paved Shoulder / Bike Lane / Parking Lane	646	5.9%
	Unpaved Right-of-Way	393	3.6%
Driver characteristics			
Driving Under the Influence (DUI)	**No**	9,567	87.6%
	Yes	341	3.1%
	Unknown	1,010	9.2%
Gender	**Male**	5,275	48.3%
	Female	3,491	32.0%
	Unknown	2,152	19.7%
Vehicle Type	**Buses & Emergency Vehicles**	95	0.9%
	Motorcycles & Small Vehicles	79	0.7%
	Passenger Vehicles	9329	85.4%
	Trucks & Commercial Vehicles	234	2.1%
	Unknown	1181	10.8%

Table 3. Summary of Continuous Variables

	Min	Max	Avg	SD	Range
Driver Age	11	98	42	18	87
Pedestrian Age	1	121	38	18	120
Street Network Density	1,385	61,987	20,347	11,585	60,602

Note: Street Network Density = (Total length of all street segments within a census tract divided by the total land area of that tract, in square miles)

Census Tract-Level Descriptive Analysis

Tables 4 and 5 summarize key socio-demographic, economic, transportation, and built environment variables at the census tract level across 2660 North Carolina. This table shows that burdens related to transportation, both financial and accessibility, are among the most widespread challenges in North Carolina census tracts, each affecting nearly half of all tracts. These burdens include high transportation costs relative to household income, limited access to reliable transit services, long commute times, and a notable share of households without access to a vehicle, all of which can severely limit opportunities for employment, education, and essential services. At the same time, social and health burdens also have a substantial share (approximately 40% and 30%, respectively). Environmental and weather burdens are less prevalent, appearing in approximately one in six and one in seven tracts. Just over 60% of tracts are classified as urban.

Additionally, street network density, derived from the NANDA dataset as the total length of street segments within each census tract, varies dramatically, ranging from sparsely connected rural areas to highly grid-like urban neighborhoods. Similarly, tract populations cluster around 4,000 residents on average but span from virtually unpopulated to more than 13,000 people, illustrating the state's broad diversity in both demographic and built-environment characteristics. This variable is used as an exposure in the model to adjust for raw crash counts by population size—essentially modeling crashes per resident—so that we estimate a true crash rate rather than simply the count, thereby controlling for the fact that larger tracts naturally tend to experience more crashes.

Figure 4 shows that most census tracts in North Carolina experienced very few crashes: over 1,100 tracts reported zero crashes. Other regions recorded at least one crash per census tract. Frequencies then decline sharply, with only a handful of tracts experiencing more than five crashes, and a small tail extends to a maximum of 92 crashes in a single tract. On average, the number of crashes per census tract is about 3, with a standard deviation of about 5. This heavy "zero inflation" and long right tail mean simple Poisson regression models may not fit well. Thus, to address both the excess zeros and the overdispersion of high-crash tracts, a ZNHB approach is used.

Table 4. Descriptive Statistics (N = 2,660 Census Tracts)

Categorical Variables	Frequency	percentage
Transportation Cost Burden: (1 = Yes)	1,267	47.6%
Transportation Access Burden: (1 = Yes)	1,226	46.1%
Social Burden: (1 = Yes)	1,087	40.9%
Health Burden: (1 = Yes)	838	31.5%
Environmental Burden: (1 = Yes)	429	16.1%
Weather Burden: (1 = Yes)	357	13.4%
Total Burden indicator: (1 = Yes)	838	31.5%
Urban Tract Indicator: (1 = Yes)	1,657	62.3%

Table 5. Summary of Continuous Variables (Census Tracts)

	Mean	Max	Min	St
Street Network Density	14,476.8	6,1987.9	4.0	9,879.0
Total Population	3,904.6	1,3021.0	0.0	1,654.5
Number of Crashes	3.1	92.0	0.0	5.3

Note: Street Network Density = (Total length of all street segments within a census tract divided by the total land area of that tract, in square miles)

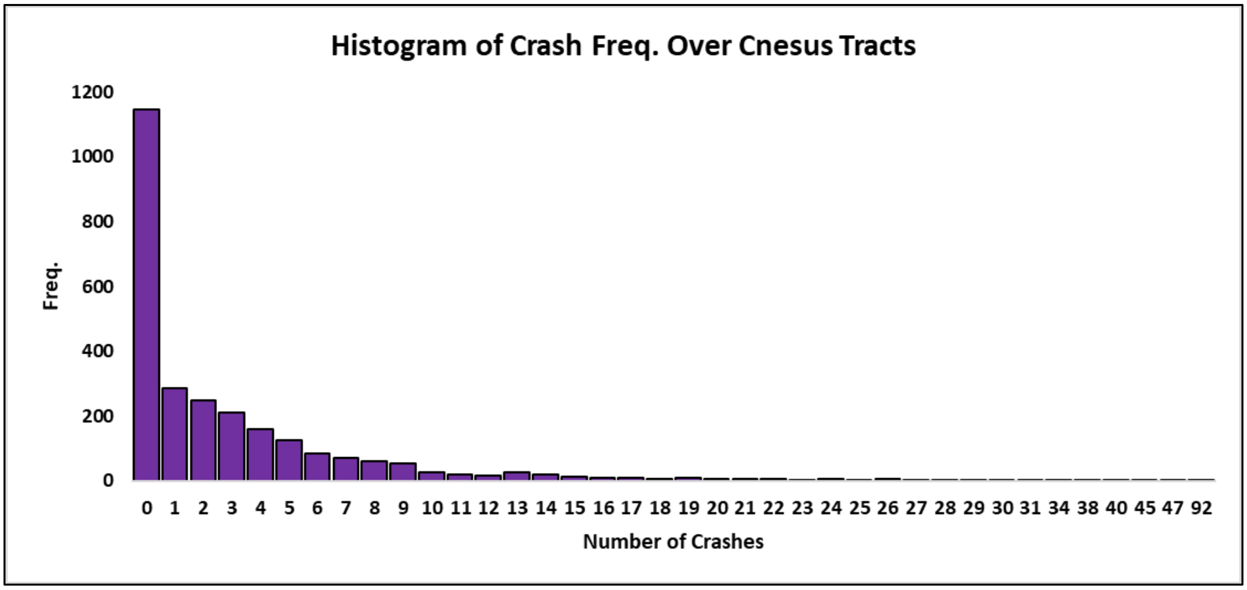

Figure 4. Crash Frequency Across 2,660 Census Tracts in NC

Figure 5 visually compares population distribution and pedestrian crash locations across North Carolina. The map illustrates total population by census tract, with darker shades representing more densely populated areas. Urban centers such as Charlotte, the Triangle, and the Triad exhibit higher population concentrations. The spatial distribution of pedestrian crashes is marked by blue points overlaid on the same population background. While many crashes occur in highly populated areas, the presence of crashes in less populated regions suggests that population density alone does not fully explain crash patterns. This comparison underscores the importance of considering additional environmental and contextual factors when assessing pedestrian crash risk.

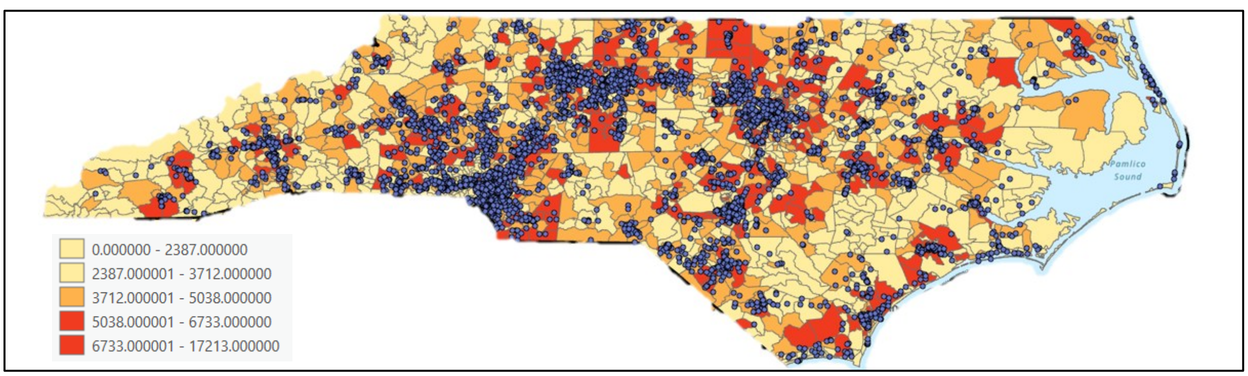

Figure 5. Population and Pedestrian Crash Distribution in NC census tracts (2018–2022)

Methodology

This study adopts a two-tiered methodological framework to assess the role of built environment factors in pedestrian safety. The analysis is divided into two components: the first focuses on the crash level, evaluating how individual crash characteristics and roadway contexts are associated with severity outcomes; the second addresses crash frequency using spatial-level modeling to capture associations at the area-wide or census tract level.

Crash-Level: Bayesian Cumulative Ordered Logit Model

To examine the factors associated with pedestrian crash severity and account for uncertainty in data and parameter estimates, we estimated a Bayesian cumulative ordered logit model using crash data from PBCAT (2019–2022), focusing on built environment and roadway attributes (Bürkner, 2017). This analysis method enables a nuanced understanding of correlates that traditional statistical methods cannot provide. Crash severity was categorized into three ordered levels based on the KABCO scale: (1) no injury (code "O"), (2) minor injury (including "B" non-incapacitating and "C" possible injuries), and (3) serious injury (a combination of "A" severe injuries and "K" fatalities). We also tested a five-level categorization aligned with the full 5-category KABCO scale, reported in the Appendix. The ordinal outcome was modeled using the cumulative logit specification:

$$\boldsymbol{Logit}\left(\boldsymbol{P}(\boldsymbol{Y} \leq \boldsymbol{k} \mid \boldsymbol{X})\right) = \boldsymbol{\theta_k} - \boldsymbol{X\beta}, \boldsymbol{k} = \mathbf{1}, \mathbf{2} \qquad 1$$

where:
θ_k are threshold parameters separating severity levels, and β are regression coefficients for the predictors X.

Explanatory variables included pedestrian position, crash type, roadway configuration, speed limit, lighting condition, weather, rural/urban status, and Burden Indicators, along with the continuous covariate of street-network density. All categorical predictors were dummy-coded, and continuous variables were standardized.

We used informative Normal priors for select coefficients based on a previous 2018 model to stabilize estimation in sparse-data scenarios (e.g., pedestrian positions or crash types):

$$\boldsymbol{\beta j} \sim \boldsymbol{Normal}(\boldsymbol{\mu}, \boldsymbol{\sigma}^{\mathbf{2}}) \qquad 2$$

Flat (weakly informative) priors were used for Burden Indicators, network density, and interaction terms:

$$\boldsymbol{\beta j} \sim \boldsymbol{Normal}(\mathbf{0}, \mathbf{10}^{\mathbf{2}}) \qquad 3$$

Model estimation was performed using the brms package with Hamiltonian Monte Carlo (HMC), using four chains and 2,000 iterations per chain. Convergence was confirmed with R^ ~1.00. Posterior predictive checks and Leave-One-Out cross-validation (LOO) (all Pareto < 0.7) were used to indicate model fit and predictive performance (which was found to be good).

