GUIDE FOR SCALABLE RISK ASSESSMENT METHODS FOR PEDESTRIANS AND BICYCLISTS

Publication No. FHWA-SA-18-032
July 2018

U.S. Department of Transportation
Federal Highway Administration

Safe Roads for a Safer Future
Investment in roadway safety saves lives
http://safety.fhwa.dot.gov

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-SA-18-032		

4. Title and Subtitle GUIDE FOR SCALABLE RISK ASSESSMENT METHODS FOR PEDESTRIANS AND BICYCLISTS	5. Report Date July 2018
	6. Performing Organization Code
7. Author(s) Shawn Turner, Ipek Sener, Michael Martin, L.D. White, Subasish Das, Robert Hampshire, Mike Colety, Kay Fitzpatrick, and Ravi Wijesundera	8. Performing Organization Report No.
9. Performing Organization Name and Address Texas A&M Transportation Institute, College Station, TX 77843-3135 University of Michigan Transportation Research Institute Kimley-Horn and Associates, Inc.	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. DFTH6116D00004, TTI Task Order #2
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Highway Administration Office of Safety 1200 New Jersey Ave SE Washington, DC 20590	13. Type of Report and Period Covered Task 4 Guide: March 2017 - July 2018
	14. Sponsoring Agency Code

15. Supplementary Notes
Government Task Manager: Tamara Redmon

16. Abstract

This guide describes scalable risk assessment methods for pedestrians and bicyclists, wherein risk is a measure of the probability of a crash to occur given exposure to potential crash events. This guide outlines eight sequential steps to develop risk values at various desired geographic scales, and describes the scope and nature of each step, including any guiding principles.

Exposure to risk is an integral element of risk, and as such, an integral element of the pedestrian and bicyclist risk assessment methods in this guide. Exposure is a measure of the number of potential opportunities for a crash to occur, and is often directly related to the number of people who walk and bike. This guide provides information on analytic methods to estimate pedestrian and bicyclist exposure, and identifies other relevant guides and resources that provide supplemental information for exposure estimation. Depending upon the desired geographic scale, one or more of these three analytic methods can be used to estimate pedestrian and bicyclist exposure: 1) site counts; 2) travel demand estimation models (several different types); and, 3) travel surveys.

17. Key Word	18. Distribution Statement
Pedestrian and bicyclist risk assessment, exposure, geographic scale, risk factors, pedestrian and bicyclist counts, demand estimation, national travel surveys	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	122	

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

TABLE OF CONTENTS

LIST OF FIGURES iii
LIST OF TABLES iv
LIST OF ABBREVIATIONS v
GLOSSARY vi
SUMMARY ix
INTRODUCTION 1
Motivation for the Guide 1
Intended Audience 1
Organization of the Guide 2
STEP 1. DETERMINE USE(S) OF RISK VALUES 4
Identifying Objectives of Risk Assessment 4
Typical Uses of Risk Values for Pedestrians and Bicyclists 4
Selecting Analysis Parameters in Remaining Steps of Risk Assessment 10
Assessing Available Data Resources and Analytic Capabilities 12
STEP 2. SELECT GEOGRAPHIC SCALE 13
Considerations When Selecting Geographic Scale 14
Multiple Analysis Scales and Aggregation Considerations 15
STEP 3. SELECT RISK DEFINITION 20
Definition 1: Observed Crash Rate 21
Definition 2: Expected Crashes 23
Definition 3: Additional Risk Indicators 24
STEP 4. SELECT EXPOSURE MEASURE 26
STEP 5. SELECT ANALYTIC METHOD TO ESTIMATE EXPOSURE 32
Site Counts 34
Demand Estimation Models 34
Travel Surveys 40
STEP 6. USE ANALYTIC METHOD TO ESTIMATE SELECTED EXPOSURE MEASURE 42
Site Counts 42
Demand Estimation Models 46
Travel Surveys 65
STEP 7. COMPILE OTHER REQUIRED DATA 89
Observed Crash Rate 89
Expected Crashes 89
Additional Risk Indicators 90
STEP 8. CALCULATE RISK VALUES 92
Case Study: Pedestrian Risk Assessment in Michigan 92
Case Study Example for Arizona DOT Bicycle Safety Action Plan 97
BIBLIOGRAPHY 101

LIST OF FIGURES

Figure S-1. Eight Steps for Scalable Risk Assessment for Pedestrians and Bicyclists.................................. x
Figure 1. Eight Steps for Scalable Risk Assessment for Pedestrians and Bicyclists...................................... 3
Figure 2. Example of Pedestrian Fatality Risk in 50 Cities.. 6
Figure 3. Example of Pedestrian Risk Analysis in California Using Household Travel Surveys 7
Figure 4. Example of Site Evaluation: Exclusive Pedestrian Phase at Single Intersection 10
Figure 5. Illustration of Geographic Scales for Exposure Estimation 14
Figure 6. Illustration of Combining Individual Crosswalks for Intersection Exposure 16
Figure 7. Illustration of Estimating Intersection Exposure from Segment Counts 16
Figure 8. Illustration of Multiple Parallel Facilities in a Corridor 17
Figure 9. Illustration Showing Count Segments on Multiple Parallel Facilities in a Corridor 17
Figure 10. Illustration of Multiple Segments Being Combined into Single Longer Segment...................... 18
Figure 11. Illustration of Multiple Census Tracts Being Combined into Single Region 19
Figure 12. Historical Trends for Observed Bicycle Crash Rates in Wisconsin 22
Figure 13. Locations in Seattle with High Risk of Pedestrian Crashes based on Expected Crashes 24
Figure 14. Calculating Distance Traveled Exposure Measure from Segment Counts................................ 29
Figure 15. Calculating Time Traveled by Expanding Household Travel Surveys 29
Figure 16. Estimating Count-Based Exposure at Intersections 30
Figure 17. Direct Demand Model Development Process 48
Figure 18. Direct Demand Model Development Detailed Flow Chart 49
Figure 19. Household Survey Data Relationships 69
Figure 20. Expanding Household Samples to Represent All Households in a Region................................ 69
Figure 21. Annual Non-Motorized Exposure Estimates, Non-Motorized Fatalities, and Risk 70
Figure 22. Interface of the Statewide Exposure Estimates Tool Component 73
Figure 23. Interface of the MPO Area Exposure Estimates Tool Component 76
Figure 24. Geography Comparison 77
Figure 25. ACS Data for Texas 82
Figure 26. Joining of ACS Commuting Data with ACS Shapefile Data 82
Figure 27. Map for Bicycle Commuting in Texas Census Tracts.. 83
Figure 28. State-Level Annual Person Trips in NHTS Online Analysis Tool 85
Figure 29. Table Designer Custom Query Tool for the 2009 NHTS... 86
Figure 30. Table Designer Custom Query Tool for the 2017 NHTS... 87
Figure 31. Step 5 Demand Estimation Using Pedestrian Trip Generation and Flow Models 93
Figure 32. Pedestrian Trip Generation Model Results for home-based other trips 94
Figure 33. Daily pedestrian trips made per PAZ for Wayne county Michigan.. 94
Figure 34. Step 7: Compile other required data 95
Figure 35. Risk measured as the expected number of crashes per PAZ 95
Figure 36. Level-of-Service-of-Safety (LOSS) map derived from the risk values.. 96
Figure 37. Higher-Risk Locations for Bicyclists on the Arizona State Highway System 100

LIST OF TABLES

Table S-1. Four Scale Categories in the Scalable Risk Assessment Process xii
Table S-2. Three Possible Definitions of Pedestrian and Bicyclist Risk xiii
Table S-3. Selection Matrix for Exposure Measures xiv
Table S-4. Selection Matrix for Analytic Methods to Estimate Exposure xv
Table 1. Example of Pedestrian Fatality Risk by Age and Gender 5
Table 2. Example of Systemic Safety Analysis that Includes Exposure Variables 8
Table 3. Example of Multi-Criteria Project Prioritization 8
Table 4. Example of Exposure Measures Included in Countermeasure Evaluation 9
Table 5. Selecting Key Parameters Based on Use(s) of Risk Values 11
Table 6. Four Scale Categories in the Scalable Risk Assessment Process 13
Table 7. Illustration of Combining BMT for Multiple Parallel Facilities in a Corridor 17
Table 8. Three Possible Definitions of Pedestrian and Bicyclist Risk 20
Table 9. Historical Trends for Observed Bicycle Crashes in Wisconsin 22
Table 10. Example of Ranking Signalized Intersections based on Multiple Risk Indicators 25
Table 11. Selection Matrix for Exposure Measures 27
Table 12. Strengths and Limitations for Exposure Measure Categories 28
Table 13. Extracting Population-Based Exposure from ACS Data Tables 31
Table 14. Selection Matrix for Analytic Methods to Estimate Exposure 33
Table 15. Key Guidance Documents for Site Counts 43
Table 16. Advantages and Disadvantages of Direct Demand Models 47
Table 17. Sampling Techniques 51
Table 18. Key Explanatory Variables of Pedestrian and Bicyclist Direct Demand Models 56
Table 19. Examples of Recent Direct-Demand Models to Estimate Pedestrian and Bicycle Volumes 59
Table 20. Strengths and Limitations of the ACS in Estimating Pedestrian and Bicyclist Travel 65
Table 21. 2017 NHTS State DOT & MPO Add-on Household Sample Sizes 68
Table 22. Bike - CBSA Peer Groupings 79
Table 23. Walk - CBSA Peer Groupings 80
Table 24. ACS Data Attributes for Pedestrian and Bicycle Commute Estimates from Table B08301 81
Table 25. Structure of 2009 and 2007 NHTS Data Files 84
Table 26. 2017 NHTS Annual Person Miles of Travel (millions) by Mode and State 88
Table 27. Eight Steps for the Scale Risk Assessment Methodology as Applied to this Case 92
Table 28. Summarizes the Factors and Scoring for the Risk Assessment Process 98
Table 29. Bicyclist/Motor-Vehicle Risk Assessment Levels 99

LIST OF ABBREVIATIONS

AADT	Annual average daily traffic
ACS	American Community Survey
BMT	Bicycle-miles of travel
CBSA	Core based statistical areas
CHTS	California Household Travel Survey
DOT	Department of Transportation
FARS	Fatality Analysis Reporting System
FHWA	Federal Highway Administration
GIS	Geographic information system
HAWK	High intensity Activated crossWalK
HCM	Highway Capacity Manual
HSM	Highway Safety Manual
MEP	Million entering pedestrians
MEV&P	Million entering vehicles and pedestrians
MnDOT	Minnesota Department of Transportation
MoPeD	Model of Pedestrian Demand
MPO	Metropolitan planning organization
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NHTS	National Household Travel Survey
PBCAT	Pedestrian and Bicycle Crash Analysis Tool
PMT	Pedestrian-miles of travel
SCAG	Southern California Association of Governments
TAZ	Traffic analysis zone
TIGER/Line	Topologically integrated geographic encoding and referencing
TDM	Travel demand model
T-MAP	Trail Modeling and Assessment Platform
VMT	Vehicle-miles of travel

GLOSSARY

Term	Definition	Source
Areas	Areas consist of an interconnected set of transportation facilities serving movements within a specified geographic space, as well as movements to and from adjoining areas. The primary factor distinguishing areas from corridors is that the facilities within an area need not be parallel to each other.	2010 Highway Capacity Manual (HCM)
Areawide	Generic term that includes all geographic scales that are not facility-specific, such as neighborhood, network, system, region, city, state, etc.	This Guide
Census block	The smallest entity for which the Census Bureau collects and tabulates decennial census information; bounded on all sides by visible and nonvisible features shown on Census Bureau maps.	U.S. Census Bureau
Census block group	A combination of Census blocks that is a subdivision of a census tract.	U.S. Census Bureau
Census tract	A small, relatively permanent statistical subdivision of a county in a metropolitan area or a selected nonmetropolitan county. Census tract boundaries normally follow visible features, but may follow governmental unit boundaries and other nonvisible features in some instances; they always nest within counties. Designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions, census tracts usually contain between 2,500 and 8,000 inhabitants.	U.S. Census Bureau
Corridors	Corridors are generally a set of parallel transportation facilities designed to move people between two locations. For example, a corridor may consist of a freeway facility and one or more parallel urban facilities.	2010 HCM
Direct demand model	A statistical model that estimates facility-specific pedestrian and bicyclist volumes based on observed volumes at a sample of locations and nearby context (such as land use and form, street type, etc.). Direct demand models are often based on regression analysis.	NCHRP Report 770, Research Team
Exposure	Measure of the number of potential opportunities for a crash to occur. This theoretical definition has been quantified or estimated many different ways in practice.	This Guide
Exposure scale	The granularity of the geographic level for which an exposure measure is desired.	This Guide
Facilities	Facilities are lengths of roadways, bicycle paths, and pedestrian walkways composed of a connected series of points and segments. The HCM defines freeway facilities, multilane highway facilities, two-lane highway facilities, urban street facilities, and pedestrian and bicycle facilities.	2010 HCM
Network	A geographic scale (mentioned in the original FHWA Statement of Work) that is most comparable to the term Area as defined in the 2010 HCM.	This Guide

Term	Definition	Source
Points	Points are places along a facility where (a) conflicting traffic streams cross, merge, or diverge; (b) a single traffic stream is regulated by a traffic control device; or (c) there is a significant change in the segment capacity (e.g., lane drop, lane addition, narrow bride, significant upgrade, start or end of a ramp influence area).	2010 HCM
Region	A geographic scale that is most comparable to the term System as defined in the 2010 HCM.	This Guide
Risk	Measure of the probability of a crash to occur given exposure to potential crash events. This theoretical definition has been quantified or estimated by dividing the expected or measured number of crashes by exposure.	This Guide
Risk factor	Any attribute or characteristic that increases the likelihood of a negative safety outcome (e.g., crash or fatality).	This Guide
Segment	A segment is the length of roadway between two points. Traffic volumes and physical characteristics generally remain the same over the length of a segment, although small variations may occur (e.g., changes in traffic volumes on a segment resulting from a low-volume driveway).	2010 HCM
Sketch planning	Methods to estimate existing or future demand that are simpler alternatives to developing complex travel demand models. Often, these methods are implemented in spreadsheets or geographic information systems (GIS) using existing travel survey and other data.	This Guide
System	Systems are composed of all the transportation facilities and modes within a particular region.	2010 HCM
Traffic analysis zone (TAZ)	A common areawide geography that are defined by metropolitan planning organizations (MPO) for use in their travel demand forecasting models. TAZ are typically composed of multiple Census blocks.	This Guide
Travel demand model	A computerized process that estimates existing and future travel demand (often on a citywide or regional basis) given numerous inputs, such as the transportation network, population and demographic characteristics, and trip-making behavior. The end result of a travel demand model is traffic volume estimates on individual transportation network links.	This Guide
Travel survey	A systematic effort to collect information about individual travel behavior. Travel surveys are typically collected from a statistical sample of travelers for a specified day or days (not an entire month or year), and typically gather aggregate trip information (travel mode, trip purpose, trip start and end location, trip length or time, etc.).	This Guide
Work trip	Travel from home to work (also known as commuting). In their Journey to Work surveys, the U.S. Census Bureau collects trip information for only work trips. Trips that have a non-work purpose are collected by FHWA's National Household Travel Survey (NHTS) and other regional household travel surveys (when administered).	This Guide

ACKNOWLEDGMENTS

The authors acknowledge this project's Expert Panel members, who provided feedback and comments on several drafts of this guide:

- Jennifer Boldry, People for Bikes.
- David Henderson, Miami-Dade Metropolitan Planning Organization.
- Betsy Jacobsen, Colorado Department of Transportation.
- Greg Krykewycz, Delaware Valley Regional Planning Commission (Philadelphia).
- Greg Lindsey, University of Minnesota.
- Krista Nordback, University of North Carolina Highway Safety Research Center.
- Michael Medina, El Paso Metropolitan Planning Organization.
- Jamie Parks, San Francisco Metropolitan Transportation Authority.
- Bob Schneider, University of Wisconsin-Milwaukee.

SUMMARY

This guide describes scalable risk assessment methods for pedestrians and bicyclists, wherein risk is a measure of the probability of a crash to occur given exposure to potential crash events. This guide:

- Outlines eight sequential steps to develop risk values at various desired geographic scales.
- Describes the scope and nature of each step, including any guiding principles.
- Provides information on analytic methods to estimate pedestrian and bicyclist exposure.
- Identifies other relevant guides and resources that provide supplemental information.

Motivation for the Guide

Many transportation agencies are placing more emphasis on improving pedestrian and bicyclist safety and reducing the risk of a fatality or serious injury to pedestrians and bicyclists. Practitioners need a methodical approach to assess pedestrian and bicyclist risk for the purposes of identifying high-priority areas and transportation facilities for safety improvement, evaluating specific countermeasures and locations before and after improvements are made, and tracking safety performance measures over time to gauge progress toward established goals. The motivation for this guide is to provide this methodical approach for assessing pedestrian and bicyclist risk for these and other applications.

Exposure to risk is an integral element of risk, and as such, an integral element of risk assessment methods in this guide. Exposure is a measure of the number of potential opportunities for a crash to occur, and is often directly related to the number of people who walk and bike. In the past, exposure has seldom been included in pedestrian and bicyclist safety analysis because of the practical challenges of collecting or estimating exposure data. Another motivation for this guide is to outline feasible methods to calculate or estimate exposure, such that exposure will be included more often in pedestrian and bicyclist risk assessment.

Intended Audience

The main audience for this guide is practitioners who want to assess pedestrian and bicyclist risk. Some elements of the risk assessment methods in this guide are straightforward and should not pose significant difficulty for most practitioners, such as tabulating observed crashes from existing databases or collecting counts of pedestrians and bicyclists. Some analytic methods in this guide are more complex and may require specialized knowledge and skills, such as estimating expected crashes or estimating pedestrian and bicyclist exposure using a travel demand model. However, the process for assessing pedestrian and bicyclist risk in this guide provides flexibility, such that practitioners may select simpler methods that are consistent with their analysis capabilities and resources.

Organization of the Guide

The organization of this guide is based on eight steps (Figure S-1) that should be followed to develop pedestrian and bicyclist risk values at various geographic scales. These eight steps have been generalized to account for a variety of uses, geographic scales, and analytic methods. The following pages in this Summary include high-level guidance on each step. The full report provides more detailed information and guidance on each of the eight steps.

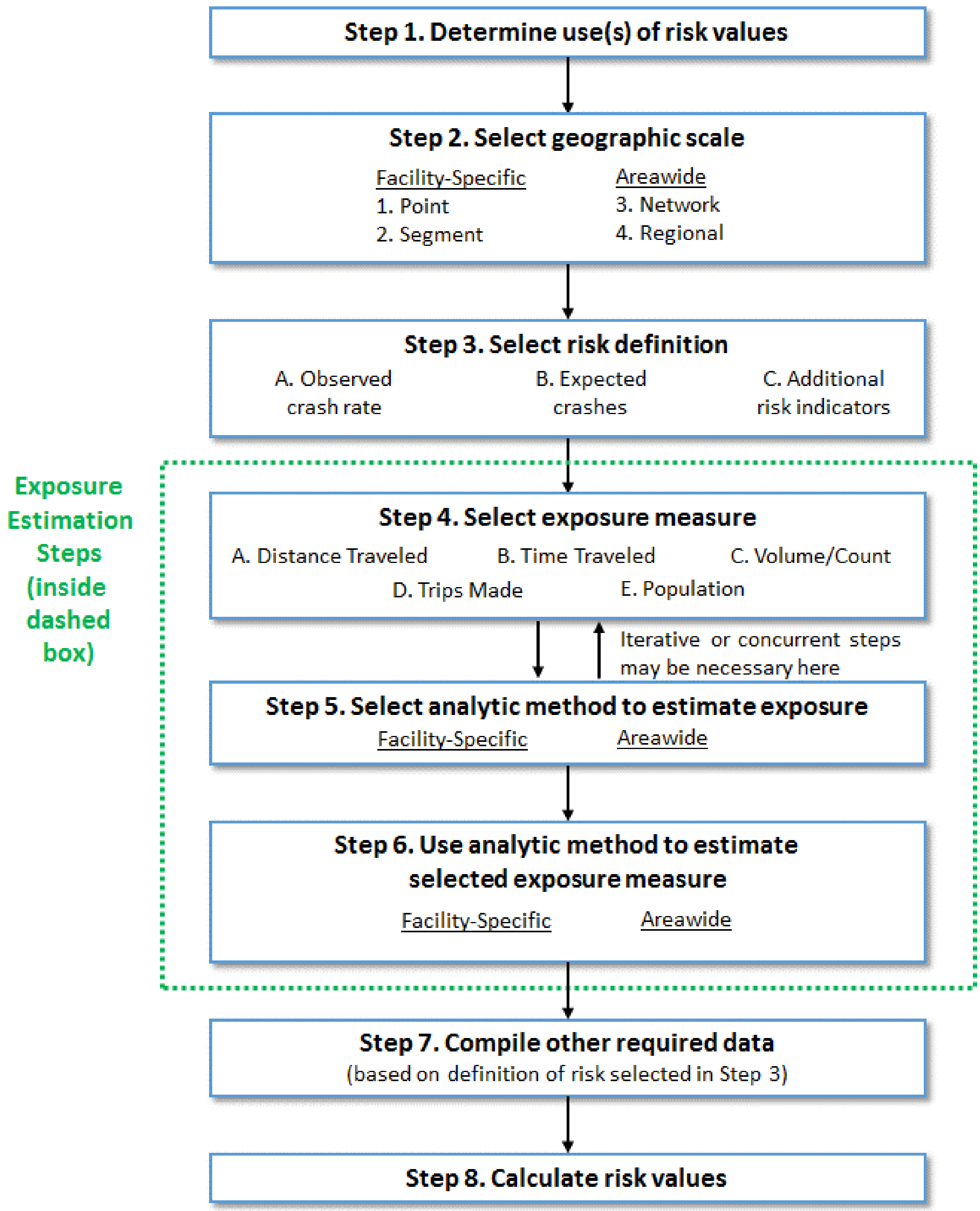


Figure S-1. Eight Steps for Scalable Risk Assessment for Pedestrians and Bicyclists

Step 1. Determine Use(s) of Risk Values

The first step in developing risk values for pedestrians and bicyclists is to define clearly the use(s) for the risk values. The use(s) of the risk values will establish key parameters (such as geographic scale) in later steps of the risk assessment process. This step is analogous to selecting the destination for a trip, such that one can pick the best route to reach the desired destination. Typical uses of pedestrian and bicyclist risk values are summarized below.

