RESEARCH

Keeping Utah Moving

Utah Department of Transportation - Research Division
4501 South 2700 West - P.O. Box 148410 - SLC,UT 84114-8410

Report No. UT-21.08

SAFETY IN NUMBERS? DEVELOPING IMPROVED SAFETY PREDICTIVE METHODS FOR PEDESTRIAN CRASHES AT SIGNALIZED INTERSECTIONS IN UTAH USING PUSH BUTTON-BASED MEASURES OF EXPOSURE

Prepared For:

Utah Department of Transportation
Research & Innovation Division

Final Report
May 2021

DISCLAIMER

The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind, and assumes no liability therefore.

Crash data are protected under 23 USC 409.

ACKNOWLEDGMENTS

The authors acknowledge the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Robert Chamberlin (consultant)
- Travis Evans
- Heidi Goedhart
- Andrea Guevara
- Peter Tang
- Mark Taylor

The authors also thank the following students and faculty at Utah State University (USU): Caleb Child and Devan Peterson for their help collecting the intersection and road network characteristics data; Doo Hong Lee and Dr. Keunhyun Park for their help assembling land use and sociodemographic data; and Seth Thompson for suggesting and pilot-testing this research idea in Patrick Singleton's CEE 5255 Transportation Data and Safety Analysis course in Fall 2018.

TECHNICAL REPORT ABSTRACT

1. Report No. UT- 21.08	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A
4. Title and Subtitle Safety in Numbers? Developing Improved Safety Predictive Methods for Pedestrian Crashes at Signalized Intersections in Utah Using Push Button-Based Measures of Exposure		5. Report Date May 2021 6. Performing Organization Code
7. Author(s) Patrick Singleton, Michelle Mekker, Ahadul Islam		8. Performing Organization Report No.
9. Performing Organization Name and Address Utah State University Department of Civil and Environmental Engineering 4110 Old Main Hill Logan, UT 84322-4110		10. Work Unit No. 7009 XEC 6137 5537901D 11. Contract or Grant No. 208179
12. Sponsoring Agency Name and Address Utah Department of Transportation 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410		13. Type of Report & Period Covered Final Aug 2019 to May 2021 14. Sponsoring Agency Code UT19.03.16

15. Supplementary Notes

Prepared in cooperation with the Utah Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration

16. Abstract

The focus of this study was threefold: (1) to estimate models of pedestrian crash frequency at signalized intersections; (2) to develop safety performance functions and crash modification factors for better interpretation and application of models' results; and (3) to examine whether the "safety in numbers" effect applies to pedestrian safety in the US. Specifically, the analysis used robust measures of pedestrian exposure: pedestrian crossing volumes estimated from one year of pedestrian push-button traffic signal data. Multiple negative binomial models – predicting 10-year counts of pedestrian crashes at nearly 1,606 signals in Utah – were estimated, to account for different levels of data availability and different needs for applying the models' results. The models showed almost similar results, indicating that signals with the following characteristics saw more pedestrian crashes:

- Higher volumes of pedestrian and motor vehicle traffic,
- Longer average crossing distances,
- Fewer approaches with pedestrian crossing restrictions,
- More crosswalks with high-visibility longitudinal markings instead of standard transverse markings,
- No prohibitions of right-turns-on-red,
- No bike lanes,
- More bus stops (and more far-side as compared to near-side bus stops),
- Greater shares of vacant land uses,
- Less employment density,
- No schools or places of worship, and
- Greater shares of people with a disability or people of Hispanic or non-White race/ethnicity – saw more pedestrian crashes.

The study also found strong support for the "safety in numbers" effect, in which pedestrian-vehicle crash rates decrease with an increase in pedestrian volumes. The authors suggest potential countermeasures, recommend various actions, and discuss future research opportunities for improving pedestrian safety at signalized intersections.

17. Key Words Pedestrian, Exposure, Crash frequency, Safety in numbers, Intersections.		18. Distribution Statement Not restricted. Available through: UDOT Research Division 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410 www.udot.utah.gov/go/research		23. Registrant's Seal N/A
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 88	22. Price N/A	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

UNIT CONVERSION FACTORS

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)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		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)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		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

*SI is the symbol for the International System of Units. (Adapted from FHWA report template, Revised March 2003)

LIST OF ACRONYMS

AADP	Average Annual Daily Pedestrians
AADT	Average Annual Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
ACS	American Community Survey
AGRC	Automated Geographic Reference Center
ATSPM	Automated Traffic Signal Performance Measures
CMF	Crash Modification Factor
FHWA	Federal Highway Administration
HSM	Highway Safety Manual
LEHD	Longitudinal Employer Household Dynamics
MUTCD	Manual on Uniform Traffic Control Devices
NB	Negative Binomial
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
SPF	Safety Performance Function
TRB	Transportation Research Board
UDPS	Utah Department of Public Safety
UDOT	Utah Department of Transportation
USDOT	United States Department of Transportation
ZINB	Zero Inflated Negative Binomial

EXECUTIVE SUMMARY

Pedestrian safety is a critical transportation issue. A person walking is at a higher risk of death or injury than other road users, if hit by a vehicle. Data reveals that there has been an increase in the number and shares of pedestrian injuries and fatalities both nationally and in Utah. These troubling trends necessitate studies to develop improved pedestrian crash prediction methods to better understand factors associated with pedestrian crashes and help in the selection and prioritization of countermeasures to improve pedestrian safety. Specifically, crash frequency models, safety performance functions (SPFs), and crash modification factors (CMFs) can greatly benefit from the inclusion of robust pedestrian exposure measures. The overall goal of this project was to identify factors associated with pedestrian safety and pedestrian crash frequency at signalized intersections in Utah. To achieve this goal, this project had three objectives: first, to estimate models of pedestrian crash frequency at signalized intersections using pedestrian volumes, motor vehicle traffic volumes, and other predictor variables relating to road network attributes, land use and built environment factors, and sociodemographic characteristics; second, to develop SPFs and CMFs from these models for better interpretation and application of results; and third, to examine the "safety in numbers" hypothesis, which suggests that pedestrian crash rates decrease with increasing volumes of people walking. One notable feature of this study was the inclusion of a measure of pedestrian exposure estimated from archived pedestrian push-button traffic signal data.

First, the potential factors correlated with pedestrian crashes were identified from a thorough literature review. Research has found links between pedestrian safety and pedestrian/vehicle exposure, built environment characteristics, and neighborhood sociodemographics. The literature review also included a brief discussion on the "safety in numbers" hypothesis and previously adopted methods to test the hypothesis. Unfortunately, research on pedestrian safety at intersections has been limited by the lack of pedestrian volume data as a measure of exposure.

Second, data relating to pedestrian crashes, pedestrian volumes, and other relevant factors – transportation characteristics, land use and built environment, and sociodemographics – were collected and assembled from different sources for 1,606 signalized intersections in Utah. In total, 2,597 pedestrian crashes that occurred at those signals from 2010 through 2019 were assembled from UDOT's (Utah Department of Transportation) Numetric website. Other data came from existing Utah geospatial databases, the U.S. Census Bureau, or were calculated using Google Maps and Streetview. A notable contribution of this research was the inclusion of unique pedestrian exposure data – specifically, annual average daily pedestrian (AADP) crossing volumes estimated from pedestrian push-button information – obtained from traffic signal controller logs archived in UDOT's Automated Traffic Signal Performance Measures (ATSPM) system.

Third, for crash frequency analysis, data collected/assembled from the signals were used to estimate a series of count data models (Poisson, negative binomial [NB], and zero-inflated negative binomial [ZINB] models) for different levels of data availability. These models were then tested for overdispersion and zero-inflation; ZINB models showed the best fit for the data sample. NB models with only transportation characteristics data were also estimated in order to develop SPFs and CMFs for application in safety predictive methods. Results in the form of SPFs and CMFs were presented in an actionable form for UDOT. The model results also confirmed the "safety in numbers" hypothesis for pedestrian safety.

Overall, the research provided several key findings and recommendations. Pedestrian crashes occurred more frequently at signalized intersections and in areas with the following characteristics:

- Higher volumes of pedestrian and motor vehicle traffic,
- Longer average crossing distances,
- Fewer approaches with pedestrian crossing restrictions,
- More crosswalks with high-visibility longitudinal markings instead of standard transverse markings,
- No prohibitions of right-turns-on-red,
- No bike lanes,
- More bus stops (and more far-side as compared to near-side bus stops),
- Greater shares of vacant land uses,
- Less employment density,

- Greater intersection density,
- No schools or places of worship, and
- Greater shares of people with a disability or people of Hispanic or non-White race/ethnicity.

Based on these results, UDOT should consider the following actions to improve pedestrian safety at signalized intersections:

- Shorten pedestrian crossing distances,
- Implement complete streets treatments,
- Prohibit right-turns-on-red in some cases,
- Continue efforts in school zones,
- Focus pedestrian safety treatments in at-risk communities, and
- Encourage walking.

Finally, this study found strong support for a "safety in numbers" effect for walking: Pedestrian crash rates decrease with increased pedestrian volumes.

1.0 INTRODUCTION

1.1 Problem Statement

The primary motivation for this research is the troubling trend of *increasing numbers and shares of pedestrian injuries and fatalities*, both nationally and in Utah. According to the National Highway Traffic Safety Administration, there were nearly 6,300 pedestrian deaths in traffic crashes in the US in 2018, representing about 17% of all traffic fatalities (NHTSA, 2019b). This was an increase from 4,700 and 11% in 2007 (NHTSA, 2019a). Utah is not immune to this issue and has also seen increases in the number and rate of pedestrian fatalities and injuries over the last 10 years. In 2019, 45 deaths and nearly 900 crashes involving people walking on Utah streets and highways were reported (UDPS, 2020). As vulnerable road users, pedestrians are more likely than other road users to be injured or killed when involved in a collision.

Given these trends, there is a need for *improved pedestrian crash prediction methods* to better understand factors (i.e., geometric, traffic, operational, and other) associated with pedestrian safety, and also to assist in the prioritization and selection of countermeasures to improve pedestrian safety at signalized intersections. Specifically, safety predictive methods – safety performance functions (SPFs) and crash modification factors (CMFs) – traditionally require the use of exposure data for estimation and application. While motor vehicle volumes are often available, pedestrian volumes rarely are, thus limiting the development, use, and accuracy of pedestrian safety predictive methods. SPFs and CMFs could greatly benefit from the inclusion of more robust data on pedestrian exposure, which is typically the biggest barrier to overcome for pedestrian safety analysis.

A secondary motivation for this work is to *examine the "safety in numbers" hypothesis* for walking. The "safety in numbers" hypothesis for walking has been considered over the last three decades. This concept suggests that pedestrian (and bicycle) crash rates decrease with increasing volumes of people walking (and bicycling). Although research has yet to clearly identify the specific causes of this observed relationship, it is assumed that the more often drivers see pedestrians and bicyclists, the more likely they are to anticipate them and have more experience driving safely around them. As with safety predictive methods, the challenge with studying the "safety in numbers" concept is the lack of pedestrian exposure data. Most research on the topic was conducted with surrogate measures of pedestrian exposure. For example, for the estimation of pedestrian volumes, researchers have: taken a "Space Syntax" modeling approach relying on street network characteristics (Raford & Ragland, 2006; Geyer et al., 2006); used travel survey data (Xu et al., 2019; Jacobsen, 2015); or generated random numbers (Elvik, 2013). An authentic dataset on pedestrian exposure would provide more reliable information for understanding whether the "safety in numbers" concept applies to pedestrian safety, knowledge that could promote more walking and bicycling through policy and planning.

This research project addressed both of these motivations – improved safety methods and the "safety in numbers" concept – by incorporating new measures of pedestrian exposure into pedestrian safety predictive methods at signalized intersections. A previous UDOT research project (Singleton et al., 2020) utilized archived pedestrian traffic signal data as a proxy for walking activity at signalized intersections and developed methods for converting pedestrian push-button actuations into pedestrian volumes. The project described in this report utilized archived traffic signal data, and pedestrian-involved crash data to develop Utah-specific SPFs and CMFs for pedestrian-vehicle collisions at signalized intersections. These locally calibrated models and methods can be incorporated into UDOT's safety performance management processes.

1.2 Objectives

The primary objective of this research project was to develop improved pedestrian crash prediction models (SPFs and CMFs) at signalized intersections using pedestrian push-button measures of exposure. A secondary objective of this research project was to test the "safety in numbers" concept for walking in a US context. Overall, these two objectives contributed to the larger goal of understanding factors affecting pedestrian-vehicle crashes at signalized intersections, which suggested recommendations to improve pedestrian safety.

1.3 Scope

This project accomplished these research objectives through the following major tasks:

- Reviewing literature on: pedestrian crash risk factors, associations between motor vehicle and pedestrian traffic volumes and pedestrian crash frequency, modeling techniques adopted for pedestrian crash frequency analysis, and the "safety in numbers" concept for walking.
- Selecting study locations: signalized intersections in Utah.
- Assembling pedestrian crash data for the study locations from existing Utah crash databases.
- Collecting data on intersection and road network characteristics – including information about pedestrian crossing distances, crosswalk marking types, and the presence of bike lanes and bus stops near signalized intersections – from aerial and street-level imagery.
- Assembling other information about study locations – including information about motor vehicle traffic volumes, transportation system characteristics, land use and built environment data, and sociodemographic characteristics – from existing UDOT, Utah, and US Census data sources.
- Calculating measures of pedestrian exposure (estimates of pedestrian volumes) at study locations, by applying the factoring methods developed in a previous UDOT project to archived pedestrian push-button data from traffic signal controllers.
- Performing crash data modeling – following best-practice guidelines – to generate SPFs and CMFs for pedestrian crashes at signalized intersections, including measures of pedestrian exposure to test the "safety in numbers" hypothesis.
- Providing recommendations regarding implementable actions and potentially effective countermeasures to improve pedestrian safety at signalized intersections.

1.4 Outline of Report

This report is organized into the following chapters:

- Chapter 1.0 includes an introduction to the research, project objectives, project scope, and the organization of the report.

- Chapter 2.0 includes a literature review of studies investigating factors associated with pedestrian crashes, the “safety-in-numbers” phenomenon, and notes on limitations in earlier studies.
- Chapter 3.0 includes details on the study locations, the data collection and assembly processes, and the types of data collected.
- Chapter 4.0 includes a descriptive and correlative analysis of pedestrian crashes and rates, a summary of crash frequency models and safety predictive methods, results from and comparisons between multiple crash frequency models that account for different levels of data availability and different needs for applying model results, and interpretation of those models into the development of SPFs and CMFs.
- Chapter 5.0 summarizes the report by highlighting the major findings, comparing those findings with earlier research, noting limitations, and outlining potential steps for future work.
- Chapter 6.0 provides recommendations for implementation of the research findings.
- References follow the main chapters.

2.0 LITERATURE REVIEW

2.1 Overview

First, factors studied in context of pedestrian crash frequency and injury severity were evaluated to select important risk factors in pedestrian crashes to be analyzed by the research team. In this chapter, key factors associated with pedestrian safety (based on previous literature) are first organized into categories, including: traffic exposure, built environment characteristics, sociodemographic characteristics, site-specific characteristics, and other spatial variables. Second, earlier studies investigating the suitability of the “safety in numbers” concept with respect to pedestrian crashes are explored. This literature review then concludes with notes about the limitations of previous research, as well as a summary of key findings. A knowledge of past research on pedestrian crash risk factors enabled the research team to select a set of appropriate explanatory variables required for data collection and analysis so that this project could build upon previous findings as well as address limitations and knowledge gaps on these topics.

2.2 Factors Affecting Pedestrian Crash Frequency

For the improvement of pedestrian safety at intersections, a detailed exploration of crash-related factors is required in order to develop effective countermeasures (Lee and Abdel-Aty, 2005; Stutts, Hunter, and Pein, 1996). Factors studied in the past regarding pedestrian crashes include traffic exposure, built environment characteristics, sociodemographic characteristics, site-specific characteristics, and other spatial variables, as summarized in the following paragraphs.

Exposure, an important predictor of crash frequencies, is typically operationalized using average volumes of motorized and/or non-motorized traffic. Increased volumes of vehicles or pedestrians at an intersection increase the chances of conflicts and hence the probability of vehicle-pedestrian collisions. Several studies found positive associations between vehicle volume and pedestrian crashes (El-Basyouny and Sayed, 2013; Cottrill and Thakuriah, 2010). El-Basyouny et al. (2013) applied a log-normal model to data from 51 signalized intersections in British Columbia to predict conflicts using traffic volume and other related variables as covariates. The results indicated a highly significant and positive relationship between traffic volumes and predicted vehicle-pedestrian conflicts: i.e., predicted conflicts were observed to be increasing with traffic volume. Research by Brüde and Larsson (1993) and Zegeer et al. (2005) also found that the number of motor vehicles per day approaching an intersection was a significant and positive predictor of pedestrian crashes.

While many studies investigated the relationships between pedestrian crashes and vehicle volume, only a few studies explored the link with pedestrian volumes due to the difficulty in obtaining such data. When included, the volume of pedestrians was the single-most important variable to explain variations in pedestrian crashes. Zegeer et al. (1985) conducted an analysis of pedestrian crashes with data from 1,297 signalized intersections across 15 US cities. The analysis found that the volume of pedestrians crossing at an intersection was the most important variable to explain pedestrian crashes and had a direct relationship to pedestrian crash occurrence. The number of pedestrian crashes generally increased with an increase in pedestrian volume. Overall, both pedestrian and vehicular traffic exposure show positive associations with pedestrian-vehicle crashes (Harwood et al., 2008; Dumbaugh and Li, 2010).

Built environment characteristics – including population and job density and local land use types – may also be linked to pedestrian crashes. Population density showed a positive association with pedestrian crash occurrence in a few studies (Dumbaugh and Li, 2010; Gladhill and Monsere, 2012). In contrast, Loukaitou-Sideris et al. (2007) and Graham and Glaister (2003) found a negative relationship between population density and pedestrian crashes. They argued that due to the lower vehicular traffic speeds in congested areas of extremely densely populated cities, there is a decrease in expected collision rates. So, results are mixed over the link between population density and pedestrian crashes. Job or employee density was found to be positively associated with pedestrian crashes (Loukaitou-Sideris et al., 2007). Also, increased proportions of land used for commercial, mixed use, park, retail, or community use has been associated with increased vehicle-pedestrian collisions in some studies (Loukaitou-Sideris et al., 2007; Wier et al., 2009). Such neighborhoods are generally lively with greater amounts of street activity and pedestrian crossings; hence, these areas may see increased pedestrian crashes.

Examples of sociodemographic characteristics that may be associated with pedestrian crashes are household income, population by age, race/ethnicity, and number of children, typically measured for residents of the surrounding neighborhood. In several studies, pedestrian crashes have been linked to population demographics such as income, race/ethnicity, and the presence of children in households. Loukaitou-Sideris et al. (2007) investigated the influence of sociodemographic and land use characteristics on pedestrian collisions in Los Angeles. The results from the study supported the assumption that the pedestrian crashes were more likely to occur in low-income, minority neighborhoods, when the other aspects of risk are controlled for. People in low-income and minority neighborhoods may be more exposed to the dangers of motor vehicle traffic, as they are more likely to walk, bike, or use transit (Ernst and McCann, 2002). Children and elderly are more at risk as they take a longer time to cross the road, increasing their exposure to motor vehicle traffic (Demetriades et al., 2004). Particularly, children in low-income neighborhoods with restricted access to playgrounds and higher traffic may be more prone to experiencing pedestrian crashes or injuries (Rivara and Barber, 1985).

Different road and intersection characteristics – including the number of lanes, signal conditions, and lighting conditions – have also been investigated in relation to pedestrian safety. Zegeer et al. (2005) explored five years of pedestrian crashes to understand the safety effects of marked versus unmarked crosswalks. The study found that a greater number of lanes was related to higher pedestrian crash frequency, whereas speed limit, crosswalk marking conditions and crosswalk marking types (e.g., continental, ladder, zebra stripes) had no significant effect on pedestrian crash rates. Lee and Abdel-Aty (2005) analyzed over four years of pedestrian crashes at intersections in Florida and found that pedestrian crash risk was observed to be reduced by the presence of beacons and improved lighting conditions at intersections and roadway segments.