The Bayesian framework was selected because it accounts for uncertainty through repeated sampling with replacement. Furthermore, the estimated model accounts for the ordinal nature of the outcome, allows the incorporation of prior knowledge, and ensures inference. Bayesian models are similar in structure to traditional frequentist regression models but differ in that they estimate a full distribution of parameter values rather than providing single-point estimates. This approach is particularly useful in handling uncertainty, as mentioned (Rezapour et al., 2021). The Bayesian cumulative logit model offers a flexible and interpretable approach to understanding the relationship between environmental factors and pedestrian crash severity.

Crash-Level: Ordered Logit Model

The ordered logit model (Grilli & Rampichini, 2014), also known as ordinal logistic regression, was used to model ordinal dependent variables when multiple categories have a natural order. In the context of our research, the Ordered Logit Model could be employed to analyze crash severity levels, categorized into levels such as 'minor,' 'moderate,' and 'severe.' Each severity level served as an ordered response, and the model helped understand how predictor variables were associated with the likelihood of crashes falling into these three categories. The following formula represents the equations of this model:

$$\boldsymbol{Y = i \; if \; \alpha^{j-1} < Y^* < \alpha^{j}} \qquad 4$$
$$\boldsymbol{Y^* = \beta X + \varepsilon}$$

Here, Y is the response variable
Y^* is a continuous latent variable
j is an ordinal response
α^{j} are the cut points or thresholds

$$\boldsymbol{Logit\,[P(Y < j)] = \log(\frac{P(Y<j)}{1-P(Y<j)})} \qquad 5$$

The final formula for the ordered logit model is as follows, where the effect of each variable is captured by β:

$$\boldsymbol{Logit\,[P(Y < j)] = \alpha^{j} + \beta^{i}X^{i}} \qquad 6$$

This estimation allows for a direct comparison between the results of the Bayesian models, with and without informative priors, and the traditional ordered logit model.

Census Tract-Level: Zero-Hurdle Negative Binomial Model

To examine area-level factors associated with pedestrian crash frequency, we estimated a Zero-Hurdle Negative Binomial model (Zeileis et al., 2008). This count regression approach is particularly suitable for overdispersed data with an excess number of zero counts, a common feature in areal crash datasets where many tracts experience no crashes over the study period. Unlike the standard negative binomial or Poisson models, the hurdle model treats zero counts as arising from a different process than positive counts. It assumes that all zero outcomes occur through a binary process (i.e., whether any crash occurs at all). The positive counts (i.e., number of crashes, conditional on at least one crash occurring) are modeled through a truncated count process.

The zero-hurdle model consists of two components: (1) a Hurdle component (binary process), which involves a logistic regression predicting the probability that a census tract experiences at least one crash. (2) Count component (truncated process): A zero-truncated negative binomial model estimating the number of crashes, given that at least one crash has occurred. Formally, let y_i denote the observed number of pedestrian crashes in census tract *i*. The model is specified as:
Overall Probability Function:

$$\boldsymbol{f(y_i|X_i,\beta,\alpha)} = \begin{cases} \boldsymbol{P_i} & \boldsymbol{, if\ y_i = 0} \\ \boldsymbol{(1-P_i)\ g(y_i|\mu_i,\alpha), if\ y_i > 0} \end{cases} \qquad 7$$

Logistic Hurdle Component:

$$\boldsymbol{P_i} = \frac{\boldsymbol{e^{\delta'\omega_i}}}{\boldsymbol{1+e^{\delta'\omega_i}}} \qquad 8$$

Where:

$\boldsymbol{P_i}$: Probability of observing a zero count.

$\boldsymbol{\omega_i}$: Covariates in the hurdle (logit) model.

δ: Coefficients for the hurdle component.

Mean Function of the Count Model:

$$\boldsymbol{ln(\mu_i) = X'_i\beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_k x_{ik}} \qquad 9$$

Where:
μ_i: Expected crash count, conditional on yi>0y_i > 0yi>0.
X'_i: Covariates for the count model.
β: Coefficients for the count component.

ZHNB models are estimated using maximum likelihood estimation (MLE). Model fit is assessed using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), where lower values indicate better model performance.

Akaike Information Criterion (AIC):

$$AIC = -2LL + 2K \qquad 10$$

Bayesian Information Criterion (BIC):

$$\boldsymbol{BIC = -2LL + 2(ln\ N)\ K} \qquad 11$$

Where:
LL: The log-likelihood,
K: The number of estimated parameters,
N: The total number of observations.

Results

Crash Level Results

Bayesian Ordinal Model

To examine the factors contributing to pedestrian crash injury severity, we conducted an ordinal Bayesian regression analysis using crash data from 2019 to 2022, which included 8,494 crashes, and compared two prior settings. In the informative-prior model, each coefficient was assigned a Normal prior with a mean and standard deviation based on a 2018 Bayesian analysis of 2,424 crashes. In contrast, the flat-prior model used vague Normal (0, 10) priors for all coefficients. By borrowing from the 2018 posteriors, the informative-prior model improved the stability of estimates for rare crash scenarios and enabled the assessment of whether earlier patterns persisted in more recent years.

The estimation results are presented in Table 6, which displays the model estimates for both the non-informative prior (parameters similar to a conventional frequentist ordered logit model) and the 2018-informed priors, along with their corresponding 95% credible intervals. To interpret the table, the positive coefficients indicate an increased likelihood of more severe injury outcomes, while negative coefficients suggest a reduced likelihood. The 95% credible intervals reflect the uncertainty around each estimate; if the interval does not include zero, the variable is considered statistically significant. Additionally, Table 7 presents the distribution of each variable derived from the Bayesian model using the 2018 dataset, which served as the prior for the 2019–2022 model.

According to the results, pedestrian position played a critical role in shaping crash outcomes. When compared to crashes occurring on travel lanes, crashes involving pedestrians on sidewalks, shared-use paths, or driveways were significantly less likely to result in higher severity outcomes. This aligns with expectations, as these areas typically offered some degree of separation from vehicular traffic. However, paved shoulders, bike lanes, and parking lanes, as well as unpaved rights-of-way, did not show a substantial association with injury severity. At the same time, this was not significant in non-informative data. Notably, these effects were more stable under the informative prior model, suggesting that leveraging past data helped to improve precision in estimates for categorical subgroups with lower frequency.

Crash type also emerged as a key determinant of severity. Crashes involving pedestrians crossing the roadway while turning or not showed significantly elevated odds of more severe outcomes than the base case of “backing” crashes. In contrast, crash types such as crossing the driveway or alley and expressway were less significant than those in the base case. Situations like dash-outs, working or playing on the road, or unusual pedestrian behaviors could intensify injury severity further. These findings underscore the risks associated with unpredictable pedestrian behavior and the complexities of vehicle-pedestrian interactions. The results for the non-informative data showed similarities with higher coefficients.

Given a crash, the posted speed limit served as an important environmental factor in injury severity. Crashes occurring on roadways with speed limits above 50 mph were consistently associated with substantially higher severity, likely due to the reduced time available to react and prevent a potential crash. Lighting conditions also significantly contribute to outcomes; crashes during daylight and dusk/dawn were generally associated with lower severity than nighttime crashes, likely due to improved visibility and earlier hazard recognition. The importance of visibility is further highlighted by the fog, smog, and smoke conditions, which were associated with an increased risk of injury compared to clear weather. These findings were consistent across both prior-informed and flat-prior models.

Several roadway design characteristics contributed significantly to crash outcomes. For example, crashes on divided highways, both protected and unprotected, tended to show greater severity relative to one-way roadways, potentially due to higher travel speeds or more complex traffic dynamics, as also reflected in the flat prior model. However, curved road segments (as opposed to straight segments) did not show significance in either model.

The model also included interaction terms between crash type and road curvature, while many of these interactions were not statistically significant. Using informative priors helped keep these interaction estimates more stable and understandable, especially when little data existed. Additionally, the interaction between "Pedestrian in Roadway – Circumstances Unknown, Working or Playing in Roadway" and curved road segments showed a negative coefficient compared to the baseline. This suggested that, in these specific situations, the likelihood of severe injury may have been lower on curved roads than expected, possibly due to lower speeds or increased driver caution in such areas. Additionally, the results from the flat-prior model suggested that the crash type "waiting to cross" may have been significantly more dangerous on curved roads. However, the corresponding coefficient is unusually large compared to others, which may indicate instability due to the lack of prior information. This highlighted the importance of incorporating informative priors, as the same interaction was not statistically significant when prior knowledge was used.

Road characteristics also were associated with crash outcomes. Crashes occurring in rural areas were associated with significantly higher severity than those in urban settings, possibly reflecting greater vehicle speeds, longer emergency response times, and limited pedestrian infrastructure. However, the interaction between rural areas and features like paved shoulders, bike lanes, or parking lanes showed a negative association, suggesting these features may have helped reduce severity in rural settings. Additionally, greater street density was associated with lower crash severity, likely due to the presence of traffic calming features, higher driver expectancy of pedestrians, and enhanced visibility. Non-intersection was positively associated with higher injury severity compared to intersection, possibly due to more control at the intersection. Two-way divided roads—whether with a positive median barrier or an unprotected median—were associated with a higher risk than one-way roads.

Including burden indicators shed light on how a census tract's socioeconomic and environmental characteristics may have influenced the severity of crash-related injuries. Among these indicators, the Environmental Burden was a significant (negative) predictor: crashes in tracts with greater environmental stressors, such as proximity to hazardous sites or aging, high-impact infrastructure, tended to produce less severe injuries. A likely explanation was that these tracts are older, densely built areas where drivers may have driven slowly, and emergency medical services are nearby, both of which would have helped limit the severity of injuries. Weather burdens, which reflected increased disasters, heat extremes, heavy rainfall, drought-driven fires, sea-level flooding, and hardened surfaces, amplified the risks of heat stress, drownings, burns, structural injuries, and slips and falls, and showed a positive association with injury severity. The remaining burden indicators, such as transportation costs and environmental factors, were not statistically linked to severity.

Driver and pedestrian characteristics, including age, pedestrian alcohol involvement, and driver DUI status, were included in both models, which produced consistent results. Alcohol use significantly increased the likelihood of injury for both parties, especially when the driver was under the influence. Older drivers were associated with a lower risk of injury, likely due to their greater driving experience. At the same time, younger pedestrians faced a higher risk of injury, possibly due to limited awareness or engaging in risky behaviors.

In this model, the expected log predictive density (elpd_loo) for the 2019–2022 model was –6193.4 with a standard error of 66.4, and the leave-one-out information criterion (looic) was 12,387.5 with a standard error of 132.9. In comparison, the flat model for the same period had an elpd_loo of –6193.8 (SE: 66.8) and a looic of 12,393 (SE: 133.5). Although the differences were minor and within the range of the standard errors, the slightly higher elpd_loo and lower looic values for the 2019–22 model indicated marginally better predictive performance, justifying its selection over the flat model.