Safety Performance Measures

The objective in this use is to estimate risk for non-motorized modes as safety performance measures, which are then used to gauge progress toward safety improvement targets at an aggregate level (e.g., city, region, state).

Network Screening, Area-Based: Identifying high-risk sub-areas for more focused analysis

The objective in this use is to estimate pedestrian and bicyclist risk (separately) to identify high-risk sub-areas for more focused analysis. This use is most common where large geographic areas are being considered (such as a metropolitan planning area or an entire state), and there is a need to focus analysis resources in those sub-areas where pedestrian and bicyclist risk is highest in relation to other sub-areas within the entire geographic area.

Network Screening, Facility-Based: Identifying high-risk facilities for project development

The objective in this use is to estimate risk for specified facilities within a given jurisdiction (e.g., city, county, etc.) to identify those facilities or facility types with the highest risk. Depending upon the mode of interest (pedestrians or bicyclists), one could consider intersections or defined street segments (or both).

Project Prioritization: Prioritizing/ranking a defined project list

The objective in this use is to estimate risk for defined improvement projects (that include specific facilities) that are being considered for funding. In this case, one uses exposure estimates to normalize crashes (i.e., account for different levels of use/activity). Also, several other ranking criteria besides safety are often considered, such as stakeholder input, constraints, opportunities, existing conditions, demand, connectivity, equity, and compliance.

Countermeasure Evaluation: Evaluating the effectiveness of a safety countermeasure

The objective in this use is to evaluate the effectiveness of a specific countermeasure in reducing pedestrian and bicyclist risk. Exposure is needed for a specific facility (or facilities) both before and after the countermeasure was implemented. In many cases, the implementation of the countermeasure will have an effect on the exposure.

Site Evaluation

The goal in this use is to evaluate pedestrian and/or bicyclist risk at a specific site or at multiple sites. A site evaluation can be conducted on a one-time basis or on a before-after basis. This use differs from network screening in that a site evaluation focuses on a limited number of selected sites, whereas network screening includes all locations or sites within a defined network or area. This use differs from a countermeasure evaluation in that a site evaluation could assess the risk of multiple combined countermeasures, whereas a countermeasure evaluation typically tries to isolate the effectiveness of a single countermeasure at numerous locations.

Step 2. Select Geographic Scale

Step 2 in the scalable risk assessment process is to select the geographic scale at which risk and exposure values are desired. The desired geographic scale is based on, and sometimes dictated by, the use(s) of the risk values as defined in Step 1. The scalable risk assessment process includes four geographic scale categories, shown in Table 6S-1.

The desired geographic scale for exposure estimates is an important parameter that will be used in several subsequent steps. For example, selection of exposure measures (Step 4) are informed by the desired geographic scale (i.e., certain exposure measures are better suited to detailed geographic scale, whereas other exposure measures are better suited to an areawide geographic scale). Similarly, the selection of analytic methods to estimate exposure (Step 5) are also based on the desired geographic scale.

Table S-1. Four Scale Categories in the Scalable Risk Assessment Process

Scale Group	Scale Category	Description	Examples
Facility-Specific	Point	Specific location where conflicting traffic streams cross, merge, or diverge.	• Single intersection or mid-block crossing • All crossings at an intersection • Conflict zone (e.g., merge area)
	Segment	Length of street or roadway between two points. Traffic volumes and physical characteristics generally remain the same along the length of a segment, although small variations may occur.	• Street segment between major intersections • Multiple street segments along a single facility, or on parallel facilities (e.g., corridor) • Street segment of defined length (e.g., one mile)
Areawide	Network	A mid-sized geographic area that includes an interconnected set of transportation facilities.	• Census tracts • Census block groups • Traffic analysis zones
	Regional	A large geographic area that includes all transportation facilities within a defined political boundary. Because of the large geographic size, land use at this scale can be heterogeneous within a defined area.	• City • County • Metropolitan Statistical Area • State

Step 3. Select Risk Definition

Step 3 in the scalable risk process is to select a specific definition of risk that will be used to calculate quantitative risk values. S-2 shows three possible definitions of risk, with two of the three definitions closely related (i.e., one uses observed/reported crashes, while the other uses expected crashes).

Table S-2. Three Possible Definitions of Pedestrian and Bicyclist Risk

	1. Observed Crash Rate	**2. Expected Crashes**	**3. Additional Risk Indicators**
Description	• Risk = Observed crashes divided by exposure • Obtain **observed crashes** from available crash database(s). • Estimate **exposure** with this guide.	• Risk = Expected crashes • Estimate **expected crashes** with Highway Safety Manual or other statistical models, using exposure as input variable. • Estimate **exposure** with this guide.	• Risk = Function of one or more risk indicators: observed crashes, facility type or condition, motor vehicle speed & volume, adjacent land use, exposure, etc. • Estimate **exposure** (if included) with this guide.
Strengths	• Common use among many practitioners. • Use with other crash analysis tools (e.g., Pedestrian and Bicycle Crash Analysis Tool, PBCAT).	• Use of expected crashes overcomes issues with low (or no) observed crash frequency. • Permits evaluation of implemented countermeasures.	• Compatible with FHWA Systemic Safety Analysis. • Approach geared to practitioners.
Limitations	• Low exposure or low (or no) frequency of observed crashes may not accurately represent risk.	• Requires advanced statistical methods to estimate expected crashes. • Highway Safety Manual pedestrian and bicyclist tools still in early stages, may not address all site locations.	• Risk is a dimensionless numeric score or rating, not a crash frequency or crash rate value.

Step 4. Select Exposure Measure

Step 4 in the scalable risk assessment process is to select a specific exposure measure to be used in the calculation of risk values. There are several different categories of exposure measures that attempt to quantify the level of contact that pedestrians and bicyclists have with potentially harmful safety outcomes. The selection of an exposure measure will depend upon several criteria, such as the use of the risk values (Step 1), the geographic scale (Step 2), and other criteria. Table 11S-3 contains a selection matrix to help analysts choose an exposure measure best suited for their analysis. The full report contains additional guidance on the strengths and limitations of each category of exposure measure.

Table S-3. Selection Matrix for Exposure Measures

Category of Exposure Measure	Typical measures	Typical scale				Typical data sources
		Point	**Segment**	**Network**	**Region**	
Distance Traveled	Miles of travel	○	●	●	●	• Site counts or demand estimation models, multiplied by segment length • Sometimes travel surveys
	Miles crossed per entering vehicle	◐				
Time Traveled	Hours of travel	○	○	●	●	• Travel surveys • Sometimes site counts combined with crossing time or average travel speed data.
	Product of crossing time and vehicle volume	○	○			
Volume/ Count	Volume/count	●	●			• Site counts • Demand estimation models
	Product of pedestrian /bicyclist volumes and motor vehicle volumes	◐	◐			
Trips Made	Number of trips			●	●	• Travel surveys
Population	Number of people that walk or cycle on regular basis			●	●	• U.S. Census data products
	Percent of the population that walk or cycle on regular basis			●	●	

Legend: ○ = to a small extent; ◐ = to a moderate extent; ● = to a great extent.
Note: Each exposure measure will be for a defined time period that matches the risk definition.

Source: Partially adapted from Greene-Roesel et al., Estimating Pedestrian Accident Exposure: Protocol Report, March 2007.

Step 5. Select Analytic Method to Estimate Exposure

Step 5 in the scalable risk assessment process is to select an analytic method (or methods) to estimate exposure. There are numerous analytic methods that can be used to estimate exposure, and the most appropriate method(s) depends upon several criteria, such as desired geographic scale (Step 2), desired exposure measures (Step 4), analysis scope, data availability, staff technical capabilities, and available analysis resources.

Note that there may be some iteration or concurrency between this step (selecting an analytic method) and the previous step (Step 4, selecting an exposure measure). For example, an analyst may want to calculate a specific exposure measure, but have no expertise in the most common analytic methods used to estimate that exposure measure. Or, a specific analytic method may not be able to estimate accurately a specific of exposure measure. For example, site counts cannot be used to estimate accurately pedestrian or bicyclist hours traveled. Therefore, many analysts are likely to consider both the desired exposure measure and the most feasible analytic method in a concurrent or iterative manner.

Table 14S-4 provides a selection matrix to help analysts make informed choices about which analytic method(s) is best suited for them. It is important to note that local customization may be required for all these models to be useable.

Table S-4. Selection Matrix for Analytic Methods to Estimate Exposure

Analytic Method		Input Data Requirements	Technical Complexity	Popularity in Practice	Direct Usability	Accuracy
Site counts		○	○	●	●	○/◐/●
Demand Estimation Models	Direct demand models	◐	○/◐	●	◐	○/◐
	Regional TDM	◐/●	◐/●	○	○/◐/●	○/◐/●
	Trip generation and flow models	◐/●	◐/●	◐	●	◐/●
	GIS-based models	◐	◐	◐	●	◐/●
	Discrete choice models	◐/●	◐/●	◐	○	◐/●
	Simulation-based traffic models	●	●	○	●	●
	Data fusion	●	◐/●	○	●	◐/●
Travel surveys		○	○	●	●	○/◐/●

Legend: ○ = low suitability; ◐ = moderate suitability; ● = high suitability.

Note: For some categories, multiple ranges (e.g., ◐/●) are used since the corresponding criteria might vary significantly based on the specific characteristics of the model developed.

Step 6. Use Analytic Method to Estimate Selected Exposure Measure

Step 6 in the scalable risk assessment process is to use the analytic method selected in Step 5 to estimate the desired exposure measure(s). All of the previous steps involve making scoping or planning decisions about how to estimate exposure. Step 6 in the process is when the detailed analysis for exposure estimation occurs. As a result, the Step 6 section in this guide is the largest and has the most content.

The full report includes a section for each of the three primary methods to estimate exposure:

- Site counts
- Demand estimation models
- Travel surveys

Step 7. Compile Other Required Data

Step 7 in the scalable risk assessment process consists of compiling other data besides exposure that is required based upon the risk definition selected in Step 3. The three possible risk definitions are:

1. Observed crash rate
2. Expected crashes
3. Additional risk indicators

Detailed instructions for compiling other required data for these three risk definitions is beyond the scope of this guide. There is extensive guidance and examples in several other reports, manuals, and guides. Therefore, these sections in the guide provide summary information and pointers to these other guidance documents.

Step 8. Calculate Risk Values

Step 8 in the scalable risk assessment process is to calculate risk values based on the outputs from previous steps. That is, Step 6 provides exposure estimates and Step 7 provides observed crashes, expected crashes, or additional risk indicators that are then used to calculate final risk values at the geographic scale chosen in previous steps.

Case studies are provided in this chapter to tie together the eight steps described in this guide. The case studies are based on actual examples of risk assessment for pedestrians and bicyclists.

INTRODUCTION

This guide describes scalable risk assessment methods for pedestrians and bicyclists, wherein risk is a measure of the probability of a crash to occur given exposure to potential crash events. This guide:

- Outlines eight sequential steps to develop risk values at various desired geographic scales.
- Describes the scope and nature of each step, including any guiding principles.
- Provides information on analytic methods to estimate pedestrian and bicyclist exposure.
- Identifies other relevant guides and resources that provide supplemental information.

Motivation for the Guide

Many transportation agencies are placing more emphasis on improving pedestrian and bicyclist safety and reducing the risk of a fatality or serious injury to pedestrians and bicyclists. Practitioners need a methodical approach to assess pedestrian and bicyclist risk for the purposes of identifying high-priority areas and transportation facilities for safety improvement, evaluating specific countermeasures and locations before and after improvements are made, and tracking safety performance measures over time to gauge progress toward established goals. The motivation for this guide is to provide this methodical approach for assessing pedestrian and bicyclist risk for these and other applications.

Exposure to risk is an integral element of risk (see equation below), and as such, an integral element of risk assessment methods in this guide.

$$Risk = \frac{Expected\ or\ measured\ crashes, by\ kind\ and\ severity}{Exposure}$$

Exposure is a measure of the number of potential opportunities for a crash to occur, and is often directly related to the number of people who walk and bike. Improvements for bicyclists and pedestrians often result in more people walking and biking, which in turn can increase exposure. However, an increase in exposure may not lead to an increase in risk if crashes increase less proportionately than exposure increases (see equation above for the relationship between risk, crashes, and exposure). In fact, several studies have found that as more people walk and bike, their risk may actually decrease (i.e., the safety-in-numbers hypothesis). In the past, exposure has seldom been included in pedestrian and bicyclist safety analysis because of the practical challenges of collecting or estimating exposure data. Another motivation for this guide is to outline feasible methods to calculate or estimate exposure, such that exposure will be included more often in pedestrian and bicyclist risk assessment.

Intended Audience

The main audience for this guide is practitioners who want to assess pedestrian and bicyclist risk. Some elements of the risk assessment methods in this guide are straightforward and should not pose significant difficulty for most practitioners, such as tabulating observed crashes from existing databases or collecting counts of pedestrians and bicyclists. Some analytic methods in this guide are more complex and may require specialized knowledge and skills, such as estimating expected crashes or estimating pedestrian and bicyclist exposure using a travel demand model. However, the process for assessing pedestrian and bicyclist risk in this guide provides flexibility, such that practitioners may select simpler methods that are consistent with their analysis capabilities and resources.

Organization of the Guide

The organization of this guide is based on eight steps (Figure 1) that should be followed to develop pedestrian and bicyclist risk values at various geographic scales. These eight steps have been generalized to account for a variety of uses, geographic scales, and analytic methods. The following is a brief annotation of these steps to help illustrate what is involved in each step. The following chapters of this guide describe each step in more detail.

1. **Determine use(s) of risk values:** What is the main objective of my safety analysis? How am I going to use the risk values? What decisions do I want to make with these risk values?

2. **Select geographic scale:** What geographic scale do I need to support the use(s) of risk values determined in Step 1?

3. **Select risk definition:** How will I quantify risk at the defined geographic scale in this analysis? Will I use observed crashes or expected crashes? Or will I combine multiple other indicators to estimate risk? All three risk definitions require exposure estimates, so Steps 4 through 6 focus on exposure estimation.

4. **Select exposure measure:** What measure should I use to quantify exposure? The choice of exposure measure is typically based on scale, data availability, and analytic methods used. Iteration or concurrent selection of exposure measures and analytic methods (Step 5) may be necessary to ensure that an analytic method can provide the required data for the exposure measure.

5. **Select analytic method to estimate exposure:** What method(s) should I use to estimate exposure? The choice of analytic method is typically based on scale, data availability, staff expertise, scope of analysis, resources available, etc. In addition, the decision tree typically splits into two distinct paths based on scale: 1) facility-specific; 2) areawide.

6. **Use analytic method to estimate selected exposure measure:** How do I use the selected analytic method to estimate the exposure measure? Analytic methods for facility-specific scales are typically distinct from areawide methods for exposure estimates.

7. **Compile other required data:** Once exposure has been estimated in Step 6, what other data are needed to calculate the risk values? Based on the definition of risk selected in Step 3: A) observed crashes are compiled from existing crash databases, B) expected crashes are estimated, or C) additional risk indicators are compiled.

8. **Calculate risk values:** In this last step, the final risk values are calculated based on the definition of risk selected in Step 3. The calculation uses the exposure estimates from Step 6 with the other required data compiled in Step 7.

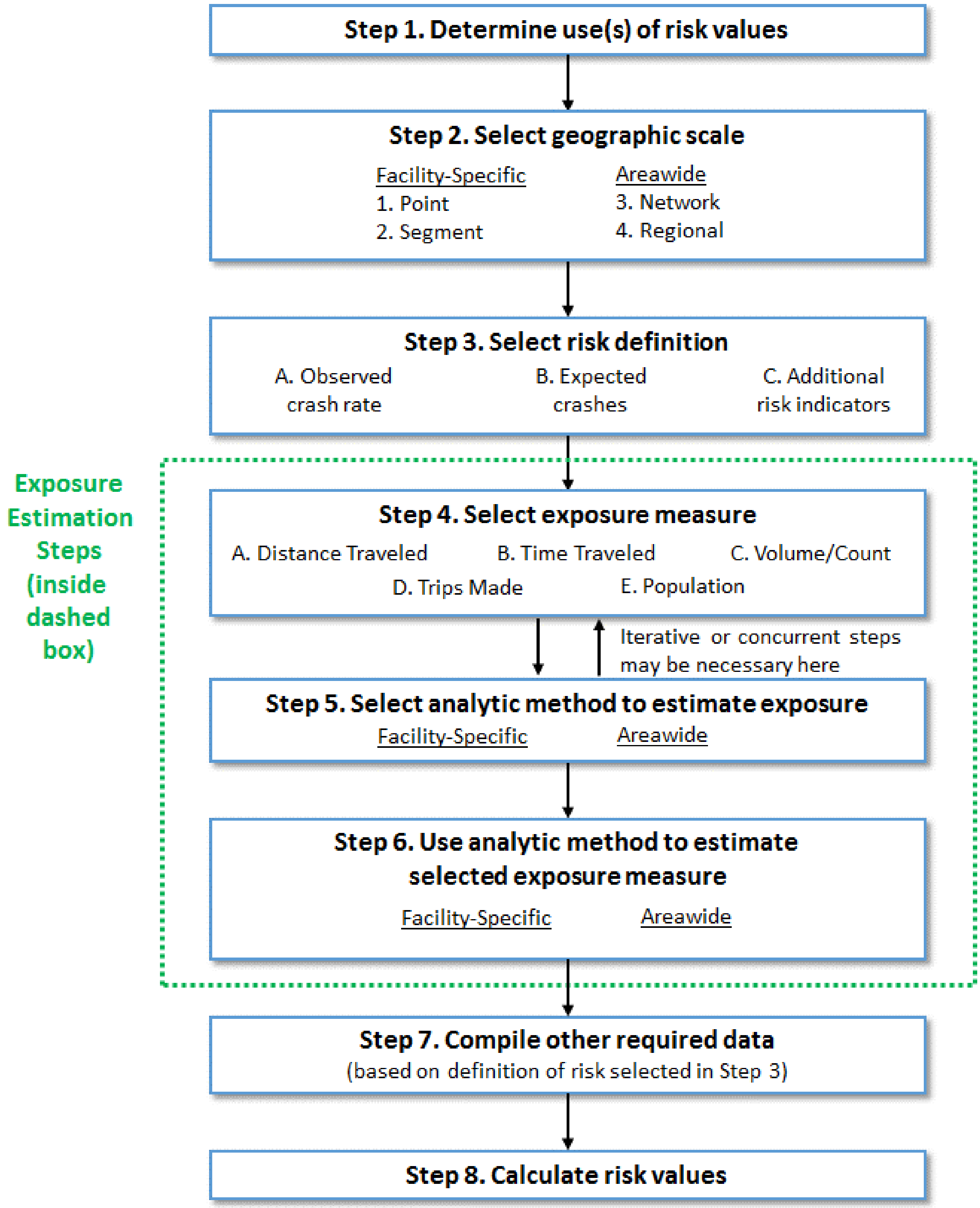


Figure 1. Eight Steps for Scalable Risk Assessment for Pedestrians and Bicyclists

STEP 1. DETERMINE USE(S) OF RISK VALUES

The first step in developing risk values for pedestrians and bicyclists is to define clearly the use(s) for the risk values. The use(s) of the risk values will establish key parameters (such as geographic scale) in later steps of the risk assessment process. This step is analogous to selecting the destination for a trip, such that one can pick the best route to reach the desired destination.