2.3 Safety in Numbers

Although a positive relationship has been found between pedestrian/bicycle crash frequency and measures of exposure (Lindsey et al., 2019), researchers have argued that it is a non-linear relationship. Specifically, they suggest that crash rates – the number of crashes per unit of exposure, or the slope of the crash frequency vs. exposure relationship – actually decline with higher levels of pedestrian/bicycle traffic. This phenomenon is popularly known as the "safety in numbers" concept (Carlson et al., 2018; Jacobsen, 2015; Elvik, 2013). To demonstrate the "safety in numbers" idea, we constructed two conceptual figures with hypothetical data. Figure 2-1 shows a non-linear relationship between pedestrian exposure and pedestrian crash frequency, capturing the positive association found between pedestrian crashes and exposure. Figure 2-2 shows a reduction in pedestrian crash rates with an increase in pedestrian exposure, demonstrating the "safety in numbers" concept.

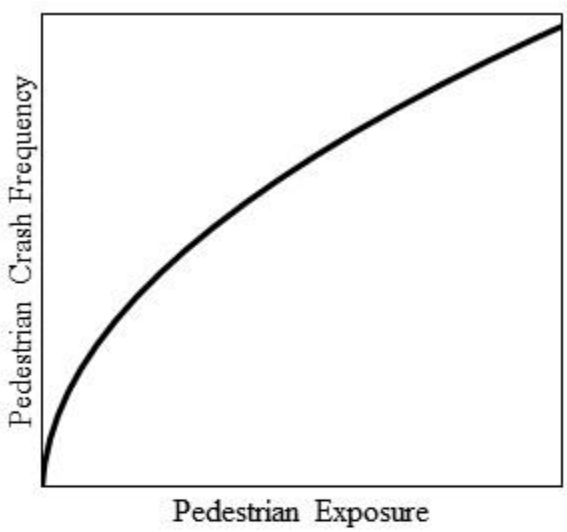


Figure 2-1 Pedestrian crash frequency increases non-linearly with pedestrian exposure

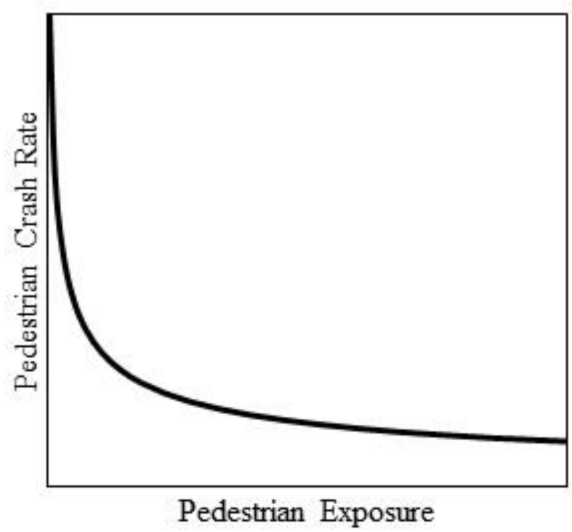


Figure 2-2 The "safety in numbers" concept shows the pedestrian crash rate decreasing with pedestrian exposure

As Elvik (2013) explains, the risk of injury to each pedestrian or cyclist becomes lower with a greater number of pedestrians and cyclists. In a meta-analysis of estimates of the "safety-in-numbers" effect including 45 studies on the topic, Elvik (2019) reported that all studies follow a common form of a multivariate crash prediction model. Although the studies share a common form, the explanatory factors considered in those studies vary considerably. Some models consider only pedestrian/vehicle volumes, while others consider a wide range of variables describing infrastructure, traffic, and/or spatial characteristics. However, the investigation reported that although there is considerable variation in estimates, nearly all studies support a "safety in numbers" effect. It was also found that the "safety in numbers" effect is stronger for pedestrians than for cyclists or motorists, and newer investigations support this concept more than earlier studies.

2.4 Limitations in Earlier Studies

Most research on pedestrian safety has been limited by the unavailability of pedestrian exposure data. Raford and Ragland (2005) note that, while police reports have made pedestrian crash data readily available for many American cities, very few municipalities have arrangements to estimate pedestrian volumes. This is due to the fact that pedestrian routes are numerous and not well defined, and often pedestrian trips are part of larger trips, e.g., walking to the bus stop (Kerridge et al., 2001). Without pedestrian volume counts or estimates, cities are left with an incomplete picture of pedestrian risk. For example, high-volume intersections may face higher pedestrian crashes per year than intersections with low pedestrian volumes. Yet, the high-volume intersection may be relatively safer to use for each pedestrian. In the absence of pedestrian volume data, authorities often end up prioritizing locations with more collisions instead of higher-risk locations (Raford and Ragland, 2005).

Efforts have been made by researchers to overcome this challenge by applying different techniques for estimating average annual pedestrian volumes as a measure of exposure. For example: Raford et al. (2005) and Geyer et al. (2005) used a Space Syntax method to predict pedestrian volumes. This method translates population density and other land use data using a network analysis of pedestrian routes and street network structure for pedestrian volume estimation. Some studies have made use of adjusted short-duration pedestrian flow profiles available from travel survey databases in their analyses (Xu et al., 2019; Jacobsen, 2015). Elvik (2013) used randomly generated numbers between threshold values as a proxy for pedestrian volume at marked crossings. These studies – including ones on the "safety in numbers" effect – have greatly increased our understanding of factors influencing pedestrian crashes, but the pedestrian volumes used as proxy measures of exposure may not be accurate and are based on assumptions that potentially limit their validity.

While earlier studies explored the effects of traffic exposure, land use and built environment attributes, and sociodemographic characteristics on pedestrian crashes, most studies did not examine in their analysis the transportation facilities (e.g., crosswalks, refuge islands, street lights, and push-buttons) used by pedestrians. The review by Harwood et al. (2008) revealed that only two crash-based studies in a European context examined the effect of narrowing the crossing width on pedestrian crashes, and these studies were limited by several factors. The review also mentioned that it was still uncertain whether crossing width had a significant impact on pedestrian safety. Harwood et al. (2008) reported that the crash-based studies on crosswalk markings had conflicting findings. As is also clear from the report, studies examining the effects of other pedestrian facilities – such as crosswalk illumination, right-turn treatments, raised islands, and bus stop locations – on pedestrian crashes are rare.

2.5 Summary

Most research on pedestrian safety at intersections has been limited by the unavailability of pedestrian exposure data. The few studies which included pedestrian exposure – including those on the "safety in numbers" concept – mostly used surrogate measures. Additionally, the studies which examined the effects of explanatory variables on pedestrian crashes mostly ignored the characteristics of different facilities used by pedestrians in the analysis. This study addresses several of these limitations by:

- Incorporating stronger measures of pedestrian exposure,
- Including key intersection variables, and
- Examining whether the "safety in numbers" concept applies to pedestrian safety in the US, specifically Utah.

3.0 DATA COLLECTION

3.1 Overview

To investigate significant factors contributing to pedestrian crashes at signalized intersections, this research included pedestrian crashes that have occurred over a time period of 10 years, from 2010 through 2019, at signalized intersections in Utah. The research team collected data on potential factors contributing to pedestrian crashes identified from the literature review. Datasets regarding pedestrian crashes, pedestrian and vehicle exposure, road and intersection characteristics, land use and built environment characteristics, and sociodemographic characteristics were formulated covering the factors for inclusion in the model (discussed in Chapter 4.0).

This chapter includes information about the sources of and the procedures used for collecting and assembling the data. First, the site selection process is described. Second, the procedure used for assembling data regarding pedestrian crashes at study locations is provided. Third, the procedure used for collecting data on intersection and road network characteristics data from aerial and street-level imagery is noted. Fourth, the assembly of other information about study locations (including land use, built environment, and sociodemographic data) from existing databases is described. Fifth, details about the calculation of measures of pedestrian exposure – including the assembly of pedestrian signal data and the application of factoring methods developed in a previous UDOT research project (Singleton et al., 2020) to estimate pedestrian volumes – are provided. Each subsection includes statistics summarizing and describing the data that were assembled.

3.2 Study Locations

This research aimed at analyzing factors affecting pedestrian crashes at signalized intersections in Utah. At the time of this study, there were 2,214 traffic signals in use across Utah. Among these, about 2,066 were conventional traffic signals with three-or-more legs, and around 148 were pedestrian-actuated flashers or pedestrian hybrid beacons. A pedestrian hybrid beacon (PHB) is a type of pedestrian-activated beacon used to stop road traffic at an unsignalized location to allow safer pedestrian crossings at a marked crosswalk (FHWA, 2009). [PHBs may also be called high-intensity activated crosswalks or HAWK signals.] The research team opted to collect data on pedestrian and vehicle exposure, road and intersection characteristics, land use and built environment characteristics, and sociodemographic characteristics for all existing signalized intersections. Some signals were not connected to the central network or did not have pedestrian push-buttons (the source of the pedestrian exposure data). Other signals – those outside of the six most populous counties in Utah – did not have detailed information about the surrounding location (e.g., land use and built environment data, and sociodemographic characteristics). So, this lack of data in the source databases limited study locations to 1,606 signalized intersections. Figure 3-1 and Figure 3-2 represent the locations of studied traffic signals at state and county levels of Utah, respectively.

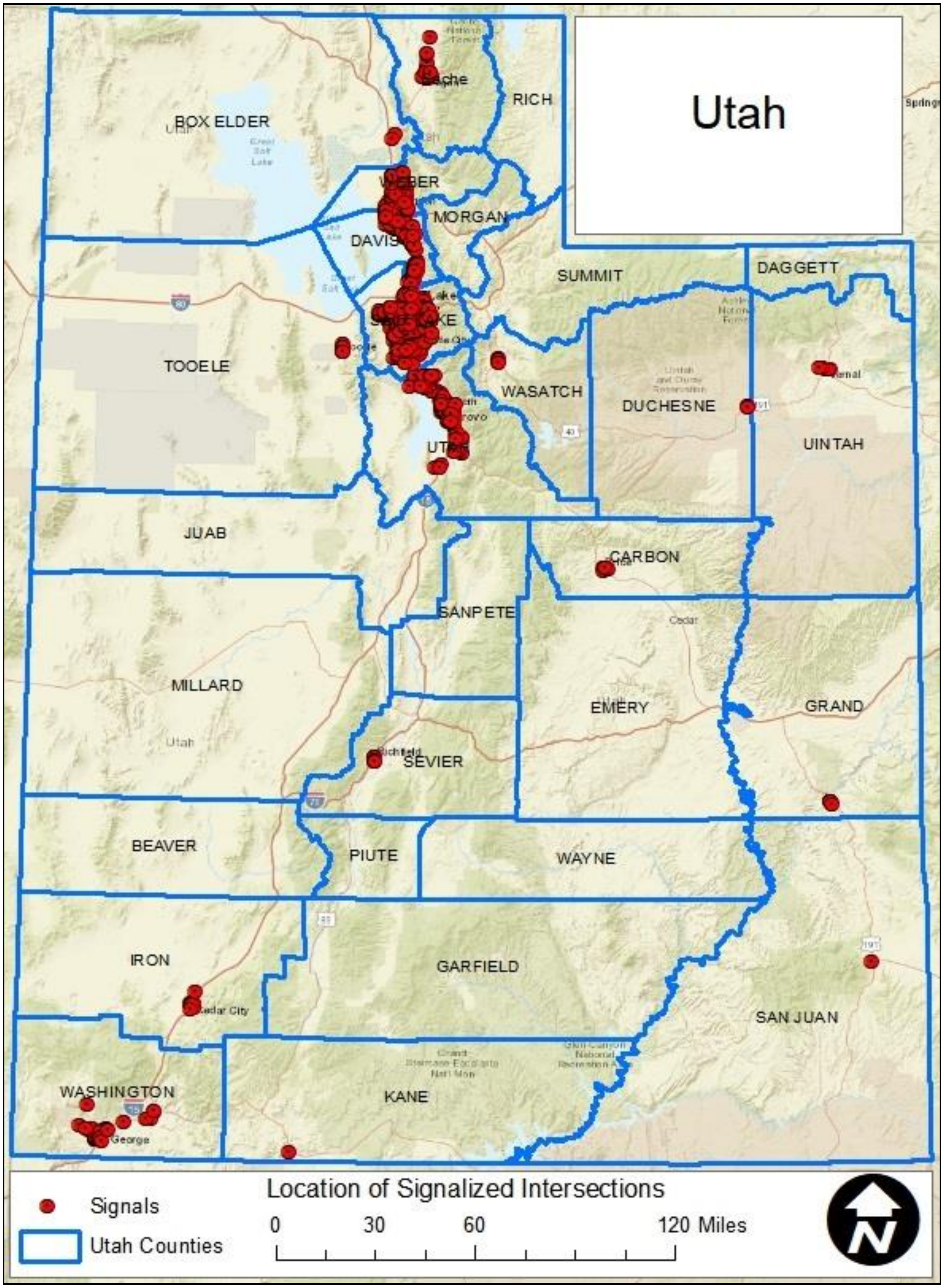


Figure 3-1 Location of studied signalized intersections: state view

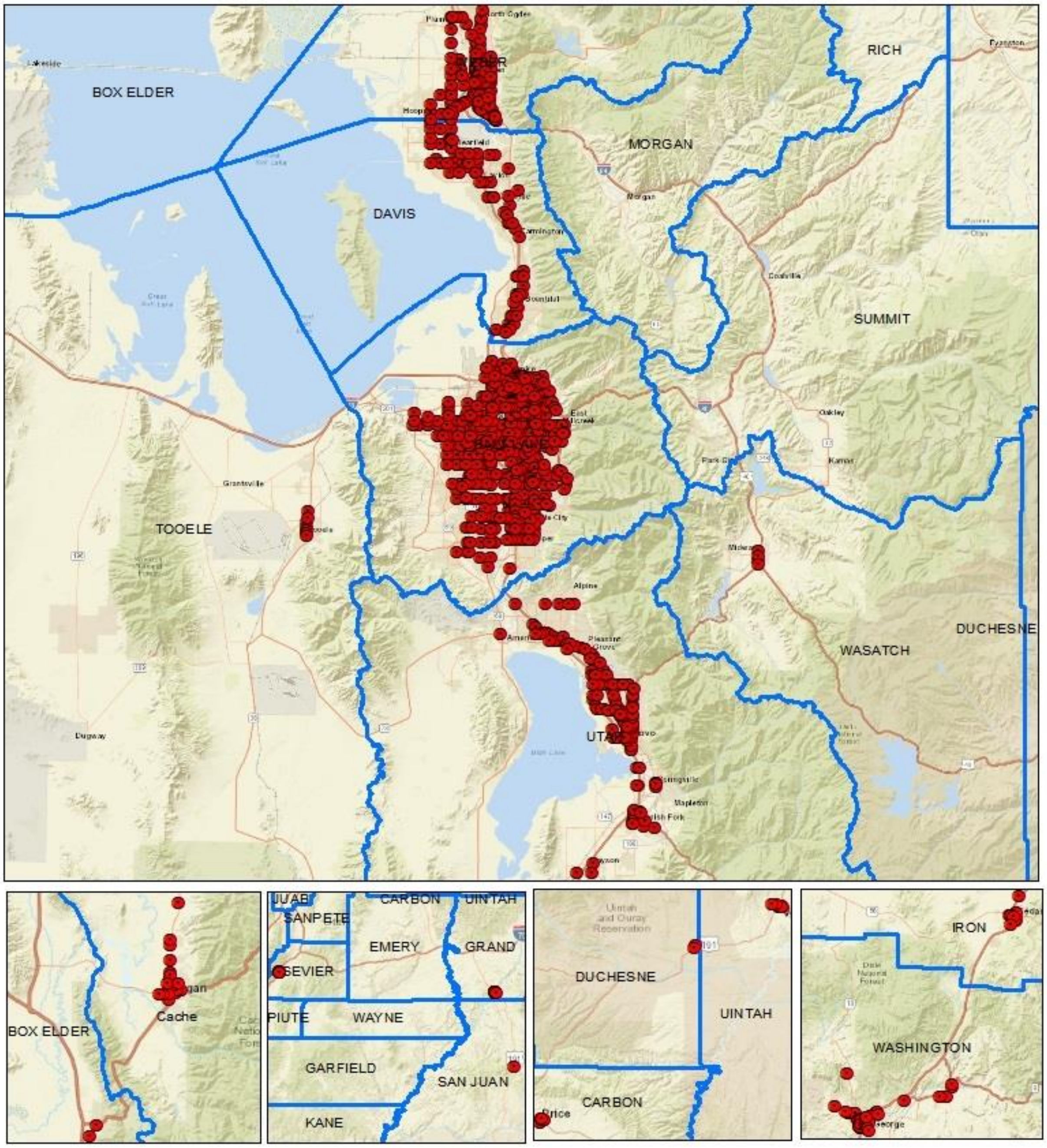


Figure 3-2 Location of studied signalized intersections: county view

3.3 Pedestrian Crash Data

Crash data for all study locations from 2010 through 2019 were obtained from the Utah Department of Transportation (UDOT) through the Numetric website (Numetric, n.d.). Each crash record contained information on temporal characteristics, spatial characteristics, contributing factors, crash severity, weather conditions, and crash participants. This information was extracted from police crash reports. No personally-identifying information was included. Crash data are protected under 23 USC 409.

To determine which pedestrian crashes occurred at or near (and related to) signalized intersections, pedestrian-involved and intersection-related crashes – as specified in the UDOT dataset – were spatially joined to the nearest intersection, and a number of heuristics were applied. First, all crashes that were closest to a signalized intersection (2,799) were considered to be signalized intersection crashes. Second, for remaining crashes that were reported as occurring at a “traffic control signal” (208), those that were less than 492 ft (150 m) from a signal (and no more than 246 ft [75 m] further from a signal than any other intersection) [91] were also considered to be signalized intersection crashes. Also, those crashes located at a “ramp intersection with crossroad” within 984 ft (300 m) of a signal or at a “4-leg intersection” and within 410 ft (125 m) of a signal (19) were also assigned to the nearest signalized intersection. Third, for the remaining crashes that were not reported as occurring at a “traffic control signal” (1,633), those that were less than 246 ft (75 m) from a signal (and no more than 82 ft [25 m] further from a signal than any other intersection) [19] were considered to be signalized intersection crashes. Fourth, any remaining crashes further than 328 ft (100 m) from any intersection but less than 656 ft (200 m) from a signal (11) were also considered to be signalized intersection crashes. (All thresholds were determined through trial and error and visual inspection of maps and crash records. Distances were measured from the crash location to the center of the intersection) The application of these heuristics resulted in 2,939 pedestrian crashes identified as occurring at or near (and related to) signalized intersections. Figure 3-3 depicts a flowchart showing these steps to determine pedestrian crashes at signals.

Figure 3-3 Flowchart of determination of pedestrian crashes at signalized intersections

Next, the 2,939 pedestrian crashes associated with signalized intersections in the 10-year study period were filtered for the study sites of 1,606 signals. In total, 2,598 pedestrian crashes were found to have occurred at or near (and related to) the study intersections after filtering. Of the 1,606 study intersections, a plurality (42%) of the signalized intersections had zero pedestrian crashes during the study period. Nineteen signalized intersections had 10 or more pedestrian crashes in the study period, including one location with the highest frequency: 23 pedestrian-involved crashes. Also, pedestrian crashes were found to occur only once at 363 signalized intersection locations and twice at 209 signalized intersections during the 10-year study period (2010 – 2019). Figure 3-4 shows the distribution of pedestrian crash frequencies at the study intersections. The mean and standard deviation of pedestrian crash frequency in the study dataset were 1.234 and 1.988 respectively.