Comparing the models revealed that the informative prior approach slightly enhanced the precision of the estimates. The use of historical information from the 2018 dataset helped stabilize effect estimates without introducing noticeable bias, thereby improving transferability and interpretability. While the direction and magnitude of effects were generally similar across models, the prior-informed approach reduced uncertainty for many predictors, especially those with limited contemporary data.

Overall, these results highlighted the complex interplay between pedestrian behaviors, roadway design, environmental context, and community vulnerability in shaping crash severity outcomes. The findings underscored the need for geographically and behaviorally targeted interventions, particularly in high-speed or rural environments, and supported the broader integration of social considerations into pedestrian safety planning.

Table 6. Bayesian Ordinal Model Results (2019–2022, N = 8,494 Crashes; O, (C+B), (A+K))

Variables	Informative prior			Non-informative prior (flat)		
	Estimate (mean)	l-95% CI (2.5%)	u-95% CI (97.5%)	Estimate (mean)	l-95% CI (2.5%)	u-95% CI (97.5%)
Intercept [1]	-2.05	-2.38	-1.71	-1.70	-2.22	-1.18
Intercept [2]	2.04	1.71	2.39	2.42	1.89	2.95
Pedestrian Position (Base = Travel Lane)						
Ped Position 2: Sidewalk / Shared Use Path / Driveway Crossing	-0.36*	-0.50	-0.24	-0.46*	-0.60	-0.32
Ped Position 3: Paved Shoulder / Bike Lane / Parking Lane	-0.05	-0.30	0.19	0.11	-0.18	0.39
Ped Position 4: Unpaved Right-of-Way	-0.09	-0.39	0.21	-0.17	-0.56	0.21
Crash Type (Base = Backing)						
Crash Type 2: Crossing Driveway or Alley and Expressway	0.33*	0.02	0.64	0.75*	0.23	1.26
Crash Type 3: Crossing Roadway - Vehicle Turning/Not Turning	0.50*	0.28	0.71	0.81*	0.36	1.23
Crash Type 4: Pedestrian in Roadway - Circumstances Unknown, Working or Playing in Roadway	0.45*	0.21	0.67	0.74*	0.30	1.18
Crash Type 5: Dash / Dart-Out	0.62*	0.38	0.85	0.92*	0.47	1.36
Crash Type 6: Walking Along Roadway	0.05	-0.18	0.28	0.34	-0.11	0.77
Crash Type 7: Unusual Circumstances	0.29*	0.06	0.52	0.59*	0.14	1.03
Crash Type 8: Others	-0.01	-0.29	0.27	0.33	-0.16	0.82

Crash Type 9: Waiting to Cross	0.46	-0.23	1.13	0.70	-0.20	1.57
Crash Type 10: Bus-Related	-0.21	-0.72	0.31	-0.13	-0.83	0.60
Speed Limit (Under 50 mph)						
Over 50 mph	0.77*	0.63	0.91	0.73*	0.58	0.89
Unknown	-0.22*	-0.38	-0.05	-0.21*	-0.38	-0.03
Lighting condition (Base = Dark)						
Dusk or Dawn	-0.51*	-0.71	-0.30	-0.49*	-0.72	-0.26
Daylight	-0.63*	-0.73	-0.54	-0.66*	-0.77	-0.55
Road Character (Base = Straight)						
Curve – Level	0.41	-0.02	0.83	0.73	-1.11	2.52
Unknown	-0.34	-0.89	0.22	-0.46	-1.32	0.38
Road Configuration (Base = One-Way, Not Divided)						
Two-Way, Divided, Positive Median Barrier	0.59*	0.37	0.81	0.70*	0.39	1.00
Two-Way, Divided, Unprotected Median	0.39*	0.19	0.60	0.49*	0.21	0.78
Two-Way, Not Divided	0.10	-0.10	0.29	0.20	-0.08	0.46
Others	0.30	-0.25	0.85	0.54	-0.34	1.41
Weather (Base = Clear)						
Cloudy	0.11	-0.07	0.28	0.06	-0.09	0.21
Rain, Sleet, Hail, Freezing Rain/Drizzle	0.00	-0.14	0.14	-0.02	-0.19	0.14
Snow	-0.39	-1.10	0.32	-0.26	-1.46	0.88
Fog, Smog, Smoke	0.80*	0.28	1.34	0.69*	0.06	1.33
Burden Indicators						
Health Burden	-0.01	-0.11	0.10	-0.01	-0.11	0.09
Environmental Burden	-0.13*	-0.26	-0.01	-0.13	-0.26	0.00
Weather Burden	0.15*	0.02	0.28	0.14*	0.00	0.28

Cost Burden	0.07	-0.04	0.17	0.06	-0.04	0.16
Location Context						
Rural (Base = Urban)	0.52*	0.39	0.66	0.59*	0.43	0.76
log (Street Network Density (miles^2))	-0.16*	-0.22	-0.09	-0.14*	-0.21	-0.06
Non-intersection (Base = Intersection)	0.20*	0.10	0.31	0.70*	0.07	1.31
Interaction Variables						
Curve - Level * Crash Type 2	0.51	-0.53	1.56	0.28	-1.97	2.66
Curve - Level * Crash Type 3	-0.07	-0.62	0.48	-0.47	-2.33	1.42
Curve - Level * Crash Type 4	-0.85*	-1.45	-0.25	-1.24	-3.04	0.63
Curve - Level * Crash Type 5	0.34	-0.41	1.07	0.17	-1.82	2.20
Curve - Level * Crash Type 6	0.02	-0.51	0.55	-0.36	-2.19	1.47
Curve - Level * Crash Type 7	0.16	-0.39	0.71	0.00	-1.87	1.88
Curve - Level * Crash Type 8	-0.01	-0.88	0.86	-0.54	-2.54	1.52
Curve - Level * Crash Type 9	0.86	-0.77	2.49	8.58*	0.28	22.13
Curve - Level * Crash Type 10	0.05	-1.46	1.57	0.04	-2.95	2.95
Ped Position 2* Rural	0.06	-0.36	0.48	0.23	-0.28	0.71
Ped Position 3 * Rural	-0.70*	-1.03	-0.37	-0.91	-1.31	-0.52
Ped Position 4 * Rural	-0.31	-0.72	0.13	-0.07	-0.60	0.47
People-involved Features						
Pedestrian Age	0.01*	-0.01	0.00	0.00	-0.01	0.00
Driver Age	-0.01*	0.01	0.02	0.01*	0.01	0.02
Drive DUI	0.92*	0.69	1.15	0.87*	0.63	1.13

Pedestrian Alcohol Flag	0.27*	0.14	0.41	0.28*	0.13	0.43
Summary Statistics						
Model (Priors from the 2018 data distribution):			Model (Non-informative prior (flat)):			
elpd_loo = –6193.8 (SE 66.4)			elpd_loo = –6196.5 (SE 66.8)			
looic = 12387.5 (SE 132.9)			looic = 12393.0 (SE 133.5)			
N = 8,494 Crashes						
Three Injury categories: O, (C + B), (A + K)						
l-95% CI (2.5%): Lower bound of 95% CI (2.5th percentile)						
u-95% CI (97.5%): Upper bound of 95% CI (97.5th percentile)						
*Statistically Significant (5%)						

Table 7. Prior Parameter Distributions (2019–2022 Model, N = 2,424 Crashes; O, (C+B), (A+K))

Variables	Distribution from 2018	
	Mean	SD
Intercept [1]	0.00	10.00
Intercept [2]	0.00	10.00
Ped Position 2: Sidewalk / Shared Use Path / Driveway Crossing	-0.14	0.15
Ped Position 3: Paved Shoulder / Bike Lane / Parking Lane	-0.49	0.32
Ped Position 4: Unpaved Right-of-Way	-0.18	0.29
Crash Type 2: Crossing Driveway or Alley and Expressway	-0.17	0.36
Crash Type 3: Crossing Roadway - Vehicle Turning/Not Turning	0.31	0.26
Crash Type 4: Pedestrian in Roadway - Circumstances Unknown, Working or Playing in Roadway	0.55	0.28
Crash Type 5: Dash / Dart-Out	0.46	0.29
Crash Type 6: Walking Along Roadway	0.12	0.27
Crash Type 7: Unusual Circumstances	0.22	0.27
Crash Type 8: Others	-0.23	0.33
Crash Type 9: Waiting to Cross	0.19	0.69
Crash Type 10: Bus-Related	0.39	0.51
Over 50 mph	0.83	0.16
Unknown	-0.50	0.33
Dusk or Dawn	-0.64	0.23
Daylight	-0.51	0.11
Curve – Level	0.52	0.42
Unknown	-0.59	0.52
Two-Way, Divided, Positive Median Barrier	0.39	0.26
Two-Way, Divided, Unprotected Median	0.35	0.24
Two-Way, Not Divided	0.10	0.21
Others	-0.11	0.52
Rural (Base = Urban)	0.35	0.14
Cloudy	0.19	0.12
Rain, Sleet, Hail, Freezing Rain/Drizzle	0.05	0.15

Snow	-0.46	0.46
Fog, Smog, Smoke	1.02	0.53
Non-intersection	0.51	0.12
Health Burden	0.00	10.00
Environmental Burden	0.00	10.00
Weather Burden	0.00	10.00
Cost Burden	0.00	10.00
log (Street Network Density (miles^2))	-0.19	0.06
Curve - Level * Crash Type 2	0.22	0.86
Curve - Level * Crash Type 3	0.41	0.56
Curve - Level * Crash Type 4	-0.26	0.67
Curve - Level * Crash Type 5	-0.10	0.66
Curve - Level * Crash Type 6	0.39	0.51
Curve - Level * Crash Type 7	-0.54	0.54
Curve - Level * Crash Type 8	0.61	0.79
Curve - Level * Crash Type 9	0.00	1.00
Curve - Level * Crash Type 10	0.00	1.01
Ped Position 2* Rural	-0.55	0.46
Ped Position 3 * Rural	-0.30	0.42
Ped Position 4 * Rural	-0.99	0.40
Pedestrian Age	0.01	0.00
Driver Age	-0.01	0.00
Drive DUI	1.09	0.27
Pedestrian Alcohol Flag	0.19	0.27

N = 2,424 Crashes,

Three Injury categories: O, (C + B), (A + K)

Ordered Logit Model

To clarify the difference between the results of the simple (Frequentist) ordered logit model and the Bayesian models, with both informative and non-informative (flat) priors, Table 8 presents the results of the ordered model with three categories: no injury, moderate injury, and severe injury.

In most cases, the estimated coefficients and significance levels from this (Frequentist) ordered logit model were very similar to those obtained from the Bayesian model with non-informative priors (as expected), indicating consistency across modeling approaches when no prior information was used.