This chapter includes a list of questions that can help to clarify your objectives for risk assessment and how you plan to use the risk values. This chapter outlines and describes common uses of risk values, with tabular and graphical examples from current practice that illustrate these uses.

This chapter also includes a section on assessing currently available data, as well as the feasibility of gathering additional data, for use in the remaining steps in the exposure estimation process. Some exposure estimation methods are very data-intensive, and analysts should be aware of existing data resources and limitations as they work through the exposure estimation process.

Identifying Objectives of Risk Assessment

Several questions can help to clarify your use(s) and objectives of risk values:

- What do I need to make progress with respect to pedestrian and bicyclist safety?
- What is the result that I am trying to reach?
- How much detail do I need in this result?
- Does my result feed into an existing process? If so, what does this existing process require?
- Am I interested in just a few locations, or an entire defined area?

Typical Uses of Risk Values for Pedestrians and Bicyclists

The following sections describe typical uses or applications of pedestrian and bicyclist risk values, as well as an example to illustrate the application of risk values. Typical uses include:

- Safety performance measures.
- Network screening: area-based.
- Network screening: facility-based.
- Project prioritization.
- Countermeasure evaluation.
- Site evaluation.

Safety Performance Measures

The objective in this use is to estimate risk for non-motorized modes as safety performance measures. For example, the Federal Highway Administration (FHWA) Safety Performance Management Program (https://safety.fhwa.dot.gov/hsip/spm/) currently requires that state departments of transportation (DOTs) and metropolitan planning organizations (MPOs) report five safety performance measures, one of which is the Number of Non-motorized Fatalities and Non-motorized Serious Injuries. Two of the motor vehicle safety performance measures (Rate of Fatalities and Rate of Serious Injuries) include motor vehicle exposure (expressed in 100 million vehicle miles traveled (VMT)), and there is a desire to have a similar exposure estimate for the non-motorized safety performance measure. Because there is no commonly accepted exposure metric for pedestrians and bicyclists, the current non-motorized safety performance measures are based on numbers of fatalities and serious injuries rather than including exposure. Many cities and small communities continue to improve infrastructure and other conditions for walking and bicycling and may desire to monitor safety performance at more granular levels.

As an example, Table 1 shows an excerpt of the National Highway Traffic Safety Administration (NHTSA) Traffic Safety Facts for pedestrians, which shows the risk of pedestrian fatalities by age and gender. Note that the exposure measure in this table is per 100,000 population, as measured by the U.S. Census.

Table 1. Example of Pedestrian Fatality Risk by Age and Gender

Pedestrians Killed/Injured in Traffic Crashes and Fatality/Injury Rates, by Age and Gender, 2015

Age (Years)	Male			Female			Total		
	Killed	Population (thousands)	Fatality Rate*	Killed	Population (thousands)	Fatality Rate*	Killed	Population (thousands)	Fatality Rate*
0–4	44	10,178	0.43	33	9,730	0.34	77	19,907	0.39
5–9	43	10,459	0.41	30	10,028	0.30	73	20,487	0.36
10–14	58	10,520	0.55	25	10,102	0.25	83	20,622	0.40
Children (≤14)	*145*	*31,157*	*0.47*	*88*	*29,860*	*0.29*	*233*	*61,016*	*0.38*
15–19	140	10,798	1.30	83	10,311	0.80	223	21,109	1.06
20–24	313	11,668	2.68	98	11,071	0.89	411	22,739	1.81
25–29	303	11,409	2.66	104	11,052	0.94	407	22,462	1.81
30–34	243	10,890	2.23	100	10,786	0.93	344	21,676	1.59
35–39	276	10,173	2.71	108	10,201	1.06	384	20,375	1.88
40–44	265	10,030	2.64	104	10,185	1.02	370	20,215	1.83
45–49	288	10,335	2.79	134	10,519	1.27	422	20,854	2.02
50–54	406	10,964	3.70	164	11,370	1.44	571	22,334	2.56
55–59	386	10,598	3.64	142	11,210	1.27	529	21,808	2.43
60–64	317	9,117	3.48	112	9,953	1.13	430	19,070	2.25
65–69	206	7,596	2.71	97	8,471	1.15	304	16,067	1.89
70–74	134	5,296	2.53	82	6,187	1.33	216	11,483	1.88
75–79	119	3,611	3.30	61	4,513	1.35	180	8,124	2.22
80+	173	4,587	3.77	129	7,500	1.72	302	12,087	2.50
Seniors (65+)	*632*	*21,090*	*3.00*	*369*	*26,671*	*1.38*	*1,002*	*47,761*	*2.10*
Total[a]	*3,749*	*158,229*	*2.37*	*1,617*	*163,190*	*0.99*	*5,376*	*321,419*	*1.67*

* Rate per 100,000 population
Source: NHTSA Report DOT-HS-812-375, Traffic Safety Facts: 2015 Data, Pedestrians, February 2017.

As another example, Figure 2 shows an excerpt from the 2016 Benchmarking Report, which shows the risk of pedestrian fatalities in 50 cities. The exposure measure in this chart is per 10,000 walking commuters. One should note that an ideal exposure measure includes all walking trips, not just walking commutes or the total population. However, accurately quantifying the total number of walking trips on a national level is challenging given available data sources.

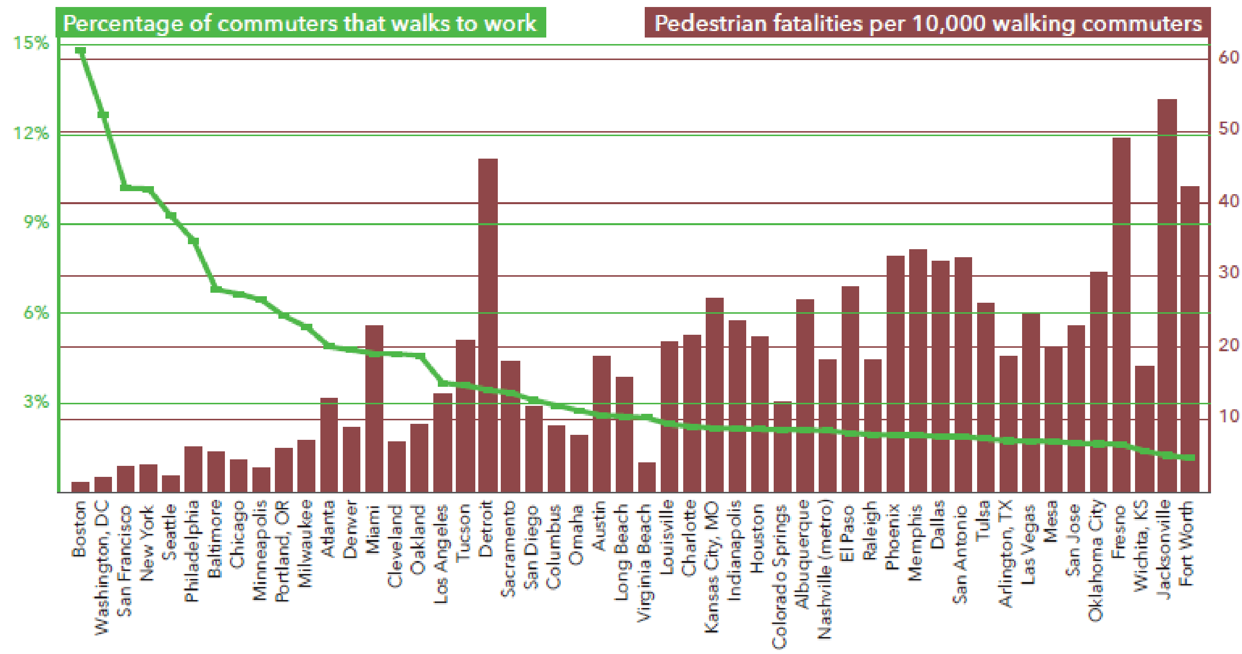


Figure 2. Example of Pedestrian Fatality Risk in 50 Cities

Source: 2016 Benchmarking Report, Alliance for Biking & Walking.

Network Screening, Area-Based: Identifying high-risk sub-areas for more focused analysis

The objective in this use is to estimate pedestrian and bicyclist risk (separately) to identify high-risk sub-areas (e.g., census geography, traffic analysis zones (TAZs), or other areawide zones) for more focused analysis. This use is most common where large geographic areas are being considered (such as a metropolitan planning area or an entire state), and there is a need to focus analysis resources in those sub-areas where pedestrian and bicyclist risk is highest in relation to other sub-areas within the entire geographic area. In some cases, there may be more detailed follow-up analysis in these high-risk sub-areas to identify facility-specific exposure, contributing risk factors, and possible improvements.

As an example, Figure 3 shows an example of a pedestrian risk analysis that classified all census tracts in California into a neighborhood type: central city, urban, suburb and rural. Pedestrian risk was quantified in each of these neighborhood types using two different travel surveys: the 2010-2012 California Household Travel Survey (CHTS) and the California travel statistics from the 2009 National Household Travel Survey (NHTS). The analysis included two different exposure measures calculated from the travel surveys: 1) million miles walked or biked on a weekday; 2) per 100,000 population. The authors concluded that the crash rates normalized by population are somewhat misleading, since pedestrian and bicyclist activity levels (and therefore exposure) are not consistent across these neighborhood types.

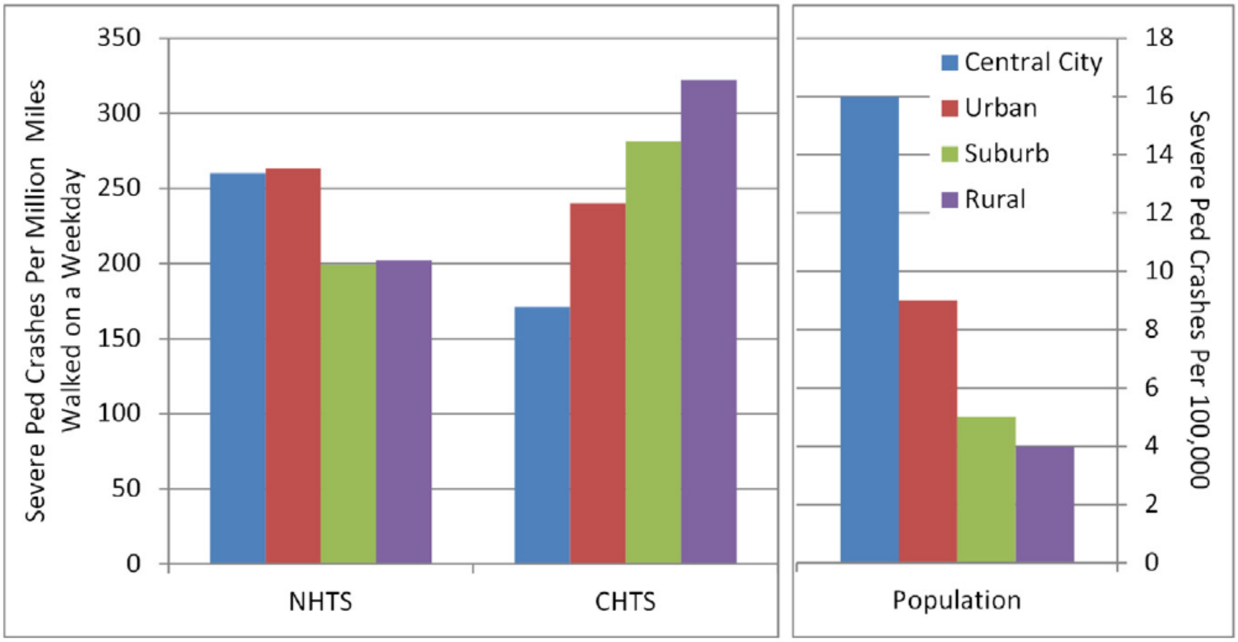


Figure 3. Example of Pedestrian Risk Analysis in California Using Household Travel Surveys
Source: Salon 2016.

Network Screening, Facility-Based: Identifying high-risk facilities for project development

The objective in this use is to estimate risk for specified facilities within a given jurisdiction (e.g., city, county, etc.) to identify those facilities or facility types with the highest risk. Depending upon the mode of interest (pedestrians or bicyclists), one could consider intersections or defined street segments (or both). In some cases, one may perform risk analysis on an ad hoc basis and not for the entire street network (perhaps only a few streets or locations).

As an example, Table 2 shows the results of a systemic safety analysis to identify high-risk intersections for pedestrians and bicyclists in the north central part of Minnesota. In this systemic analysis, motor vehicle volume (called Cross Product in the table) and Primary Land Use are used to represent exposure.

Table 2. Example of Systemic Safety Analysis that Includes Exposure Variables

#	Intersection ID	Route System	Route No.	Description	Speed Limit	Cross Product[b]	Traffic Control	Major Corridor Speed	Skew	On/Near Curve	Primary Land Use	Severe Ped/Bike Crash Density	Total Stars	Crash Cost
34	3.210.025	MN	210	4TH ST NWCSAH20 MSAS103/BRNRD	35	★	★	★		★	★		★★★★★	$1,050,200
35	3.024.009	MN	24	CSAH 75/CLEARWATER	40	★	★	★	★		★		★★★★★	$747,600
36	3.023.028	MN	23	19 1/2 AV/ST CLD	35	★		★	★	★	★		★★★★★	$574,800
37	3.023.050	MN	23	TH 25/FOLEY	45	★	★	★	★		★		★★★★★	$558,000
38	3.027.015	MN	27	4TH ST MSAS 106/LITTLE FALLS	30	★	★		★	★	★		★★★★★	$366,400
39	3.023.011	MN	23	RED RVR AVCSAH 2/COLD SPRING	35	★	★	★		★	★		★★★★★	$292,800
40	3.023.020	MN	23	6TH AV S MSAS107 M95/WAITPK	40	★	★	★		★	★		★★★★★	$0
41	3.210.021	MN	210	ELDER DR SM140/BAXTER	55	★		★	★		★		★★★★	$10,558,200
42	3.012.003	US	12	JOHNSON AVE M-54 LT/COKATO	35	★		★			★	★	★★★★	$10,418,000
43	3.015.011	MN	15	N JCT TH 23 DIV ST/ST CLOUD	45	★	★	★			★		★★★★	$5,838,400
44	3.015.012	MN	15	3RD ST N CSAH81 MSAS 114/STC	45	★	★	★			★		★★★★	$4,310,200
45	3.169.004	US	169	197TH AV MSAS116 M118/ELKRV	55	★	★	★			★		★★★★	$1,696,200
46	3.015.019	MN	15	CSAH 29/SAUK RAPIDS	60	★	★	★			★		★★★★	$1,671,800
47	3.010.011	US	10	E JCT TH 210 LT/MOTLEY	30	★	★			★	★		★★★★	$1,612,200
48	3.210.026	MN	210	4TH ST N MSAS114/BRAINERD	35	★	★	★			★		★★★★	$1,241,800
49	3.210.027	MN	210	TH 371B RTM 60 LT/BRAINERD	35	★	★	★			★		★★★★	$1,186,600
50	3.023.022	MN	23	WAITE AVEMSAS101/WAITEPARK	40	★	★	★			★		★★★★	$1,146,000
53	3.025.030	MN	25	RIVER ST MSAS112/MONTICELLO	30	★	★			★	★		★★★★	$891,400
54	3.012.020	US	12	BUFFALO AVCSAH 12TH 25/MONTR	35	★	★	★			★		★★★★	$641,000
55	3.023.088	MN	23	N JCT TH 65 CSAH 6/MORA	30	★	★		★		★		★★★★	$622,200
56	3.025.029	MN	25	BROADWAY CSAH75/MONTICELLO	30	★	★			★	★		★★★★	$619,600

Source: Report FHWA-SA-17-002, Systemic Safety Project Selection Tool Supplemental Case Studies, December 2016.

Project Prioritization: Prioritizing/ranking a defined project list

The objective in this use is to estimate risk for defined improvement projects (that include specific facilities) that are being considered for funding. In this case, one uses exposure estimates to normalize crashes (i.e., account for different levels of use/activity). Also, several other ranking criteria besides safety are often considered (see National Cooperative Highway Research Program (NCHRP) Report 803, Active Trans Priority Tool), such as stakeholder input, constraints, opportunities, existing conditions, demand, connectivity, equity, and compliance.

As an example, Table 3 shows an example from NCHRP Report 803 that includes Safety as one of multiple project scoring criteria. In the example, a weight is applied to each criterion, and then a composite Prioritization Score is calculated. Note that this example uses a safety criterion, but it could also include a risk or exposure criterion as well.

Table 3. Example of Multi-Criteria Project Prioritization

	A	B	C	D	I	J	M	N	U
1	Step 10A: Calculate Priority Score								
3									
4									
5	ID	GAP LOCATION	Stakeholder Input SCORE	Stakeholder Input WEIGHTED SCORE	Safety SCORE	Safety WEIGHTED SCORE	Demand SCORE	Demand WEIGHTED SCORE	Prioritization Score
7	1	CENTRAL AVE	6.3	62.5	0.0	0.0	8.1	32.5	95.0
8	2	WASHINGTON/JEFFERSON CORRIDOR	4.2	41.7	7.1	57.1	8.4	33.6	132.4
9	3	3RD ST	9.6	95.8	4.3	34.3	3.8	15.0	145.2
10	4	12TH ST	0.8	8.3	1.4	11.4	2.5	10.0	29.8
11	5	15TH AVE	0.4	4.2	4.3	34.3	3.6	14.6	53.0
12	6	ENCANTO BLVD	6.3	62.5	4.3	34.3	7.7	30.9	127.7
13	7	OSBORN RD	8.8	87.5	2.9	22.9	5.2	20.6	131.0
14	8	OAK ST	3.8	37.5	2.9	22.9	4.0	16.0	76.4
15	9	20TH ST	2.1	20.8	0.0	0.0	3.1	12.6	33.4
16	10	3RD/5TH	1.3	12.5	10.0	80.0	3.1	12.5	105.0
17	11	DEER VALLEY DR	3.3	33.3	0.0	0.0	5.4	21.5	54.8
18	12	UNION HILLS DR	5.0	50.0	7.1	57.1	9.9	39.8	146.9
19	13	19TH AVE	5.8	58.3	7.1	57.1	3.5	14.0	129.5
20	14	32ND ST	8.8	87.5	10.0	80.0	6.8	27.3	194.8
21	15	40TH ST	3.3	33.3	5.7	45.7	3.1	12.6	91.6

Source: NCHRP Report 803, Pedestrian and Bicycle Transportation Along Existing Roads—ActiveTrans Priority Tool Guidebook, 2015.

Countermeasure Evaluation: Evaluating the effectiveness of a safety countermeasure

The objective in this use is to evaluate the effectiveness of a specific countermeasure in reducing pedestrian and bicyclist risk. Exposure is needed for a specific facility (or facilities) both before and after the countermeasure was implemented. In many cases, the implementation of the countermeasure will have an effect on the exposure. Exposure typically increases in the after period because the countermeasure improves conditions for pedestrians and/or bicyclists, thereby attracting more pedestrian and bicyclists. In addition, the normal growth in motor vehicle traffic may increase exposure if that is included in exposure calculations. Therefore, it is important that exposure be accurately measured or estimated in both the before and after periods.

As an example, Table 4 shows a before-after evaluation of pedestrian hybrid beacons (formerly called High intensity Activated crossWalK (HAWKs)) in Tucson, Arizona. Table 4 shows crash frequencies and crash rates for the treatment sites as well as three reference groups. Two exposure measures are shown: 1) million entering vehicles and pedestrians, and 2) million entering pedestrians. The Empirical Bayes method is used to compare statistically the observed crash frequency during the after period (with treatment installed) to an estimated crash frequency if the treatment had not been installed.