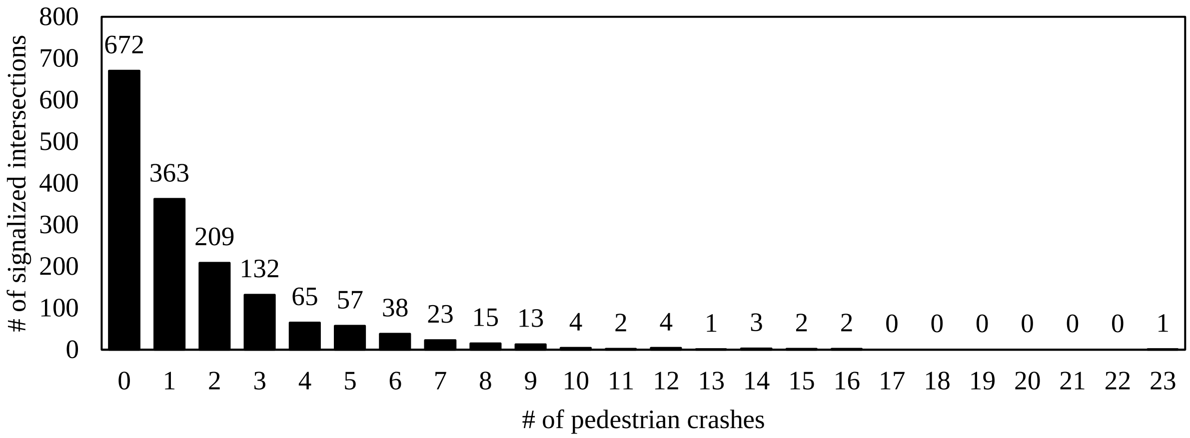


Figure 3-4 Frequency distribution of pedestrian crashes (2010-2019)

3.4 Intersection and Road Network Characteristics Data

As one of the objectives of this study was to identify intersection and road network characteristics that are directly related to pedestrian crash frequency at signalized intersections, detailed data regarding different features at selected sites were gathered from aerial and street-level imagery. The intersection and road network characteristics that have been examined were intersection type, crossing distances, crosswalk marking types, the presence of no-right-turn-on-red signs, the presence of a channelized right-turn lane, and the presence of bike lanes and nearby bus stops along the roads approaching and leaving the intersections. The following sections detail how the data related to intersection and road network characteristics were collected from aerial and street-level imagery.

3.4.1 Intersection Type

The intersection type – or the number or configuration of legs (approaches) that join to form an intersection – is often observed to influence crash risk conditions for pedestrians at signals. Pedestrian crashes generally increase with the number of approaches at an intersection; i.e., pedestrian crash risk is higher at intersections with more legs/approaches when compared to intersections with fewer legs/approaches (Pulugurtha and Sambhara, 2011). This is likely due to both greater opportunities for exposure and increased intersection complexity.

The vast majority (1,397, 87%) of the 1,606 signalized intersections in this study were standard 4-leg intersections. Most of the remaining signals (158, 10%) were 3-leg intersections; only six were 2-leg and three were 5-leg intersections. Two-leg intersections were usually mid-block traffic signals for pedestrian crossings, rather than PHB/HAWK signals. Most PHBs/HAWKs were not connected to the central network or did not have pedestrian push-buttons (the source of the pedestrian exposure data) and so were eliminated during filtering. There were few other special intersection types present in the final dataset. A diverging diamond interchange (DDI) is a type of diamond freeway interchange, where the two directions of non-freeway road traffic cross to the opposite (left) side of the road on both sides of the freeway interchange, to simplify the traffic signal phasing and turns to/from the freeway ramps (FHWA, 2014). A single-point urban interchange (SPUI) is a freeway interchange built with a large overpass or underpass, providing space where all the ramps and cross-street approaches meet at a single traffic signal-controlled intersection (FHWA, 2010). There were nine DDIs and 33 SPUIs in the final dataset. Table 3-1 shows the composition of intersections by type.

Table 3-1 Intersections by type

Intersection type	*# (%)*
2-leg (mid-block)	6 (0%)
3-leg	158 (10%)
4-leg	1,397 (87%)
5-leg	3 (0%)
Diverging diamond interchange (DDI)	9 (1%)
Single-point urban interchange (SPUI)	33 (2%)

3.4.2 Crossing Distances

Longer street crossings mean that it takes pedestrians longer to cross the street, which increases their exposure to potential conflicts or crashes with motor vehicles. To measure crossing distances at signalized intersections, satellite imagery from Google Earth was used (see Figure 3-5). For each crossing, the measuring tool was utilized to measure the curb-to-curb distance along the center of each crossing (rounded to the nearest foot) and thus obtain the distance to cross a particular leg of the signalized intersection. The crosswalk lengths for each leg of all signalized intersections were recorded in a spreadsheet. The mean crosswalk distance for a particular signalized intersection was then obtained by summing crosswalk distances for all legs and dividing by the number of legs. Across all signals, the average mean crosswalk length was around 82 ft, reflecting both the location of many signals along multi-lane arterials as well as the fact that Utah city streets are generally wider than elsewhere in the US (Smith, 2015).

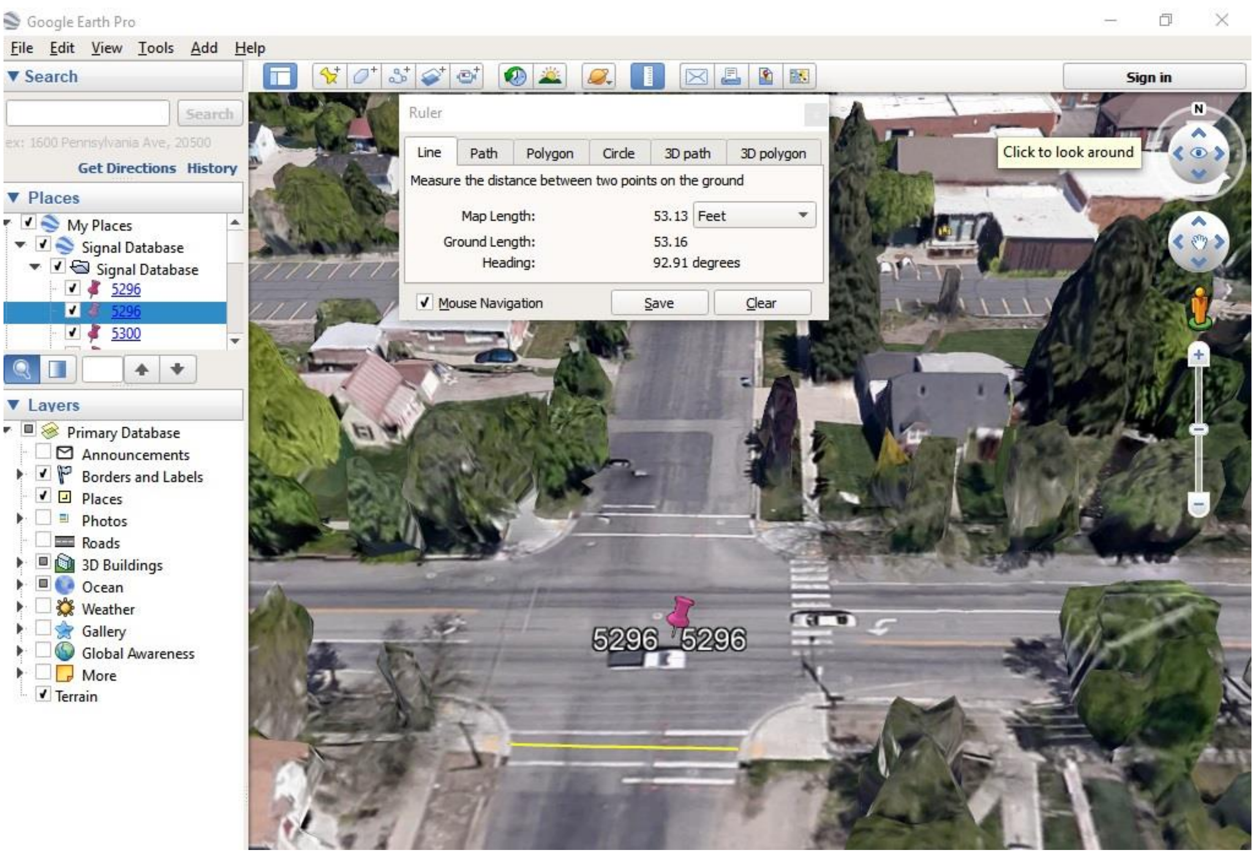


Figure 3-5 Measuring crosswalk length in Google Earth

3.4.3 Crosswalk Markings

Crosswalk markings can alert drivers to the presence of a crossing location where they may expect the presence of pedestrians. To determine the presence and type of crosswalk markings, aerial imagery was inspected for all studied signalized intersection crossings. The presence or absence of a marked crosswalk on each leg of the intersection was determined. Marked crosswalks were also categorized by their marking patterns, which are shown in Figure 3-6. The legs of signalized intersections were summed according to crosswalk type (standard, continental, zebra, and ladder) to obtain the number of legs with a particular crosswalk type. The following variables were prepared to feed into the model: the number of marked crosswalks (total), as well as the number of legs with standard, continental, zebra, and ladder markings. Some agencies may give crosswalks with longitudinal markings different names (e.g., high-visibility crosswalks) or use them in certain typical situations (e.g., at school crossings).

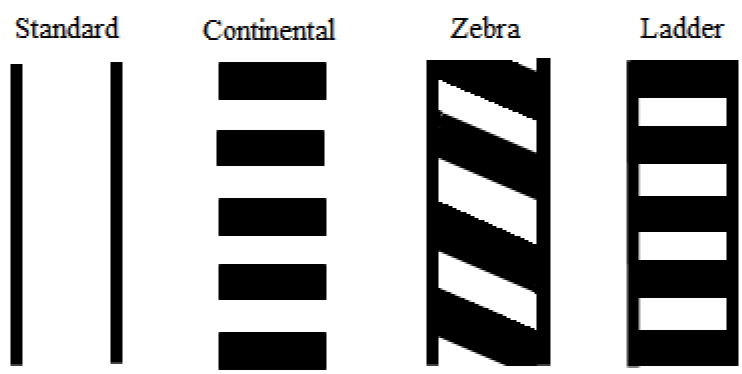


Figure 3-6 Crosswalk marking types

Table 3-2 shows the number and proportion of each of the variables. Other than for a small fraction of locations, all crosswalks were marked in some way. Only 13 out of 1,606 signalized intersections had no markings in their crosswalks. Most intersections were observed to have marked crosswalks on either four (1,132, 70%) or three (151, 10%) legs. Crosswalks with standard transverse markings were the most common, while crosswalks with zebra markings were rarely seen. Of the study intersections, 1,550 (97%) had pedestrian crossings with standard markings in at least one leg of the intersection, and 930 signalized intersections (58%) had four crosswalks with standard markings. In contrast, no study intersection had crosswalks with zebra or ladder markings across all four legs. Only one signalized intersection was observed to have any crosswalk with zebra markings. Continental markings were the second most commonly observed type of crosswalk marking, with 261 (16%) intersections having at least one crosswalk with continental markings.

Table 3-2 Intersections by crosswalk marking types

	# (%) of intersections with (0-4) of each characteristic				
Description of characteristic	*0 legs*	*1 leg*	*2 legs*	*3 legs*	*4 legs*
# with marked crosswalks	13 (1%)	96 (6%)	214 (13%)	151 (10%)	1,132 (70%)
with Standard markings	56 (3%)	135 (9%)	265 (16%)	220 (14%)	930 (58%)
with Continental markings	1,344 (84%)	127 (8%)	93 (6%)	34 (2%)	8 (0%)
with Zebra markings	1,605 (100%)	1 (0%)	-	-	-
with Ladder markings	1,598 (100%)	6 (0%)	1 (0%)	1 (0%)	-

3.4.4 Right-Turn Treatments

Right-turn geometries and operations can affect pedestrian safety. Channelized right turns allow easier movements for right-turning vehicles, which may lead to faster turning speeds, longer stopping distances, and more severe crashes. Drivers making permissive right turns on red may be distracted by watching for gaps in motor vehicle movements and may miss crossing pedestrians. For each approach to a signalized intersection, aerial and street-level imagery from Google Earth and Google StreetView were visually investigated for the presence of channelized right-turn lanes or permanent no-right-turn-on-red signs. (Due to our data collection method, we did not capture information about time/condition-dependent no-right-turn-on-red signs, such as electronic blank-out signs. However, we suspect that these signs are used in only a few locations in Salt Lake City, usually to warn of conflicts with light-rail vehicles, not conflicts with pedestrians.) Figure 3-7 and Figure 3-8 represent right-turn lane conditions: a no-right-turn-on-red sign and a channelized right-turn lane, respectively.

Figure 3-7 No-(right) turn-on-red signs (FHWA, 2009)

Figure 3-8 Channelized right-turn lane

Table 3-3 shows the number and percentage of intersections with various numbers of channelized right-turn lanes and right-turn lanes with "no right turn on red" signs in the recorded dataset. Of the study intersections, 1,434 intersections (89%) had no channelized right-turn lanes. In contrast, 172 (11%) had at least one and 29 (2%) had four channelized right-turn lanes.

Additionally, data shows that 1,581 (around 98% of the study intersections) had no lanes with “no right turn on red” sign. Only 25 intersections (2%) had at least one approach with a “no right turn on red” sign.

Table 3-3 Intersections by right-turn conditions

	# (%) of intersections with (0-4) of each characteristic				
Description of characteristic	*0 legs*	*1 leg*	*2 legs*	*3 legs*	*4 legs*
# with channelized right turn	1,434 (89%)	86 (5%)	45 (3%)	12 (1%)	29 (2%)
# with “no right turn on red” signs	1,581 (98%)	25 (2%)	-	-	-

3.4.5 Bike Lanes and Bus Stops

To understand the potential effects of the presence of bus stops and bike lanes on pedestrian crashes, related information was collected. The presence of bike lanes (of any type) on the portion of each leg approaching and leaving the intersections were identified and recorded from satellite imagery. Figure 3-9 presents an example of a bike lane near an intersection.

Similarly, the presence of at least one transit stop located within 300 ft of the signalized intersection was recorded for the approaching and leaving portions of each leg. Bus stops placed immediately prior to the intersection (approaching) were designated as near-side bus stops, whereas those placed immediately after passing through the intersection (leaving) were designated as far-side bus stops in the dataset. Figure 3-10 presents an illustration of near-side and far-side bus stops.

Figure 3-9 Example of a bike lane

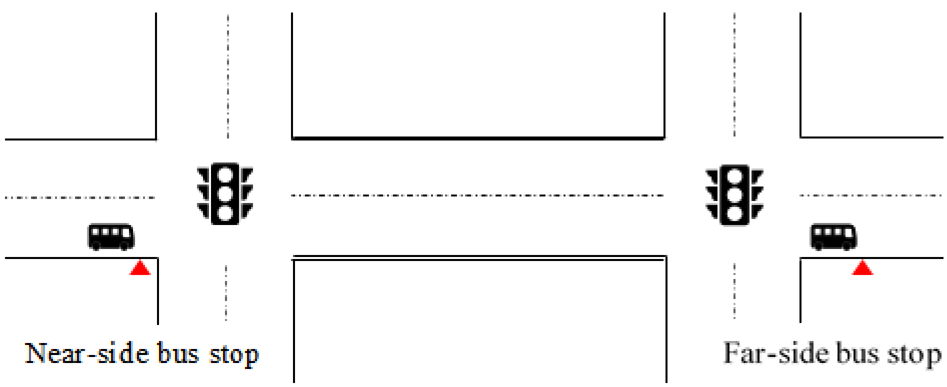


Figure 3-10 Illustration of near-side and far-side bus stops

The number of inbound and outbound bike lanes, and the number of near-side and far-side bus stops, were obtained by summing up the corresponding features present in the legs of each intersection. Table 3-4 presents the numbers and percentages of intersections for these variables. Near-side bus stops were not present at 1,215 intersections (76%), but 391 intersections (24%) had near-side bus stops in at least one approach. In contrast, 962 intersections (60%) had no far-side bus stop. The remaining 644 (40%) intersections had at least one leg with a far-side bus stop. Only 19 intersections had far-side bus stops on all four legs of the intersection, while just two intersections had near-side bus stops on all four legs. Of the study intersections, 1,165 (73%) had no legs with bike lanes in the inbound directions. The other 441 intersections (27%) had at least one approach with an inbound bike lane. Bike lanes on legs in the outbound direction were not present at 1,156 (72%) study intersections. The other 450 intersections (28%) had at least one bike lane in the outbound direction.

Table 3-4 Intersections by variables related to bus stops and bike lanes

	# (%) of intersections with (0-4) of each characteristic				
Description of characteristic	*0 legs*	*1 leg*	*2 legs*	*3 legs*	*4 legs*
# with near-side bus stops	1,215 (76%)	292 (18%)	91 (6%)	6 (0%)	2 (0%)
# with far-side bus stops	962 (60%)	361 (22%)	227 (14%)	37 (2%)	19 (1%)
# with bike lanes (inbound)	1,165 (73%)	78 (5%)	291 (18%)	32 (2%)	40 (2%)
# with bike lanes (outbound)	1,156 (72%)	92 (6%)	283 (18%)	34 (2%)	41 (3%)

3.4.6 Street Lighting Conditions

Street lighting condition has a direct influence on the perceptions and reactions of both pedestrians and motor vehicle operators. This has been examined in some previous studies to understand its effect on the occurrence and severity of pedestrian crashes (Hu et al., 2020).

In this study, data related to the lighting conditions of the study intersections were collected for the analysis of pedestrian crash frequency. Specifically, the satellite imagery was investigated to determine the presence of overhead street lights at the corners of the study intersections. Figure 3-11 shows an example of an overhead street light at a signalized intersection. Data regarding lighting conditions were coded solely based on the presence of the lights; no field investigation was carried out to examine whether the lighting was functional or not.

Figure 3-11 Overhead street light at a signalized intersection

All but 43 intersections (out of 1,606 studied intersections) had overhead street lights installed in at least one or more corners of the intersections. Table 3-5 describes the variables related to street lights. Among the 1,563 signalized intersections with street lighting, 1,411 were four-leg and 146 were three-leg intersections. Overhead street lights were missing from 31 four-leg and 10 three-leg signalized intersections.

Table 3-5 Intersections by variables related to street lighting condition

	# (%) of intersections with (0-4) of the characteristic				
Description of characteristic	*0 legs*	*1 leg*	*2 legs*	*3 legs*	*4 legs*
Presence of street lights					
Yes	0 (0%)	0 (0%)	6 (0%)	146 (9%)	1,411 (88%)
No	0 (0%)	0 (0%)	2 (0%)	10 (1%)	31 (2%)

3.5 Intersection Data from Existing Databases

Several other signalized intersection attributes relevant for the study of factors affecting pedestrian crashes were obtained from existing databases, including: motor vehicle traffic volumes, transportation system characteristics, land use and built environment data, and sociodemographic characteristics. When appropriate, these data were calculated for the area within a quarter-mile of each intersection. The assembly of each of these types of data is described in the paragraphs below.

Vehicle exposure data – i.e., the 2017 annual average daily traffic (AADT) volume for major legs and minor legs of signalized intersections – were processed from Road Centerlines data from Utah AGRC (Automated Geographic Reference Center). First, for all roadway segments approaching each signal, characteristics including 2017 AADT, roadway class, and route number were assembled. Second, heuristics were applied to determine the major approaches (max of two) and minor approaches: based first on roadway class, second on larger traffic volumes, and third on lower route number. Third, AADT values were averaged within the major/minor roadway segments to obtain major and minor AADT for the signalized intersections. Fourth, some signals with properly missing minor AADT – because they had no minor legs (e.g., PHBs/HAWKs or midblock crossings) – were assigned a minor AADT value of zero to increase the valid sample size.

Additional information about land use and built environment characteristics nearby each signal were obtained from a variety of sources and processed. Each variable was calculated for a quarter-mile street network-based buffer around each signalized intersection. The percentage of different types of land use (residential, commercial, industrial, and vacant) around each signal were calculated from parcel-level land use maps obtained from the Utah AGRC website. Population and employment density variables were calculated using block group-level data from the 2013-2017 American Community Survey (ACS) and block-level from the 2017 Longitudinal Employer Household Dynamics (LEHD) datasets, respectively. Using similar data from Utah AGRC, the acreage of parks and number of schools and places of worship within a quarter-mile network distance of each signal were also calculated. Intersection density (a measure of connectivity) was also calculated from information about the location of road and street intersections, also from Utah AGRC.