For categorical variables, such as pedestrian position, road characteristics, and crash type, a positive coefficient indicated a higher likelihood of more severe injury compared to the reference category. In comparison, a negative coefficient indicates a lower likelihood. For continuous variables, such as street network connectivity and age, positive coefficients suggested that increases in the variable were associated with greater injury severity.

In the Frequentist model, one category (Curve-Level * Crash Type 9) with only two observations, all at a single severity level, produced an extremely high t-value due to quasi-complete separation, where a predictor nearly perfectly predicted the outcome, resulting in unstable estimates. The Bayesian model with a flat prior produced a more reasonable but still uncertain estimate, while the model with informative priors stabilized the coefficient considerably. This highlighted the advantage of incorporating prior knowledge to regularize effects, especially in sparse data settings.

Table 8. Ordered Logit Model Results (2019–2022, N = 8,494 Crashes; O, (C+B), (A+K))

Variables	Estimate	Std Error	T value	P value
Intercept [1]	-2.27	0.28	-8.14	0.00
Intercept [2]	1.83	0.28	6.54	0.00
Pedestrian Position (Base = Travel Lane)				
Ped Position 2: Sidewalk / Shared Use Path / Driveway Crossing	-0.46*	0.08	-6.13	0.00
Ped Position 3: Paved Shoulder / Bike Lane / Parking Lane	0.11	0.14	0.75	0.45
Ped Position 4: Unpaved Right-of-Way	-0.17	0.20	-0.86	0.39
Crash Type (Base = Backing)				
Crash Type 2: Crossing Driveway or Alley and Expressway	0.76*	0.26	2.94	0.00
Crash Type 3: Crossing Roadway - Vehicle Turning/Not Turning	0.81*	0.22	3.73	0.00
Crash Type 4: Pedestrian in Roadway - Circumstances Unknown, Working or Playing in Roadway	0.75*	0.22	3.36	0.00
Crash Type 5: Dash / Dart-Out	0.93*	0.23	4.10	0.00
Crash Type 6: Walking Along Roadway	0.35	0.22	1.59	0.11
Crash Type 7: Unusual Circumstances	0.60*	0.22	2.69	0.01
Crash Type 8: Others	0.35	0.25	1.40	0.16
Crash Type 9: Waiting to Cross	0.71	0.44	1.61	0.11
Crash Type 10: Bus-Related	-0.13	0.36	-0.35	0.73
Speed Limit (Under 50 mph)				
Over 50 mph	0.73*	0.08	9.14	0.00
Unknown	-0.21*	0.09	-2.35	0.02
Lighting condition (Base = Dark)				
Dusk or Dawn	-0.49*	0.12	-4.02	0.00
Daylight	-0.66*	0.05	-12.12	0.00
Road Character (Base = Straight)				
Curve - Level	0.69	0.98	0.71	0.48

Unknown	-0.46	0.43	-1.09	0.28
Road Configuration (Base = One-Way, Not Divided)				
Two-Way, Divided, Positive Median Barrier	0.70*	0.15	4.54	0.00
Two-Way, Divided, Unprotected Median	0.49*	0.14	3.40	0.00
Two-Way, Not Divided	0.20	0.14	1.43	0.15
Others	0.55	0.44	1.26	0.21
Weather (Base = Clear)				
Cloudy	0.06	0.08	0.76	0.48
Rain, Sleet, Hail, Freezing Rain/Drizzle	-0.02	0.08	-0.29	0.77
Snow	-0.26	0.58	-0.45	0.65
Fog, Smog, Smoke	0.69*	0.32	2.17	0.03
Burden Indicators				
Health Burden	-0.01	0.05	-0.13	0.90
Environmental Burden	-0.13*	0.07	-2.00	0.05
Weather Burden	0.14*	0.07	2.06	0.04
Cost Burden	0.06	0.05	1.19	0.24
Location Context				
Rural (Base = Urban)	0.59*	0.08	7.21	0.00
log (Street Network Density (miles^2))	-0.13*	0.04	-3.50	0.00
Non-intersection (Base = Intersection)	0.10	0.06	1.72	0.09
Interaction Variables				
Curve - Level * Crash Type 2	0.22	1.18	0.19	0.85
Curve - Level * Crash Type 3	-0.44	1.01	-0.43	0.66
Curve - Level * Crash Type 4	-1.20	1.02	-1.18	0.24
Curve - Level * Crash Type 5	0.20	1.07	0.19	0.85
Curve - Level * Crash Type 6	-0.33	1.01	-0.32	0.75
Curve - Level * Crash Type 7	0.03	1.01	0.02	0.98
Curve - Level * Crash Type 8	-0.52	1.09	-0.47	0.63

Curve - Level * Crash Type 9	10.91*	0.00	22867.34	0.00
Curve - Level * Crash Type 10	0.07	1.64	0.04	0.97
Ped Position 2* Rural	0.23	0.25	0.92	0.36
Ped Position 3 * Rural	-0.91*	0.20	-4.58	0.00
Ped Position 4 * Rural	-0.07	0.27	-0.25	0.80
People-involved Features				
Driver Age	0.00*	0.00	-2.55	0.01
Pedestrian Age	0.01*	0.00	9.86	0.00
Drive DUI	0.87*	0.13	6.55	0.00
Pedestrian Alcohol Flag	0.28*	0.08	3.76	0.00
Summary Statistics Log-likelihood (fitted model): −6146.19 Log-likelihood (null model): −6839.65 McFadden pseudo-R^2: 0.10 (acceptable fit)	Cox & Snell R^2: 0.15 Nagelkerke R^2: 0.19 Likelihood ratio (G^2): 1386.93			
N = 8,894 Crashes, Three Injury categories: O, (B + C), (A + K) Statistically Significant (5%)				

Census Tract Level Results

Using a zero-hurdle negative binomial model to examine pedestrian crash frequency across 2,660 census tracts in North Carolina, we identified several key patterns related to community features and characteristics of the built environment as shown in Table 9. The total population in each tract was included as an exposure by adding the log(population) as an offset in the count component. This way, we modeled crashes per person rather than raw counts, ensuring tracts of different sizes were comparable. The model consisted of two components: the count component, which estimated the number of pedestrian crashes within tracts where crashes occurred, and the zero-hurdle component, which modeled the likelihood of having zero crashes.

In the count part of the model, several community-burden indicators were significantly associated with higher pedestrian crash frequencies. Specifically, tracts designated as having a high transportation cost burden, environmental burden, and weather burden were more likely to experience elevated crash counts. These relationships suggested that economic stressors and environmental conditions may have created environments where pedestrians were at greater risk, possibly due to inadequate infrastructure, lower-quality urban design, or higher exposure.

Additionally, tracts with greater street network length (as captured by the log of total segment length) were positively associated with crash frequency, suggesting that more extensive networks, potentially indicating greater development or higher traffic volumes, led to an increased risk of pedestrian crashes. In contrast, tracts with access burden (difficulty accessing jobs, services, or transportation) were associated with slightly fewer pedestrian crashes, possibly due to reduced pedestrian activity or isolation from major roadways.

In the zero-hurdle part of the model, which estimates the likelihood of a tract having no crashes, tracts with higher cost burdens and weather burdens were more likely to record at least one crash (i.e., less likely to be crash-free). In contrast, tracts with greater access burden and those identified as urban were more likely to remain crash-free. Environmental burden did not significantly affect whether a tract ever experienced a crash, even though it raised counts where crashes did occur.

Together, these findings suggested that pedestrian safety risk was not evenly distributed but was closely tied to the intersection of socioeconomic vulnerability, environmental stressors, and urban form. While burdened areas often bore a heavier burden of pedestrian safety concerns, some types of these issues—particularly those related to access—may also have coincided with reduced pedestrian activity and, in turn, fewer reported crashes. These insights underscored the importance of place-based strategies that accounted for exposure and vulnerability in enhancing pedestrian safety.

Overall, the model showed a reasonable fit for explaining pedestrian crash counts across tracts. The McFadden pseudo-R^2 of 0.10 indicated that the model explained approximately 10% of the variation, which was considered acceptable for count models. The dispersion value 1.365 indicated slight overdispersion, and the hurdle structure effectively managed this. The model fit the data well and captured key patterns without major issues.

Table 9. ZHNB Model Results (N = 2660 Census Tracts, NC)

Count Section	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-11.35*	0.53	-21.52	0.00
Cost Burden (1 = Yes)	0.48*	0.05	9.88	0.00
Access Burden (1 = Yes)	-0.22*	0.07	-2.81	0.01
Environmental Burden (1 = Yes)	0.73*	0.07	11.11	0.00
Weather Burden (1 = Yes)	0.38*	0.07	5.72	0.00
log (Street Network Density (miles^2))	0.42*	0.06	8.69	0.00
Log (theta)	0.65	0.08	8.69	0.00
Zero Section	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	0.92*	0.12	7.50	0.00
Cost Burden (1 = Yes)	0.34*	0.09	3.66	0.00
Access Burden (1 = Yes)	-1.48*	0.10	-14.77	0.00
Environmental Burden (1 = Yes)	0.12	0.15	0.81	0.42
Weather Burden (1 = Yes)	1.27*	0.19	6.69	0.00
Urban Tract Indicator (1 = Yes)	-0.35*	0.11	-3.33	0.00

Summary Statistics
Exposure variable: Total population per census tract
Log-likelihood (at convergence): –5,179.71
AIC: 10,385.43
BIC: 10461.89
McFadden pseudo-R^2: 0.10
Dispersion (mean squared Pearson): 1.365
Statistically Significant (5%)

Association of Community Burden Indicators with Crash Severity

The pairwise correlations among six burden indicators across all 8,494 crashes are also examined and shown in Table 10. Notably, the social burden index showed a very high correlation (0.70) with the transportation cost burden; therefore, we excluded the social burden to avoid redundancy. Next, we ran a cumulative-logit crash severity model, including health, environmental, weather, cost, and access burdens, to find the raw effect of each burden on crash severity. Table 11 presents these results: environmental and weather burdens both significantly reduced the odds of more severe injuries, while Transportation access and cost burdens significantly increased them (OR ≈ 1.61 and 1.15, respectively); health burden shows a modest positive effect. This approach enabled us to isolate the direct influence of indicators, particularly access burden, on injury severity, without mixing overlapping signals from the built environment.

By modeling burden indicators separately at the crash level, we observed that tracts with an access burden had 1.6 times greater odds of a more severe crash outcome compared to those without, holding other factors constant, even though access burden was associated with lower pedestrian crash frequency in the earlier ZHNB model with crash frequency as the dependent variable. In contrast, areas with environmental or weather burdens were associated with reduced severity (OR ≈ 0.6 and 0.8, respectively). The health burden modestly increased severity (OR ≈ 1.2), while the cost burden showed a marginally positive association (OR ≈ 1.1). By excluding the highly collinear social burden index and focusing on these five binary indicators, we gained clearer, non-redundant insights into how different community burdens relate to crash severity. These findings can inform targeted safety interventions, such as improving transit access or mitigating health vulnerabilities, in the most at-risk neighborhoods.