Table 4. Example of Exposure Measures Included in Countermeasure Evaluation

Treatment Group	Measure	All Crashes with Intersecting Street Name			Crashes coded as Intersection-Related Crashes		
		Before	After	Percent Change	Before	After	Percent Change
HAWK sites (21)	Frequency	11.0	9.2	-17	5.0	3.3	-34
	Total crashes/MEV&P	0.748	0.618	-17	0.341	0.223	-35
	Severe crashes/MEV&P	0.265	0.210	-21	0.138	0.094	-32
	Pedestrian crashes/MEV&P	0.029	0.005	-83	0.017	0.002	-86
	Pedestrian crashes/MEP	3.081	0.511	-83	1.826	0.255	-86
Reference group 1: signalized intersections (36)	Frequency	44.9	41.9	-7	19.6	16.8	-14
	Total crashes/MEV&P	1.953	1.788	-8	0.854	0.716	-16
	Severe crashes/MEV&P	0.549	0.503	-8	0.294	0.241	-18
	Pedestrian crashes/MEV&P	0.020	0.016	-23	0.010	0.008	-16
	Pedestrian crashes/MEP	2.051	1.546	-25	1.025	0.839	-18
Reference group 1: unsignalized intersections (35)	Frequency	4.2	4.3	3	1.6	1.3	-17
	Total crashes/MEV&P	0.285	0.292	2	0.108	0.090	-17
	Severe crashes/MEV&P	0.098	0.088	-10	0.043	0.038	-10
	Pedestrian crashes/MEV&P	0.006	0.009	52	0.003	0.004	42
	Pedestrian crashes/MEP	1.383	2.078	50	0.615	0.866	41
Reference group 2: unsignalized intersections (102)	Frequency	5.9	6.1	3	2.4	2.1	-9
	Total crashes/MEV&P	0.418	0.430	3	0.166	0.150	-9
	Severe crashes/MEV&P	0.140	0.141	0	0.060	0.056	-6
	Pedestrian crashes/MEV&P	0.006	0.011	93	0.001	0.003	143
	Pedestrian crashes/MEP	1.233	2.297	86	0.257	0.602	134

Crashes/MEV&P = Type of given crash (total, severe, or pedestrian crashes) per million entering vehicles and pedestrians.
Pedestrian crashes/MEP = Pedestrian crashes per million entering pedestrians.
Note: Frequency is expressed as the average annual number of total crashes for a site with the given intersection control and study period.

Source: Report FHWA-HRT-10-042, Safety Effectiveness of the HAWK Pedestrian Crossing Treatment, July 2010.

Site Evaluation

The goal in this use is to evaluate pedestrian and/or bicyclist risk at a specific site or at multiple sites. A site evaluation can be conducted on a one-time basis or on a before-after basis. This use differs from network screening in that a site evaluation focuses on a limited number of selected sites, whereas network screening includes all locations or sites within a defined network or area. This use differs from a countermeasure evaluation in that a site evaluation could assess the risk of multiple combined countermeasures, whereas a countermeasure evaluation typically tries to isolate the effectiveness of a single countermeasure at numerous locations.

As an example, Figure 4 shows a site evaluation conducted at a single intersection in Ft. Lauderdale, Florida. The Florida DOT implemented exclusive pedestrian phases in the intersection signal timing, and evaluated the site before and after implementation of the phases. Because Florida DOT wanted to monitor the improvement soon after implementation, additional risk indicators in the form of conflict rates (normalized by pedestrian and bicyclist exposure) were used in the evaluation.

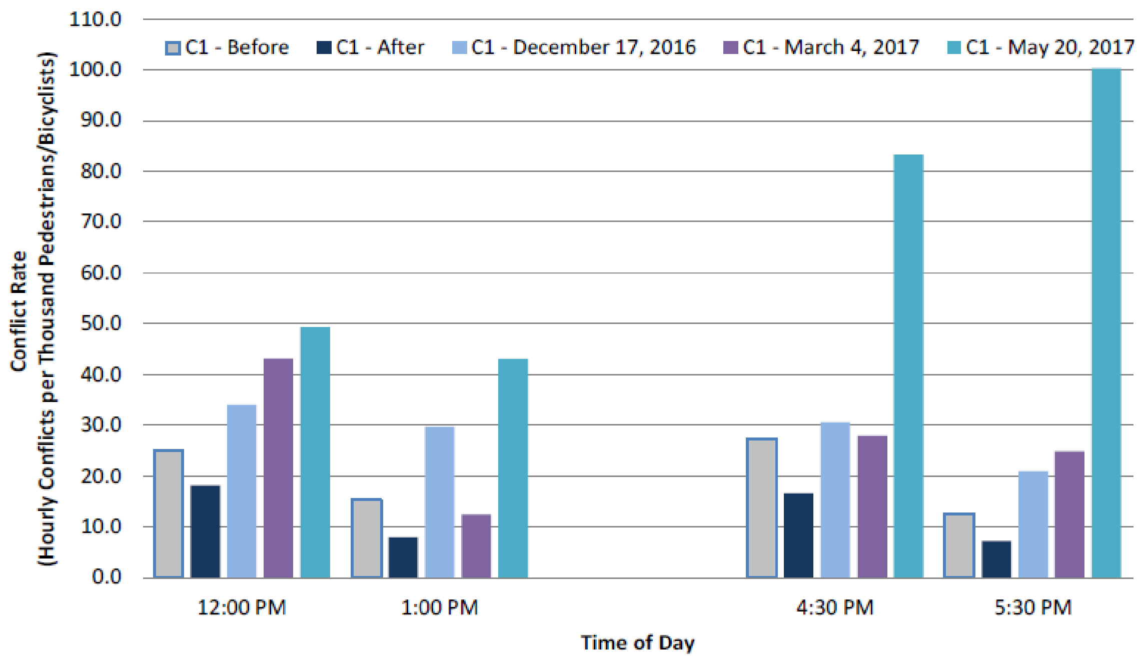


Figure 4. Example of Site Evaluation: Exclusive Pedestrian Phase at Single Intersection
Source: Florida DOT Internal Tech Memo, November 2017.

Selecting Analysis Parameters in Remaining Steps of Risk Assessment

As indicated earlier in this chapter, the defined use(s) of the risk values establish key parameters in later steps of the risk assessment process. Table 5 outlines key parameters that must be selected in Steps 2 through 5, and provides some initial guidance for parameter selection based on each of the defined uses presented earlier. Note that each step, and the considerations for selecting key parameters in each step, is covered in more detail in the upcoming chapters of this guide. Table 5 is intended as a preview and overview of decisions that must be made in these upcoming risk assessment steps.

Table 5. Selecting Key Parameters Based on Use(s) of Risk Values

Step 1. Define Use(s) of Risk Values	Step 2. Select Geographic Scale				Step 3. Select Risk Definition			Step 4. Select Exposure Measure					Step 5. Select Analytic Method		
	Point	Segment	Network	Regional	Observed Crash Rate	Expected Crashes	Risk Indicators	Distance Traveled	Time Traveled	Volume/Counts	Trips Made	Population	Site Count	Estimation Model	Travel Survey
Safety performance measures: Track changes in risk over time	X	X	✓	✓	✓	X	X	✓	✓	X	O	O	O	X	✓
Network screening: area-based: Identify high-risk areas for possible improvement	NA	NA	✓	✓	✓	O	O	✓	✓	X	O	O	X	O	✓
Network screening: facility-based: Identify high-risk facilities for possible improvement	✓	✓	NA	NA	✓	O	O	O	O	✓	O	O	✓	✓	X
Project prioritization: Rank projects based on existing risk or expected risk reduction	✓	✓	O	O	✓	O	✓	O	O	✓	X	X	✓	✓	X
Countermeasure evaluation: Evaluate if a specific countermeasure reduces risk (and by how much)	✓	✓	NA	NA	✓	O	O	O	O	✓	X	X	✓	O	X
Site evaluation: Evaluate if risk was reduced after site improvements (and by how much)	✓	✓	NA	NA	✓	✓	O	O	O	✓	X	X	✓	O	X

Legend: ✓ = Yes, preferred; O = yes, as a secondary preference; X = Not likely; NA = Not applicable

Assessing Available Data Resources and Analytic Capabilities

In the next few steps of the scalable risk assessment process, you will have to make several decisions about risk definitions, exposure measures and supporting data, and analytic methods to use. These decisions are strongly influenced by available data resources and analytic capabilities. Therefore, you should assess available data resources and capabilities within your agency or region to determine the type and scope of scalable risk assessment that is feasible. The following checklist should help in key decision areas:

Selecting a Risk Definition

Determine what definition of risk can be supported.

Risk definition of Observed Crash Rate:

- ☐ Familiar with traditional crash data and reporting methods?
- ☐ Have access to pedestrian and/or bicyclist crash data of sufficient accuracy and quality?

Risk definition of Expected Crashes:

- ☐ Familiar with Highway Safety Manual (HSM) or other statistical methods of estimating expected crashes?
- ☐ Can most recent version of HSM address all road and location types of interest? If not, can other statistical models be developed?

Risk definition of Additional Risk Indicators:

- ☐ Familiar with Systemic Safety Analysis?
- ☐ Have road inventory and land use data (or can collect) for Systemic Safety Analysis?

Exposure Data Resources

Determine what data resources currently exist for calculating or estimating exposure. Do you have:

- ☐ Recent pedestrian and bicyclist counts at all locations of interest? Or only some locations?
- ☐ Road inventory and land use data (or can collect) for estimating exposure? In a geographic information system (GIS)?
- ☐ Familiarity with recent regional household travel survey that includes walking and biking trips?
- ☐ Familiarity with American Community Survey (ACS) or NHTS?
- ☐ Bike share system data, either origin-destination or route-based?
- ☐ Pedestrian or bicyclist route choice data (e.g., Strava or other trip recording apps)?

Analytic Capabilities

Determine your analytic capabilities to estimate exposure. Are you familiar with:

- ☐ GIS and basic geospatial analysis?
- ☐ Appropriate statistical models?
- ☐ Regional travel demand models (TDMs)?
- ☐ Trip generation and flow models?
- ☐ Discrete choice models?
- ☐ Simulation-based traffic models?
- ☐ Data fusion and combining disparate raw data sources?

Staff and Funding Resources

- ☐ Ability to hire contractor?
- ☐ Funds available to conduct analysis of interest?
- ☐ Staff available to conduct analysis of interest?

STEP 2. SELECT GEOGRAPHIC SCALE

Step 2 in the scalable risk assessment process is to select the geographic scale at which risk and exposure values are desired. The desired geographic scale is based on, and sometimes dictated by, the use(s) of the risk values as defined in Step 1.

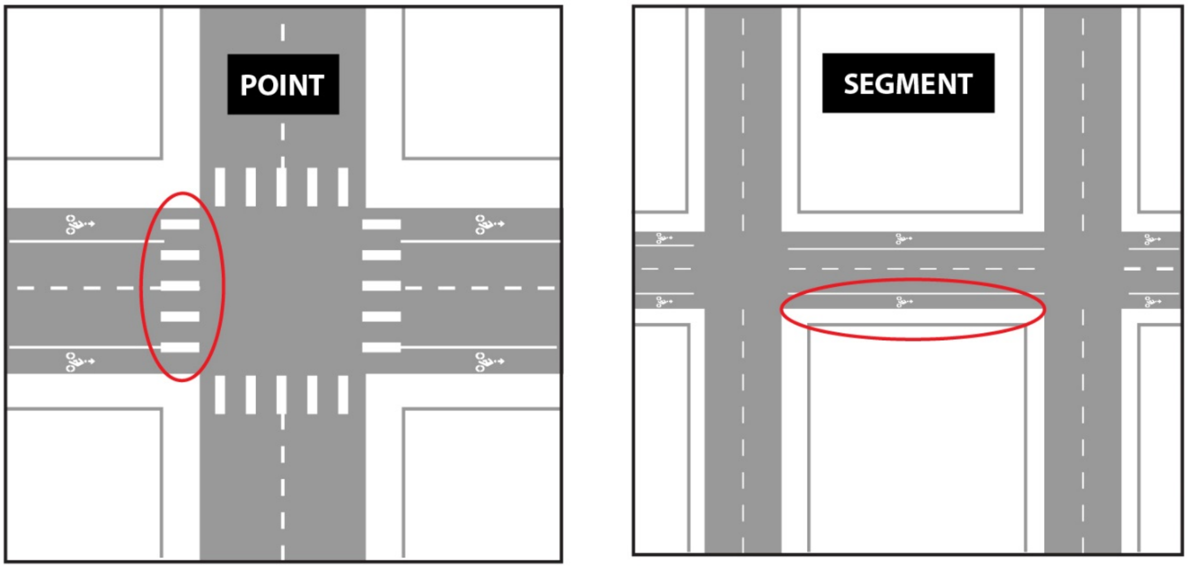
The desired geographic scale for exposure estimates is an important parameter that will be used in several subsequent steps. For example, selection of exposure measures (Step 4) are informed by the desired geographic scale (i.e., certain exposure measures are better suited to detailed geographic scale, whereas other exposure measures are better suited to an areawide geographic scale). Similarly, the selection of analytic methods to estimate exposure (Step 5) are also based on the desired geographic scale. These steps later in the guide provide more detail on how scale informs these decisions.

The scalable risk assessment process includes four geographic scale categories, listed in Table 6 and shown graphically in Figure 5 in order of most granular to most aggregate. Note that the four scale categories can be grouped into two general scale groups: 1) facility-specific; and 2) areawide. These two scale groups typically use fundamentally different analytic methods to estimate exposure. This is discussed further in Step 5.

Table 6. Four Scale Categories in the Scalable Risk Assessment Process

Scale Group	Scale Category	Description	Examples
Facility-Specific	Point	Specific location where conflicting traffic streams cross, merge, or diverge.	• Single intersection or mid-block crossing • All crossings at an intersection • Conflict zone (e.g., merge area)
	Segment	Length of street or roadway between two points. Traffic volumes and physical characteristics generally remain the same along the length of a segment, although small variations may occur.	• Street segment between major intersections • Multiple street segments along a single facility, or on parallel facilities (e.g., corridor) • Street segment of defined length (e.g., one mile)
Areawide	Network	A mid-sized geographic area that includes an interconnected set of transportation facilities.	• Census tracts • Census block groups • TAZs
	Regional	A large geographic area that includes all transportation facilities within a defined political boundary. Because of the large geographic size, land use at this scale can be heterogeneous within a defined area.	• City • County • Metropolitan Statistical Area • State

Facility-Specific Scales



Areawide Scales

Figure 5. Illustration of Geographic Scales for Exposure Estimation

Considerations When Selecting Geographic Scale

When selecting a geographic scale for risk assessment, one should consider several practicalities and limitations:

- **Small or zero numbers:** Granular scales may be susceptible to the small numbers problem, wherein the number of crashes or exposure may be very small or even zero for some locations or areas. These small or zero values for crashes or exposure can provide results that are misleading or not useful. Caution should be used in selecting scales that are too granular for the input data.

- **Applying areawide characteristics to specific facilities:** For facility-specific scales, some context characteristics can be derived from the encompassing area. For example, demographic or socioeconomic variables from census tracts or block groups can be applied to specific intersections or street segments within that geography. This assignment assumes that these variables are relatively constant within the defined area or geography. The assignment of area characteristics to a street segment can become more complicated if the defined segment traverses multiple areas with different characteristics. In these cases, a weighted average (based on portion of overall segment length) can be used, but one should note that these weighted averages for long segments could mask interesting contextual differences at a given location.

- **Major streets on defined area boundaries:** Major streets and roads often form the boundaries for defined areas (such as census geographies or TAZs). Depending on the locational precision of reported crashes and street geometry, pedestrian and bicyclist crashes that occur on major streets could be split between the defined areas on either side of the major street, thereby diluting the number of crashes that are assigned to each area on either side. In turn, this could lower the calculated crash frequency and risk values for these areas where a major street forms a boundary, providing misleading results for both areas.

Multiple Analysis Scales and Aggregation Considerations

In some cases, several geographic scales may be desired for the exposure and risk values. For example, one may want to develop risk values at the segment level as well as for TAZs. In cases when multiple geographic scales are desired, one should select the most granular desired scale for estimating exposure and risk, and the more aggregate desired scales can be calculated by combining the granular exposure estimates (assuming that the more granular exposure estimates are feasible to calculate). Later sections in this chapter discuss this aggregation process in more detail. Note that it is much less feasible to estimate exposure at an aggregate scale and then decompose the aggregate values to a more granular scale.

As shown in the rightmost column of Table 6, there are multiple possible scales within each of the four scale categories. For example, the Segment scale category could include several possible scenarios such as short segments between intersections, longer segments that traverse multiple intersections, or even multiple parallel segments in a miles-long corridor. However, in all of these possible scenarios, the exposure will be estimated at its base scale unit (in this case, a segment), and then aggregated up to the desired analysis, reporting or presentation level. The following sections provide several examples of aggregating to slightly different scales within the same scale category (as shown in Table 6).

Combining Multiple Crosswalks to Intersection

Figure 6 illustrates that multiple crosswalks can be aggregated up to the intersection point (I) by summing the total crossings per crosswalk (A+B+C+D) (Radwan et al. 2016). Pedestrians or bicyclists that turn a corner on the sidewalk without crossing the street are not counted since they did not enter a shared space with motor vehicle traffic. Each crosswalk is treated as a separate location with its own crossing count, thereby capturing the number of times a pedestrian or cyclist is exposed to motor vehicle traffic.

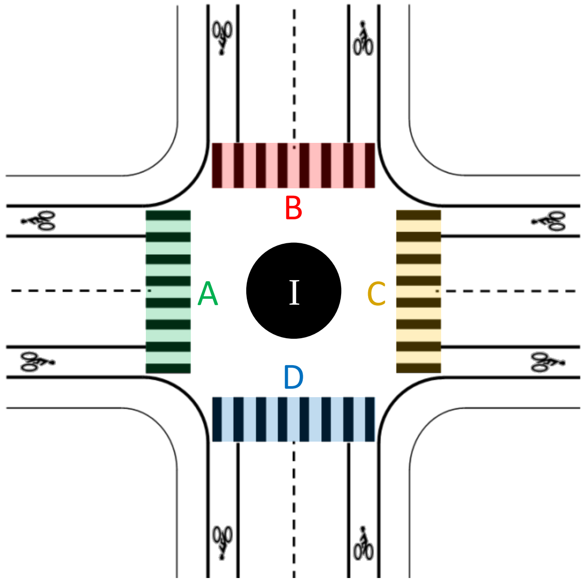


Figure 6. Illustration of Combining Individual Crosswalks for Intersection Exposure

Segments to Intersection (Point)

Two-way segment-level volume data can be used to estimate intersection volume if dedicated intersection count data do not exist (Wang et al. 2016). As shown in Figure 7, each leg of an intersection (A, B, C, D) represented by individual segments and should be summed and then divided by two to equal the total intersection volume represented as a point (I). Depending on the segment-level volume data, this method may exclude specific types of traffic, such as pedestrian traffic on sidewalks that does not cross a street at the designated intersection.

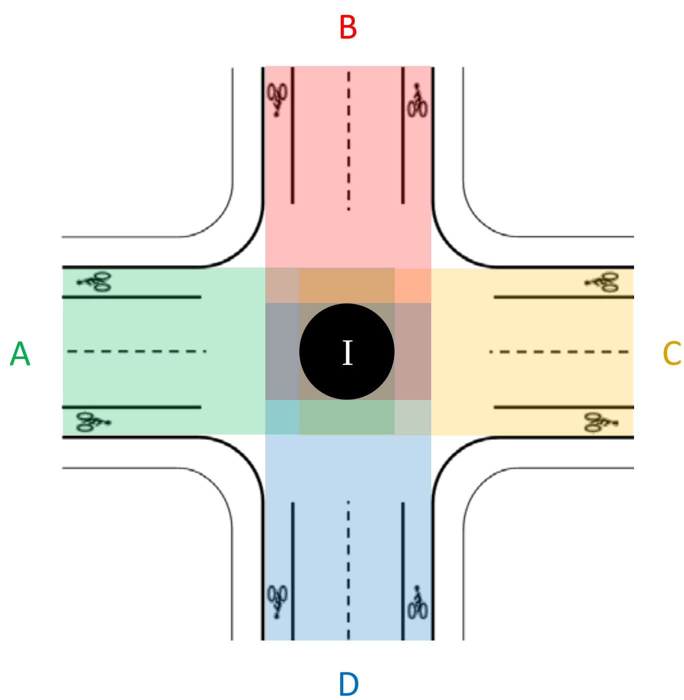


Figure 7. Illustration of Estimating Intersection Exposure from Segment Counts

Segments to Corridor

Corridor volumes can be calculated by aggregating the volume multiplied by length for each of the parallel segments that make up the corridor, e.g. roadways, bikeways, multi-purpose paths, etc. The example below shows the Cesar Chavez Blvd. corridor in downtown Austin, Texas along Lady Bird Lake, which is comprised of a roadway, a multi-purpose path, and a bikeway, as shown in Figure 8.