Sociodemographic characteristics of nearby neighborhoods were calculated using the same quarter-mile network buffers around each signal. Specifically, 2013-2017 ACS data from the US Census were used to obtain information about median household income, average vehicle ownership, mean household size, percentage of the population with a disability, and percentage of population of Hispanic or non-white race/ethnicity.

Due to data limitations, land use, built environment, and sociodemographic characteristics could only be assembled for signals in the six largest counties in Utah (Salt Lake, Utah, Davis, Weber, Washington, and Cache). Together, these counties contain more than 80% of Utah's population and the vast majority of the traffic signals in the state. Descriptive statistics for these variables are shown later in Table 3-7.

3.6 Pedestrian Exposure Data

One unique aspect and contribution of this study is the use of novel and more complete pedestrian exposure data, which (as the literature review noted) is often missing from pedestrian safety studies. The pedestrian exposure data used here came from traffic signals, specifically derived from pedestrian activity events at signalized intersections that were recorded in high-resolution traffic signal controller logs (Sturdevant et al., 2012). When a traffic signal includes walk indications and pedestrian push-buttons for detection, two relevant events can be recorded. First, pedestrian detection events occur whenever the push-button is pressed, which could happen multiple times per signal cycle. Second, a pedestrian call registered event is recorded the first time in a cycle (usually) that a push-button is pressed for a particular phase or crossing. Either (or both) of these events may be used as a proxy for pedestrian crossing volumes, which is the typical measure of pedestrian exposure, within a given time period.

Although pedestrian traffic signal data are not perfect measures of pedestrian volumes (Blanc et al., 2015; Kothuri et al., 2017), recent work in an earlier UDOT research project by Singleton et al. (2020) has demonstrated that such data can be used to predict pedestrian crossing volumes at signalized intersections with relative accuracy. Throughout 2019, more than 10,000 hours of videos of pedestrian crossing events were recorded at 90 signalized intersections throughout Utah, and more than 175,000 pedestrians were manually counted. These data were then compared to traffic signal push-button-based measures of pedestrian activity, using simple non-linear (quadratic and piece-wise linear) regression models predicting hourly pedestrian crossing volumes as a function of pedestrian signal activities. Over more than 22,500 hours of data, the correlation between observed and model-predicted hourly pedestrian crossing volumes was 0.84, with a mean absolute error of only 3.0 (Singleton et al., 2020). Overall, that research project demonstrated that pedestrian signal data can be used to estimate reasonably accurate pedestrian crossing volumes. For the purposes of this research project, these pedestrian signal data provide greater temporal and spatial coverage for measuring pedestrian exposure (more locations over longer time periods), thus improving the understanding of relationships between pedestrian crashes and pedestrian volumes.

UDOT is a national leader in the development and deployment of the Automated Traffic Signal Performance Measures (ATSPM) system (Day et al., 2016) through which pedestrian events from high-resolution traffic signal controller logs can be obtained. As of fall 2018, UDOT was centrally archiving data from more than 1,900 state- and locally-owned signals (Taylor and Mackey, 2018). For this study, one year (July 2017 through June 2018) of pedestrian data were obtained from all available traffic signals in Utah. After cleaning the data for incompleteness, the regression models developed by Singleton et al. (2020) were applied to the pedestrian signal data. These estimates (by phase of the signal and hour of the day) were then aggregated across all crossings at each signal and all hours in each day, and then the daily estimates were averaged over all days in the year to calculate the annual average daily pedestrian (AADP) crossing volumes at each signal. Descriptive statistics for AADP are shown later in Table 3-7. AADT at signals in this analysis ranged from 0.163 to 6,737 pedestrians per day. The mean and median of AADP for all the study signals was found to be 270 and 111 respectively. The histogram of pedestrian exposure is shown in Figure 3-12.

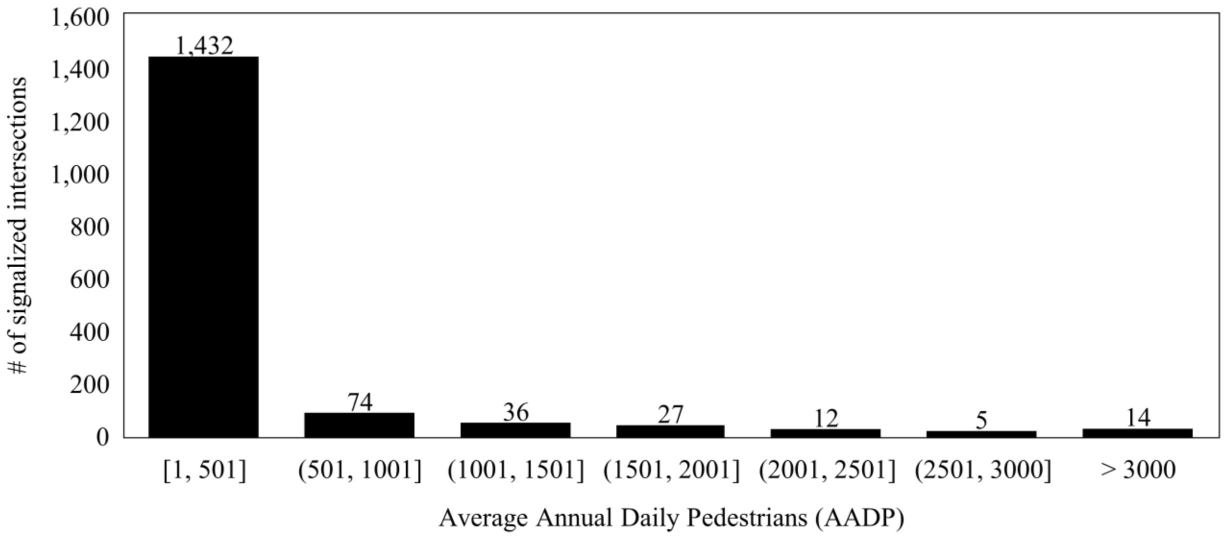


Figure 3-12 Histogram of pedestrian exposure at signals

As an example, the 10 highest (estimated) pedestrian volume signalized intersections are shown in Table 3-6. (There may actually be higher-volume pedestrian intersections in Utah, but many downtown Salt Lake City intersections always operate on pedestrian recall and have no push buttons and thus no pedestrian activity data.) The high-volume locations make intuitive sense. Most of these signals are located in a small area of downtown Salt Lake City characterized by large centers of employment, shopping, and culture, as well as frequent transit service. For example, Signal 7244 is located adjacent to the Salt Lake City Public Library, the Salt Lake City and County Building, and a light rail station. Two other signals (5807 and 6631) are located at the edge of large university campuses (Utah State University and Brigham Young University). The remaining two high pedestrian volume signals are in downtown Moab, a city in eastern Utah that sees high tourist activity due to its location adjacent to Arches and Canyonlands National Park.

Table 3-6 Signals in Utah with the highest estimated average pedestrian volumes

Rank	*Signal*	*Location*	*Estimated AADP*
1	7138	S Temple & State St, Salt Lake City	6,737
2	7244	400 S & 200 E, Salt Lake City	4,868
3	7139	100 S & State St, Salt Lake City	4,519
4	7248	400 S & 600 E, Salt Lake City	4,450
5	5807	700 N & 800 E, Logan	4,446
6	8303	100 S & Main St, Moab	4,307
7	7243	400 S & Main St, Salt Lake City	4,009
8	7142	400 S & State St, Salt Lake City	3,909
9	8302	Center St & Main St, Moab	3,544
10	6631	1230 N & Canyon Rd, Provo	3,476

3.7 Data Preprocessing

The data collected from all the sources were assembled into one complete dataset. Each observation in the dataset consisted of a signalized intersection, with pedestrian crash frequency data that occurred over a 10-year time period at that location, along with the corresponding road network facilities and the surrounding land use, built environment, and sociodemographic features. The raw dataset was comprised of 2,214 observations, one for each signalized intersection. But pedestrian volume estimates from pedestrian signal activities (Singleton et al., 2020) were obtained for only around 1,600 signalized intersections. Furthermore, road geometry and other features were not available for a few other observations. These observations without a complete set of all feature characteristics were removed from the final dataset prior to use in modeling. The final dataset included observations for 1,606 signals and 2,598 pedestrian crashes that occurred at those signals. Table 3-7 summarizes the final dataset characteristics and descriptive statistics.

Table 3-7 Descriptive statistics of variables in the final dataset (N = 1,606)

Variable	*Min.*	*Max.*	*Mean*	*Std. Dev.*
Dependent variable, frequency model				
# of pedestrian-involved crashes	0	23	1.62	2.32
Measures of exposure				
Annual average daily pedestrian volume (AADP)	0.16	6,737	269.95	572.78
Average daily traffic in major direction ($AADT_{MAJ}$)	450	186,000	23,312.09	12,900.82
Average daily traffic in minor direction ($AADT_{MIN}$)	0	57,000	8565.02	7,789.45
Transportation characteristics				
Presence of overhead street lighting	0	1	0.97	0.16
Intersection type				
2-leg (mid-block)	0	1	0.00	0.06
3-leg	0	1	0.09	0.29
4-leg	0	1	0.87	0.33
5-leg	0	1	0.00	0.04
Diverging diamond interchange (DDI)	0	1	0.00	0.07
Single point urban interchange (SPUI)	0	1	0.02	0.14
# crosswalks, total	0	4	3.45	0.96
# crosswalks with standard markings	0	4	3.14	1.17
# crosswalks with continental markings	0	4	0.27	0.71
# crosswalks with ladder, zebra, or other markings	0	3	0.01	0.11
# crosswalks with continental, ladder, or zebra markings	0	4	0.29	0.72
Crosswalk length (mean, ft)	20	185	81.83	19.89
# approaches with no pedestrian crossing	0	4	0.44	0.83
# approaches with no right-turn-on-red	0	1	0.15	0.12
# approaches with channelized right turns	0	4	0.20	0.69
# approaches with bike lanes	0	4	0.59	1.03
# of bus stops within 300 ft of intersection	0	6	0.93	1.18
# approaches with near-side bus stops	0	4	0.31	0.60
# approaches with far-side bus stops	0	4	0.62	0.89
Intersection density (# per mi^2) [a]	6.07	313.17	97.66	49.12
Land use and built environment characteristics [a]				
% land use residential	0	84	31	23.51
% land use commercial	0	92	28	20.75
% land use industrial	0	83	2.41	10.51
% land use vacant	0	100	4.54	8.74
Population density (1,000 per mi^2)	0.08	23.51	4.51	3.02
Employment density (1,000 per mi^2)	0.02	216.03	7.30	11.51
Park area (acre)	0	37.15	1.45	3.61
# of schools	0	5	0.31	0.61
# of places of worship	0	6	0.51	0.78
Sociodemographic characteristics [a]				
Household income (median, $1,000)	20.5	144.61	61.33	21.87
Vehicle ownership (mean)	0.55	3.00	1.81	0.45
Household size (mean)	1.41	13.72	3.11	0.85
% of the population with a disability	2.51	27.06	10.64	4.12
% of the population of Hispanic or non-white race/ethnicity	0.00	75.66	17.26	13.50

[a] These variables were measured using a quarter-mile network buffer.

3.8 Summary

The data for this project were collected and assembled from a number of different sources. Satellite and street-level imagery were used to collect different intersection and road network features. Data to investigate the effect of land use type and built environment characteristics on pedestrian crashes were collected from the Utah AGRC website. US Census data were used to assemble sociodemographic data for the neighborhoods surrounding each signalized intersection. Pedestrian crash data for the 10-year study period were available from the UDOT Numetric website. Also, other important traffic and road characteristics information – namely motor vehicle volumes – were assembled. The pedestrian exposure data were collected and estimated from pedestrian traffic signal data archived in UDOT's ATSPM system. All the collected data were interpreted and processed using the software environment R to obtain appropriate data to feed into the model. To ensure consistency of categories across the study, all labels were compared to identify potential misnomers due to human error and resolved through careful revision and revisiting of the source databases when necessary. Processing of raw data allowed the research team to better quantify and calibrate models to assess all potential factors affecting pedestrian crashes.

4.0 DATA ANALYSIS

4.1 Overview

To better understand all the factors that contribute to pedestrian crash frequencies at signalized intersections in Utah, the collected data (described in Chapter 3.0) for 1,606 signalized intersections and 2,598 pedestrian crashes were analyzed. Specifically, a series of count data models were estimated, culminating in several negative binomial models whose results were also interpreted in terms of safety performance functions and crash modification factors.

This chapter contains information about how the data were analyzed and details about the results of the data analysis. First, some descriptive analyses highlight pedestrian crash hot-spots and characteristics associated with those hotspots. Second, the more complex statistical modeling procedures are described, including the estimation of Poisson, negative binomial, and zero-inflated models, as well as the generation of safety performance functions and crash modification factors. Third, the model results are presented, described, and interpreted.

4.2 Descriptive Analysis of Pedestrian Crash Frequencies and Rates

In order to develop a preliminary idea of the relationship between pedestrian crashes and area characteristics, the project team first identified pedestrian crash hot spots. Table 4-1 shows 20 signals (out of the 1,606 studied signals) with the highest pedestrian crash frequencies over the 10-year study period. Among the studied signals, these 20 signals had 10 or more pedestrian crashes (including one signal with 23) in the study period. These signals with the highest pedestrian crashes make intuitive sense. All but one of these intersections are located in Salt Lake County, which is the most populous county in Utah. The surrounding neighborhoods of these signals are mostly characterized by large hubs of employment, business and cultural centers, shopping malls, and grocery stores, and are on major arterials with frequent transit services. For example, the signal with the highest number of crashes (7104) is adjacent to a few supermarkets and shopping malls, a bank, and a number of restaurants. The only signal in this group outside of Salt Lake County (5019, with 11 pedestrian crashes) is located in Ogden, an urban center north of Salt Lake City.

The signal is close to a state office (Utah Department of Workforce Services), a school, a few markets, and a number of restaurants.

Table 4-1 Signals with the highest pedestrian crash frequencies

Rank	*Signal*	*Location*	*Pedestrian crashes (#)*	*AADP*
1	7104	4100 S & Redwood Rd, West Valley City	23	960
2	7194	7800 S & 700 E, Sandy	16	281
3	7157	4500 S & State St, Murray	16	658
4	5118	700 S & State St, Clearfield	15	325
5	7168	7200 S & State St, Midvale	15	451
6	7207	Fort Union Blvd & 900 E, Midvale	14	492
7	7282	3500 S & 4000 W, West Valley City	14	527
8	7155	3300 S & State St, South Salt Lake	14	931
9	7102	3500 S & Redwood Rd, West Valley City	13	585
10	4100	3900 S & 700 W, South Salt Lake	12	292
11	7279	3500 S & 4800 W, West Valley City	12	360
12	7328	5400 S & 4015 W, Taylorsville	12	382
13	7283	3500 S & 3600 W, West Valley City	12	1,180
14	5019	28th St & Washington Blvd, Ogden	11	466
15	1107	2100 S & 900 E, Salt Lake City	11	1,081
16	1120	400 S & 1300 E, Salt Lake City	10	NA
17	4114	4715 S & 4800 W, West Valley City	10	187
18	7107	4700 S & Redwood Rd, Taylorsville	10	466
19	7295	3300 S & West Temple, South Salt Lake	10	857
20	7142	400 S & State St, Salt Lake City	10	3,909

Crash data are protected under 23 USC 409.

Pedestrian crash rate is frequently used as a measure of pedestrian crash risk. Pedestrian crash frequency at a signal is the number of pedestrian crashes that occurred at or near the signals. In contrast, pedestrian crash rate is the number of pedestrian crashes and pedestrians normalized by dividing by pedestrian volume. (This rate calculation only accounts for pedestrian volumes, not motor vehicle volumes.) Therefore, the project team identified another set of locations with the highest pedestrian crash rates (crashes over 10 years per average daily pedestrians). Table 4-2 lists the signals with the highest pedestrian crash rates. Since pedestrian crashes are rare events, this list mostly contains locations with only one pedestrian crash over ten years but low pedestrian volumes too, thus generating very high crash rates. As a result, we filtered the list of intersections for those that contain at least 5 crashes. Table 4-3 lists the signals (with 5+ crashes) with the highest pedestrian crash rate. This produces a ranking of signals that is less sensitive to random chance and more indicative of a systematic pedestrian safety issue. Figure 4-1 and Figure 4-2 present maps indicating the 20 signal locations with the highest pedestrian crashes and the 20 signals with the highest crash rates in Utah.

Table 4-2 Signals with the highest pedestrian crash rates (with any # of crashes)

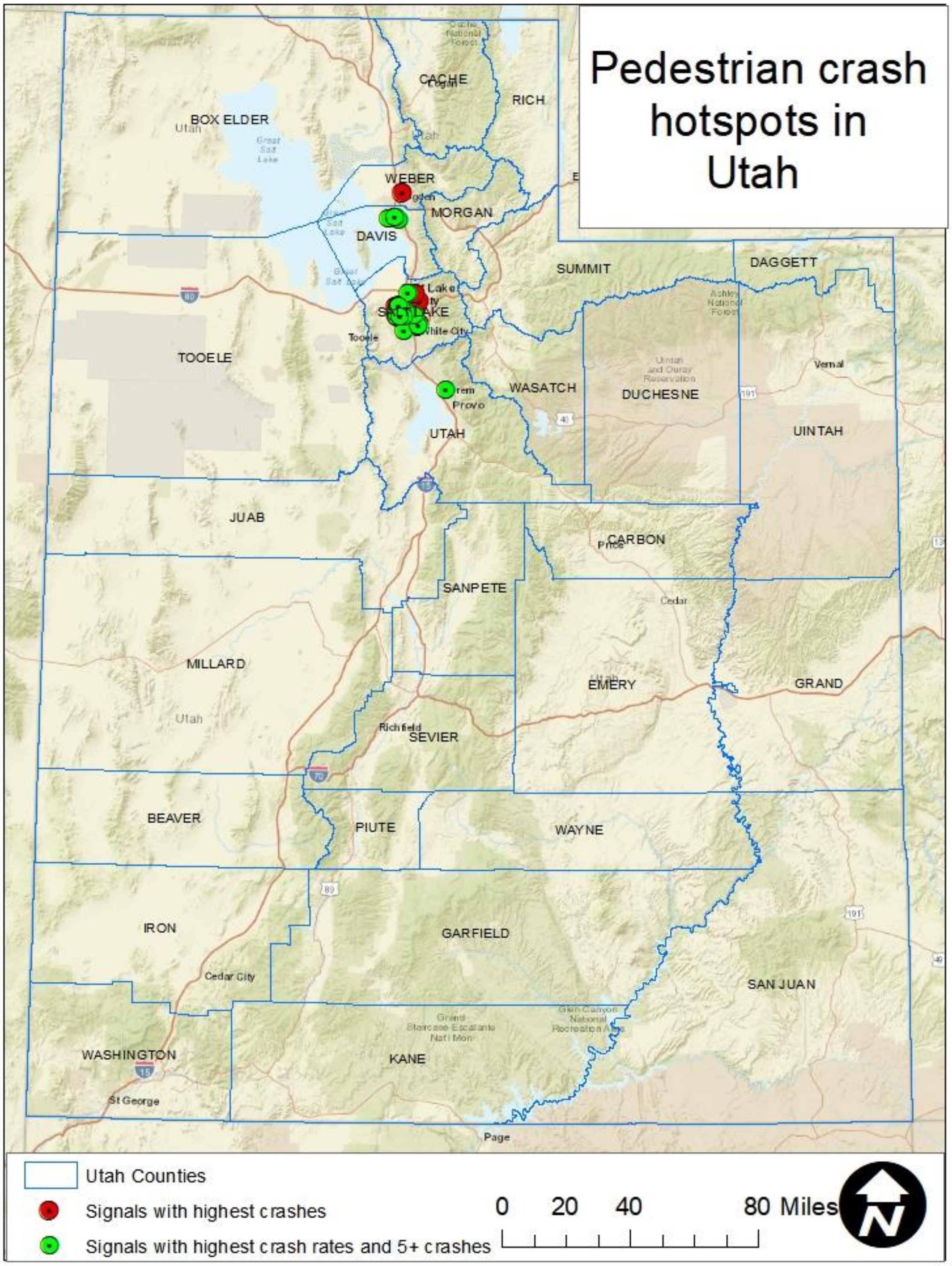
Rank	*Signal*	*Location*	*Pedestrian crashes (#)*	*AADP*	*Pedestrian crash rate*
1	5171	200 N & I-15 NB Ramps, Kaysville	4	2	2.404
2	5057	Shadow Valley & Harrison Blvd, Ogden	1	2	0.665
3	6044	I-15 NB Ramps & Main St, Spanish Fork	1	3	0.308
4	7812	2000 N, SR-36 & Tooele	1	4	0.241
5	5096	Hinckley Dr & 1900 W, Roy/Weber County	1	8	0.132
6	5091	1800 N & Main St, Sunset	4	36	0.112
7	4121	6200 S & 2700 W, Taylorsville	1	10	0.103
8	7015	9000 S & 700 W, Sandy	2	20	0.098
9	5160	700 S & 1500 E, Clearfield	1	10	0.098
10	7030	700 N & I-215 W SB Ramps, Salt Lake City	1	10	0.096
11	5079	5000 S & Washington, South Ogden	1	10	0.096
12	5077	4300 S & Washington, South Ogden	1	10	0.096
13	8405	1300 S & Main St, Richfield	1	11	0.091
14	5086	5600 S & Freeway Park Dr, Roy	1	12	0.087
15	7053	1820 S & Bangerter Hwy, Salt Lake City	1	12	0.082
16	8620	Red Cliffs Dr & 2450 E, St. George	3	36	0.082
17	7418	9000 S & 2700 W, West Jordan	8	101	0.079
18	4883	7800 S & 4450 W, West Jordan	1	13	0.078
19	7608	4700 S & I-215 W NB Off-ramp, Taylorsville	1	13	0.077
20	8111	700 S & Bluff St, St. George	4	52	0.077

Crash data are protected under 23 USC 409.