Table 10. Correlation Matrix of Burden Indicators

Variables	Health Burden	Envir. Burden	Social Burden	Weather Burden	Cost Burden	Access Burden
Health Burden	1.00	0.11	0.36	0.26	0.32	–0.34
Environmental Burden	0.11	1.00	0.19	0.49	0.14	–0.39
Social Burden	0.36	0.19	1.00	0.11	0.70	–0.07
Weather Burden	0.26	0.49	0.11	1.00	0.06	–0.38
Cost Burden	0.32	0.14	0.70	0.06	1.00	–0.07
Access Burden	–0.34	–0.39	–0.07	–0.38	–0.07	1.00

Table 11. Ordered Logit Model Results Using Only Burden Indicators (N = 8,494 Crashes)

Variables	Estimate	Std. Error	t value	p value	OR
Health Burden (1 = Yes)	0.20*	0.05	3.7	0.00	1.2
Environmental Burden (1 = Yes)	−0.42*	0.06	−7.3	0.00	0.6
Weather Burden (1 = Yes)	−0.13*	0.06	−2.2	0.03	0.8
Cost Burden (1 = Yes)	0.13*	0.06	2.7	0.05	1.1
Access Burden (1 = Yes)	0.47*	0.06	7.8	0.00	1.6

Summary Statistics
McFadden's R^2: 0.096
McFadden's adjusted (R^2ML): 0.340
Cox & Snell (R^2CU): 0.345
Statistically Significant (5%)

Discussion

This study provides a comprehensive assessment of pedestrian crash risk and injury severity by integrating crash-level severity modeling and census tract-level crash frequency modeling, utilizing three datasets, with a primary focus on PBCAT. At the crash level, Bayesian modeling incorporate uncertainty and informative priors derived from a 2018 analysis to improve estimate stability, particularly for rare crash scenarios and low-frequency pedestrian behaviors, ensuring more reliable inference in the presence of sparse data. The results provide insights into how various built environment features contribute to pedestrian safety outcomes, how these relationships change across different lighting conditions, and which behavioral patterns can be targeted to improve safety.

The analysis highlights several built environment factors that significantly shape both the likelihood and severity of pedestrian crashes, consistent with previous studies (Clifton et al., 2009; Ukkusuri et al., 2011; Yu, 2015). At the crash level, the location of the crash in relation to pedestrian infrastructure plays a critical role in injury severity. Crashes that occur on designated pedestrian spaces, such as sidewalks, shared-use paths, or driveways, are significantly less likely to result in serious injuries compared to those that occur on travel lanes, where vehicle-pedestrian conflicts are more direct (Schneider et al., 2004). This suggests that the presence of pedestrian infrastructure can help reduce the severity of injuries in the event of a crash, emphasizing the importance of separating pedestrian spaces from vehicle travel lanes to enhance pedestrian safety. Additionally, high injury severity on roads with speed limits above 50 mph suggests an urgent need for engineering interventions (e.g., lower posted speed limits, speed humps, narrower lanes, road diets) and enforcement strategies (e.g., automated speed cameras) in high-speed or rural corridors (Zahabi et al., 2011).

At the neighborhood level, the ZHNB model showed that tracts with longer total street segment lengths had higher crash counts, indicating a relationship between infrastructure density and pedestrian exposure. Land use factors, captured indirectly through burdens such as transportation cost burden and access burden, also influenced crash frequency. High transportation cost burden tracts, likely reflecting car-dependent, infrastructure-poor neighborhoods, saw more crashes, while high access burden areas experienced fewer, possibly due to lower pedestrian volumes (Patwary & Khattak, 2024). These results suggest that certain built environment features may elevate crash frequency while concurrently reducing the severity of injuries when crashes occur, supported by previous studies (Clifton et al., 2009; Ewing & Dumbaugh, 2009).

Lighting conditions showed a clear and significant relationship with injury severity. Crashes occurring during daylight or at dawn/dusk were associated with lower severity compared to those at night. This suggests that visibility plays a critical role in preventing high-severity injury outcomes. Even with the same road or land use characteristics, the protective effect of the built environment may be diminished in low-light conditions (Ferenchak et al., 2022; Liu et al., 2019). Nighttime conditions likely amplify the dangers of certain built environment features, such as curved roads or high-speed corridors, because pedestrians are harder to detect and drivers may be less alert (Ferenchak & Abadi, 2021). This powerfully highlights the importance of lighting in enhancing safety for all road users. Targeted lighting improvements at high-risk locations (hot spots) can serve as an effective intervention to reduce crash severity and improve pedestrian visibility, particularly during low-light conditions.

Our findings suggest multiple, interrelated pathways through which the built environment influences crash outcomes. Crashes in rural areas had higher severity, likely reflecting delayed medical responses, higher average vehicle speeds, and fewer safety-enhancing features, such as lighting or sidewalks (Travis et al., 2012). In contrast, higher street network density, a proxy for finer-grained, more walkable networks, was associated with lower crash severity, likely due to slower traffic and increased driver awareness of pedestrian activity (Marshall & Garrick, 2011).

In terms of burden conditions, tracts with higher environmental burden (e.g., dense urban areas with aging infrastructure) showed lower injury severity, potentially due to slower traffic speeds and proximity to emergency services, despite these areas' overall social vulnerability (Ewing & Dumbaugh, 2009). Weather burden, on the other hand, may increase injury severity, possibly due to more extreme weather conditions, heat exposure, or inadequate pedestrian infrastructure that is not resilient to weather stressors (Gu et al., 2025).

Tracts with transportation cost burden, measured as higher transportation expenses for residents in the tract, experienced higher pedestrian crash frequencies, underscoring how socioeconomically vulnerable populations may be forced to walk in unsafe environments. In contrast, the access burden, which refers to limited transportation access often characterized by poor street connectivity, long travel times to key destinations, low transit availability, and a higher share of households without vehicles, was associated with lower crash frequencies, likely due to fewer walking trips in such areas. However, when focusing on crash severity, access burden showed the opposite pattern: tracts with higher access burden had increased odds of severe injury outcomes, even though crashes were less frequent. This suggests that when crashes do occur in these areas, they may be more dangerous, possibly due to longer emergency response times or riskier road environments (Patwary et al., 2024).

This study also examined crash types in detail. While curved roadways did not show a significant main effect compared to straight roads, possibly due to their lower frequency in our dataset, prior research (e.g., (Swanson et al., 2016)) has found that crashes on curves are often more severe. In our analysis, the crash type "Pedestrian in Roadway – Circumstances Unknown, Working or Playing in Roadway" was associated with an increased likelihood of injury compared to the reference type (backing), which involves a vehicle moving in reverse, aligning with national trends that identify pedestrian roadway presence as high risk (Chang, 2008). However, the interaction between this crash type and curved roadways was associated with lower injury severity, suggesting a complex relationship: while the behavior itself is risky, its occurrence on curved roads may moderate severity, possibly due to reduced vehicle speeds or increased driver vigilance in such settings. This finding highlights the complex and context-dependent relationship between roadway geometry and crash circumstances, suggesting that certain high-risk behaviors may pose less danger in specific roadway settings, possibly due to lower vehicle speeds or increased driver caution on curves.

Additionally, several crash types, such as dash/dart, unusual circumstances, and crossings, including driveways, alleys, and expressways, were significantly associated with higher injury severity compared to the reference category, backing. These results highlight the significant associations of pedestrians and drivers involved in crashes, particularly their behavior, with the severity of injuries. This influence may also be shaped by other factors, such as age and alcohol involvement, which showed modest associations in the analysis (Chang, 2008; Vanlaar et al., 2016). However, the interactions between these crash types and pedestrian positions (e.g., sidewalk, shoulder, or driveway) were not statistically significant, suggesting that location alone may not moderate the risk posed by these crash types (Van Houten et al., 2001). This highlights the importance of well-marked crossings, pedestrian signals, refuge islands, and traffic-calming measures at locations where sight distance is limited.

Overall, this study highlights the multilayered nature of pedestrian crash severity, shaped by both the built environment and the characteristics of pedestrians and drivers, as well as overall census tract features. By integrating crash-level and tract-level analyses, we identified key factors, including pedestrian behavior, crash type, community burdens, and roadway features, that jointly influence outcomes. The findings underscore the importance of targeted interventions, such as improving infrastructure (well-marked crossings, pedestrian signals, refuge islands, and traffic-calming measures) in areas with high burdens, enhancing pedestrian visibility, and addressing behavioral risks. While some results revealed complex or context-dependent interactions, they point to the need for place-based strategies that consider both exposure and vulnerability. These insights can inform planning, policy, and design efforts aimed at mitigating the severity of pedestrian injuries and enhancing safety across various contexts.

Conclusions and Recommendations

This study provides a comprehensive, multiscale analysis of pedestrian crash outcomes by creating a unique database and conducting crash-level severity and census tract-level crash frequency modeling. Specifically, using a Bayesian framework that accounts for uncertainty in data and parameter estimates enriched with informative priors, the analysis improved estimate stability and uncovered meaningful relationships even in low-frequency scenarios. This analysis enables a nuanced understanding of correlates over traditional methods.

Key findings reveal that several neighborhood-level built environment features were significantly associated with pedestrian crash frequency and injury severity. These include location context (e.g., non-intersection and rural areas), pedestrian infrastructure (e.g., sidewalks and shared-use paths), and roadway design elements such as street network length and curvature. Community-level burdens, such as transportation costs, environmental burdens, and climate vulnerability, were associated with higher crash frequency and severity. In contrast, access burden was associated with fewer crashes but greater injury severity when they did occur, suggesting context-specific risk pathways.

To examine the correlation between the built environment and crash outcomes under different lighting conditions, we tested interaction terms between lighting and various built environment features. Most of these interaction terms were not statistically significant. However, when lighting and certain built environment features were examined separately, all showed significant associations with crash severity and produced interpretable results. For instance, crashes occurring under dark conditions were associated with significantly higher severity compared to daylight or dusk/dawn, highlighting the importance of enhanced visibility in areas with frequent pedestrian activity. Rural locations, including small towns—especially along travel corridors—may benefit from targeted lighting upgrades. These findings underscore the impact of the built environment on both exposure and injury outcomes, highlighting the need for targeted infrastructure and behavioral interventions tailored to specific settings.

The results also emphasize that crash severity is not solely a function of physical infrastructure or pedestrians' and drivers' behaviors but is instead shaped by a complex interplay of contextual and environmental conditions. Risky pedestrian behaviors, such as dashing, darting, or playing in the roadway, were key contributors to injury severity and often outweighed driver actions, although their effects varied by road context (e.g., lower severity on curved roads). Paved shoulders, bike lanes, and parking lanes in rural areas are associated with reduced crash severity.