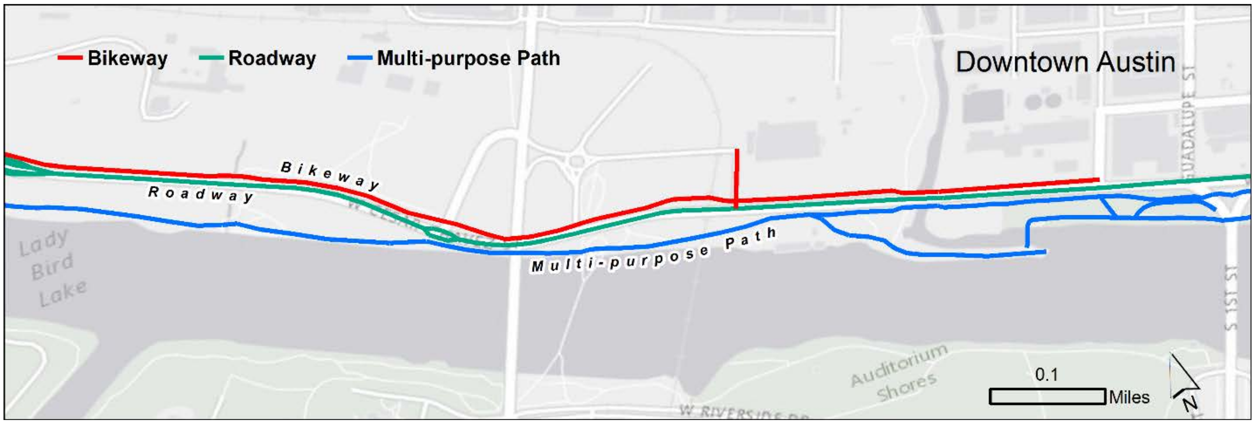


Figure 8. Illustration of Multiple Parallel Facilities in a Corridor

Figure 9 shows each facility divided into segments with dots as segment endpoints. In this example, bicycle-miles of travel (BMT) is calculated using crowd-sourced data as bicycle counts for each of the facility segments and then summed to equal the total BMT for per facility in the corridor. Table 7 shows the BMT totals per facility, which were summed to equal the total BMT for the entire corridor.

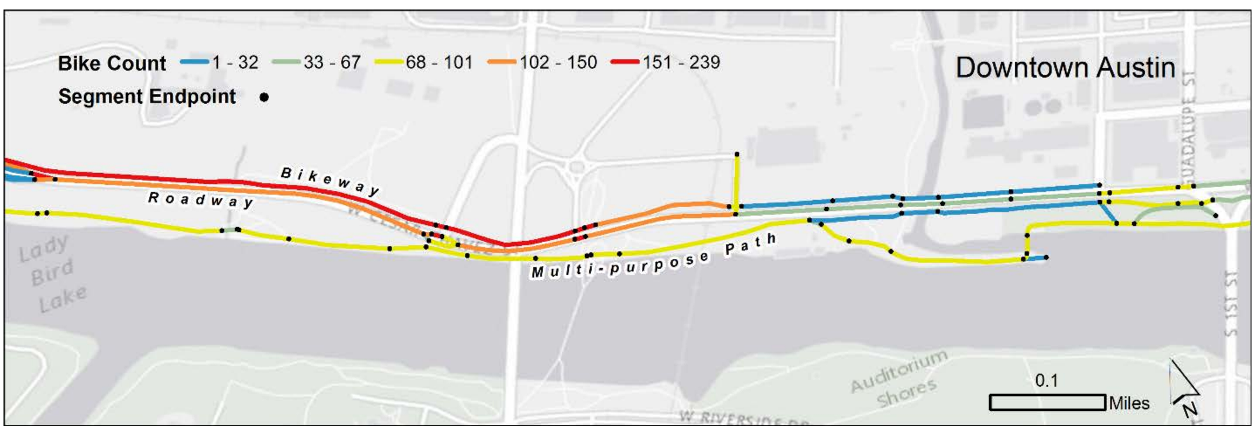


Figure 9. Illustration Showing Count Segments on Multiple Parallel Facilities in a Corridor

Table 7. Illustration of Combining BMT for Multiple Parallel Facilities in a Corridor

Facility	Number of Segments	BMT
Bikeway	14	183
Roadway	24	148
Multi-purpose Path	33	149
Corridor Total	**71**	**480**

Segments to Multiple Combined Segments

Figure 10 illustrates how smaller individual street segments could be combined into a longer segment. The example is for University Drive in College Station, Texas that runs along a major university, of which Wellborn Road bisects into Central Campus and West Campus districts. Each district has its own unique urban character that could possibly affect pedestrian traffic, and therefore, the roadway is treated as separate segments. In this case, pedestrian-miles of travel (PMT) would be calculated per roadway segment in each district and summed to equal total PMT for a both West Campus and Central Campus University Drive.

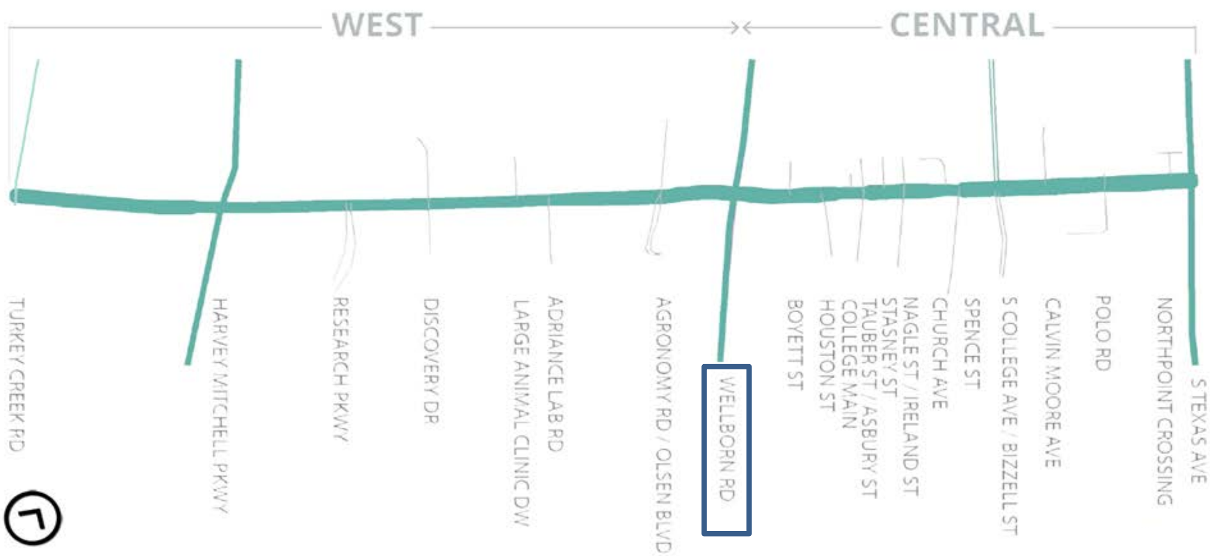


Figure 10. Illustration of Multiple Segments Being Combined into Single Longer Segment

Network to Regional

The example below (Figure 11) shows the region, represented by Travis County, Texas, subdivided by Census tracts to represent a network level scale. Each tract contains the average daily BMT, which can be summed to total the regional average daily BMT.

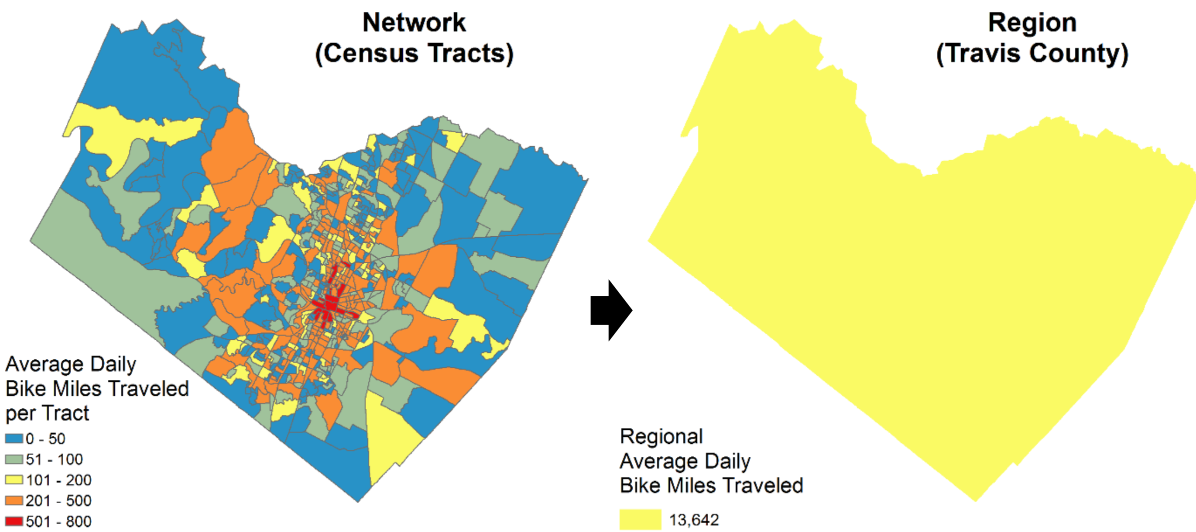


Figure 11. Illustration of Multiple Census Tracts Being Combined into Single Region

STEP 3. SELECT RISK DEFINITION

In the safety literature, the term risk has been defined as a measure of the probability of a crash to occur given exposure to potential crash events (Turner et al., 2017). This conceptual definition, however, does not provide enough specific detail for one to calculate a risk value.

Therefore, Step 3 involves selecting a specific definition of risk that will be used to calculate quantitative risk values. Table 8 shows three possible definitions of risk, with two of the three definitions closely related (i.e., one uses observed crashes, while the other uses expected crashes).

Table 8. Three Possible Definitions of Pedestrian and Bicyclist Risk

	1. Observed Crash Rate	**2. Expected Crashes**	**3. Additional Risk Indicators**
Description	• Risk = Observed crashes divided by exposure • Obtain **observed crashes** from available crash database(s). • Estimate **exposure** with this guide.	• Risk = Expected crashes • Estimate **expected crashes** with HSM or other statistical models, using exposure as input variable. • Estimate **exposure** with this guide.	• Risk = Function of one or more risk indicators: observed crashes, facility type or condition, motor vehicle speed & volume, adjacent land use, exposure, etc. • Estimate **exposure** (if included) with this guide.
Strengths	• Common use among many practitioners. • Use with other crash analysis tools (e.g., Pedestrian and Bicycle Crash Analysis Tool, PBCAT).	• Use of expected crashes overcomes issues with low (or no) observed crash frequency. • Permits evaluation of implemented countermeasures.	• Compatible with FHWA Systemic Safety Analysis. • Approach geared to practitioners.
Limitations	• Low exposure or low (or no) frequency of observed crashes may not accurately represent risk.	• Requires advanced statistical methods to estimate expected crashes. • HSM pedestrian and bicyclist tools still in early stages, may not address all site locations.	• Risk is a dimensionless numeric score or rating, not a crash frequency or crash rate value.

The first risk definition is a crash rate (i.e., observed crashes per exposure unit), whereas the second risk definition is a crash frequency (i.e., expected crashes per time unit). There is ongoing discussion in the transportation safety community about the appropriate use of crash rate versus crash frequency in different safety analyses. Crash frequency reflects the overall magnitude of crashes for a given location and time period, whereas crash rate reflects a normalized rate of crashes for a given exposure. If a crash rate is used to quantify risk, a location with just a few crashes but also very low exposure could be classified as a high-risk location, due to the low exposure value in the denominator of the crash rate equation. Similarly, a location with many crashes but also a very high exposure could be classified as a low-risk location, due the high exposure value in the denominator of the crash rate equation. Therefore, one should carefully consider whether the appropriate measure is crash frequency, crash rate, or a combination of these and other safety performance measures. In some safety analyses, both crash frequency and crash rate measures are used to understand better the different dimensions of pedestrian and bicyclist safety.

The remaining sections of this chapter describe examples of these three risk definitions that have been used in practice. These examples are from existing literature and are used to illustrate similarities and differences between the uses of these three risk definitions.

Definition 1: Observed Crash Rate

The first risk definition is an observed crash rate, which is calculated by dividing observed (i.e., reported) pedestrian or bicyclist crashes by the corresponding exposure measure. Observed pedestrian or bicyclist crashes are compiled from available crash databases. The exposure measure in the denominator is estimated or calculated as described in this guide.

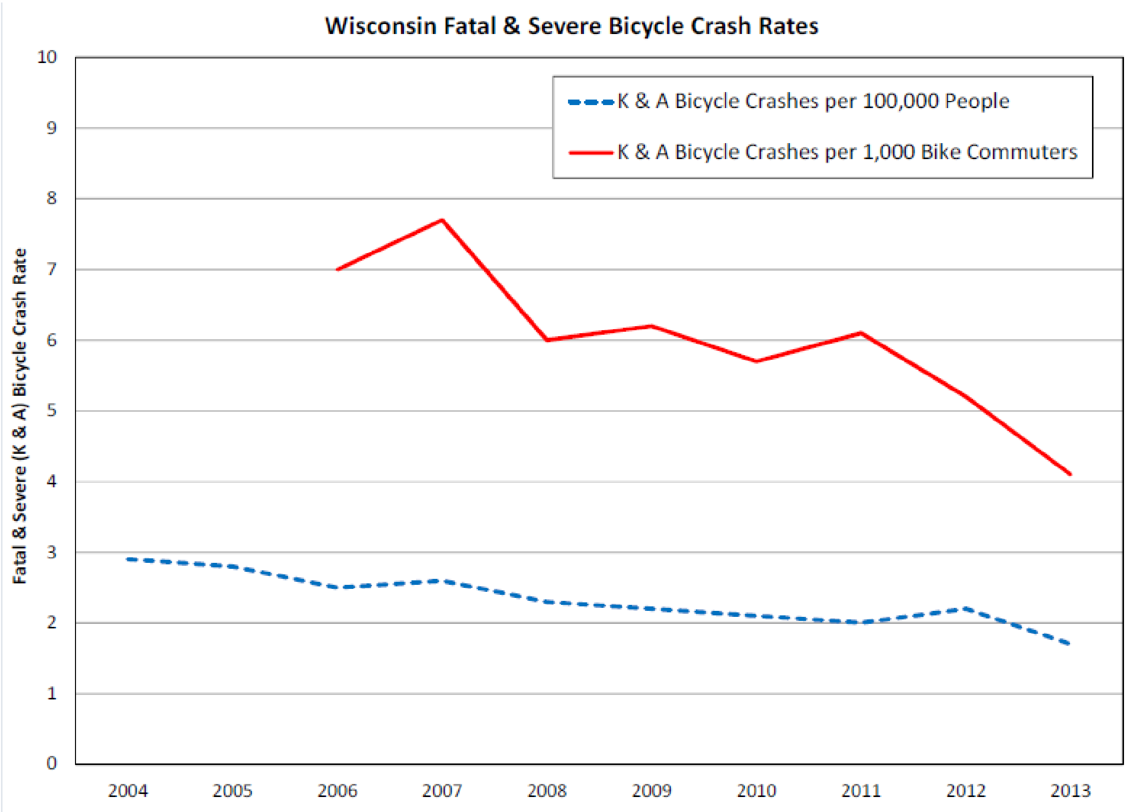
Example: Observed Crash Rate

Schneider et al. (2015) explored the characteristics of pedestrian and bicycle crashes to assist the Wisconsin DOT in providing safe pedestrian and bicycle travel through education, enforcement, and engineering treatments. The authors calculated observed crash rates for the entire state of Wisconsin to determine the historical trend of pedestrian and bicycle crash risk over time. Crash data were obtained from the WisTransPortal Database for years 2011-2013.

Population and level of activity (i.e., walking, bicycling, and driving) were used as measures of exposure to account for any impact that a changing population size or travel behavior may have on risk (i.e., observed crash rate). The authors used the U.S. Census Bureau's ACS intercensal population estimates and commuters by mode to calculate crashes per 100,000 people and crashes per 1,000 walk/bike commuters. Wisconsin DOT VMT data were used to account for the level of driving in the state over the same time by calculating the crashes per million VMT. These measures were used in the absence of statewide direct pedestrian and bicycle counts.

Note: For “K & A”, “K” are fatal crashes and “A” are crashes with incapacitating injuries.

Figure 12 shows the historical trends for the number of fatal (K-level) and incapacitating (A-level) bicycle crashes relative to population and the number of bicycle commuters. The graph shows the difference in rates between the two forms of exposure, although they both indicate a downward trend over the years. The same is true for the pedestrian crash rates.



Note: For "K & A", "K" are fatal crashes and "A" are crashes with incapacitating injuries.

Figure 12. Historical Trends for Observed Bicycle Crash Rates in Wisconsin
Source: Schneider et al., 2015.

Despite their differences, these exposure measures allowed the authors to understand better the change in pedestrian and bicycle crash risk by accounting for exposure at the state level. Table 9 shows that the number of bicycle crashes has only declined slightly between the years 2004-2013, whereas the bicycle commuter crash rate was nearly halved. Therefore, an increase in the number of bicycle commuters was associated with a reduction in bicycle crash risk in Wisconsin.

Table 9. Historical Trends for Observed Bicycle Crashes in Wisconsin

Severity Level	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Fatality (K)	14	14	8	10	9	7	9	12	11	10	**104**
Incapacitating Injury (A)	145	144	132	136	119	118	109	105	115	87	**1,210**
Other/No Injury (B, C, or O)	1,091	1,064	1,034	1,093	1,005	990	1,055	1,003	1,098	924	**10,357**
Total	**1,250**	**1,222**	**1,174**	**1,239**	**1,133**	**1,115**	**1,173**	**1,120**	**1,224**	**1,021**	**11,671**

Source: WisTransPortal Database (Wisconsin TOPS Laboratory 2014a)
Source: Schneider et al., 2015.

Definition 2: Expected Crashes

The second risk definition is expected crashes, which are typically estimated using statistical models described in the HSM and elsewhere. These statistical models use exposure as an independent variable (along with other variables) to estimate expected crashes. The exposure input to statistical models can be estimated or calculated as described in this guide.

Example: Expected Crashes

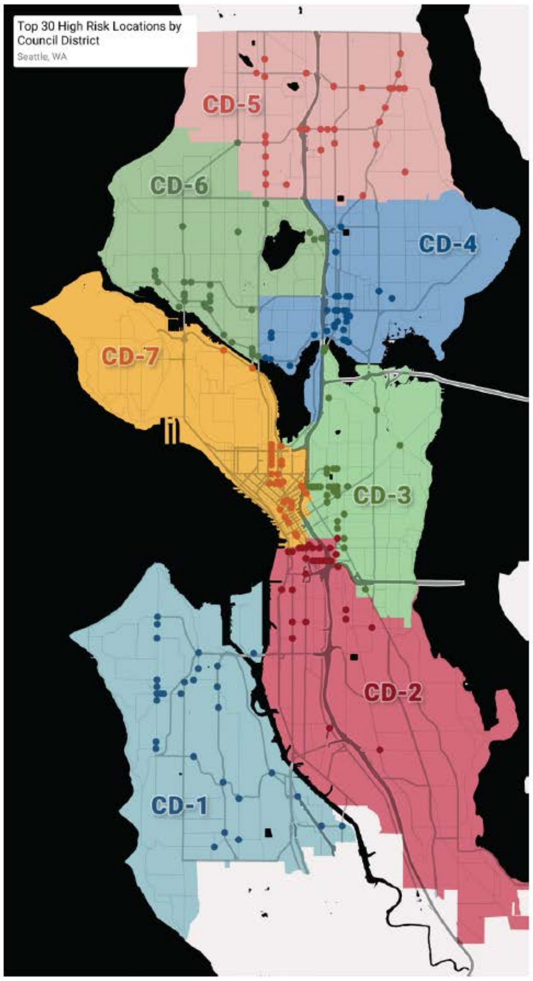
As part of its Vision Zero initiative, the Seattle DOT and others conducted a pedestrian safety analysis to proactively identify locations and prioritize safety improvements with the goal of preventing future crashes (Seattle DOT 2016, Thomas et al. 2017). The Seattle analysis used a risk definition of expected crashes to overcome several limitations of observed crashes at specific locations, such as regression-to-the-mean and a reactive focus only where crashes have occurred in the past rather than where they are likely to occur in the future.