Table 4-3 Signals with the highest pedestrian crash rates (with 5+ crashes)

Rank	*Signal*	*Location*	*Pedestrian crashes (#)*	*AADP*	*Pedestrian crash rate*
1	7418	9000 S & 2700 W, West Jordan	8	101	0.079
2	7209	7400 S & 900 E, Midvale	6	93	0.064
3	5115	2000 N & 1700 W, Layton	9	145	0.062
4	4505	3100 S & 4000 W, West Valley City	7	120	0.058
5	7610	5415 S & 4800 W, Salt Lake County	8	138	0.058
6	7194	7800 S & 700 E, Sandy	16	281	0.057
7	7348	5300 S & Woodrow (120 W), Murray	7	129	0.054
8	4114	4715 S & 4800 W, West Valley City	10	187	0.053
9	5132	Bernard Fisher Hwy & 1000 E, Clearfield	8	150	0.053
10	5707	2600 S & 500 W, Bountiful	6	117	0.051
11	5111	1700 S & 1000 E, Clearfield	5	98	0.051
12	7605	4610 S & Redwood Rd, Taylorsville	5	103	0.049
13	4511	4100 S & 3200 W, West Valley City	9	188	0.048
14	6394	1200 N & State St, Orem	7	151	0.046
15	7090	400 S & Redwood Rd, Salt Lake City	6	129	0.046
16	6901	400 S & 400 E, Springville	6	130	0.046
17	5118	700 S & State St, Clearfield	15	325	0.046
18	7331	5400 S & 2700 W, Taylorsville	5	109	0.046
19	7281	3500 S & 4155 W, West Valley City	6	133	0.045
20	7329	5400 S & 3600, Taylorsville	5	114	0.044

Crash data are protected under 23 USC 409.



Crash data are protected under 23 USC 409.

Figure 4-1 Map showing signals with the highest pedestrian crash frequencies and rates

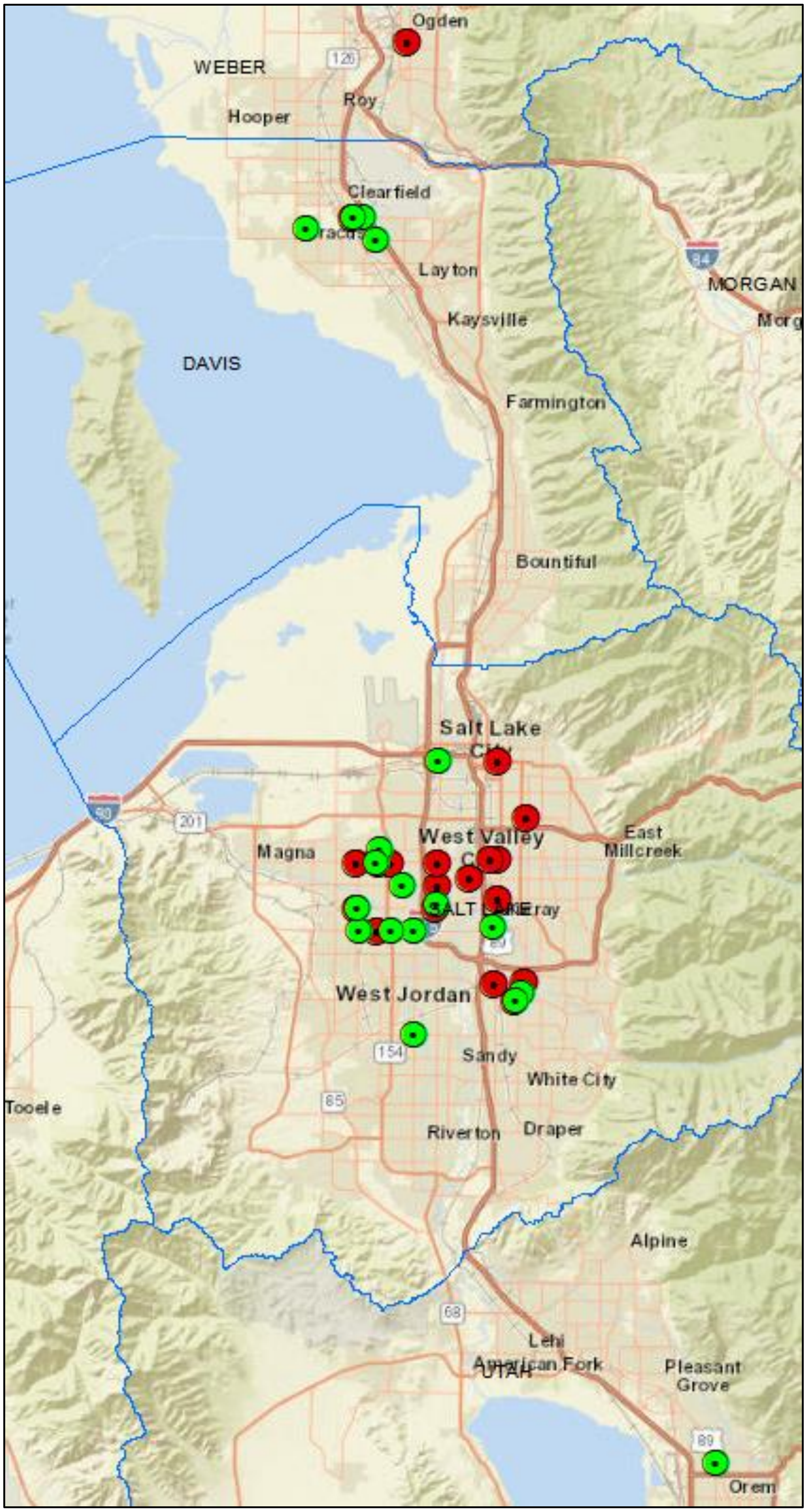


Figure 4-2 Hot spot locations for pedestrian crashes near Salt Lake, Ogden, and Orem

After preliminary investigations, it was found that three of the 20 signal locations with the highest crash rate (filtered for sites with 5+ crashes) matched three of the 20 locations with the highest crash frequency: signals 7194, 4114, and 5118. These signals, with the highest pedestrian crash rates, also had 10 or more pedestrian crashes over the 10-year study period. A look into the characteristics of the areas surrounding these three signals reveals that they are commonly located in residential areas and have at least a supermarket, gas station, or parking area on the corner of the intersection. Additionally, transit services are also frequent in these locations. For signals (4511, 7418, 5115, 5132 and 7610) with the highest crash rates and where pedestrian crashes occurred eight or nine times in the 10-year study period, large parking areas and frequent transit service are also common. At least one large medical/cultural/recreation center was present near these signals. Most other signals with the highest crash rates but seven or fewer crashes are located in cities in Salt Lake County (Sandy, Murray, Salt Lake City, Midvale, West Valley City, and Taylorsville). The surrounding areas are often characterized by large businesses, medical centers, shopping malls, grocery stores, schools, and religious establishments. For example: signals 4505, 7605, 7348, 7331 and 7281 are located near large/small hospitals, religious establishments, and schools or community colleges. A few signals (5707, 5111, 6901 and 6394) are also in this category of high pedestrian crash rate and seven or fewer crashes. These are located outside Salt Lake County (in Clearfield, Bountiful, Orem, and Springville), but their surrounding areas are also characterized by business, education, and recreation centers.

4.3 Pedestrian Crash Frequency Modeling

Like most crash frequency data, the pedestrian crash frequency data used in this study were discrete, random, and non-negative. The modeling framework of generalized linear models (GLMs) is more suited to such count data than ordinary linear regression, which can predict negative, non-integer values of the dependent variable. The Poisson regression model has been widely used as a starting point to model count data (Lord and Mannering, 2010), but it assumes that the variance of the count data is equal to the mean. When the count data used are over-dispersed (i.e., the variance is greater than the mean), a negative binomial (NB) regression model is usually more appropriate for the dataset. An additional term in the NB model allows the variance to be different from the mean of the dataset; thus, the Poisson model is a special case of the NB model.

Although this NB model may be a statistically significantly better fit to the data than the more restrictive Poisson model – as tested empirically using a likelihood ratio test – it does not account for any excess zeros in the dataset. This phenomenon of excess zeros (also known as zero-inflation) – which is rather common in crash frequency data – refers to the presence of more zero count observations (zero crash locations) than would otherwise be predicted by the assumed statistical distribution (either Poisson or negative binomial). Perhaps some signalized intersections may be so safe (and/or low volume) that a crash would be expected only once every 100 years, while others are more dangerous but may still see zero crashes during the observation period. Zero-inflated models can account for excess zeros by including a first-stage model predicting the probability of the observation belonging to a zero-count group, followed by a regular Poisson or NB model predicting the count if belonging to the regular-count group.

As stated earlier, there were no pedestrian crashes during the study period at 42% of the signalized intersections. Hence, the adoption of a zero-inflated version of the NB model (ZINB) was plausible, as it can accommodate overdispersion arising from both unobserved heterogeneity and excess zeros (Miranda-Moreno and Fu, 2006). The probability density function for the ZINB model is as follows:

$$P(Y = y_{it}) = \begin{cases} P_{it} + (1 - P_{it}) \frac{1}{(1+\alpha \mu_{it})^{\frac{1}{\alpha}}} & y_{it} = 0 \\ (1 - P_{it}) \frac{\Gamma\left(y_{it} + \left(\frac{1}{\alpha}\right)\right)}{\Gamma(y_{it}+1)\Gamma\left(\frac{1}{\alpha}\right)} \frac{(\alpha \mu_{it})^{y_{it}}}{(1+\alpha \mu_{it})^{y_{it} + \left(\frac{1}{\alpha}\right)}} & y_{it} > 0 \end{cases} \quad \text{(Eq. 4-1)}$$

where α is the dispersion parameter and Γ is the gamma function for the ZINB model.

Since the criteria to compare and select appropriate models depends on the presence and the source of overdispersion in the crash data, a non-nested likelihood ratio test can be used to check for the existence of overdispersion (Isgin et al., 2008). Specifically, the Vuong test can be used to examine the contribution of excess zeros in overdispersion (Vuong, 1989); it compares the zero-inflated models with single count models (Poisson and NB). When the value of the test is significant for the Poisson-based models, it indicates that only zero counts contribute to overdispersion and that the zero-inflated Poisson (ZIP) model is more appropriate than the regular Poisson model (Hosseinpour et al., 2013). When the value of the Vuong test is significant in the case of the NB-based model, it indicates that both excess zeros and heterogeneity account for overdispersion and a ZINB model is appropriate.

Estimation of the count frequency models in this fashion allowed the research team to better quantify the factors contributing to pedestrian crashes in two primary ways. First, as the NB or ZINB models are based on negative binomial distribution, these can better accommodate the high, natural variability of crash data than traditional modeling techniques based on the normal distribution. Second, using more years of data in the model allows the method to concentrate on the long-term expected crash frequency rather than short-term observed crash frequency, thus mitigating regression-to-the-mean bias (the issue of crash frequencies increasing or decreasing in years subsequent to low or high frequencies, respectively).

In this project, both Poisson and NB models were initially estimated. When comparing the two using a likelihood ratio test, the NB model was found to be a significantly better fit to the data. Therefore, both NB and ZINB models were subsequently estimated. When comparing the two using a Vuong test, the ZINB model was found to be a significantly better fit to the data. Therefore, this study reports on the results estimated from the calibrated ZINB model. Results for the NB model are also presented for the ease of developing other interpretations, as described in the following section.

4.4 Development of SPFs and CMFs

The Highway Safety Manual (HSM) [AASHTO, 2010] includes a series of methods for predicting crash frequencies for different transportation facilities. Safety performance functions (SPFs) and crash modification factors (CMFs) are two fundamental elements of the crash predictive methods described in HSM. Both SPFs and CMFs can be obtained by re-interpreting coefficients resulting from the estimation of crash frequency models, specifically NB models.

SPFs are regression equations that calculate the baseline predicted average crash frequency for a location, given a small set of basic characteristics about the location, including traffic volumes and – for road segments – the segment length. In the case of pedestrian crashes, the traffic volume measures of "exposure" included in an SPF are annual average daily pedestrian (AADP) volumes, as well as annual average daily traffic (AADT) volumes for the major and minor approaches, as shown in Eq. 4-2:

$$\boldsymbol{N_{base} = exp\left(\alpha + \beta_1 \times \ln(AADP) + \beta_2 \times \ln\left(AADT_{maj}\right) + \beta_3 \times \ln(AADT_{min})\right)} \quad \text{(Eq. 4-2)}$$

where:

- N_{base} = predicted pedestrian crash frequency at intersection for base conditions;
- $AADP$ = average annual daily pedestrian volume (pedestrians/day);
- $AADT_{maj}$ = average daily traffic volume for major road (vehicles/day);
- $AADT_{min}$ = average daily traffic volume for minor road (vehicles/day); and
- $\alpha, \beta_1, \beta_2, \beta_3$ = regression coefficients, obtained from an estimated NB model.

CMFs are ratios (centered around 1.00) representing how much crash frequencies could change (multiplicatively) with a change in a specific condition. These values are expressed as the ratio of the crash frequency for a location with specific characteristics divided by the crash frequency for a location with baseline characteristics, as shown in Eq. 4-2. A CMF is usually defined for a specific characteristic or change in characteristic: e.g., roadways with 10-foot lanes versus 12-foot lanes. In the HSM's predictive methods, it is assumed that multiple CMFs (each corresponding to a specific characteristic) can be multiplied together and by the baseline predicted average crash frequency to obtain the site-specific predicted average crash frequency, given a location's specific characteristics, as shown in Eq. 4-4.

$$\boldsymbol{CMF = \frac{N_{crash}\, for\ location\ with\ specific\ characteristic}{N_{crash}\, for\ location\ with\ baseline\ characteristic}} \quad \text{(Eq. 4-3)}$$

$$\boldsymbol{N_{crash} = N_{base} \times (CMF_1 \times CMF_2 \times \ldots CMF_n)} \quad \text{(Eq. 4-4)}$$

where,

- N_{base} = predicted pedestrian crash frequency at intersection for base conditions;
- N_{crash} = predicted pedestrian crash frequency at intersection for specific conditions;
- CMF_i = crash modification factor for characteristic i, obtained from an estimated NB model.

One way to obtain the coefficients in the SPF and the specific CMF values is to estimate a crash frequency model, specifically a negative binomial model. The SPF can be thought of as a restricted version of an NB model, where the only estimated coefficients are the measures of exposure (for an intersection model), and all other variables/coefficients are collapsed into the intercept term α using the baseline values of the variables. Thus, by assuming baseline values for other variables (e.g., number of lanes = 4), one can generate an SPF from the results of an estimated NB model. Similarly, CMF values can be obtained from the estimated coefficients of an NB model. Because of the functional form of the NB model, taking e (the exponential constant) to the power of the coefficient yields the estimated proportional change in the outcome (crash frequency) as a result of a unit change in the variable, which is itself a CMF. Thus, by assuming baseline values and specific changes in other variables (e.g., the number of lanes decreases from 4 to 2), one can generate a CMF from the results of an estimated NB model. Based on the results of the NB models, the next section interprets those results in terms of estimated SPFs and CMFs.

4.5 Model Estimation Results

This section reports on the results obtained from estimating crash frequency models of pedestrian crashes at signalized intersections in Utah, as a function of various explanatory variables. First, multiple ZINB and NB model estimation results are presented to account for different levels of data availability and different needs for applying the models' results. Second, the model results are interpreted by developing SPFs and CMFs following the predictive methods described in the HSM.

4.5.1 ZINB Model Results

As previously described in Chapter 3.0, complete data were not available for all 1,606 signalized intersections. One of the biggest sources of missing data were traffic volumes on the minor approach ($AADT_{MIN}$). Therefore, the research team decided to estimate two sets of models: one with as many explanatory factors as possible but fewer locations, and one with as many locations as possible but fewer explanatory factors. In the end, the only difference between the two models ended up being the use of the minor AADT variable, the elimination of which allowed for several hundred more locations to be included in the model.

For both datasets, a series of Poisson, negative binomial, and zero-inflated crash frequency models were estimated following a specific estimation process. First, all models were estimated using all possible explanatory variables. Second, the best fit model type was determined using tests for overdispersion and zero-inflation. In both cases, the data were significantly over-dispersed, indicating that NB models were better than Poisson models, and the ZINB models fit significantly better than the NB models as measured by a Vuong test. Third, the researchers used backwards elimination to remove variables that were not statistically significant from the model one by one, starting with the zero-inflated portion and moving on to the negative binomial portion. Elimination was stopped when all variables were at least marginally significant ($p < 0.10$). Thus, the results of the final ZINB models are presented in Table 4-4 and Table 4-5.

The following information may be useful when interpreting count data regression model results, like those in the following tables. The dispersion parameter α represents the degree to which the data are over-dispersed. A common measure of the goodness of fit of a model is the log-likelihood value, which is the natural logarithm of the likelihood function. The likelihood function is what is optimized when estimating a statistical model using maximum likelihood estimation, while adjusting the parameters (coefficients) so that they reproduce the observed data as best as possible. Log-likelihood values are always negative (indicating less than perfect fit), but their value has no interpretation on its own, only when compared to a "null" model that contains no independent (predictor) variables. McFadden's pseudo-R^2 value – one minus the ratio of the estimated model log-likelihood to the null model log-likelihood – is a way to measure the improvement in explained variability of the dependent (outcome) variable of the estimated model (containing many independent variables) over the null model. Like a regular R^2 value, it ranges from 0 (worst fit) to 1 (best fit), but it cannot be interpreted in exactly the same way. Typical pseudo-R^2 values for crash frequency models are typically less than 0.50, indicating that crashes are somewhat random events that cannot be perfectly predicted.