The findings of this study have important practical implications for improving pedestrian safety through both infrastructure design and policy. Targeted investments in nighttime lighting, especially in under-resourced communities, can reduce injury severity in crashes occurring after dark. Expanding pedestrian-designated spaces, such as sidewalks and shared-use paths, can also help prevent severe injuries. Additionally, identifying and addressing risky pedestrian behaviors, particularly in high-exposure environments, can inform public education campaigns and enforcement strategies. Finally, the context-specific relationships between built environment features and crash outcomes support a place-based approach to safety planning, where interventions are tailored to the unique needs and risks of each neighborhood.

Despite the study's valuable modeling approach and multi-level perspective, several limitations should be acknowledged. First, the data are primarily drawn from observational police-reported crash records, which may underreport certain variables, such as pedestrians' exact intent or actions during the crash. Second, the built environment measures were primarily captured at the census tract level, which may not reflect the exact conditions at the crash site. Additionally, model assumptions, such as the distributional forms in the Bayesian model, may not fully reflect the complexities of the real world. The accuracy and completeness of data across sources may impact the validity of certain estimates. Lastly, while interaction effects were explored, some may have gone undetected due to limited sample sizes for specific combinations of conditions.

This research offers critical insights into the intersection of pedestrian behavior, the built environment, and community vulnerability, providing a data-driven foundation for enhancing pedestrian safety strategies. Future research should investigate dynamic, temporally sensitive modeling incorporating pedestrian activity data, land-use changes, and real-time traffic conditions. The use of detailed crash narrative data, such as officer-reported descriptions or coded behavioral factors, may provide more localized insights into crash circumstances and contributing factors, especially pedestrian behaviors. Additionally, future studies should consider longitudinal assessments to evaluate how improvements to the built environment or policy interventions affect pedestrian crash outcomes over time. By incorporating additional data, future work can continue to support evidence-based decision-making that enhances pedestrian safety in transportation systems.

Reference

Ackaah, W., Apuseyine, B. A., & Afukaar, F. K. (2020). Road traffic crashes at night-time: characteristics and risk factors. *Int J Inj Contr Saf Promot*, *27*(3), 392-399. https://doi.org/10.1080/17457300.2020.1785508

Ahmad, N., & Khattak, A. J. (2024). The role of human errors and violations in pedestrian-related crashes: Harnessing a unique database and accounting for heterogeneity. *Journal of safety research*, *91*, 136-149.

Ailshire, J., Melendez, R., Chenoweth, M., & Gypin, L. (2023). *National Neighborhood Data Archive (NaNDA): Street Connectivity by Census Tract and ZIP Code Tabulation Area, United States, 2010 and 2020* (Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR38580.v2

Auerbach, J., Eshleman, C., & Trangucci, R. (2021). A hierarchical Bayes approach to adjust for selection bias in before–after analyses of vision zero policies. *Computational Statistics*, *36*, 1577-1604.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, *80*, 1-28.

Chang, D. (2008). *National pedestrian crash report*.

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

Clifton, K. J., Burnier, C. V., & Akar, G. (2009). Severity of injury resulting from pedestrian–vehicle crashes: What can we learn from examining the built environment? *Transportation Research Part D: Transport and Environment,*, *14*(6), 425-436. https://doi.org/10.1016/j.trd.2009.01.001

Congiu, T., Sotgiu, G., Castiglia, P., Azara, A., Piana, A., Saderi, L., & Dettori, M. (2019). Built environment features and pedestrian accidents: An Italian retrospective study. *Sustainability*, *11*(4), 1064.

Das, S., Le, M., & Dai, B. (2020). Application of machine learning tools in classifying pedestrian crash types: A case study. *Transportation safety and environment*, *2*(2), 106-119.

Evenson, K. R., LaJeunesse, S., & Heiny, S. (2018). Awareness of Vision Zero among United States' road safety professionals. *Injury Epidemiology*, *5*(1), 21. https://doi.org/10.1186/s40621-018-0151-1

Ewing, R., & Dumbaugh, E. (2009). The built environment and traffic safety: a review of empirical evidence. *Journal of Planning Literature*, *23*(4), 347-367.

Ewing, R., Hamidi, S., & Grace, J. B. (2014). Urban sprawl as a risk factor in motor vehicle crashes. *Urban Studies*, *53*(2), 247-266. https://doi.org/10.1177/0042098014562331

Ferenchak, N. N., & Abadi, M. G. (2021). Nighttime pedestrian fatalities: A comprehensive examination of infrastructure, user, vehicle, and situational factors. *Journal of safety research*, *79*, 14-25. https://doi.org/10.1016/j.jsr.2021.07.002

Ferenchak, N. N., Gutierrez, R. E., & Singleton, P. A. (2022). Shedding light on the pedestrian safety crisis: An analysis across the injury severity spectrum by lighting condition. *Traffic Injury Prevention*, *23*(7), 434-439.

Fonsecaa, F., Ribeiroa, P. J. G., Conticellib, E., Jabbaria, M., Papageorgiouc, G., Tondellib, S., & Ramosa, R. A. R. (2022). Built environment attributes and their influence on walkability. *International Journal of Sustainable Transportation*, *16*(7), 660–679. https://doi.org/https://doi.org/10.1080/15568318.2021.1914793

Gayah, V. V., Guler, S. I., Liu, H., Blackburn, L., & Hamilton, I. (2022). *Quantification of systemic risk factors for pedestrian safety on North Carolina*.

Grilli, L., & Rampichini, C. (2014). Ordered logit model. *Encyclopedia of quality of life and well-being research*, 4510-4513.

Gu, Z., Peng, B., & Xin, Y. (2025). Higher traffic crash risk in extreme hot days? A spatiotemporal examination of risk factors and influencing features. *International Journal of Disaster Risk Reduction*, *116*, 105045.

Hu, W., & Cicchino, J. B. (2022). Relationship of pedestrian crash types and passenger vehicle types. *Journal of safety research*, *82*, 392-401.

Kim, D.-G., Washington, S., & Oh, J. (2006). Modeling crash types: New insights into the effects of covariates on crashes at rural intersections. *Journal of Transportation Engineering*, *132*(4), 282-292.

Kitali, A. E., Kidando, E., Sando, T., Moses, R., & Ozguven, E. E. (2017). Evaluating aging pedestrian crash severity with Bayesian complementary log–log model for improved prediction accuracy. *Transportation Research Record*, *2659*(1), 155-163.

Liu, J., Hainen, A., Li, X., Nie, Q., & Nambisan, S. (2019). Pedestrian injury severity in motor vehicle crashes: an integrated spatio-temporal modeling approach. *Accident Analysis & Prevention*, *132*, 105272.

Lopez, D., Malloy, L. C., & Arcoleo, K. (2022). Police narrative reports: do they provide end-users with the data they need to help prevent bicycle crashes? *Accident Analysis & Prevention*, *164*, 106475.

M Rahman, KM Kockelman, & Perrine, K. (2022). Investigating risk factors associated with pedestrian crash occurrence and injury severity in Texas. *Traffic Injury Prevention*, *23*(5), 283-289 https://doi.org/https://doi.org/10.1080/15389588.2022.2059474

Mahmoud, N., Abdel-Aty, M., Cai, Q., & Zheng, O. (2021). Vulnerable road users' crash hotspot identification on multi-lane arterial roads using estimated exposure and considering context classification. *Accident Analysis & Prevention*, *159*, 106294. https://doi.org/10.1016/j.aap.2021.106294

Mansfield, T. J., Peck, D., Morgan, D., McCann, B., & Teicher, P. (2018). The effects of roadway and built environment characteristics on pedestrian fatality risk: A national assessment at the neighborhood scale. *Accid Anal Prev*, *121*, 166-176. https://doi.org/10.1016/j.aap.2018.06.018

Marshall, W. E., & Garrick, N. W. (2011). Evidence on why bike-friendly cities are safer for all road users. *Environmental Practice*, *13*(1), 16-27.

Miranda-Morenoa, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, *43*, 1624–1634.

Munira, S., Sener, I. N., & Dai, B. (2020). A Bayesian spatial Poisson-lognormal model to examine pedestrian crash severity at signalized intersections. *Accident Analysis & Prevention*, *144*, 105679.

NHTSA. (2023). *Pedestrian Safety: Prevent Pedestrian Crashes*. National Highway Traffic Safety Administration,. Retrieved February 21 from https://www.nhtsa.gov/road-safety/pedestrian-safety

Patwary, A. L., Haque, A. M., Mahdinia, I., & Khattak, A. J. (2024). Investigating transportation safety in disadvantaged communities by integrating crash and Environmental Justice data. *Accident Analysis & Prevention*, *194*, 107366.

Patwary, A. L., & Khattak, A. J. (2024). *Exploring How Urban Form, Demographics, and Disadvantaged Communities are Linked with Pedestrian and Bicycle Safety* Transportation Research Board, Washington DC.

Prato, C. G., Kaplan, S., Patrier, A., & Rasmussen, T. K. (2018). Considering built environment and spatial correlation in modeling pedestrian injury severity. *Traffic Injury Prevention*, *19*(1), 88-93.

Qiu, L., & Nixon, W. A. (2008). Effects of adverse weather on traffic crashes: systematic review and meta-analysis. *Transportation Research Record*, *2055*(1), 139-146.

Rezapour, M., Wulff, S. S., Mehrara Molan, A., & Ksaibati, K. (2021). Application of Bayesian ordinal logistic model for identification of factors to traffic barrier crashes: considering roadway classification. *Transportation Letters*, *13*(4), 308-314.

Salon, D., & McIntyre, A. (2018). Determinants of pedestrian and bicyclist crash severity by party at fault in San Francisco, CA. *Accid Anal Prev*, *110*, 149-160. https://doi.org/10.1016/j.aap.2017.11.007

Schneider, R. J., Ryznar, R. M., & Khattak, A. J. (2004). An accident waiting to happen: a spatial approach to proactive pedestrian planning. *Accident Analysis & Prevention*, *36*(2), 193-211.

Siddiqui, C., Abdel-Aty, M., & Choi, K. (2012). Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accident Analysis & Prevention*, *45*, 382-391.

Soto, F., Kitali, A. E., Asif Raihan, M., & Alluri, P. (2022). Influence of built environment on pedestrian crashes: a case study of Miami-Dade County. *Transportation Research Record*, *2676*(9), 677-692.

Swanson, E. D., Yanagisawa, M., Najm, W., Foderaro, F., Azeredo, P., & John, A. (2016). *Crash avoidance needs and countermeasure profiles for safety applications based on light-vehicle-to-pedestrian communications*.

Travis, L. L., Clark, D. E., Haskins, A. E., & Kilch, J. A. (2012). Mortality in rural locations after severe injuries from motor vehicle crashes. *Journal of safety research*, *43*(5-6), 375-380.

Ukkusuri, S., Hasan, S., & Aziz, H. M. A. (2011). Random Parameter Model Used to Explain Effects of Built-Environment Characteristics on Pedestrian Crash Frequency. *Transportation Research Record: Journal of the Transportation Research Board*, *2237*(1), 98-106. https://doi.org/10.3141/2237-11

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

Van Houten, R., Louis Malenfant, J., & McCusker, D. (2001). Advance Yield Markings: Reducing Motor Vehicle—Pedestrian Conflicts at Multilane Crosswalks with Uncontrolled Approach. *Transportation Research Record*, *1773*(1), 69-74.