The Seattle analysis developed statistical models for two crash types:

- All reported pedestrian-motor vehicle collisions at intersections.
- Motor vehicle traveling straight, pedestrian crossing at intersection.

The analysis included a range of roadway inventory, motorized and non-motorized traffic (i.e., exposure), land use, and socioeconomic variables to estimate expected pedestrian crashes.

The statistical crash models were used to estimate expected crashes for the defined Seattle street network, for the purposes of identifying those locations with the highest risk for pedestrian crashes in the future. The network screening results are shown in Figure 13, which illustrates the top 30 ranked locations in each Seattle City Council District. Note that although this hotspot map looks similar to a traditional crash hotspot map, it is based on expected crashes (i.e., where crashes are most likely to occur in the future) rather than observed crashes. The sidebar in this figure summarizes the key differences between risk definitions based on observed historical crashes versus expected crashes.



Systemic vs. High Crash Approach to Reducing Crashes

A systemic approach 1) proactively identifies sites based on risk factors associated with a particular crash type, and 2) uses cost-effective strategies to address potential safety issues system-wide. These strategies might include locations with and without a crash history. This allows us to address future safety risks before they become an issue.

This approach complements our traditional high-collision analysis, which identifies and recommends safety improvements for locations with a high number of crashes ('hot spots'). We'll also continue to address safety concerns at bicycle or pedestrian crash 'hot spot' locations.

Figure 13. Locations in Seattle with High Risk of Pedestrian Crashes based on Expected Crashes
Source: Seattle DOT, City of Seattle Bicycle and Pedestrian Safety Analysis, 2016

Definition 3: Additional Risk Indicators

The third risk definition is using one or more risk indicators that are combined into a composite risk score. The risk indicators are those traffic, roadway, and land use features that are thought (or proven) to contribute to higher levels of pedestrian and bicyclist crashes. The risk indicators may be identified through a formal analysis (such as FHWA's System Safety Analysis process) or through an informal locally defined process. In this approach, pedestrian and bicyclist exposure (or those features that lead to higher exposure, such as walkable streets and mixed land use) is often used as a contributing risk factor. Like the previous two risk definitions, this guide is intended for use to estimate exposure in this Additional Risk Indicators definition of risk.

Example: Additional Risk Indicators

The Minnesota DOT (MnDOT) 2014 Strategic Highway Safety Plan identified pedestrian and bicyclist crashes as a statewide priority. After completing county-specific safety plans, MnDOT decided to update the safety plans for each of the districts using the systemic approach. Initial analysis showed that about 67 percent of all severe pedestrian and bicyclist crashes occurred in the Minneapolis-St. Paul metropolitan area. As a result, a systemic evaluation of this area was conducted with a focus on pedestrian and bicyclist crashes. The St. Paul case study shows how MnDOT applied the systemic safety process to pedestrian and bicyclist crashes in the urban areas across the State's system in Greater Minnesota. Databases used in this analysis are: 1) MnDOT crash record system, 2) video logs, 3) Google Earth for geometric details, and 4) MnDOT database on traffic volumes.

For the Minnesota district safety plans, the systemic safety process identified the following risk indicators as significantly contributing to pedestrian and bicyclist crashes:

- Presence of traffic signal control
- High speed limit (major approach)
- Presence of pedestrian trip generators
- Motor vehicle traffic volume
- Intersection skew
- Roadway curvature

Table 10 shows a table excerpt of signalized intersections for pedestrian and bicyclist risk. Note that a simple ranking score (i.e., one star for each risk indicator) was used to rank and prioritize signalized intersections.

Table 10. Example of Ranking Signalized Intersections based on Multiple Risk Indicators

#	Intersection ID	Route System	Route No.	Description	Speed Limit	Cross Product[b]	Traffic Control	Major Corridor Speed	Skew	On/Near Curve	Primary Land Use	Severe Ped/Bike Crash Density	Total Stars	Crash Cost
34	3.210.025	MN	210	4TH ST NWCSAH20 MSAS103/BRNRD	35	★	★	★		★	★		★★★★★	$1,050,200
35	3.024.009	MN	24	CSAH 75/CLEARWATER	40	★	★	★	★		★		★★★★★	$747,600
36	3.023.028	MN	23	19 1/2 AV/ST CLD	35	★		★	★	★	★		★★★★★	$574,800
37	3.023.050	MN	23	TH 25/FOLEY	45	★	★	★	★		★		★★★★★	$558,000
38	3.027.015	MN	27	4TH ST MSAS 106/LITTLE FALLS	30	★	★		★	★	★		★★★★★	$366,400
39	3.023.011	MN	23	RED RVR AVCSAH 2/COLD SPRING	35	★	★	★		★	★		★★★★★	$292,800
40	3.023.020	MN	23	6TH AV S MSAS107 M95/WAITPK	40	★	★	★		★	★		★★★★★	$0
41	3.210.021	MN	210	ELDER DR SM140/BAXTER	55	★		★	★		★		★★★★	$10,558,200
42	3.012.003	US	12	JOHNSON AVE M-54 LT/COKATO	35	★		★			★	★	★★★★	$10,418,000
43	3.015.011	MN	15	N JCT TH 23 DIV ST/ST CLOUD	45	★	★	★			★		★★★★	$5,838,400
44	3.015.012	MN	15	3RD ST N CSAH81 MSAS 114/STC	45	★	★	★			★		★★★★	$4,310,200
45	3.169.004	US	169	197TH AV MSAS116 M118/ELKRV	55	★	★	★			★		★★★★	$1,696,200
46	3.015.019	MN	15	CSAH 29/SAUK RAPIDS	60	★	★	★			★		★★★★	$1,671,800
47	3.010.011	US	10	E JCT TH 210 LT/MOTLEY	30	★	★			★	★		★★★★	$1,612,200
48	3.210.026	MN	210	4TH ST N MSAS114/BRAINERD	35	★	★	★			★		★★★★	$1,241,800
49	3.210.027	MN	210	TH 371B RTM 60 LT/BRAINERD	35	★	★	★			★		★★★★	$1,186,600
50	3.023.022	MN	23	WAITE AVEMSAS101/WAITEPARK	40	★	★	★			★		★★★★	$1,146,000
53	3.025.030	MN	25	RIVER ST MSAS112/MONTICELLO	30	★	★			★	★		★★★★	$891,400
54	3.012.020	US	12	BUFFALO AVCSAH 12TH 25/MONTR	35	★	★	★			★		★★★★	$641,000
55	3.023.088	MN	23	N JCT TH 65 CSAH 6/MORA	30	★	★		★		★		★★★★	$622,200
56	3.025.029	MN	25	BROADWAY CSAH75/MONTICELLO	30	★	★			★	★		★★★★	$619,600

Source: Report FHWA-SA-17-002, Systemic Safety Project Selection Tool Supplemental Case Studies, December 2016.

STEP 4. SELECT EXPOSURE MEASURE

Step 4 in the scalable risk assessment process is to select a specific exposure measure to be used in the calculation of risk values. There are several different categories of exposure measures that attempt to quantify the level of contact that pedestrians and bicyclists have with potentially harmful safety outcomes. The five categories of exposure measures included in the guide are briefly defined below.

- **Distance Traveled:** Exposure measures in this category are based on the cumulative distance traveled by all pedestrians and bicyclists at the specified geographic scale. The most common measures in this category are pedestrian miles traveled (PMT) or bicyclist miles traveled (BMT). In cases when the exposure measure values are numerically large, the measures are expressed in thousands or millions of miles traveled.

- **Time Traveled:** Exposure measures in this category are based on the cumulative time traveled by all pedestrians and bicyclists at the specified geographic scale. The most common measures in this category are pedestrian hours traveled or bicyclist hours traveled.

- **Volume/Count:** Exposure measures in this category are based on the volume or count of pedestrians or bicyclists for a specified time period and geographic scale. In some cases, the counts may be annualized for a typical day (such as annual average daily traffic (AADT)), or may represent shorter periods of time (such as an hour). Exposure measures in this category do not capture the distance or time traveled, only the number of pedestrians or bicyclists. Count-based exposure measures are typically used for facility-specific geographic scales, since counts are readily associated with specific street crossings or segments.

- **Trips Made:** Exposure measures in this category are based on the cumulative number of trips made by all pedestrians and bicyclists at the specified geographic scale. Trip-based exposure measures are typically used for areawide geographic scales, since a pedestrian or bicyclist trip typically includes numerous streets segments and crossings.

- **Population**: Exposure measures in this category are based on the population (or specified sub-population) at a specified geographic scale. The most common measure in this category is the number of people (or percent of the population) that walk or cycle. By definition, population-based exposure measures must be used for areawide geographic scales (i.e., no population associated with street segments or crossings).

The selection of an exposure measure will depend upon several criteria, such as the use of the risk values (Step 1), the geographic scale (Step 2), and other criteria. This chapter of the guide introduces the various categories of exposure measures and provides guidance on selecting the most appropriate exposure measure given criteria (such as geographic scale and analytic method). Table 11 contains a selection matrix to help analysts choose an exposure measure best suited for their analysis. Note that each exposure measure will be for a defined time period that matches other variables in the risk definition (such as crashes or other risk indicators). Table 12 provides guidance on the strengths and limitations of each category of exposure measure.

Table 11. Selection Matrix for Exposure Measures

Category of Exposure Measure	Typical measures	Typical scale				Typical data sources
		Point	Segment	Network	Region	
Distance Traveled	Miles of travel	○	●	●	●	• Site counts or demand estimation models, multiplied by segment length • Sometimes travel surveys
	Miles crossed per entering vehicle	◐				
Time Traveled	Hours of travel	○	○	●	●	• Travel surveys • Sometimes site counts combined with crossing time or average travel speed data.
	Product of crossing time and vehicle volume	○	○			
Volume/ Count	Volume/count	●	●			• Site counts • Demand estimation models
	Product of pedestrian /bicyclist volumes and motor vehicle volumes	◐	◐			
Trips Made	Number of trips			●	●	• Travel surveys
Population	Number of people that walk or cycle on regular basis			●	●	• U.S. Census data products
	Percent of the population that walk or cycle on regular basis			●	●	

Legend: ○ = to a small extent; ◐ = to a moderate extent; ● = to a great extent.
Note: Each exposure measure will be for a defined time period that matches the risk definition.

Source: Partially adapted from Greene-Roesel et al., Estimating Pedestrian Accident Exposure: Protocol Report, March 2007.

Table 12. Strengths and Limitations for Exposure Measure Categories

Category	Strengths	Limitations
Distance Traveled	• Most commonly used measure for motor vehicle exposure. • Can be used at facility-specific and areawide geographic scales. • Calculation typically based on simple data inputs (counts and segment lengths). • Better represents quantity of travel than population or volume/count-based measures.	• Not the best measure for comparing risk between different modes, due to shorter distances travelled for pedestrians and bicyclists.
Time Traveled	• Better measure for comparing risk among different travel modes, as it accounts for different prevailing travel speeds (and therefore travel times) for each mode. • Can be used at facility-specific and areawide geographic scales. • Better represents quantity of travel than population or volume/count-based measures.	• The number of walking and bicycling trips are often underreported in travel surveys. • The amount of time traveled for walking or bicycling can be overestimated in surveys, especially if based on self-reporting or recall.
Volume/ Count	• Simple exposure measures that are based entirely on site counts. • No assumptions are made about distance or time traveled.	• Does not account for the distance or time traveled by pedestrians or bicyclists, which is an important weighting mechanism in exposure measures. • Less meaningful when aggregated to areawide geographic scales.
Trips Made	• Trip reporting is common in travel surveys and also emerging GPS-based smartphone applications. • Provides meaningful exposure measures for areawide geographic scales. • Trip-based measures can be subdivided (i.e., by trip purpose) for more detailed exploratory analysis.	• Walking and bicycling trips are often underreported in travel surveys. • A large number of survey respondents are needed to adequately represent the full population. • Trip-based measures are not meaningful for facility-specific geographic scales.
Population	• Typically easy and low-cost to obtain; available for most areawide geographic scales. • Travel surveys can be used to subdivide the total population to the portion that walk or cycle on a regular basis.	• Does not account for the number of trips or intensity (i.e., distance or time traveled). • Simple population-based measures may not even account for the portion of the population that walks or cycles (e.g., available census data are limited to commute trips made by the working population). • Does not capture visitor or tourism trips/travel.

Source: Partially adapted from Greene-Roesel et al., Estimating Pedestrian Accident Exposure: Protocol Report, March 2007.

- **Distance Traveled:** Exposure is often measured in terms of distance traveled by multiplying the segment count by the segment length. For example, if a community is interested in the total BMT for a specific bikeway that is comprised of several segments, then bicycle count data is necessary for each roadway segment along with their individual lengths in terms of miles. BMT could then be calculated per segment by applying the following equation:

$$Segment\ Length * Mode\ Volume = \ Mode\ Distance\ Traveled$$

In Figure 14, segments AB and BC have individual BMT totals since they are different lengths and have their own bicycle count data. Each segment's BMT can then be summed together to equal the total BMT for the complete length of the bikeway from A to C to equal 355 BMT. Mode-specific distance traveled is usually expressed in terms of the measurement units used for segment length (e.g., 1.25 *miles* * 200 bicycles = 250 Bicycle *Miles* Traveled).

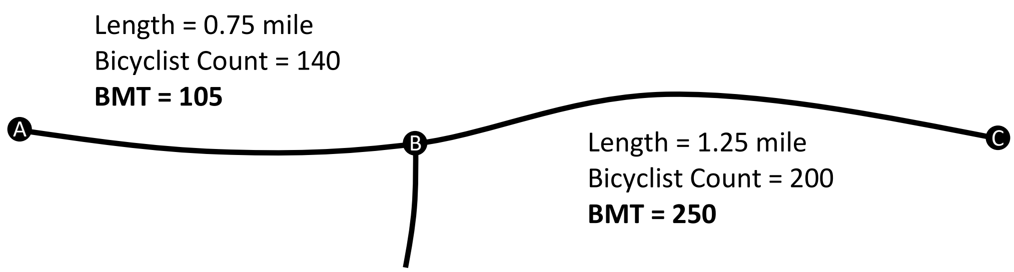


Figure 14. Calculating Distance Traveled Exposure Measure from Segment Counts

- **Time Traveled:** Household travel surveys can be used to measure the amount of time traveled by a specific mode. Survey respondents fill out a travel diary indicating origins and destinations with the start and end times of trips along with the mode that was used. Since the survey represents only a stratified sample of the population, weights must be applied to expand the survey sample so that it represents the entire population of the study area. Survey weights indicate how many households each survey observation represents of the total population of households – these weights are typically provided along with the survey data. For this purpose, the analyst needs to enumerate the total duration of trips by mode per household type as defined in the survey stratification. They then need to multiply each household type's total duration of trips by its corresponding survey weight to equal the total daily duration of trips by mode for the entire study area (Figure 15).

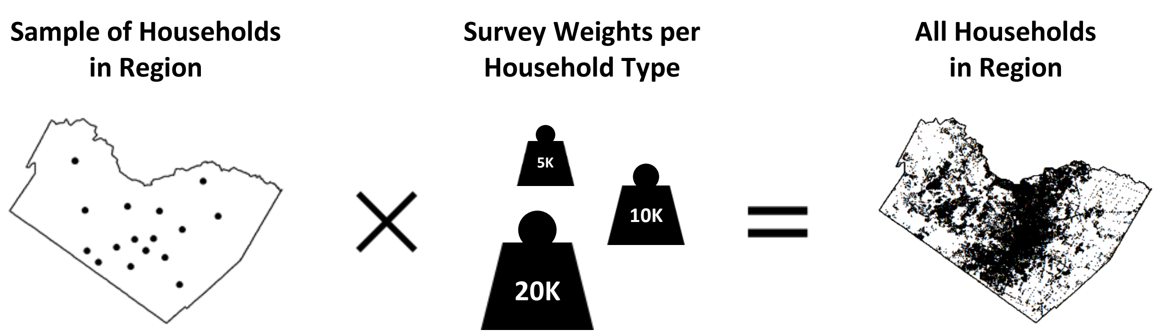


Figure 15. Calculating Time Traveled by Expanding Household Travel Surveys

For example, the regional household travel survey for Austin, Texas can be used to estimate the total amount of time traveled by walking for the five-county region. The survey sample is comprised of 3,000 households and 8,100 persons and can be expanded to represent the population of the study area by applying the survey weights. The result is a total of 189,256 daily walk trips with an average trip duration of 16 minutes equaling approximately 50,437 hours of walking per day. The NHTS also provides these details for various geographies at a regional scale.

- **Volume/Count:** Exposure measured in terms of volume or counts is common for intersection studies. Such studies may only include bicycle or pedestrian crossing volumes; however, this approach ignores the potential conflicts caused by motor vehicles also passing through the intersection. A more comprehensive exposure measure for intersections is the product of pedestrian/bicyclist volumes and motor vehicle volumes. For example, the following intersection has two counters (C) on opposing corners of the intersection so that they each observe two legs of the intersection, as indicated in Figure 16.

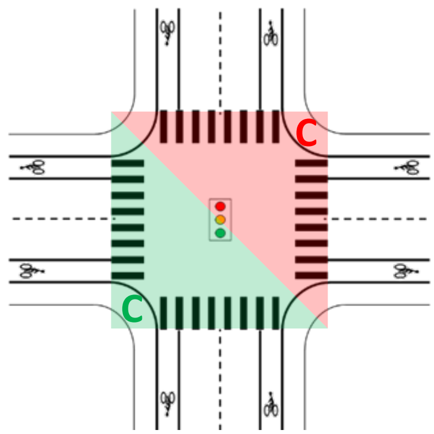


Figure 16. Estimating Count-Based Exposure at Intersections

Each counter counts the total number of entering motor vehicles and bicyclists, as well as the pedestrians using the crosswalks. In this example, pedestrians are counted each time they cross the street and are excluded if they remain on the sidewalk to turn the corner. Since this study is focused on the intersection as a whole, the two sets of counts are combined to equal the total number of bicyclists, pedestrians, and motor vehicles entering the intersection. From these totals, an analyst can apply the following equations to calculate both the bicycle and pedestrian exposures for the intersection:

$$Bicycle\ Volume * Motor\ Vehicle\ Volume = Bicycle\ Exposure$$

$$Pedestrian\ Volume * Motor\ Vehicle\ Volume = Pedestrian\ Exposure$$

The above example illustrate one possible way to count non-motorized and motorized volume at an intersection. There are several ways to count user volumes at intersections, and the preferred method may depend on overall intersection user volume, user mix, common turning movements, sight lines, the potential to use automated technology, and data collection resources.

- **Trips Made:** Similar to "Time Traveled," household travel surveys can be used to measure the number of trips made by a specific mode within a specific study area. The method of expanding the survey sample is the same as above. In this case, the analyst needs to enumerate the total number trips by mode per household type as defined in the survey stratification. They then need to multiply each household type's total number of trips by its corresponding survey weight to equal the total number of daily trips by mode for the entire study area. For example, the regional household travel survey for Austin, Texas can be used to estimate the total number of trips by mode for the five-county region. The survey sample is comprised of 3,000 households and 8,100 persons and can be expanded to represent the population by applying the survey weights. The result is 189,256 daily walk trips.

- **Population:** Population estimates can be used as an alternative to direct exposure measurements when it is impractical to collect exposure data due to cost constraints. Some of the most commonly used population estimates are those provided by the U.S. Census Bureau, such as the decennial census and ACS. Each of these possess varying levels of data currency and underlying sample sizes. For example, if an agency wanted to know the approximate total number of people who walk to work within Travis County, then they could use the 2016 ACS 1-year estimate for the "Commuting Characteristics by Sex" (see table S0801 and B08301). The ACS data in Table 13 states that Travis County has approximately 662,881 workers 16 years and over, of which 1.8% or 19,932 workers commute to work by walking. Margins of errors are also provided so that confidence intervals can be calculated around the estimates.

Table 13. Extracting Population-Based Exposure from ACS Data Tables

Subject	Estimate	Margin of Error
Workers 16 years and over	662,881	7,656
MEANS OF TRANSPORTATION TO WORK (%)		
Car, truck, or van	83	1
Drove alone	73.3	1.1
Carpooled	9.7	0.8
In 2-person carpool	7.6	0.8
In 3-person carpool	1.2	0.3
In 4-or-more person carpool	0.9	0.3
Workers per car, truck, or van	1.07	0.01
Public transportation (excluding taxicab)	3.1	0.4
Walked	**1.8**	**0.4**
Bicycle	1.2	0.3
Taxicab, motorcycle, or other means	1.3	0.3
Worked at home	9.5	0.8

STEP 5. SELECT ANALYTIC METHOD TO ESTIMATE EXPOSURE

Step 5 in the scalable risk assessment process is to select an analytic method (or methods) to estimate exposure. There are numerous analytic methods that can be used to estimate exposure, and the most appropriate method(s) depends upon several criteria, such as desired geographic scale (Step 2), desired exposure measures (Step 4), analysis scope, data availability, staff technical capabilities, and available analysis resources.