Table 4-4 ZINB Model A (N^1 = 1,038)

Variables	B^2	SE^3	z^4	p^5
Negative binomial portion				
(Intercept)	-6.8573	0.6995	-9.804	0.000
Measures of exposure				
Annual average daily pedestrian volume, estimated (AADP) [a]	0.4005	0.0387	10.352	0.000
Annual average daily traffic, major approaches ($AADT_{MAJ}$) [a]	0.4063	0.0722	5.624	0.000
Annual average daily traffic, minor approaches ($AADT_{MIN}$) [a]	0.0607	0.0212	2.866	0.004
Transportation system characteristics				
Intersection type (ref. = 4-leg)				
2-leg (mid-block)	-1.2396	0.7981	-1.553	0.120
3-leg	-0.2217	0.1507	-1.472	0.141
5-leg	-0.4915	0.5316	-0.925	0.355
Diverging diamond interchange (DDI)	-1.0314	1.0947	-0.942	0.346
Single point urban interchange (SPUI)	-0.5658	0.4457	-1.269	0.204
# crosswalks with continental, ladder, or zebra markings	0.1157	0.0360	3.219	0.001
Crosswalk length, mean (ft)	0.0041	0.0018	2.230	0.026
# approaches with no right-turn-on-red	-0.4995	0.2694	-1.854	0.064
# approaches with bike lanes	-0.0775	0.0288	-2.692	0.007
# of bus stops within 300 ft of intersection	0.1060	0.0237	4.472	0.000
Land use and built environment characteristics				
% land use vacant [b]	0.0099	0.0055	1.813	0.070
Employment density (1,000 per mi^2) [b]	-0.0099	0.0031	-3.176	0.002
Sociodemographic characteristics				
% of population with a disability [b]	0.0208	0.0079	2.648	0.008
% of population of Hispanic or non-white race/ethnicity [b]	0.0127	0.0025	5.007	0.000
Zero-inflated portion				
(Intercept)	4.0533	0.8469	4.786	0.000
Annual average daily pedestrian volume, estimated (AADP) [a]	-0.9666	0.2167	-4.462	0.000
Population density (1,000 per mi^2) [b]	-0.8187	0.1769	-4.627	0.000
% of population of Hispanic or non-white race/ethnicity [b]	0.0517	0.0169	3.062	0.002

[a] The natural log of these variables (+1) entered the model.
[b] These variables were measured using a quarter-mile network buffer.
Notes for this and future model results tables:
[1] *N* denotes the number of observations used in the model.
[2] *B* is the model estimated parameter used to infer about unknown population characteristics.
[3] *SE* denotes the standard error of the *B* estimate.
[4] *z* value is a Wald test statistic, which divides *B* by *SE*.
[5] *p*-value is the statistical significance of the Wald test.

Table 4-5 ZINB Model B (*N* = 1,441)

Variables	*B*	*SE*	*z*	*p*
Negative binomial portion				
(Intercept)	-6.3563	0.5582	-11.387	0.000
Measures of exposure				
Annual average daily pedestrian volume, estimated (AADP) [a]	0.4076	0.0337	12.108	0.000
Annual average daily traffic, major approaches ($AADT_{MAJ}$) [a]	0.4015	0.0558	7.194	0.000
Transportation system characteristics				
Intersection type (ref. = 4-leg)				
2-leg (mid-block)	-1.7309	0.7654	-2.261	0.024
3-leg	-0.1455	0.1272	-1.144	0.253
5-leg	-0.4678	0.5314	-0.880	0.379
Diverging diamond interchange (DDI)	-0.8080	1.1036	-0.732	0.464
Single point urban interchange (SPUI)	0.0010	0.2802	0.004	0.997
# crosswalks with continental, ladder, or zebra markings	0.1267	0.0330	3.843	0.000
Crosswalk length, mean (ft)	0.0044	0.0016	2.690	0.007
# approaches with no pedestrian crossing	-0.2087	0.0676	-3.087	0.002
# approaches with no right-turn-on-red	-0.4394	0.2472	-1.777	0.076
# approaches with bike lanes	-0.0680	0.0259	-2.632	0.008
# of bus stops within 300 ft of intersection	0.1465	0.0274	5.353	0.000
# approaches with near-side bus stops	-0.0917	0.0485	-1.892	0.058
Land use and built environment characteristics				
% land use vacant [b]	0.0105	0.0045	2.328	0.020
Employment density (1,000 per mi^2) [b]	-0.0089	0.0028	-3.168	0.002
# of schools [b]	-0.0806	0.0440	-1.833	0.067
# of places of worship [b]	-0.0787	0.0343	-2.297	0.022
Sociodemographic characteristics				
% of population with a disability [b]	0.0297	0.0068	4.342	0.000
% of population of Hispanic or non-white race/ethnicity [b]	0.0100	0.0022	4.634	0.000
Zero-inflated portion				
(Intercept)	5.3043	0.9371	5.661	0.000
Annual average daily pedestrian volume, estimated (AADP) [a]	-1.1678	0.2235	-5.226	0.000
# approaches with no pedestrian crossing	-0.6540	0.3406	-1.920	0.055
% land use industrial [b]	-0.0601	0.0229	-2.622	0.009
Population density (1,000 per mi^2) [b]	-0.8581	0.1550	-5.537	0.000
% of population of Hispanic or non-white race/ethnicity [b]	0.0637	0.0164	3.893	0.000

[a] The natural log of these variables (+1) entered the model.
[b] These variables were measured using a quarter-mile network buffer.

For ZINB Model A, with all possible explanatory variables but fewer locations (*N* = 1,038), the model yielded a dispersion parameter of $\alpha = 0.261$. The estimated model's log-likelihood was -1,626.1, compared to a null model (intercept-only Poisson model) log-likelihood of -2,414.8, yielding a McFadden pseudo-R^2 value of 0.327. This indicates that the ZINB Model A explains substantially more of the variance in vehicle-pedestrian collision frequency than an intercept-only Poisson model. For ZINB Model B, with more locations (*N* = 1,441) but without minor AADT, the model yielded a similar dispersion parameter of $\alpha = 0.251$. The estimated model's log-likelihood was -2,164.9, compared to a null model (intercept-only Poisson model) log-likelihood of -3,153.7, yielding a McFadden pseudo-R^2 value of 0.314. While the goodness of fit is not as strong as for the ZINB Model A, the goodness of fit is still substantially better than an intercept-only model. Since ZINB Model A was a slightly better fit to the data than ZINB Model B, the following describes the results of Model A primarily, with some mention of where Model B's results differ.

A distinctive feature of these models was the inclusion of pedestrian volumes (AADP), in addition to vehicular volumes (i.e. $AADT_{MAJ}$ and $AADT_{MIN}$), to account for measures of exposure. This specification of the models yielded notable results. The results suggested that pedestrian volume and both major and minor traffic volumes were significantly associated with pedestrian crashes. The associations of all the exposure measures were positive but less than one, indicating that pedestrian-vehicle collisions occurred more frequently at signalized intersections where the volumes of pedestrians and motor vehicle traffic on major and minor approaches were higher. The result implied that an increase in vehicle volumes on major and minor roads by 10% would be expected to increase the number of pedestrian crashes by 4.0% and 0.6% respectively. The result also implied that a 10% increase in pedestrian crossing volumes would be expected to increase the number of pedestrian crashes by around 4.0%. This supports the "safety in numbers" hypothesis because the pedestrian crash rate would go down with increasing pedestrian volumes (pedestrian crashes increase slower than pedestrian volumes). Overall, these findings are consistent with the existing literature, which suggest that both pedestrian and vehicular traffic exposure show positive associations with pedestrian-vehicle crashes (Harwood et al., 2008).

The model results also suggested several transportation system characteristics that were significantly associated with the frequency of pedestrian-vehicle collisions. In addition to other predictor variables, a dummy variable was introduced to investigate the variation of pedestrian crashes at different intersection types (e.g., two-/three-/five-leg intersection, DDI, SPUI) with respect to standard four-leg signalized intersections. Only two-leg intersections in Model B showed a significant and negative association (although it was also negative but not significant in Model A), which means that there were comparatively fewer pedestrian crashes at two-leg (e.g., mid-block crossing) signals compared to four-legged signalized intersections.

One expected finding is that mean crosswalk distance was significantly and positively associated with pedestrian crash frequency. Specifically, pedestrian crashes increased about 5% for every 12 ft of crossing distance; alternatively, pedestrian crashes would be expected to decrease by about 9–10% if a crossing were shortened by two lanes (24 ft), such as through the use of curb extensions. This finding is expected, since longer crossings expose pedestrians to more traffic lanes and for a longer amount of time, thus increasing the chances of a collision.

One group of significant predictors generating unexpected results was related to crosswalks and crossings. Intersections that had more approaches with pedestrian crossing restrictions saw fewer pedestrian crashes (in Model B). Specifically, an increase of one approach with no pedestrian crossings (i.e., imposing a road crossing restriction on a currently used approach) at an intersection would be expected to decrease pedestrian crashes by around 19% (Model B). Crosswalks are sometimes provided with high-visibility (longitudinal) continental, ladder, or zebra markings instead of standard (transverse) markings. But our results indicated that intersections with more continental, ladder, or zebra marked crosswalks instead of standard marked crosswalks saw more pedestrian crashes. Converting one standard marked crosswalk to other high-visibility markings (continental, ladder, or zebra) might increase pedestrian crashes by 12-14%, according to the models. However, the most comprehensive study on the topic by Zegeer et al. (2002) did not find any association between crosswalk marking pattern and pedestrian crashes. But, as discussed in Section 2.4, crash-based studies examining the effects of crossing markings are rare and offer conflicting results. This result could be a statistical artifact specific to this study's data and may not be reproduced in a different or future study.

The second group of significant transportation system characteristics was related to turn restrictions and other modes on the approaches. Signalized intersections where right turns on red were prohibited had fewer pedestrian crashes than would otherwise be expected; even one approach signed with a no-right-turn-on-red sign would be expected to decrease pedestrian crashes by around 37%; doing the same for all four approaches of a 4-leg signal could decrease pedestrian crashes by around 83-86%, according to the model results (however, no signals in the dataset exhibited this characteristic). This finding matches research about the benefits of reducing right-turn conflicts, especially when the right-turning vehicles have a red light, since drivers may not be looking for pedestrians in their path. Having bike lanes on approaches seemed to also decrease pedestrian crashes. Adding bike lanes to two approaches would be expected to reduce pedestrian crashes by 13-14%, depending on the model. This finding matches that of the crosswalk length, indicating that the presence of bike lanes could reduce the "effective" crossing distance for pedestrians, or at least the distance and time they are exposed to higher-speed and higher-mass motor vehicles.

The model results (in both Models A and B) suggest that intersections with more bus stops within 300 ft of the intersection saw more pedestrian crashes. This matches previous research finding a positive association between transit stops and pedestrian crashes. However, results from Model B shows that far-side bus stops were more strongly associated with pedestrian crashes than near-side bus stops. Moving two far-side bus stops to be near-side bus stops could reduce pedestrian crashes by 17%, according to the model results (Model B). This could be a finding specific to this study's dataset; perhaps Utah transit agencies are more likely to put far-side (instead of near-side) bus stops at signals on larger, higher-speed, and busier roadways, where there are higher volumes of right-turning traffic. However, when near-side bus stops are placed close enough to the intersection, right-turning vehicles cannot merge in front of stopped vehicles. Near-side stops also prohibit vehicles from entering opposing lanes to pass stopped transit vehicles. Both situations enable simpler access at crosswalks. Note that streets in Utah (especially those with traffic signals) tend to be wider than in many other locations in the US, so these findings may be different than in other states or regions.

Several land use and built environment characteristics were found to be significant in the models. Pedestrian crashes were more frequent at signals in areas with larger shares of vacant land uses. Specifically, 10% increases in vacant land uses would be expected to increase pedestrian crashes by 10–11%. There were slightly fewer pedestrian crashes in areas with greater concentrations of jobs (employment density). The presence of schools and places of worship within a quarter-mile walking distance of the signal were associated with fewer pedestrian crashes (only in Model B); specifically, a 7-8% reduction in pedestrian crashes for each additional nearby school or place or worship.

Among sociodemographic characteristics, a couple of variables were significantly associated with pedestrian crashes. There were more pedestrian crashes in neighborhoods with a greater share of people with disabilities and in areas with more people of Hispanic or non-White race/ethnicity. Specifically, neighborhoods with 1% more people with disabilities or Hispanic/non-White populations would be predicted to have 1-3% more pedestrian crashes.

Since these are ZINB models, they also contain a zero-inflated portion, which helps to predict the signals that would be expected to have zero pedestrian crashes by default. Several factors seemed to predict whether or not signalized intersections would see no crashes. Specifically, intersections with lower pedestrian volumes (in particular) were more likely to have no pedestrian crashes. In both models, signals with lower population density but greater shares of people of Hispanic or non-White race/ethnicity were also more likely to have zero crashes involving pedestrians. Finally, only in Model B, having zero pedestrian crashes was also associated with fewer approaches with crossing restrictions and lower percentages of industrial land uses.

4.5.2 NB Model Results

In order to provide more actionable results and findings for transportation agencies, the researchers estimated several additional pedestrian crash frequency models using a limited number of explanatory variables. Although many land use, built environment, and sociodemographic characteristics were significantly associated with pedestrian crash frequencies, most transportation agencies do not have the ability to manipulate or adjust those characteristics. As a result, CMFs developed using such information would be less actionable. Also, SPFs and CMFs should be developed from NB models, not ZINB models.

Therefore, the researchers estimated another set of Poisson and negative binomial crash frequency models (still on the same two datasets), following a similar estimation process as before. First, all models were estimated using the restricted set of explanatory variables (only measures of exposure and transportation system characteristics). Second, the best fit model type was determined using tests for overdispersion. In both cases, the data were significantly over-dispersed, indicating that NB models were better than Poisson models. Third, backwards elimination was used to remove variables that were not statistically significant from the model one by one. Elimination was stopped when all variables were at least marginally significant ($p <$

0.10). Thus, the results of the final restricted NB models are presented in Table 4-6 and Table 4-7.

Table 4-6 NB Model C (*N* = 1,111)

Variables	*B*	*SE*	*z*	*p*
Negative binomial portion				
(Intercept)	-7.6600	0.6293	-12.172	0.000
Measures of exposure				
Annual average daily pedestrian volume, estimated (AADP) [a]	0.4699	0.0289	16.282	0.000
Annual average daily traffic, major approaches ($AADT_{MAJ}$) [a]	0.4988	0.0631	7.900	0.000
Annual average daily traffic, minor approaches ($AADT_{MIN}$) [a]	0.0750	0.0199	3.760	0.000
Transportation system characteristics				
# crosswalks with continental, ladder, or zebra markings	0.1776	0.0368	4.820	0.000
# approaches with no pedestrian crossing	-0.2216	0.0696	-3.183	0.001
# approaches with bike lanes	-0.0711	0.0302	-2.356	0.018
# of bus stops within 300 ft of intersection	0.1765	0.0311	5.678	0.000
# approaches with near-side bus stops	-0.1173	0.0575	-2.039	0.041

[a] The natural log of these variables (+1) entered the model.

Table 4-7 NB Model D (*N* = 1,528)

Variables	*B*	*SE*	*z*	*p*
Negative binomial portion				
(Intercept)	-7.3251	0.5370	-13.641	0.000
Measures of exposure				
Annual average daily pedestrian volume, estimated (AADP) [a]	0.4967	0.0250	19.879	0.000
Annual average daily traffic, major approaches ($AADT_{MAJ}$) [a]	0.4851	0.0565	8.590	0.000
Transportation system characteristics				
# crosswalks with continental, ladder, or zebra markings	0.1722	0.0345	4.985	0.000
Crosswalk length, mean (ft)	0.0029	0.0016	1.822	0.068
# approaches with no pedestrian crossing	-0.1711	0.0538	-3.178	0.001
# approaches with bike lanes	-0.0664	0.0273	-2.432	0.015
# of bus stops within 300 ft of intersection	0.1555	0.0226	6.871	0.000

[a] The natural log of these variables (+1) entered the model.

For NB Model C, with a restricted set of explanatory variables but fewer locations (*N* = 1,111), the model yielded a dispersion parameter of $\alpha = 0.395$. The estimated model's log-likelihood was -1,748.8, compared to a null model (intercept-only Poisson model) log-likelihood of -2,543.9, yielding a McFadden pseudo-R^2 value of 0.312. This indicates that the NB Model C explains substantially more of the variance in vehicle-pedestrian collision frequency than an intercept-only Poisson model. For NB Model D, with more locations (*N* = 1,528) but without minor AADT, the model yielded a similar dispersion parameter of $\alpha = 0.427$. The estimated model's log-likelihood was -2,326.8, compared to a null model (intercept-only Poisson model) log-likelihood of -3,298.7, yielding a McFadden pseudo-R^2 value of 0.295. Since NB Model C was a slightly better fit to the data than NB Model D, the following describes the results of Model C primarily, with some mention of where Model D's results differ.

The NB models produced similar results to the ZINB models. Pedestrian volume (AADP) and traffic volume for both major and minor approaches ($AADT_{MAJ}$, $AADT_{MIN}$) were both significantly and positively associated with pedestrian crashes. Compared to ZINB models, these measures of exposure in the NB model had slightly stronger associations with pedestrian crashes. An increase in pedestrian/vehicle volumes by 10% would be expected to increase pedestrian crashes by around 4.7–5.0% (pedestrian volumes), 4.9–5.0% (traffic volumes on the major road), and 0.7% (traffic volumes on the minor road, Model A only). For other transportation system variables, the results from Models C and D were quite similar to those from Models A and B. Pedestrian crashes were more frequent at signals with longer average crossing distances, more crosswalks containing high-visibility continental, ladder, or zebra markings, fewer approaches with crossing restriction, no bike lanes, more bus stops.

4.5.3 Developed SPFs and CMFs

As discussed in Section 4.4, based on the NB model estimation results from the previous section, the equations and coefficients were adapted into the outputs used in the HSM's predictive methods into SPFs and CMFs. This involved assuming some baseline characteristics for variables other than measures of exposure. Specifically, the following baseline characteristics were assumed for a generic signalized intersection in Utah:

- Crosswalk length, mean (ft): 84 ft, corresponding to 5 lanes (12 ft each) plus 2 parking or turn lanes (12 ft each), also roughly corresponding to the average value of the mean crosswalk length at signals in the sample;
- # crosswalks with continental, ladder, or zebra markings: 0, so assuming that all 4 crosswalks have standard markings;
- # approaches with no pedestrian crossing: 0, assuming all approaches have pedestrian crossing (i.e., no restriction for pedestrians to cross the road) ;
- # approaches with bike lanes: 0, assuming no bike lanes;
- # of bus stops within 300 ft of intersection: 0, assuming no bus stops; and
- # approaches with near-side bus stops: 0, assuming no near-side bus stops.

Next, these baseline values were applied to the NB model coefficients estimated in Table 4-6 and Table 4-7 to adjust the intercept coefficient as described in Section 4.4. This generated the coefficients for the SPFs, as shown in Table 4-8 and the following equations. The first equation can be used if information on pedestrian volumes as well as traffic volumes on both major and minor approaches are available. The second equation can be used if information on pedestrian volumes is available, but information on traffic volumes is only available for the major approaches. The reason for adding +1 to the pedestrian/traffic volumes values is to ensure that when AADP or AADT is zero, the contribution to crash frequency will be zero ($\ln 1 = 0$). Also, recall that the output that these models are predicting is the 10-year pedestrian crash frequency, not the number of pedestrian crashes per year. (To obtain the long-run average pedestrian crash frequency per year, one would divide the output of these functions by 10.)

Table 4-8 SPF coefficients obtained from NB Models C and D

Coefficient	*Variable*	*SPF, Model C*	*SPF, Model D*
α	Intercept	-7.6600	-7.0815
β_1	Annual average daily pedestrian volume, estimated (AADP) [a]	0.4699	0.4967
β_2	Annual average daily traffic, major approaches ($AADT_{MAJ}$) [a]	0.4988	0.4851
β_3	Annual average daily traffic, minor approaches ($AADT_{MIN}$) [a]	0.0750	—

[a] Use the natural log of these variables (+1).

$$\boldsymbol{N_{baseC}} = \boldsymbol{e}^{[-7.66+0.4699\times\ln(AADP+1)+0.4988\times\ln(AADT_{maj}+1)+0.0750\times\ln(AADT_{min}+1)]} \quad \text{(Eq. 4-5)}$$

$$\boldsymbol{N_{baseD}} = \boldsymbol{e}^{[-7.0815+0.4967\times\ln(AADP+1)+0.4851\times\ln(AADT_{maj}+1)]} \quad \text{(Eq. 4-6)}$$

Next, for each of the explanatory variables that was not a measure of exposure, the researchers defined units of change that would be interpretable and convertible into a CMF. Following the procedures described in Section 4.4, the researchers multiplied the NB model coefficients estimated in Table 4-6 and Table 4-7 by these units of change, and then raised e to that power. The assumed units of change and the resulting CMFs are shown in Table 4-9. The CMFs from the two models were similar, so the averaged CMF values are also presented.