Vanlaar, W., Mainegra Hing, M., Brown, S., McAteer, H., Crain, J., & McFaull, S. (2016). Fatal and serious injuries related to vulnerable road users in Canada. *Journal of safety research*, *58*, 67-77. https://doi.org/10.1016/j.jsr.2016.07.001

Wali, B., & D. Frank, L. (2024). Redefining walkability to capture safety: Investing in pedestrian, bike, and street level design features to make it safe to walk and bike *Transportation Research Part A 181*. https://doi.org/https://doi.org/10.1016/j.tra.2024.103968

Wang, K., Zhang, W., Jin, L., Feng, Z., Zhu, D., Cong, H., & Yu, H. (2022). Diagnostic analysis of environmental factors affecting the severity of traffic crashes: From the perspective of pedestrian–vehicle and vehicle–vehicle collisions. *Traffic Injury Prevention*, *23*(1), 17-22.

WHO. (2023). *New WHO manual aims to strengthen pedestrian safety*. World Health Organization. Retrieved July 5, 2025 from https://www.who.int/news/item/03-05-2023-new-who-manual-aims-to-strengthen-pedestrian-safety

Wu, Y., Abdel-Aty, M., & Lee, J. (2018). Crash risk analysis during fog conditions using real-time traffic data. *Accident Analysis & Prevention*, *114*, 4-11.

Yu, C.-Y. (2015). Built environmental designs in promoting pedestrian safety. *Sustainability*, *7*(7), 9444-9460.

Yu, C.-Y. (2024). The impact of built environments on pedestrian safety around bus stops. *Journal of Urban Affairs*, *46*(2), 267-280. https://doi.org/10.1080/07352166.2022.2049447

Yue, L., Abdel-Aty, M., Wu, Y., Zheng, O., & Yuan, J. (2020). In-depth approach for identifying crash causation patterns and its implications for pedestrian crash prevention. *Journal of safety research*, *73*, 119-132. https://doi.org/10.1016/j.jsr.2020.02.020

Zahabi, S. A. H., Strauss, J., Manaugh, K., & Miranda-Moreno, L. F. (2011). Estimating Potential Effect of Speed Limits, Built Environment, and Other Factors on Severity of

Pedestrian and Cyclist Injuries in Crashes. *Transport Res Rec.*(2247), 81-90. https://doi.org/https://doi.org/10.3141/2247-10.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of statistical software*, *27*, 1-25.

Appendix A

Crash severity was re-estimated using the complete 5-level KABCO classification without merging the categories: K (fatal), A (serious injury), B (minor injury), C (possible injury), and O (no injury). Bayesian models with both informative and non-informative (flat) priors were applied to this expanded classification. Access burden is also included in these models. Table 13 also presents the prior distribution derived from 2018 data, while Table 12 summarizes the results for the 2019–2022 period.

The findings show that results are generally consistent with those from the earlier three-category model. Model fit is slightly improved when using informative priors. However, the three-category model demonstrated better overall performance based on summary statistics; for example, the five-category model had a higher LOOIC. Moreover, while the access burden variable was included in the five-category model, none of the burden indicators were statistically significant. In contrast, the three-category model (which excluded access burden) identified two significant burden predictors. These differences suggest that the three-category model provides a better overall fit and more stable estimates, supporting its use in the final analysis.

Table 14 further explores the five-category model under a simplified crash type grouping to assess the consistency of results. In this version, only seven specific crash types are retained, with "Crash Type 2: Crossing Driveway or Alley and Expressway" serving as the base category. Other categories include Crash Type 3 (Crossing Roadway – Vehicle Turning/Not Turning), Crash Type 4 (Pedestrian in Roadway – Circumstances Unknown, Working or Playing), Crash Type 5 (Dash/Dart-Out), Crash Type 6 (Walking Along Roadway), and Crash Type 10 (Bus-Related) as well as all remaining types that are grouped under "Other." Additionally, pedestrian position is categorized into two groups: travel lane vs. non-travel lane areas, and intersection type is reduced to two broad categories to streamline model interpretation. Table 15 also shows the prior distribution derived from 2018 data for this model. The results align closely with the findings reported in the main analysis.

Table 12. Bayesian Ordinal Model Results (2019–2022, N = 8,494 Crashes, KABCO)

Variables	Informative prior			Non-informative prior (flat)		
	Estimate (mean)	l-95% CI (2.5%)	u-95% CI (97.5%)	Estimate (mean)	l-95% CI (2.5%)	u-95% CI (97.5%)
Intercept [1]	-2.04	-2.31	-1.75	-1.68	-2.09	-1.26
Intercept [2]	0.09	-0.18	0.36	0.46	0.03	0.87
Intercept [3]	1.95	1.69	2.22	2.33	1.91	2.74
Intercept [4]	3.11	2.84	3.39	3.49	3.07	3.90
Pedestrian Position (Base = Travel Lane)						
Ped Position 2: Sidewalk / Shared Use Path / Driveway Crossing	-0.33*	-0.43	-0.22	-0.41*	-0.53	-0.29
Ped Position 3: Paved Shoulder / Bike Lane / Parking Lane	-0.18	-0.39	0.02	-0.05	-0.28	0.18
Ped Position 4: Unpaved Right-of-Way	-0.20	-0.45	0.04	-0.19	-0.50	0.12
Crash Type (Base = Backing)						
Crash Type 2: Crossing Driveway or Alley and Expressway	0.29*	0.03	0.56	0.80*	0.40	1.19
Crash Type 3: Crossing Roadway - Vehicle Turning/Not Turning	0.49*	0.32	0.65	0.85*	0.51	1.18
Crash Type 4: Pedestrian in Roadway - Circumstances Unknown, Working or Playing in Roadway	0.52*	0.33	0.69	0.88*	0.52	1.23
Crash Type 5: Dash / Dart-Out	0.74*	0.55	0.94	1.12*	0.76	1.46
Crash Type 6: Walking Along Roadway	0.15	-0.04	0.33	0.49*	0.15	0.84
Crash Type 7: Unusual Circumstances	0.39*	0.20	0.57	0.73*	0.38	1.08

Crash Type 8: Others	0.15	-0.09	0.39	0.54*	0.16	0.92
Crash Type 9: Waiting to Cross	0.43	-0.12	0.97	0.68	-0.04	1.38
Crash Type 10: Bus-Related	-0.01	-0.44	0.43	0.15	-0.44	0.74
Speed Limit (Under 50 mph)						
Over 50 mph	0.76*	0.63	0.89	0.72*	0.58	0.85
Unknown	-0.14*	-0.29	0.00	-0.14	-0.29	0.00
Lighting condition (Base = Dark)						
Dusk or Dawn	-0.42*	-0.59	-0.25	-0.42*	-0.61	-0.23
Daylight	-0.52*	-0.60	-0.45	-0.55*	-0.64	-0.46
Road Character (Base = Straight)						
Curve - Level	0.43*	0.03	0.82	0.64	-1.02	2.35
Unknown	-0.41	-0.89	0.06	-0.53	-1.22	0.13
Road Configuration (Base = One-Way, Not Divided)						
Two-Way, Divided, Positive Median Barrier	0.58*	0.40	0.75	0.65*	0.41	0.90
Two-Way, Divided, Unprotected Median	0.33*	0.17	0.50	0.40*	0.17	0.63
Two-Way, Not Divided	0.06	-0.10	0.21	0.12	-0.09	0.35
Others	0.33	-0.14	0.80	0.53	-0.12	1.21
Weather (Base = Clear)						
Cloudy	0.05	-0.11	0.21	0.04	-0.08	0.17
Rain, Sleet, Hail, Freezing Rain/Drizzle	-0.09	-0.21	0.03	-0.12	-0.26	0.02
Snow	-0.31	-0.90	0.29	-0.40	-1.42	0.59
Fog, Smog, Smoke	0.72*	0.23	1.20	0.78*	0.21	1.34
Burden Indicators						
Health Burden	-0.03	-0.12	0.07	-0.03	-0.12	0.07
Environmental Burden	-0.07	-0.17	0.03	-0.07	-0.17	0.03

Weather Burden	0.08	-0.03	0.19	0.07	-0.04	0.18
Cost Burden	-0.04	-0.15	0.08	-0.04	-0.17	0.08
Access Burden	0.06	-0.03	0.15	0.06	-0.03	0.14
Location Context						
Rural (Base = Urban)	0.47*	0.35	0.59	0.54	0.39	0.68
log (Street Network Density (miles^2))	-0.13*	-0.18	-0.07	-0.11*	-0.17	-0.04
Non-intersection (Base = Intersection)	0.20*	0.11	0.28	0.11*	0.01	0.21
Interaction Variables						
Curve - Level * Crash Type 2	0.75	-0.19	1.73	0.65	-1.37	2.70
Curve - Level * Crash Type 3	0.12	-0.36	0.62	-0.15	-1.92	1.54
Curve - Level * Crash Type 4	-0.82*	-1.36	-0.27	-1.05	-2.81	0.65
Curve - Level * Crash Type 5	0.10	-0.58	0.78	-0.01	-1.88	1.75
Curve - Level * Crash Type 6	-0.01	-0.49	0.46	-0.25	-1.97	1.44
Curve - Level * Crash Type 7	-0.05	-0.54	0.45	-0.15	-1.92	1.55
Curve - Level * Crash Type 8	-0.23	-1.02	0.55	-0.76	-2.68	1.06
Curve - Level * Crash Type 9	0.87	-0.70	2.41	2.42	-0.87	6.08
Curve - Level * Crash Type 10	0.52	-0.83	1.88	0.95	-1.63	3.47
Ped Position 2* Rural	0.24	-0.10	0.59	0.44*	0.01	0.87
Ped Position 3 * Rural	-0.62*	-0.90	-0.33	-0.79*	-1.12	-0.46
Ped Position 4 * Rural	-0.23	-0.59	0.12	-0.10	-0.55	0.36
People-involved Features						
Pedestrian Age	0.01*	-0.01	0.00	0.00	-0.01	0.00

Driver Age	-0.01*	-0.01	0.00	-0.01*	-0.01	0.00
Drive DUI	0.01*	0.01	0.01	0.01*	0.01	0.02
Pedestrian Alcohol Flag	0.35*	0.24	0.46	0.33*	0.21	0.46
Summary Statistics						
Model (Priors from the 2018 data distribution): elpd_loo = –11510.3 (SE 57.0) looic = 23020.5 (SE 113.9) p_loo = 31.9 (SE 0.4)			Model (Non-informative prior (flat)): elpd_loo = –11512.5 (SE 57.3) looic = 23025.5 (SE 114.6) p_loo = 32.1 (SE 0.5)			
N = 8,494 Crashes Five Injury categories: O, C, B, A, K l-95% CI (2.5%): Lower bound of 95% CI (2.5th percentile) u-95% CI (97.5%): Upper bound of 95% CI (97.5th percentile) *Statistically Significant (5%)						