As indicated in Figure 1, there may be some iteration or concurrency between this step (selecting an analytic method) and the previous step (Step 4, selecting an exposure measure). For example, an analyst may want to calculate a specific exposure measure, but have no expertise in the most common analytic methods used to estimate that exposure measure. Or, a specific analytic method may not be able to accurately estimate a specific of exposure measure. Therefore, many analysts are likely to consider both the desired exposure measure and the most feasible analytic method in a concurrent or iterative manner.

This chapter introduces these analytic methods and provides guidance on selecting the most appropriate method(s) given various criteria. The analytic methods outlined in the guide are listed below and briefly described in the following pages.

- Site counts
- Demand estimation models
    - Direct demand models
    - Regional TDM
    - Trip generation and flow models
    - GIS-based models
    - Discrete choice models
    - Simulation-based traffic models
    - Data fusion
- Travel surveys
    - ACS
    - NHTS
    - Regional household travel survey

Table 14 provides a method selection matrix to help analysts make informed choices about which analytic method(s) is best suited for them. It is important to note that local customization may be required for all these models to be useable.

Table 14. Selection Matrix for Analytic Methods to Estimate Exposure

Analytic Method		Input Data Requirements	Technical Complexity	Popularity in Practice	Direct Usability	Accuracy
Site counts		○	○	●	●	○/◐/●
Demand Estimation Models	Direct demand models	◐	○/◐	●	◐	○/◐
	Regional TDM	◐/●	◐/●	○	○/◐/●	○/◐/●
	Trip generation and flow models	◐/●	◐/●	◐	●	◐/●
	GIS-based models	◐	◐	◐	●	◐/●
	Discrete choice models	◐/●	◐/●	◐	○	◐/●
	Simulation-based traffic models	●	●	○	●	●
	Data fusion	●	◐/●	○	●	◐/●
Travel surveys		○	○	●	●	○/◐/●

Legend: ○ = low suitability; ◐ = moderate suitability; ● = high suitability.

Note: For some categories, multiple ranges (e.g., ◐/●) are used since the corresponding criteria might vary significantly based on the specific characteristics of the model developed.

Site Counts

Site counts are a direct measurement of the number of pedestrians or bicyclists at a defined location for a defined time period. The counts may be gathered automatically from various technology-based sensors or manually from human observers. Site counts are taken at a point, but in some cases are applied along a street segment where the counts are not expected to vary along the segment length. Counts are more commonly used to estimate exposure when the desired facility coverage is limited, as count data collection for all facilities within a large network or region is cost-prohibitive (unless extensive sampling is used). In some cases, it is also a challenge to get representative pedestrian and bicyclist counts due to seasonal variation with these modes.

Demand Estimation Models

There are numerous estimation models that have been used to estimate pedestrian and bicyclist demand (i.e., counts) on specific facilities. The models range in complexity and input requirements, and some have been more commonly used than others. Several of the models rely on pedestrian and bicyclist count data for model development or calibration, representing somewhat of a hybrid approach that combines counts and models. In addition, while most of these models provide the volume estimate directly (e.g. direct demand models), some must be integrated with other methodologies to provide demand estimates (e.g. route choice models). The descriptions of the latter types of models have also been included in this step due to their potential role as non-motorized planning tools that can be used in exposure estimation.

Important to note that the models described here have an established history and been applied frequently in various fields. However, only a few of them have been used extensively in the context of estimating exposure for safety analysis, and the remaining models have been used infrequently or on a very limited basis. Direct demand models are noteworthy to mention since they played a major role in pedestrian and bicyclist volume estimation especially in supporting traffic safety studies. However, an overview of all potential models is provided here for completeness. Regardless of the model chosen, the analyst should keep in mind that the models might need to be customized and calibrated with respect to the local characteristics of the project (e.g. study area). The analysts are recommended to review the key considerations provided later in this section before model selection. This will help choose a feasible model that will provide an optimal solution with acceptable accuracy.

This section of Step 5 first provides an overview of the demand estimation models. A list of resource documents is then provided for analysts who are interested in learning more about these models. The section is concluded with discussions on key considerations when selecting a model.

Overview of Demand Estimation Models

Direct Demand Models

Direct demand models are statistical models that estimate facility-specific pedestrian and bicyclist volumes based on observed volumes at a sample of locations and nearby context (such as land use and form, street type, etc.). Direct demand models are often based on regression analysis. These models are commonly used and are appealing due to their simplicity; however, they are limited in terms of capturing the underlying behaviors and travel patterns that produce higher and lower volumes in particular locations. Many of the existing direct demand models are also based on relatively small sample sizes. A detailed discussion on direct demand models and step-by-step instructions to develop them are provided in Step 6 given their practicality and widely use in the literature for pedestrian and bicyclist volume estimation modeling.

Regional Travel Demand Models

Regional TDM are computerized systematic processes that estimates existing and future travel demand at a regional basis given numerous inputs, such as the transportation network, population and demographic characteristics, and trip-making behavior. The end result of a regional travel demand model is traffic volume estimates on individual transportation network links, but several other model outputs can also be obtained, such as vehicle miles of travel, mode shares, origin, and destination of trips. Regional TDM constitute part of long-range transportation plans developed for MPOs or state or local agencies. Traditional TDM and enhanced models of tour- and activity-based are the most commonly known model structures of this category.

Traditional Regional Travel Demand Models. Traditional regional TDM are the current state-of-practice models for regional travel demand forecasting, which are mostly based on trip-based approach. The trip-based models generally consist of four main steps: trip generation, trip distribution, mode share, and traffic assignment. TAZs are the most commonly used geographic units to inventory existing and future demographic data required for modeling purposes. Therefore, trip-based models are particularly limited in estimating non-motorized travel due to their coarse level of spatial analysis structure. To overcome these limitations, several enhancements have been made to trip-based models over the last recent decades. Such enhancements can also be beneficial to increase the sensitivity of trip-based models to pedestrian and bicyclist trips, such as developing an enhanced trip generation model sensitive to land use factors or an enhanced auto ownership model as input to non-motorized trip production. On the other hand, these models are not yet adequate in estimating exposure for bicycle and pedestrian travel in various ways, such as in capturing the fine-grained differences in intersection-level bicycle or pedestrian activity or in capturing recreation trip purpose which is a key consideration in pedestrian and bicyclist trip rates.

Advanced Regional Travel Demand Models. Advanced regional TDM represent an emerging practice in regional travel demand forecasting models. They are used to overcome the limitations of traditional regional TDM. For example, tour- or activity-based models provide superior alternatives to traditional four-step models since they are based on individuals rather than trips, and the spatial resolution can be reduced to a smaller level of geography (such as parcels instead of TAZ). Particularly, as indicated by Sener et al. (2009), “by placing primary emphasis on activity participation and focusing on sequences or patterns of activity participation and travel behavior rather than travel, the activity-based approach recognizes the spatial and temporal linkages among the various activity-travel decisions of an individual, as well as the linkages between the activity-travel patterns of different individuals within a household”. Tour- or activity-based models simulate the activity travel decisions of households and individuals, which yields to activity-travel patterns that are composed of various decisions, such as to determine when or where an individual participates in a certain activity. Therefore, these models provide a more behaviorally realistic representation of travel through detailed information on individuals’ activity-travel patterns. Despite their advantages, the adoption of tour- and activity-based models have been relatively slow in practice especially due to several considerations, particularly on staff, time, and budget resources needed to transfer from traditional trip-based models to tour- or activity-based models.

Therefore, while providing opportunities in estimating exposure for pedestrians and bicyclists, such advanced models have not yet become a practice in the safety field.

Trip Generation and Flow Models

Trip generation and flow models can be considered as variations of regional TDM with the volume estimate developed as the model output. Such specially focused models that can be particularly beneficial to be applied at the corridor and subarea planning level. Although the models are applicable to both bicycle or pedestrian travel, there have been mostly pedestrian-based models of this category. The two common examples are pedestrian trip generation and flow models and network simulation models.

Pedestrian Trip Generation and Flow Models. These models are conceptually similar to regional four-step models, but focus on specific market segments particularly to improve the sensitivity of the models to bicycle and pedestrian travel. Pedestrian trip generation and flow models are among the examples of these focused models, which uses smaller geographical zones rather than TAZs. They generally include trip generation, trip distribution and network assignment steps. However, since they only focus on pedestrian trips, the models do not include a mode choice component. An example application of this type of modeling is the Model of Pedestrian Demand (MoPeD), which were developed by the University of Maryland's National Center for Smart Growth. The model requires data on vehicle ownership, street connectivity, parcel-level land use, Census population and employment, and travel survey. MoPeD uses pedestrian analysis zones (which are block or street-level) as the level of analysis. The end result is an estimate of the numbers of pedestrians, or pedestrian volumes, which will occur on sidewalks and intersections in the study area over a 24-hr period (see for detailed procedural information: http://kellyjclifton.com/products/moped/).

Network Simulation Models. Network analysis models can also be categorized as variations of the four-step modeling approach, which are generally based on spatially driven network simulation procedures and use a representation of a network to estimate volumes for specific facility types over an entire area. The models use detailed network structures with complete links and nodes and with various other complementary data elements (e.g. street network density). Space Syntax is one of the most well-known example studies of network analysis models, which are described as "suite of modeling tools and simulation techniques used to analyze pedestrian movement and to predict pedestrian volume" (Raford and Ragland 2004). Although these models may have the potential to provide more appropriate exposure measures for non-motorized travel, Kuzmyak et al. (2014) indicated that two potential reasons for its relatively minimal usage: 1) the information on its special software is limited; and 2) the process is not intuitive to transportation planners since its process does not follow traditional trip generation and distribution steps. Instead, it uses spatial characteristics and relationships to explain the route chosen.

GIS-based Models

GIS-based models heavily use GIS tools and GIS-based measurements in determining activity levels. They can be described as spatial models of built environments and proximities. In the current context, these models are often used to estimate non-motorized travel under alternative land use and transportation investment scenarios. Important to note that, the models falling under this category are significantly dependent on GIS-based modeling and forecasting tools and use GIS as the main feature rather than as a tool to the modeling framework.

GIS-based Walk Accessibility Model. The GIS-based walk accessibility model (developed for NCHRP 770) provides an enhanced example to GIS-based tools, expanding their capability by estimating pedestrian trip tables using GIS-derived walk-accessibility scores. This model relies entirely on GIS tools and data to create relationships between land use activity, accessibility to opportunities defined by the shape and service of the transportation networks, and mode choice. It is similar to the Walk Score program, but it uses an enhanced methodology to estimate walk potential and mode choice. While generating walk trip tables, the accessibility model does not have an assignment module to estimate facility volumes. In addition, the model is limited to walk travel, but the structure might be applicable to bicycle travel if adequate bicycle data are available.

GIS-based Origin Destination Centrality Demand Model. Another example of GIS-based models is the methodology developed by McDaniel et al. (2014) to estimate bicycle volume, which has been recently automated via an online tool available at: http://uidaho.maps.arcgis.com/apps/webappviewer/index.html?id=88af7e023fd24d31965c4d0b62fdadd9. Although this model uses a combination of different techniques, it is included here due to its lineage with using GIS as the main feature of the modeling framework. Specifically, “the method uses network analysis to quantify travel patterns between origins and destinations through a new metric that is called origin–destination (O-D) centrality. The metric is then used as an explanatory variable in a direct demand model that is programmed as a tool for GIS software.” Once the input files are provided, the tool follows a systematic process including computing O-D centrality, calibrating the direct demand model and estimating bicycle volume across the entire street network.

Discrete Choice Models

Discrete choice models are used to analyze and predict decision makers’ preferences among several alternatives. The underlying principles of discrete choice models can also be used to determine bicyclists and pedestrian activity in the context of exposure analysis. In general, these models do not necessarily provide direct volume estimates but are included here due to their potential role as part of non-motorized planning tools that can be used in exposure estimation especially when integrated with other tools. For example, cross behavior models are used to develop information about the crossings and crossing behavior as an exposure measure. Route choice models provide significant information on the factors influencing an individual’s route choice (e.g., a bicyclist’s route), which can be then be integrated into a traffic assignment model to develop better volume estimates.

Simulation-based Traffic Models

Simulation-based traffic models are built through the application of advanced computerized programs and mathematical modeling of transportation systems. They mainly use outputs of regional TDM as inputs into their algorithm to determine detailed activity levels. These models can be applied at microscopic, macroscopic, or mesoscopic levels. For example, agent-based microsimulation models have been used to capture pedestrian activity in an area through simulation of individual pedestrian movement in crowds using complex behavioral rules and environmental modeling. While providing quite accurate, detailed, and visually strong traffic flow, these models are quite complex and require significant input data and special resources, such as specialized software (e.g., VISSIM, PARAMICS) and unique technical expertise.

Data Fusion

Data fusion is a process of integrating several data sources into a single one that provides a more accurate representation. Data fusion methodologies provide promising tools in estimating pedestrian and bicyclists’ activity especially considering the various data sources used in estimation as described above. The advancements in technology and heavy use of mobile devices have provided new opportunities in collecting crowdsourced data and extending the capability of data collection for non-motorized users (Lee and Sener 2017). For example, passively collected data through GPS-enabled smartphones, or actively collected data through user's initiated smartphone applications (e.g., Strava) can be integrated with direct field counts to obtain a more informative pedestrian and bicyclists' volume estimates. It is important to note that all these data sources have their own limitations and biases in sampling, and need careful processing to develop a final dataset which contains best estimates of exposure.

Resource Documents for Demand Estimation Models

Analysts who are interested in learning more on demand estimation models will find many more procedural details in the following comprehensive guidance documents:

- NCHRP Report 770, *Estimating Bicycling and Walking for Planning and Project Development: A Guidebook*, 2014.
- Aoun et al. (2015). Bicycle and Pedestrian Forecasting Tools: State of the Practice. DTFHGI-11-H-00024. Federal Highway Administration, Washington, DC.
- DKS Associates, University of California, Irvine, University of California Santa Barbara, and Utah State University. (2007). *Assessment of Local Models and Tools for Analyzing Smart-Growth Strategies*.
- NCHRP Report 765, *Analytical Travel Forecasting Approaches for Project-Level Planning and Design*, 2014.
- Travel Forecasting Resources website. *http://tfresource.org/Travel_Forecasting_Resource*
- Transportation Research Circular E-C153, *Dynamic Traffic Assignment: A Primer*, 2011.VDOT (2014). *Travel Demand Modeling Policies and Procedures.*
- Train, K. (2009) *Discrete Choice Methods with Simulation.*

The following documents also provide information and examples of application on the demand estimation models described in Step 5. Since several examples of direct demand models are included in Step 6, they are not provided here.

Regional Travel Demand Models

- Clifton et al. (2013). *Improving the Representation of the Pedestrian Environment in TDM*– Phase I.
- Sener et al. (2009). *Tour-based Model Development for TxDOT: Evaluation and Transition Steps.*
- Southern California Association of Governments (SCAG) (2016). *SCAG Regional Travel Demand Model and 2012 Model Validation.*
- SHRP 2 Report S2-C46-RR-1, *Activity-Based TDM: A Primer*, 2015.
- http://ppms.trec.pdx.edu/media/project_files/Clifton_510_final_combined.pdf

Trip Generation and Flow Models

- Clifton et al (2004). *Pedestrian Flow Modeling for Prototypical Maryland Cities*.
- Clifton et al (2008). *Pedestrian Demand Model for Evaluating Pedestrian Risk Exposure*.
- Raford, N. and Ragland, D. (2004). *Space Syntax: Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety*.
- Raford, N. and Ragland, D.R. (2006). *Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: Case of Boston, Massachusetts.*

GIS-based Models

- McDaniel et al. (2014). *Using Origin-Destination Centrality to Estimate Directional Bicycle Volumes*.
- NCHRP's Report 770 (2014). *GIS-based Walk Accessibility Model.*

Discrete Choice Models

- Hood et al. (2011). *A GPS-based Bicycle Route Choice Model for San Francisco, California*.
- Lassarre, et al. (2007). *Measuring Accident Risk Exposure for Pedestrians in Different Micro-Environments*.
- Papadimitriou, E., Yannis, G., and Golias, J. (2012). *Analysis of Pedestrian Exposure to Risk in Relation to Crossing Behavior.*
- Sener et al. (2009). *An Analysis of Bicycle Route Choice Preferences in Texas, U.S. Transportation.*
- Zimmermann et al. (2017). *Bike Route Choice Modeling Using GPS Data without Choice Sets of Paths*.

Simulation-based Traffic Models

- Abdelghany et al. (2012). Dynamic Simulation Assignment Model for Pedestrian Movements in Crowded Networks. Transportation Research Record 2316. pp. 95–105.

Data Fusion

- Proulx, F. and Pozdnoukhov, A. (2017). *Bicycle Traffic Volume Estimation using Geographically Weighted Data Fusion.*

Key Considerations in Model Selection

Analysts should consider the following points when selecting a demand estimation model:

- It is important that one carefully reviews the project goal or objectives together with the resources available (time, budget, data availability, expertise needed, etc.). While the most advanced form of modeling might be desired to answer most potential questions with relatively high level of accuracy, it is likely that there are limitations with the resources. It might be needed to focus on the most feasible and practical approaches and choose the one that will provide an optimal solution with acceptable accuracy.

- A model is as good as its input data. Using a very advanced form of modelling with inadequate data may not result in any better estimations than a simple form of a model with very good data. It is essential that pm assesses currently available data and the feasibility of gathering additional data for a practical and informative analysis. A robust dataset will not only improve model performance at present, but will also be beneficial in validating various other analyses and models that may be needed in future.

- Most of the demand estimation models may be available through different sources because of their use in other fields. For instance, an MPO of a large urban area might have already developed an activity-based travel demand model, which might provide detailed information on non-motorized trips. It is recommended to investigate all potential options available in the region to eliminate any duplicative work. While increasing the usability of what has been already done in the region, this might also help execute improvements to the existing models.

- Before using or adapting any available model in the region, it is important to clearly understand what the models have been designed for, how well they perform, and if it would be suitable to use/adapt them for the question of interest. For almost all models, local customizations are expected to be made.

- It is important to note that the end results of these models may embed unique characteristics of the region for which they were created and hence may not be directly transferable. Therefore, while the structure and general requirements are similar, the models may need to be redesigned, reimplemented, and calibrated with respect to local conditions.

The following checklist should also be helpful in evaluating the model selection/development decision:

- ☐ Familiar with the model type and/or components, such as:
    - Trip-based, activity/tour-based TDM, etc.?
    - Pedestrian flow model, network simulation model, etc.?
    - Accessibility model, O-D centrality model, etc.?
    - Bicycle route choice model, cross behavior model, etc.?
    - Microscopic, macroscopic, mesoscopic simulation traffic model?
- ☐ Familiar with theories used in modeling, such as:
    - Random utility maximization for choice modeling?
    - Methodologies to integrate the model outcomes to generate volumes (e.g. network assignment methods)?
    - Simulation techniques and software (e.g. VISSIM) for simulation-based models?
    - Data mining, advanced statistical modeling for data fusion, etc.?
- ☐ Have access to any model in the region (such as regional TDM)?
- ☐ Does the study focus (e.g. study area, facility type, etc.) show commonality with your study?
- ☐ Is the model available transferable to another region?
- ☐ Is the model sensitive to bicycle or pedestrian trips? For example, does the regional travel demand model capture the differences in facility-level bicycle or pedestrian activity?
- ☐ Does the model provide desired level of resolution needed for pedestrian and bicycle trips?
- ☐ Are there existing data to develop any model, such as:
    - Representative site counts for a direct demand model?
    - Household travel survey land-use data, transportation network, and system performance data for a regional travel demand model?
    - Pedestrian network and land use data for a trip generation and flow model?
    - Travel network, accessibility measure and land use data for a GIS-based model?
    - Bicycle route choice preference survey data for a discrete choice model?
    - Pace geometrics, demand matrix for a simulation-based model?
    - Site counts, emerging data sources (e.g. Strava), travel demand model outputs for a data fusion model?
- ☐ Familiar with sampling and site selection methodologies? Factor adjustment?
- ☐ Have resources available (budget, time, staff, etc.) to develop any model?