Table 4-9 Estimated CMFs with corresponding change in site conditions

Variable	*Change*	*CMF, Model C*	*CMF, Model D*	*CMF, averaged*
Crosswalk length, mean (ft)	−24	—	0.933	0.933
# crosswalks with continental, ladder or zebra markings [a]	+1	1.194	1.188	1.191
# approaches with no pedestrian crossing [a]	+1	0.801	0.843	0.822
# approaches with bike lanes	+2	0.867	0.876	0.872
# of bus stops within 300 ft of intersection	−2	0.703	0.733	0.718
# approaches with near-side bus stops [a]	+2	0.791	—	0.791

[a] These results are contrary to expectations or not supported by previous research. We do not recommend using these CMFs without additional research.

CMFs are centered around 1.00 and multiply the SPF-predicted number of crashes, so a number greater than 1.00 indicates an increase in crash frequency as a result of the change, while a number smaller than 1.00 indicates a decrease in crash frequency as a result of the change. The amount difference from 1.00 (in hundredths) can be interpreted as the percentage increase or decrease. Thus, the results suggest that reducing the mean crosswalk length by 24 ft (two 12-foot travel lanes) would be expected to decrease pedestrian crashes by 7%. Restricting pedestrians from crossing the road at a single approach of a signalized intersection would yield around an 18% reduction in pedestrian crashes, while converting a standard marked crosswalk into one with high-visibility continental, ladder, or zebra markings may increase pedestrian crashes by around 19%. Adding bike lanes to two of the approaches could reduce pedestrian crashes by around 13%. (Re-)moving two bus stops that were within 300 ft of a signalized intersection could reduce pedestrian crashes by 28%. Alternatively, moving two bus stops from the far-side to the near-side of the intersection might be expected to decrease pedestrian crashes by 21%.

We would urge caution when considering applying some of these CMFs in a predictive sense. As noted (previously, in Table 4-9, and in the following chapters), several of these findings – regarding crosswalk marking type, pedestrian crossing prohibitions, and near-side bus stops – are either contrary to expectations or not supported by previous research. We do not recommend using these specific CMFs at this time, and instead suggest conducting future research that investigates these findings.

4.6 Summary

The preliminary descriptive analysis, statistical model calibration, model estimation results, and their interpretation were presented in this chapter. First, the project team performed a descriptive analysis to identify the pedestrian crash hot spots and gain some preliminary insights on the characteristics associated with the hot spots. The analysis found significant differences in characteristics of signals with the most frequent pedestrian crashes and those with the highest crash rates. Then, the collected data were fed into two sets of Poisson, negative binomial, and zero-inflated models – one set with as many explanatory factors as possible but fewer locations, and one set with as many locations as possible but fewer explanatory factors – to investigate the effects of different explanatory variables related to exposure, road network characteristics, land use and built environment attributes, and sociodemographic characteristics on pedestrian crashes at signalized intersections. The chapter briefly introduced each crash frequency model and discussed the model calibration process. All of the count models (Poisson, negative binomial, and zero-inflated models) were evaluated using tests for overdispersion and zero-inflation to determine the best-fitting model. NB models showed a better fit than Poisson models to the study dataset. However, ZINB models were found the be the most appropriate for the dataset after performing the Vuong test.

Estimated results from ZINB models showed significant relationships between pedestrian crashes and different transportation characteristics, such as: crossing lengths, crossing prohibitions, crosswalk marking types, right-turning conditions, the presence of bike lanes, the placement of bus stops (far-side/near-side), several land use and built environment attributes, and sociodemographic characteristics. Another notable finding was that both the pedestrian exposure (estimated from pedestrian push-button data) and vehicle exposure showed expected positive associations with pedestrian crashes, while a "safety in numbers" effect was observed for pedestrian crashes and pedestrian volumes. The NB models only investigated the relationships between pedestrian crashes and transportation system characteristics, but they showed similar relationships as in the ZINB models. The model results suggested that pedestrian crashes were more frequent at signals with crosswalks containing high-visibility (continental, ladder, or zebra) marking, fewer approaches with crosswalk prohibition, no bike lanes, and more bus stops. Finally, the project team developed safety performance functions (SPFs) and crash modification factors (CMFs) with the results from the NB model, following the HSM safety predictive methods. This provided a set of estimated coefficients of SPFs and predicted changes in pedestrian crashes with a unit change in explanatory variables related to transportation system characteristics.

5.0 CONCLUSION

5.1 Summary

The overall goal of this research project was to explore different factors related to exposure, road network characteristics, and land use and built environment characteristics contributing to pedestrian-vehicle crashes at signalized intersections. With this primary goal, the study had three objectives:

1. To calibrate crash prediction models with a set of explanatory variables (including key road network facilities).
2. To develop improved pedestrian crash prediction models (SPFs and CMFs) at signalized intersections using pedestrian push-button measures of exposure.
3. To validate the “safety in numbers” concept for walking in a dataset consisting of a robust measure of pedestrian exposure.

Chapter 1.0 introduced the project, while Chapter 2.0 provided background material on the research topic and key limitations from previous studies. Chapter 3.0 described the data collection process, including obtaining, processing, and describing pedestrian and traffic exposure data, data related to road network characteristics, land use and built environment characteristics, and sociodemographic characteristics. Chapter 4.0 reported on the data analysis, including a brief introduction of the models, the model selection process, the ZINB and NB model results, and the development of SPFs and CMFs. The final Chapter 6.0 provides recommendations for implementation of the research findings. In this chapter, the researchers conclude by highlighting the major findings from the data collection and analyses, discussing the implication of the results for the “safety in numbers” hypothesis, and noting study limitations and challenges.

5.2 Key Findings

This study identified significant risk factors affecting pedestrian crash frequencies at 1,606 signalized intersections in Utah by analyzing 2,598 pedestrian crashes that occurred at those intersections from 2010 through 2019. Notably, the use of annual average daily pedestrian crossing volumes estimated from traffic signal data as the measure of pedestrian exposure facilitated a more robust model estimation process. After testing all of the count data models, Section 4.5 presented results from the NB and ZINB models, which accounted for different levels of data availability and different needs for applying the models' results. SPFs and CMFs were also developed to interpret the models' results following the HSM predictive methods. The following subsections highlight several key findings from these analyses, including insights about factors associated with pedestrian crashes and implications for the "safety in numbers" concept for pedestrians.

5.2.1 Factors Associated with Pedestrian Crashes at Signalized Intersections

The calibrated ZINB models developed as part of the crash frequency analysis showed that several characteristics of the road network, land use, built environment, and neighborhood sociodemographics were significantly associated with more (or fewer) pedestrian crashes. In addition, the simpler NB models created to develop SPFs and CMFs for pedestrian crashes at signalized intersections offer similar insights and ways to quantify potential impacts of changes to transportation and intersection elements. Here, we highlight findings related to: crossings and crosswalks, right-turn treatments, bike lanes, bus stops, land use / built environment characteristics, and sociodemographic characteristics.

Pedestrian crashes occurred more frequently at signalized intersections with longer average crossing distances. This is not surprising, since pedestrians are more exposed to motor vehicle traffic on longer crossings, which also take longer to cross, thus increasing the chance for a collision. The CMF for crosswalk length indicates that pedestrian crashes might decrease by 7% if the average crosswalk distance were reduced by 24 ft, the equivalent of two 12-foot travel or parking lanes. This suggests the feasibility for safety treatments like curb extensions to reduce pedestrian crashes; these countermeasures also have the secondary benefit of shortening crossing times, which can improve traffic signal performance.

Pedestrian crashes were also more frequent at signals with fewer approaches with prohibited crossings and at those with more high-visibility (longitudinal) continental, ladder, or zebra markings instead of standard (transverse) markings in the crosswalks. The CMFs would predict that restricting pedestrians from crossing a single approach of a signalized intersection would yield around an 18% reduction in pedestrian crashes, and replacement of a standard marking by a continental, ladder, or zebra marking pattern in crosswalks would yield a 19% increase in pedestrian crashes. However, previous research to support such findings is rare and has conflicting outcomes. In our opinion, this result does not mean that restricting pedestrian crossing or preventing the installation of more visible crosswalk markings (like continental, ladder, or zebra types) would be effective in reducing pedestrian crashes. Removing crosswalks would likely frustrate pedestrians, which could lead to riskier pedestrian crossing behaviors, and at a minimum this would greatly increase pedestrian delay and out-of-direction travel. Instead, this finding could be a spurious correlation or statistical artifact specific to our dataset or study area.

In the ZINB models, restrictions on right turns – the presence of no-right-turn-on-red (RTOR) signs – at intersections appeared to be greatly effective in preventing pedestrian crashes, with a model-predicted 37% reduction in pedestrian crashes for adding one right-turn restriction. We suspect that this variable's lack of significance in the NB models (resulting in no CMF) was due to a small sample size. Earlier studies suggest that prohibiting RTOR increases driver compliance with stop lines and reduces the number of drivers turning right on red without stopping (Retting et al., 2002). This finding makes sense, since right-turning drivers looking for a gap in traffic may block or not see pedestrians crossing in their path. RTOR restrictions and other countermeasures were found helpful in preventing more than 27,000 pedestrian crashes each year in the US (NHTSA, 1998).

A novel finding of this study was the link between the presence of bike lanes and pedestrian crashes. Based on both the ZINB models and the CMFs, the addition of bike lanes to two of the approaches could reduce pedestrian crashes by around 13%. In some respects, bike lanes reduce the portion of the crossing distance where pedestrians are exposed to higher-speed and higher-mass motor vehicles while crossing the road. Bike lanes may also provide better sightlines between people walking and driving, as well as a place for cars to wait and look for pedestrians before turning. Also, the presence of bike lanes could indicate other complete streets treatments, such as traffic calming devices, that have also been shown to improve pedestrian safety (LaPlante and McCann, 2008).

Another notable finding was that intersections with more bus stops (and with more far-side instead of near-side bus stops) also had more pedestrian crashes. This positive association between transit stops and pedestrian crashes is consistent with other research. According to the CMFs, removing two bus stops near a signalized intersection could reduce pedestrian crashes by 28%. However, placing bus stops near intersections is desirable from a pedestrian accessibility and walking-distance perspective. Additionally, moving two bus stops from the far-side to the near-side of the intersection would be expected to decrease pedestrian crashes by 21%. We did not find any previous research investigating the impact of bus stop placement on pedestrian crashes at signalized intersections, so we do not know whether this finding could be specific to the study dataset, or more indicative of a general trend. However, this finding could be related to other omitted variables; for example, perhaps bus stops on high-volume, high-speed roadways with large right-turn volumes (other risk factors for pedestrian crashes) are normally placed on the far-side of the intersection.

Several characteristics of land uses and the built environment were also linked to pedestrian crash frequency in one or more models. Past studies mostly focused on investigating the linkage of commercial, residential, or industrial land uses with pedestrian crashes (Ukkusuri et al., 2011). However, the project team did not find any research studying the effect of vacant land use on pedestrian crashes. Motorists might become less expectant of pedestrians in areas without any major establishments (e.g., business/cultural centers, parks, schools, places of worship). Areas with higher employment density had fewer crashes, which might similarly reflect driving behavior in places where pedestrians are expected. The presence of schools and places of worship within a quarter-mile walking distance of the signal was also indicative of fewer pedestrian crashes. This result is especially important for schools, since it could indicate that pedestrian safety treatments and initiatives in school zones – reduced speed zones, signage, flashing lights, crossing guards, and enforcement – are working and may be effective in other areas as well.

Finally, the analysis found pedestrian crashes were more frequent in areas with a greater share of people with disabilities and in areas with more people of Hispanic or non-White race/ethnicity. These findings are also supported by previous studies (Ukkusuri et al., 2011; Kay et al., 2014; Zegeer and Bushell, 2012). Specifically, pedestrians with disabilities might be at a greater risk when crossing roads. These findings suggest that a greater attention to pedestrian safety issues and a greater investment in pedestrian safety treatments may be warranted in neighborhoods with higher populations of people with disabilities or in communities of color.

5.2.2 Safety in Numbers

Recall that one of our study objectives was to examine the “safety in numbers” hypothesis for walking. As a reminder, this phenomenon occurs when the pedestrian crash rate (crashes per pedestrian) decreases as the pedestrian volume increases. In an NB or ZINB crash frequency model, a “safety in numbers” effect can usually be concluded if the coefficient for pedestrian volume is positive and less than one. But, Elvik (2013) argued that crashes involving pedestrians and vehicles depend both on pedestrian and vehicle volume. Hence, if the sum of the coefficients for pedestrian and vehicle volumes is greater than one, the data contains a partial “safety in numbers” effect: i.e., a “safety in numbers” effect is observed for pedestrians when the motor vehicle volume is kept constant. If the sum of the coefficients of pedestrian volume and vehicle volume on major and minor roads is less than one, this suggests a complete “safety in numbers” effect.

The unique use of robust measures of pedestrian exposure estimated from traffic signal data allowed this study to provide stronger insights into the “safety in numbers” concept for pedestrians at US signalized intersections. Specifically, across all models, the researchers found strong support for a “safety in numbers” effect for pedestrians: a 10% increase in pedestrian crossing volumes would be predicted to only increase pedestrian crash frequencies by around 4–5%. In other words, pedestrian crashes increase half as much as pedestrian volumes, thus leading to reduced crash rates (on a per-person basis) as pedestrian volumes increase. (Recall that pedestrian crash rates are crashes divided by volumes. If crashes increase slower than volumes, then the rate will decrease.) Figure 5-1 depicts this relationship in our dataset (compare to the theoretical Figure 2-2), where pedestrian crash rates (frequency divided by exposure) decline with increasing pedestrian volumes. Although this study’s analysis was not designed to uncover the reasons for the “safety in numbers” relationship, potential explanations assume that the more often motorists see people walking, the more likely they are to be aware of them and look out for them, and the more experience they have driving safely around non-motorized users. This key finding suggests that efforts to increase walking (through increased pedestrian volumes) will also provide greater safety and reduce the crash risk for any particular pedestrian. The non-linear slope of this curve also suggests that crash rates decrease the fastest at the lowest pedestrian volume intersections, which may suggest locations to target with pedestrian-enhancing efforts.

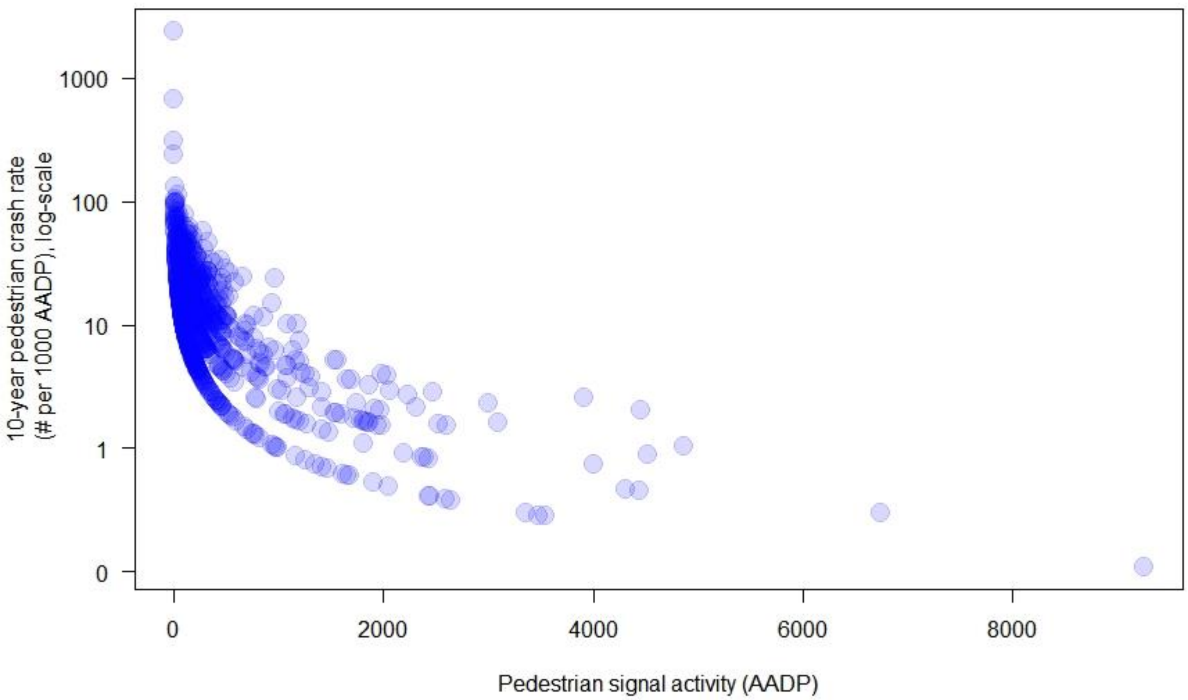


Figure 5-1 Demonstration of the "safety in numbers" effect for pedestrians at signals

5.3 Limitations and Challenges

This study was not without limitations that could be addressed through future work. Due to a lack of data in source databases (pedestrian exposure data at traffic signals, or other explanatory variables about locations), 608 signals (and 341 crashes at those signals) were excluded from the final dataset. A larger and more complete sample might have yielded slightly different results, especially if unobserved characteristics of the omitted locations were correlated with factors that contribute to pedestrian crashes. Fundamentally, the analysis method – in which the dependent/outcome variable was the frequency of pedestrian crashes over a 10-year period, but the independent/input variables were each measured at a single point in time – is a limitation driven by a lack of temporally-varying data. The built environment, sociodemographic, road characteristics, and pedestrian volume data were collected for a single time point or year, rather than over a 10-year period, and this time point was slightly different in each dataset. Factors such as household income, land use types, crosswalk marking/type/distance, the location of bus stops, or pedestrian volumes may have changed slightly (or even significantly) over the study period. Due to data limitations, we were unable to capture and account for these changes. Future work on this pedestrian safety topic should consider using multi-year data of predictor variables (if available) for a more comprehensive analysis. Finally, the results of this research are specific to Utah and its unique environment, laws, culture, and road user behaviors, which may limit the generalizability of findings to other states and regions. However, the concordance of our findings with those from previous studies suggests that this research does help provide more generalizable knowledge about factors influencing pedestrian safety at signalized intersections.

6.0 RECOMMENDATIONS AND IMPLEMENTATION

6.1 Recommendations

This research directly addresses UDOT's top strategic goal of zero fatalities by investigating safety issues regarding the most vulnerable road user group: pedestrians. This project satisfies the need for a comprehensive set of potential factors affecting pedestrian crashes at signalized intersections. The data set, comprised of transportation characteristics, land use and built environment factors, sociodemographic data, and (most notably) robust traffic signal-based measures of pedestrian exposure, allowed the researchers to develop useful crash frequency models and to draw important policy implications, including about the "safety in numbers" effect for pedestrians.

Specifically, findings related to transportation characteristics drawn from the explicit model development process provide the basis for recommended treatments and countermeasures that UDOT could prioritize to help reduce the risk of pedestrian crashes significantly in Utah. These recommendations include:

- **Shorten pedestrian crossing distances**: Consistent with expectations about longer exposure to potential conflicts with motor vehicle traffic, signals with longer average crossing distances saw slightly more pedestrian crashes, all else equal. This finding supports efforts to shorten crossing distances, such as through the use of curb extensions and other strategies. It is important to note that shorter crossing distances can also offer signal timing efficiencies in some circumstances. Shorter crossings (especially across the main street) would reduce the minimum green time needed for the side streets, thus providing flexibility for re-allocating green time to the major approaches. Future work can study the tradeoffs and ideal situations in which shorter pedestrian crossings could compensate for any negative impacts to motor vehicle mobility.

- **Implement complete streets treatments**: The finding that intersections with bike lanes also saw fewer pedestrian crashes could imply several things. First, it relates to the impact of shorter pedestrian crossing distances, since bike lanes can shorten the "effective" pedestrian crossing length of a roadway. Second, this lends support for other types of complete streets interventions and treatments to provide safe and comfortable spaces for all road users to use the street.