Table 13. Prior Parameter Distributions (2019–2022 Model, N = 2,424 Crashes, KABCO)

Variables	Distribution from 2018	
	Mean	SD
Intercept [1]	0.00	10.00
Intercept [2]	0.00	10.00
Intercept [3]	0.00	10.00
Intercept [4]	0.00	10.00
Ped Position 2: Sidewalk / Shared Use Path / Driveway Crossing	-0.18	0.11
Ped Position 3: Paved Shoulder / Bike Lane / Parking Lane	-0.52	0.26
Ped Position 4: Unpaved Right-of-Way	-0.46	0.24
Crash Type 2: Crossing Driveway or Alley and Expressway	-0.42	0.30
Crash Type 3: Crossing Roadway - Vehicle Turning/Not Turning	0.29	0.22
Crash Type 4: Pedestrian in Roadway - Circumstances Unknown, Working or Playing in Roadway	0.33	0.24
Crash Type 5: Dash / Dart-Out	0.63	0.24
Crash Type 6: Walking Along Roadway	0.18	0.23
Crash Type 7: Unusual Circumstances	0.43	0.23
Crash Type 8: Others	0.10	0.27
Crash Type 9: Waiting to Cross	0.42	0.65
Crash Type 10: Bus-Related	0.50	0.43
Over 50 mph	0.87	0.14
Unknown	-0.32	0.28
Dusk or Dawn	-0.46	0.19
Daylight	-0.43	0.09
Curve - Level	0.41	0.40
Unknown	-0.52	0.46
Two-Way, Divided, Positive Median Barrier	0.43	0.20
Two-Way, Divided, Unprotected Median	0.28	0.19
Two-Way, Not Divided	0.06	0.17
Others	0.01	0.46

Rural (Base = Urban)	0.29	0.13
Cloudy	0.19	0.11
Rain, Sleet, Hail, Freezing Rain/Drizzle	-0.01	0.12
Snow	-0.23	0.40
Fog, Smog, Smoke	0.49	0.51
Non-intersection	0.44	0.09
Health Burden	0.00	10.00
Environmental Burden	0.00	10.00
Weather Burden	0.00	10.00
Cost Burden	0.00	10.00
Access Burden	0.00	10.00
log (Street Network Density (miles^2))	-0.19	0.06
Curve - Level * Crash Type 2	0.26	0.81
Curve - Level * Crash Type 3	0.49	0.52
Curve - Level * Crash Type 4	-0.64	0.61
Curve - Level * Crash Type 5	-0.22	0.58
Curve - Level * Crash Type 6	0.26	0.47
Curve - Level * Crash Type 7	-0.58	0.49
Curve - Level * Crash Type 8	0.93	0.77
Curve - Level * Crash Type 9	0.00	0.98
Curve - Level * Crash Type 10	0.00	0.99
Ped Position 2* Rural	-0.42	0.38
Ped Position 3 * Rural	-0.31	0.36
Ped Position 4 * Rural	-0.79	0.32
Pedestrian Age	0.26	0.81
Driver Age	-0.01	0.00
Drive DUI	1.06	0.24
Pedestrian Alcohol Flag	0.01	0.00

N: 2,424 Crashes
Five Injury Severity Categories: K, A, B, C, O

Table 14. Bayesian Ordinal Model Results (N = 8,494 Crashes, KABCO, With Adjusted Variables)

Variables	Informative prior			Non-informative prior (flat)		
	Estimate (mean)	l-95% CI (2.5%)	u-95% CI (97.5%)	Estimate (mean)	l-95% CI (2.5%)	u-95% CI (97.5%)
Intercept [1]	-2.14	-2.42	-1.87	-2.48	-2.87	-2.10
Intercept [2]	-0.02	-0.29	0.25	-0.36	-0.72	0.02
Intercept [3]	1.84	1.56	2.11	1.51	1.14	1.88
Intercept [4]	2.99	2.72	3.27	2.66	2.29	3.04
Pedestrian Position (Base = Non-Travel Lane Areas[1])						
Ped Position 2: Travel Lane	0.28*	0.20	0.37	0.33	0.23	0.44
Crash Type (Base = Crash Type 3: Crossing Driveway or Alley and Expressway)						
Crash Type 3: Crossing Roadway - Vehicle Turning/Not Turning	0.37*	0.21	0.53	0.01	-0.25	0.26
Crash Type 4: Pedestrian in Roadway - Circumstances Unknown, Working or Playing in Roadway	0.41*	0.24	0.60	0.06	-0.21	0.32
Crash Type 5: Dash / Dart-Out	0.65*	0.46	0.84	0.29*	0.01	0.56
Crash Type 6: Walking Along Roadway	0.04	-0.13	0.21	-0.33*	-0.60	-0.06
Crash Type 10: Bus-Related	-0.04	-0.47	0.38	-0.60*	-1.12	-0.05
Crash Type: Others[2]	0.18*	0.01	0.34	-0.18	-0.43	0.07
Speed Limit (Under 50 mph)						
Over 50 mph	0.76*	0.64	0.88	0.71*	0.57	0.85
Unknown	-0.16*	-0.30	-0.02	-0.16*	-0.31	-0.02
Lighting condition (Base = Dark)						
Dusk or Dawn	-0.42*	-0.59	-0.25	-0.42*	-0.62	-0.23
Daylight	-0.52*	-0.60	-0.44	-0.56*	-0.65	-0.47

Road Character (Base = Straight)						
Curve - Level	0.48*	0.31	0.65	0.50*	0.30	0.69
Unknown	-0.40	-0.87	0.07	-0.45	-1.11	0.20
Road Configuration (Base = One-Way, Not Divided)						
Two-Way, Divided, Positive Median Barrier	0.61*	0.43	0.79	0.68*	0.44	0.93
Two-Way, Divided, Unprotected Median	0.34*	0.17	0.50	0.43*	0.19	0.65
Two-Way, Not Divided	0.06	-0.09	0.22	0.14	-0.08	0.37
Others	0.34	-0.15	0.84	0.50	-0.16	1.15
Weather (Base = Clear)						
Cloudy	0.09	-0.03	0.20	0.47	-0.08	0.17
Rain, Sleet, Hail, Freezing Rain/Drizzle	-0.08	-0.21	0.04	-0.11	-0.25	0.03
Snow	-0.39	-1.02	0.22	-0.46	-1.44	0.52
Fog, Smog, Smoke	0.71*	0.22	1.21	0.78*	0.20	1.36
Burden Indicators						
Health Burden	-0.03	-0.12	0.07	-0.02	-0.11	0.07
Environmental Burden	-0.07	-0.17	0.03	-0.07	-0.19	0.04
Weather Burden	0.08	-0.03	0.18	0.07	-0.04	0.18
Cost Burden	0.06	-0.02	0.14	0.06	-0.02	0.14
Access Burden	-0.06	-0.18	0.06	-0.06	-0.17	0.06
Location Context						
Rural (Base = Urban)	0.49*	0.37	0.61	0.54*	0.40	0.68
log (Street Network Density (miles^2))	-0.13*	-0.18	-0.07	-0.11*	-0.17	-0.04
Non-intersection (Base = Intersection)	0.20*	0.12	0.28	0.11*	0.02	0.21
Interaction Variables						
Curve - Level * Crash Type 4	-0.87*	-1.30	-0.43	-0.91*	-1.40	-0.41

Travel lane * Rural	0.32*	0.15	0.47	0.21	-0.01	0.42
People-involved Features						
Driver Age	-0.01*	-0.01	0.00	-0.01*	-0.01	0.00
Pedestrian Age	0.01*	0.01	0.01	0.01*	0.01	0.01
Drive DUI	0.92*	0.72	1.13	0.89*	0.66	1.13
Pedestrian Alcohol Flag	0.35*	0.23	0.45	0.32*	0.20	0.45

Summary Statistics

Model (Priors from the 2018 data distribution):

elpd_loo = –11524.6 (SE 59)

looic = 23049.2 (SE: 113.9)

p_loo = 33.3 (SE 0.5)

Model (Non-informative prior (flat)):

elpd_loo = –11525.6 (SE 56.9)

looic = 23051.2 (SE: 113.7)

p_loo = 41.0 (SE 0.8)

N: 8,494 Crashes
Five Injury Severity Categories: K, A, B, C, O
1 Sidewalk / Shared Use Path / Driveway Crossing/ Paved Shoulder / Bike Lane / Parking Lan/ Unpaved Right-of-Way
2 Backing, Unusual Circumstances, Waiting to Cross, Others

l-95% CI (2.5%): Lower bound of 95% CI (2.5th percentile)

u-95% CI (97.5%): Upper bound of 95% CI (97.5th percentile)
*Statistically Significant (5%)

Table 15. Prior Parameter Distributions (Adjusted Model, 2019–2022, N = 2,424 Crashes, KABCO)

Variables	Distribution from 2018	
	Mean	SD
Intercept [1]	0.00	10.00
Intercept [2]	0.00	10.00
Intercept [3]	0.00	10.00
Intercept [4]	0.00	10.00
Ped Position 2: Travel Lane Areas	0.26	0.10
Crash Type 3: Crossing Roadway - Vehicle Turning/Not Turning	0.50	0.21
Crash Type 4: Pedestrian in Roadway - Circumstances Unknown, Working or Playing in Roadway	0.49	0.24
Crash Type 5: Dash / Dart-Out	0.78	0.24
Crash Type 6: Walking Along Roadway	0.32	0.22
Crash Type 10: Bus-Related	0.64	0.43
Crash Type: Others[1]	0.45	0.22
Over 50 mph	0.85	0.13
Unknown	-0.37	0.28
Dusk or Dawn	-0.45	0.18
Daylight	-0.43	0.08
Curve – Level	0.48	0.18
Unknown	-0.53	0.45
Two-Way, Divided, Positive Median Barrier	0.44	0.20
Two-Way, Divided, Unprotected Median	0.29	0.19
Two-Way, Not Divided	0.07	0.17
Others	0.01	0.45
Rural (Base = Urban)	0.33	0.12
Cloudy	0.20	0.11
Rain, Sleet, Hail, Freezing Rain/Drizzle	0.00	0.13
Snow	-0.33	0.38
Fog, Smog, Smoke	0.48	0.49

Non-intersection	0.39	0.09
Health Burden	0.00	10.00
Environmental Burden	0.00	10.00
Weather Burden	0.00	10.00
Access Burden	0.00	10.00
Cost Burden	0.00	10.00
log (Street Network Density (miles^2))	-0.16	0.05
Curve - Level * Crash Type 4	-0.67	0.54
Travel Lane * Rural	0.71	0.20
Pedestrian Age	0.01	0.00
Driver Age	-0.01	0.00
Drive DUI	1.05	0.23
Pedestrian Alcohol Flag	0.38	0.12

N: 2,424 Crashes
Five Injury Severity Categories: K, A, B, C, O

1 Backing, Unusual Circumstances, Waiting to Cross, Others