- **Prohibit right-turns-on-red in some cases**: Although "no right-turn-on-red" signs were present at only a few locations, they were strongly predictive of fewer pedestrian crashes, after controlling for all other factors. When selectively applied, prohibiting right-turns-on-red (RTOR) seems to be a promising strategy for improving pedestrian safety, especially in locations or at times of day with high volumes of pedestrians or high conflicting volumes of right-turning vehicles. Pedestrian push-button actuated blank-out signs could also prohibit right-turns-on-red only when pedestrians are present. Additional work could identify types of intersections or situations (e.g., volume profiles) where no RTORs can be implemented to improve pedestrian safety in ways that would not severely compromise vehicle operations.

- **Continue pedestrian safety efforts in school zones**: The fact that pedestrian crashes were lower at signals in areas near schools suggests that school zone treatments and initiatives to improve pedestrian safety may be working. These efforts – which include speed zones, high visibility signage/markings/signals, and crossing guards or enforcement – should be maintained and perhaps even expanded to other non-school areas with high pedestrian activity at specific times of day.

- **Focus pedestrian safety treatments in at-risk communities**: The positive association of areas with higher shares of the population with a disability or Hispanic or non-White race/ethnicity is troubling from an equity perspective. Greater efforts should be made to improve pedestrian safety and install proven pedestrian safety treatments, especially in those areas. At-risk communities could be prioritized when selecting projects to improve pedestrian safety. These findings can be combined with results from other past and ongoing UDOT research projects investigating the sociodemographics of crashes in Utah.

- **Encourage walking and increase pedestrian volumes**: Strong evidence for the "safety in numbers" effect for walking supports multifaceted efforts to increase walking and promote pedestrian activity in cities and communities. Increasing pedestrian volumes, especially in the lowest-volume locations, can help to make drivers aware of and increase their expectation of pedestrians, thus reducing pedestrian crash risk.

- **Study the effectiveness of different types of crosswalk markings**: Surprisingly, the more visible continental type of crosswalk marking was associated with more pedestrian crashes. However, the project team believes this finding may be due to a statistical artifact, and previous research to support these findings are rare. There are likely other research methods that are better suited to examining the safety effectiveness of different types of crosswalk markings, including human-factors research involving laboratory and field studies of crosswalk marking visibility in different conditions.

- **Study the placement of bus stops near signalized intersections in more detail**: The finding that far-side bus stops was more strongly associated with increased pedestrian crashes than near-side bus stops could inform the placement of transit stops at intersections. But, again, other research has yet to support this finding. A more detailed observational study of pedestrian behaviors surrounding near-side and far-side stops would be better able to identify specific design issues and considerations surrounding bus stop location at signalized intersections. Also, transit operations and pedestrian accessibility should play a major role in determining whether near-side or far-side stops are more efficient and effective.

6.1.1 Future Research

The rich dataset comprising transportation characteristics, land use and built environment factors, and sociodemographic characteristics allowed the project team to develop robust crash frequency models. Specifically, the inclusion of pedestrian exposure data estimated from one year of traffic signal pedestrian push-button data and 10 years of pedestrian crash data benefited explicit model development in an actionable form. Future work could include a larger data sample of pedestrian push-button activations (i.e., over 10 years, if available) for pedestrian exposure together with the same length of crash data in order to account for variation over time and to refine the estimate of the “safety in numbers” effect. A larger-scale study involving both signalized (and non-signalized) intersections and road segments with transportation characteristics, land use and built-environment, and sociodemographic data collected for multiple points in time could be undertaken in order to provide more insights in other locations. Also, future work could involve more advanced modeling techniques such as latent class models and random parameter models to account for heterogeneity and randomness in data. Before/after observational studies could also be beneficial towards investigating the specific impacts of road safety interventions in pedestrian crash risk. Additional information about actual lighting conditions or the presence of median refuge islands could be considered in these models.

At a more fine-grained level, individual pedestrian crashes (and the narratives from those crash reports) could be compared to available traffic signal controller log data to help identify and reconstruct sequences of events preceding such crashes. In some cases, traffic signal data could help to identify if motor vehicles ran a red light or failed to yield the right-of-way to pedestrians, or if a pedestrian crossed against a don't walk signal indication. However, care should be taken when generalizing results from this type of crash-by-crash analysis. Pedestrian crashes are rare events, and the conditions that led to one particular crash may not be the same conditions that lead to many crashes or more risky situations.

As a middle ground, one could use traffic signal controller log data in a more aggregate sense to investigate specific traffic flow and signal operational characteristics related with pedestrian safety at signalized intersections. Much of this could be related to permissive left- and right-turn conflicts with crossing pedestrians. One could correlate pedestrian volumes at specific times of day with left-turn phasing (and left-turn motor vehicle volumes) during those time periods to investigate if there are any relationships with pedestrian crashes. Similarly, if radar or other detectors were able to quantify right-turning motor vehicle volumes (including during different signal phases, e.g., on red vs. green), then those volumes could be correlated with pedestrian crossing volumes for specific conflicting crosswalks and associations with pedestrian crashes can be established.

6.2 Implementation Plan

This study identified several road network, land use, built environment, and sociodemographic characteristics that were associated with pedestrian crashes at signalized intersections, and developed Utah-specific SPFs and CMFs for pedestrian crashes at signalized intersections. In order to fully implement the outcomes of this research (a set of factors associated with pedestrian crashes, and recommended countermeasures and treatments), several general steps are possible.

First, the findings of this research can be integrated into UDOT's roadway safety management process. Specifically, the Utah-specific SPFs and CMFs developed for pedestrian crashes at signalized intersections can supplement or replace those contained within the Highway Safety Manual. When considering safety improvements along a corridor, in a study area, or even statewide, the SPFs can be applied to calculate the baseline predicted average crash frequency for a location, which can be further adjusted using CMFs – from this project, other research, or the CMF Clearinghouse (FHWA, n.d.) – as well as the recommended empirical-Bayes (EB) method. Such estimates are critical as performance metrics when performing the network screening step of the traditional roadway safety management process, because they help to identify sites with more pedestrian crashes than would otherwise be expected while controlling for local site characteristics as well as regression-to-the-mean. Such sites could thereafter be investigated in more detail for proposed pedestrian safety treatments.

Second, the proposed changes suggested by the models and described in the recommendations section above could be selectively implemented in locations where they may do the most good. This can be done through safety-specific processes (as described in the previous paragraph), or more gradually whenever an intersection is considered or proposed for changes or improvements. A pedestrian crash hot-spot analysis – such as was presented in Section 4.2 – could suggest locations most in need of pedestrian safety treatments. But other treatments – such as bike lanes or curb extensions – may be options to consider when road resurfacing is planned or when curb ramps need to be installed or replaced. Other intersection changes – such as bus stop placement or prohibiting right-turns-on-red – do not need to wait for specific roadway surfacing or construction projects and can be implemented with signage in some cases. The research team encourages UDOT to consider other ways to integrate these pedestrian safety improvements into existing planning, construction, operations, and maintenance processes.

Third, the effectiveness of the countermeasures and recommended treatments described above should be investigated using robust methods, including before/after analysis with treatment and control sites. In order to truly measure the pedestrian safety effectiveness in Utah of treatments like prohibiting right-turns-on-red or adding bicycle lanes, multiple years of crash data both before and after the installation of such treatments is required. Also, studying similar sites both with and without such treatments provides greater confidence that changes in pedestrian crash frequency are due to those specific treatments and not to other trends or factors. If such treatments show significant differences in pedestrian crash frequencies, then the effectiveness can be used in cost-benefit analyses to help prioritize specific applications of those treatments statewide. Overall, these efforts can lead to a more comprehensive and more reliable set of countermeasures, which can be implemented gradually for all traffic signals leading to an overall increase in safety conditions for pedestrians in Utah.

REFERENCES

American Association of State Highway and Transportation Officials (AASHTO). (2010). *Highway Safety Manual*. American Association of State Highway and Transportation Officials. http://www.highwaysafetymanual.org/

Blanc, B., Johnson, P., Figliozzi, M., Monsere, C., & Nordback, K. (2015). Leveraging signal infrastructure for nonmotorized counts in a statewide program: Pilot study. *Transportation Research Record: Journal of the Transportation Research Board*, *2527*, 69-79. https://doi.org/10.3141/2527-08.

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

Carlson, K., Murphy, B., Ermagun, A., Levinson, D., & Owen, A. (2018). *Safety in numbers: Pedestrian and bicyclist activity and safety in Minneapolis* (CTS 18-05). Center for Transportation Studies, University of Minnesota. https://hdl.handle.net/11299/194707

Cottrill, C. D., & Thakuriah, P. V. (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis & Prevention*, *42*(6), 1718-1728. https://doi.org/10.1016/j.aap.2010.04.012

Day, C. M., Taylor, M., Mackey, J., Clayton, R., Patel, S. K., Xie, G., ... & Bullock, D. (2016). Implementation of Automated Traffic Signal Performance Measures. *ITE Journal, 86*(8), 26–34.

Demetriades, D., Murray, J., Martin, M., Velmahos, G., Salim, A., Alo, K., & Rhee, P. (2004). Pedestrians injured by automobiles: relationship of age to injury type and severity. *Journal of the American College of Surgeons*, *199*(3), 382-387. https://doi.org/10.1016/j.jamcollsurg.2004.03.027

Dumbaugh, E., & Li, W. (2010). Designing for the safety of pedestrians, cyclists, and motorists in urban environments. *Journal of the American Planning Association*, *77*(1), 69-88. https://doi.org/10.1080/01944363.2011.536101

El-Basyouny, K., & Sayed, T. (2013). Safety performance functions using traffic conflicts. *Safety Science*, *51*(1), 160-164. https://doi.org/10.1016/j.ssci.2012.04.015

Elvik, R. (2013). Can a safety-in-numbers effect and a hazard-in-numbers effect co-exist in the same data? *Accident Analysis & Prevention*, *60*, 57-63. https://doi.org/10.1016/j.aap.2013.08.010

Elvik, R., & Goel, R. (2019). Safety-in-numbers: An updated meta-analysis of estimates. *Accident Analysis & Prevention*, *129*, 136-147. https://doi.org/10.1016/j.aap.2019.05.019

Ernst, M., & McCann, B. (2002). *Mean Streets 2002*. Surface Transportation Policy Project. http://www.transact.org/PDFs/ms2002/MeanStreets2002.pdf

Federal Highway Administration (FHWA). (2009). *Manual on Uniform Traffic Control Devices for Streets and Highways*. Federal Highway Administration, U.S. Department of Transportation. https://mutcd.fhwa.dot.gov/

Federal Highway Administration (FHWA). (2010). *Alternative Intersections/Interchanges: Informational Report (AIIR)*. Federal Highway Administration, U.S. Department of Transportation. https://www.fhwa.dot.gov/publications/research/safety/09060/009.cfm

Federal Highway Administration (FHWA). (2014). *Diverging Diamond Interchange Informational Guide*. Federal Highway Administration, U.S. Department of Transportation. http://safety.fhwa.dot.gov/

Federal Highway Administration (FHWA). (n.d.). *Crash Modification Factors Clearinghouse*. Federal Highway Administration, U.S. Department of Transportation. http://www.cmfclearinghouse.org/

Fitzpatrick, K., Iragavarapu, V., Brewer, M., Lord, D., Hudson, J. G., Avelar, R., & Robertson, J. (2014). *Characteristics of Texas pedestrian crashes and evaluation of driver yielding at pedestrian treatments* (FHWA/TX-13/0-6702-1). Texas A&M Transportation Institute. http://tti.tamu.edu/documents/0-6702-1.pdf

Geyer, J., Raford, N., Ragland, D., & Pham, T. (2005). The continuing debate about safety in numbers: Data From Oakland, CA. UC Berkeley: Safe Transportation Research & Education Center. https://escholarship.org/uc/item/0v85242q

Geyer, J., Raford, N., Pham, T., & Ragland, D. R. (2006). Safety in numbers: data from Oakland, California. *Transportation Research Record: Journal of the Transportation Research Board*, *1982*(1), 150-154. https://doi.org/10.1177/0361198106198200119

Gladhill, K., & Monsere, C. M. (2012). Exploring traffic safety and urban form in Portland, Oregon. *Transportation Research Record: Journal of the Transportation Research Board*, *2318*(1), 63-74. https://doi.org/10.3141/2318-08

Graham, D. J., & Glaister, S. (2003). Spatial variation in road pedestrian casualties: the role of urban scale, density and land-use mix. *Urban Studies*, *40*(8), 1591-1607. https://doi.org/10.1080%2F0042098032000094441

Harwood, D. W., Bauer, K. M., Richard, K. R., Gilmore, D. K., Graham, J. L., Potts, I. B., ... & Hauer, E. (2008). *Pedestrian safety prediction methodology* (NCHRP Web-Only Document 129: Phase III). Transportation Research Board. https://dx.doi.org/10.17226/23083

Hosseinpour, M., Prasetijo, J., Yahaya, A. S., & Ghadiri, S. M. R. (2013). A comparative study of count models: Application to pedestrian-vehicle crashes along Malaysia federal roads. *Traffic Injury Prevention*, *14*(6), 630-638. https://doi.org/10.1080/15389588.2012.736649

Hu, L., Wu, X., Huang, J., Peng, Y., & Liu, W. (2020). Investigation of clusters and injuries in pedestrian crashes using GIS in Changsha, China. *Safety Science*, *127*, 104710. https://doi.org/10.1016/j.ssci.2020.104710

Isgin, T., Bilgic, A., Forster, D. L., & Batte, M. T. (2008). Using count data models to determine the factors affecting farmers' quantity decisions of precision farming technology adoption. *Computers and Electronics in Agriculture*, *62*(2), 231-242. https://doi.org/10.1016/j.compag.2008.01.004

Jacobsen, P. L. (2015). Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury Prevention*, *21*(4), 271-275. http://dx.doi.org/10.1136/ip.9.3.205rep

Kerridge, J., Hine, J., & Wigan, M. (2001). Agent-based modelling of pedestrian movements: the questions that need to be asked and answered. *Environment and Planning B: Planning and Design*, *28*(3), 327-341. https://doi.org/10.1068/b2696

Kothuri, S., Nordback, K., Schrope, A., Phillips, T., & Figliozzi, M. (2017). Bicycle and pedestrian counts at signalized intersections using existing infrastructure: Opportunities and challenges. *Transportation Research Record: Journal of the Transportation Research Board*, *2644*, 11–18. https://doi.org/10.3141/2644-02

LaPlante, J., & McCann, B. (2008). Complete streets: We can get there from here. *ITE Journal*, *78*(5), 24.

Lee, C., & Abdel-Aty, M. (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. *Accident Analysis & Prevention*, *37*(4), 775-786. https://doi.org/10.1016/j.aap.2005.03.019

Lindsey, G., Tao, T., Wang, J., & Cao, J. (2019). *Pedestrian and bicycle crash risk and equity: Implications for street improvement projects* (CTS 19-15). Center for Transportation Studies, University of Minnesota. https://hdl.handle.net/11299/203635

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, *44*(5), 291-305. https://doi.org/10.1016/j.tra.2010.02.001

Loukaitou-Sideris, A., Liggett, R., & Sung, H. G. (2007). Death on the crosswalk: A study of pedestrian-automobile collisions in Los Angeles. *Journal of Planning Education and Research*, *26*(3), 338-351. https://doi.org/10.1177/0739456X06297008

Miranda-Moreno, L. F., & Fu, L. (2006). A comparative study of alternative model structures and criteria for ranking locations for safety improvements. *Networks and Spatial Economics, 6*(2), 97-110. https://doi.org/10.1007/s11067-006-7695-2

National Highway Traffic Safety Administration (NHTSA). (2019a). *Traffic safety facts: Pedestrians* (DOT HS 812 681). National Highway Traffic Safety Administration: U.S. Department of Transportation. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812681

National Highway Traffic Safety Administration (NHTSA). (2019b). *2018 Fatal motor vehicle crashes: Overview* (DOT HS 812 826). National Highway Traffic Safety Administration: U.S. Department of Transportation. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812826

Numetric. (n.d.). *Numetric*. Utah Department of Transportation. https://udot.numetric.com/signin

Pulugurtha, S. S., & Sambhara, V. R. (2011). Pedestrian crash estimation models for signalized intersections. *Accident Analysis & Prevention*, *43*(1), 439-446.

Raford, N., & Ragland, D. R. (2005). *Pedestrian volume modeling for traffic safety and exposure analysis*. UC Berkeley: Safe Transportation Research & Education Center. https://escholarship.org/uc/item/9cn8d3nq

Raford, N., & Ragland, D. (2006). *Pedestrian volume modeling for traffic safety and exposure analysis: The case of Boston, Massachusetts*. UC Berkeley: Safe Transportation Research & Education Center. https://escholarship.org/uc/item/61n3s4zr

Retting, R. A., Nitzburg, M. S., Farmer, C. M., & Knoblauch, R. L. (2002). Field evaluation of two methods for restricting right turn on red to promote pedestrian safety. *ITE Journal*, *72*(1), 32-36.

Rivara, F. P., & Barber, M. (1985). Demographic analysis of childhood pedestrian injuries. *Pediatrics*, *76*(3), 375-381.

Singleton, P. A., Runa, F., & Humagain, P. (2020). *Utilizing archived traffic signal performance measures for pedestrian planning and analysis* (UT-20.17). Salt Lake City, UT: Utah Department of Transportation. https://drive.google.com/file/d/1AwLf1DZVw0Vj-btPl5eoWe0UOw9TyFwq/view?usp=sharing

Smith, S. A. (2015). The Cities of Zion? Mormon and non-Mormon town plans in the US Mountain West, 1847–1930. *Journal of Historical Geography*, *50*, 1-13. https://doi.org/10.1016/j.jhg.2015.03.003

Sturdevant, J. R., Overman, T., Raamot, E., Deer, R., Miller, D., Bullock, D. M., ... & Remias, S. M. (2012). *Indiana traffic signal hi resolution data logger enumerations*. Purdue University. http://dx.doi.org/10.4231/K4RN35SH

Taylor, M., & Mackey, J. (2018). *Automated Traffic Signal Performance Measures*. Presented at the 2018 UDOT Annual Conference, Sandy, UT.

Ukkusuri, S., Hasan, S., & Aziz, H. 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

Utah Department of Public Safety (UDPS). (2020). *Pedestrian Involved*. Utah Department of Public Safety. https://udps.numetric.net/pedestrian-involved#/

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, *57*(2), 307-333. https://doi.org/10.2307/1912557

Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., & Bhatia, R. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. *Accident Analysis & Prevention*, *41*(1), 137-145. https://doi.org/10.1016/j.aap.2008.10.001

Xu, P., Xie, S., Dong, N., Wong, S. C., & Huang, H. (2019). Rethinking safety in numbers: are intersections with more crossing pedestrians really safer? *Injury Prevention*, *25*(1), 20-25. http://dx.doi.org/10.1136/injuryprev-2017-042469

Zegeer, C. V., Opiela, K. S., & Cynecki, M. J. (1985). *Pedestrian signalization alternatives. Final report* (FHWA-RD-83-102). Federal Highway Administration, U.S. Department of Transportation. https://rosap.ntl.bts.gov/view/dot/41946

Zegeer, C. V., Stewart, J. R., Huang, H. H., & Lagerwey, P. A. (2002). *Safety effects of marked vs. unmarked crosswalks at uncontrolled locations: Executive summary and recommended guidelines* (FHWA-RD-01-075). Federal Highway Administration, U.S. Department of Transportation. https://rosap.ntl.bts.gov/view/dot/15480

Zegeer, C. V., Stewart, J. R., Huang, H. H., Lagerwey, P. A., Feaganes, J. R., & Campbell, B. J. (2005). *Safety effects of marked versus unmarked crosswalks at uncontrolled locations: Final report and recommended guidelines* (FHWA-HRT-04-100). Federal Highway Administration, U.S. Department of Transportation. https://rosap.ntl.bts.gov/view/dot/40068

Zegeer, C. V., & Bushell, M. (2012). Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis & Prevention*, *44*(1), 3-11. https://doi.org/10.1016/j.aap.2010.12.007