Travel Surveys

A travel survey is a systematic effort to collect information about individual travel behavior. Travel surveys are typically collected from a statistical sample of travelers for a specified day or days (not an entire month or year), and typically gather aggregate trip information (travel mode, trip purpose, trip start and end location, trip length or time, etc.). Depending upon the number of travelers surveyed, trip information from travel surveys are often summarized into more aggregate geographic zones (not on specific facilities) to improve the statistical precision and accuracy of the survey data.

ACS

The ACS is a national ongoing survey of a sample of U.S. households by the U.S. Census Bureau that gathers a wide variety of information (e.g., demographic, social, economic, housing) in addition to their primary travel mode from home to work. Therefore, the ACS does not have trip information for non-commute trips (whereas NHTS does, but on a five- to seven-year cycle). Because the ACS only asks about the primary travel mode, it does not include modes of travel that may be considered secondary (such as walk trips to public transit).

NHTS

The NHTS is a national survey of daily and long-distance travel that is conducted every five to seven years from a sample of U.S. households by the U.S. DOT. The survey provides estimates of trips and miles by travel mode (including walking and bicycling), trip purpose, and other household attributes and demographics.

Regional Household Travel Survey

A travel survey typically conducted by a MPO for the purpose of developing a regional travel demand forecasting model. The frequency of these surveys varies from city to city, with some planning agencies conducting household travel surveys every eight to ten years or longer.

STEP 6. USE ANALYTIC METHOD TO ESTIMATE SELECTED EXPOSURE MEASURE

Step 6 in the scalable risk assessment process is to use the analytic method selected in Step 5 to estimate the desired exposure measure(s). All of the previous steps involve making scoping or planning decisions about how to estimate exposure. Step 6 in the process is when the detailed analysis for exposure estimation occurs. As a result, the Step 6 section in this guide is the largest and has the most content.

Step 6 includes a section for each of the three primary methods that can be used to estimate exposure.

- Site counts: Overview of the manual and automated counting procedures.
- Demand estimation models: Overview of the demand estimation models, with a particular focus on direct demand models, in estimating non-motorized exposure.
- Travel surveys: Overview of the most commonly used travel surveys in estimating exposure. This section also introduces an online interactive tool that estimates exposure using national travel survey data from NHTS and ACS.

Site Counts

Site counts are direct measurements of the number of pedestrians or bicyclists at a defined location. The counts may be gathered automatically from various technology-based sensors or manually from human observers. Site counts are taken at a point and are typically used to represent two different scales: point and segment. For the point scale, counts are most commonly used for intersection crosswalks. For the segment scale, the site count (taken at a point) is applied to the entire length of a street segment (where the counts are not expected to vary significantly along the segment length).

Counts are more commonly used to estimate exposure when the desired facility coverage is limited, as count data collection for all facilities within a large network or region is cost-prohibitive (unless extensive sampling is used). In some cases, it is also a challenge to get representative pedestrian and bicyclist counts due to seasonal and day-of-week variation.

This section of Step 6 provides an overview of counting procedures for pedestrians and bicyclists. In particular, this section will highlight considerations and issues that are relevant to site counts used for exposure estimation. Analysts will find many more procedural details in the comprehensive guidance documents listed in Table 15.

Table 15. Key Guidance Documents for Site Counts

Guidance Document or Report	Useful Resources or Application
FHWA 2016 Traffic Monitoring Guide, Chapter 4 Traffic Monitoring for Non-Motorized Traffic	• Automatic counter technology and equipment • Systematic monitoring using combination of permanent and short-duration count sites • Non-motorized traffic patterns • Standardized non-motorized traffic data format
Report FHWA-HEP-17-011, Coding Nonmotorized Station Location Information in the 2016 Traffic Monitoring Guide Format, 2016	• Extensive guidance and interpretation on the standardized non-motorized traffic data format and attributes in the Traffic Monitoring Guide • Relevant for submitting non-motorized count data to FHWA's Travel Monitoring Analysis System
Report FHWA-HEP-17-012, FHWA Bicycle-Pedestrian Count Technology Pilot Project - Summary Report, 2016	• Automatic counter technology and equipment • Identifying suitable count locations • Practical lessons learned in collecting and using count data
NCHRP Report 797, Guidebook on Pedestrian and Bicycle Volume Data Collection, 2014	• Systematic monitoring of counts • Automatic counter technology and equipment, calibration and validation, and technology evaluation
NCHRP Web-Only Document 229, Methods and Technologies for Pedestrian and Bicycle Data Collection: Phase 2, 2016	• Extensive and most up-to-date evaluation of automatic counter technology
Report FHWA-HPL-16-026, Exploring Pedestrian Counting Procedures: A Review and Compilation of Existing Procedures, Good Practices, and Recommendations, May 2016	• Automatic counter technology and equipment • Automatic counter validation and calibration • Recommended count practices for various facility types • Data management procedures (quality assurance, metadata, data analysis)
Alta Planning + Design, Innovation in Bicycle and Pedestrian Counts: A Review of Emerging Technology, 2016	• Automatic counter technology and equipment • Innovative and emerging technology for non-motorized counts

Typical Applications of Site Counts

In some exposure analyses, site counts may serve as the only source of data for exposure estimation. However, the collection of site counts on all street segments or at all signalized intersections within a city or region is cost-prohibitive. Therefore, the use of site counts only is most applicable to exposure analyses that are facility-specific (i.e., point or segment scale) and focus on a limited number of intersections or street segments.

In most exposure analyses, site counts are collected at a small but representative sample of locations, and then an estimation model is developed and calibrated from these site counts and used to estimate pedestrian and bicyclist volumes at uncounted locations. The next major section in this chapter describes several different demand estimation models that use a sample of site counts and other street inventory and land use variables to predict pedestrian and bicyclist volumes at all locations citywide.

Some cities, MPOs, and state DOTs have begun to collect pedestrian and bicyclist counts as part of a routine monitoring program, so it may not be necessary to start from the beginning for exposure estimation. One should inquire about existing count data not just with transportation agencies, but also with other groups and agencies that may share your interest in pedestrian and bicyclist activity:

- City or county parks and recreation department (e.g., on shared use paths).
- National or state parks (e.g., on internal or connector paths).
- Public health departments (e.g., monitoring physical activity).
- Retail or business associations (e.g., on pedestrian malls or plazas).
- Pedestrian and/or bicyclist advocacy groups.

Even if other agencies' existing counts can be used, it is likely that additional counts will be desired for the purposes of exposure estimation. In some cities, the existing and ongoing counts are at a very limited number of locations, and may have been chosen because of high pedestrian and bicyclist usage or recent facility improvements. For purposes of exposure estimation, additional counts may be needed at high crash locations or at a broader range of locations that represent a mix of facility types and land uses (for purposes of estimation model development and calibration).

Calculating AADT from Site Counts

Site counts can be conducted for varying durations (e.g., 12 hours, 48 hours, 7 days, etc.) and at different times of the year, but the final count measure most commonly used in both motorized and non-motorized exposure estimation is AADT. In these cases, AADT is applied to a defined segment, and is then multiplied by the defined segment length to calculate average annual daily PMT or BMT. The same process can be used at intersection crossings, whereby the crossing length is used to calculate PMT or BMT at each crossing. Earlier in this guide, Step 4 described a simple example of this calculation process.

The estimation of AADT from a short duration site count is a recommended practice for pedestrian and bicyclist exposure estimation. AADT estimates are typically made using what are called factor adjustments (described in detail in Section 4.4 of FHWA's *Traffic Monitoring Guide*). Aside from being a widely accepted practice in motor vehicle analysis, there are several reasons why AADT values should be estimated and used in exposure analysis. Pedestrian and bicyclist counts can vary dramatically by month of the year, day of the week, and by prevailing weather conditions. Site counts that were collected for a single day during favorable weather are not a representative sample of all other days during the year. Therefore, in simple terms, the factor adjustment process is used to adjust the count samples to represent more accurately the true annual average pedestrian and bicyclist usage.
The factor adjustment process for AADT estimation requires continuous count data from similar location(s) to scale accurately a short duration count to represent an annual average count. However, some cities or regions may not have the required continuous count data for factor adjustment. At the time of this Guide development, many agencies are still working to implement the factor adjustment process. Several organizations are currently working to develop default seasonal adjustment factors based on climate zones.[1] If factor adjustment and AADT estimation is not feasible for your application,

[1] Rails-to-Trails Conservancy is developing nationwide seasonal adjustment factors in its Trail Modeling and Assessment Platform (T-MAP), as is Portland State in its PORTAL non-motorized count database.

every effort should be made to collect site counts during seasons/months, days of the week, and weather conditions that most closely resemble typical conditions during the year.

Collecting Counts Specifically for Exposure Estimation

The following paragraphs describe key considerations when collecting counts specifically for exposure estimation. Many of these considerations are addressed in more detail in other data collection guides, but are highlighted here to emphasize their importance.

- **Use automated counter equipment as much as possible:** Automated counter equipment allows counts to be conducted for longer periods (multiple days), which reduces error in AADT estimates. Automated counter equipment can also reduce the labor cost of data collection. The resources listed at the beginning of this Site Counts section provide detailed guidance on selecting the appropriate technologies for automated counting.

- **Avoid very short duration counts (i.e., two-hour counts):** The FHWA Traffic Monitoring Guide recommends a minimum duration of seven days for short duration counts. If this duration is not cost-feasible at all of the desired locations, then at least one to two twelve-hour periods should be counted. Two-hour counts should be avoided as much as possible, even it if means reducing the number of count locations to allow for a six-hour or twelve-hour count. Several research efforts have documented the high error rates that result from estimating AADT from two-hour counts.

- **Seek balance between number of count locations and duration:** As implied in the previous bullet, it may be necessary to balance the number of count locations with the duration of each count. This may be necessary to avoid very short duration counts yet still collect data at a representative number of count locations. In these cases, one is balancing the temporal error (from sampling short durations of time during the year) with the spatial error (from sampling very few locations of all locations to be analyzed). At this time, achieving a balance between number of count locations and count duration is considered more of an art form and not science.

- **Select representative months and days of week for your area and count location:** Site counts should be collected during typical or normal seasons/days/times, especially if adjustment factors are not feasible to use for estimating AADT values. Even if adjustment factors are planned for use, site counts should be collected during months/days/times that are considered typical or normal conditions. This helps to reduce the magnitude of the factor adjustment, and ultimately, reduce the error associated with AADT estimates.

- **Focus on balance of high-priority yet representative locations:** If site counts must be collected specifically for exposure estimation, the count collection effort should focus on a balance of high-priority yet representative locations. High priority locations for exposure analysis are likely to be those locations that have a high crash frequency. However, if an estimation model is to be developed, site counts at high-crash locations are likely to be a biased input for model development. The result is an estimation model that only predicts counts accurately at high-crash locations. Therefore, high-crash locations should be balanced with other locations that represent a range of facility types and land use patterns. Note that this balance is considered a more of an art form and not science at this time.

Demand Estimation Models

As described in Step 5, there are several estimation models to estimate pedestrian and bicyclist demand for input to an exposure measure. In this section of Step 6, we provide a detailed overview and key considerations in developing direct demand models since they are the most widely used tools in the literature for pedestrian and bicyclist volume estimation modeling. Specifically, detailed overview of direct demand models is presented first, and then step-by-step instructions are provided (with examples) to develop a direct demand model.

Detailed Overview of Direct Demand Models

Direct demand models are the most widely used tools in the literature for pedestrian and bicyclist volume estimation modeling (especially in supporting traffic safety studies). These models have been primarily used to develop facility-specific demand estimations for the local level of community, project, and facility planning and to evaluate and prioritize projects. The potential usages of direct demand models are listed as follows by Kuzmyak et al. (2014):

- Answer questions about facility use or needs that could not be addressed with traditional trip-based regional models because of limitations related to scale and ad hoc treatment of non-motorized modes.
- Address the need for estimates of walk activity on links and at intersections for safety analysis and design.
- Address the need for estimates of bicycle activity to support questions on bike network design and to support decisions on facility needs.
- Provide a better connection between the context of the given built environment and non-motorized travel behavior and demand.

The FHWA has sponsored a Non-Motorized Travel Analysis Toolkit, which includes various applications to support non-motorized transportation planning and modeling. This Toolkit includes several direct demand models to estimate pedestrian and bicycle volumes (http://nmtk.pedbikeinfo.org/ui/#/). Direct demand models have also been identified as the primary tools to measure bicyclist and pedestrian exposure for safety analysis.

Direct demand models are generally based on different versions of regression modeling to explain "demand levels as recorded in counts as a function of measured characteristics of the adjacent environment" (Kuzmyak et al. 2014). As indicated by Munira and Sener (2017), "the concept of using a direct-demand model to estimate non-motorized activity is not new. Studies dating back 50 years have forecast non-motorized traffic using count and spatial data." Schmiedeskamp and Zhao (2016) explained such models as following "a similar approach of first proposing a set of explanatory variables, fitting some form of regression model, and then interpreting and justifying the results according to the guiding theory." Direct demand models are based on variety of data sources such as activity counts, census population and employment characteristics, land use and topography and transportation network characteristics. A detailed discussion on the model variables are discussed in the next section of model development.

Direct demand models are appealing due to their simplicity and convenience in development and application, and since they are generally based on available data. These models are particularly useful for screening and preliminary analyses especially when the resources are limited and a more comprehensive (and relatively expensive) model is not available or not possible to develop. However, direct demand models are limited in terms of capturing the behavioral structure. In addition, they are usually not transferable due to their strong linkage to local context, activity levels and the characteristics that the models are built on. Aoun et al. (2015) summarized the advantages and disadvantages of direct demand models as in Table 16.

Table 16. Advantages and Disadvantages of Direct Demand Models

Advantages	Disadvantages
• Software requirements are usually limited to spreadsheets or standard statistical software packages. • Can be created largely using existing data. • Most necessary data is typically publicly available and can be found at a variety of geographic levels. • Network connectivity can be estimated, but requires additional time/ resources to quantify.	• They do not take into account individual trip choices and factors. • Activity level (count) data is costly to collect, depending on geographic scale. • They may inaccurately correlate activity levels with adjacent land uses. • Validity between datasets may not be satisfactory. • Datasets typically used (i.e. U.S. Census Data) are not frequently updated.

Source: (adapted from) Aoun et al. 2015

Kuzmyak et al. (2014) highlighted the need to be judicious in the development and application of direct demand models, and suggested the following guidelines:

- The models need to be developed from scratch for each study, and well calibrated to existing conditions with the specific area and on the specific facilities under study. Models developed for a specific area cannot be construed as transferable.
- Uncertainties developed due to unaccounted origin-destination, route choice, and trip purpose data may be narrowed down by developing counts and models focused on a specific time period.
- After model calibration, the reliability of the models to predict volume in individual locations and the overall study area should be tested.
- Need to be judicious in the types of applications or decisions to be supported by the models.

The readers are referred to Munira and Sener (2017) for an in-depth review of the available literature associated with direct-demand modeling to estimate bicycle and pedestrian activity.

Development of a Direct Demand Model

This section provides step-by-step instructions to develop a direct demand model.

The generalized approach to develop a direct demand model includes three primary phases as demonstrated in Figure 17.

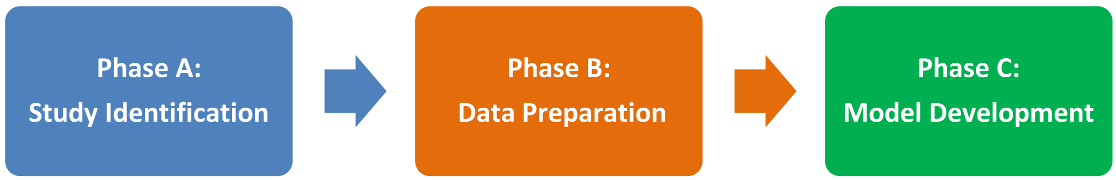


Figure 17. Direct Demand Model Development Process

In addition, Figure 18 provides a flowchart of an algorithm to help the analysts walk through the tasks needed to complete the phases when developing a direct demand model.

Below, we provide detailed instructions on each phase, and the corresponding tasks involved in each phase as consistent with the flowchart provided in Figure 18. Specifically, the main objective and primary tasks of each phase are presented first, followed by discussion of key considerations when processing the phase.

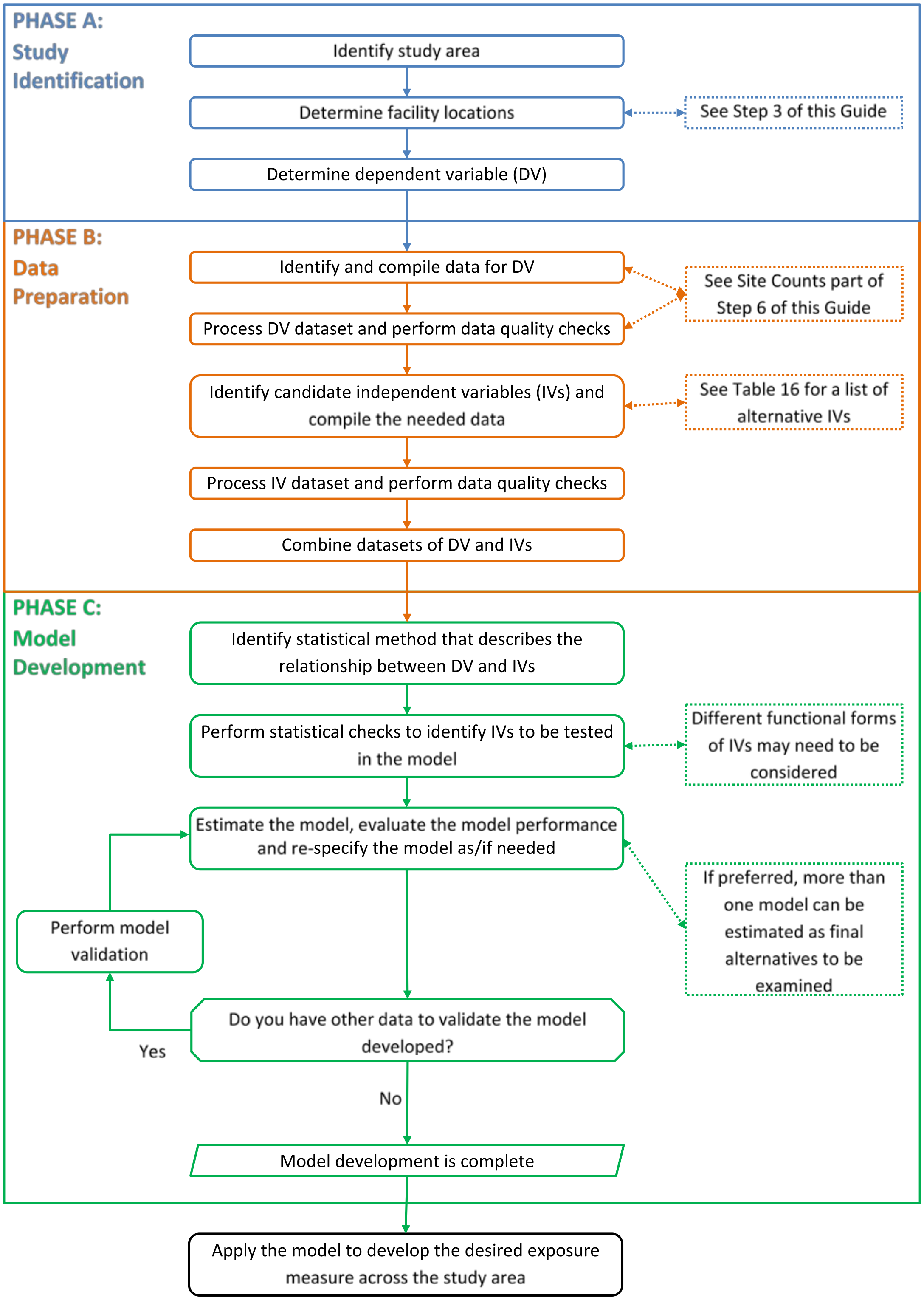


Figure 18. Direct Demand Model Development Detailed Flow Chart

Phase A: Study Identification

Objective: The main objective of Phase A is to identify the study focus.

Primary Tasks: Phase A includes three primary tasks:

- Task 1: Identify study area.
- Task 2: Determine facility locations in the identified study area.
- Task 3: Determine variable of interest (i.e. dependent variable) of the direct demand model.

Processing the Phase:

Task 1: Phase A starts with identification of the overall study area (or target population) for which the direct demand model is desired to be developed. When executing the task, it is of utmost importance to identify project objectives to avoid any unnecessary process, optimize the limited resources and limit the bias.

Task 2: The next task is to determine the facility locations at which the exposure measure is desired. Example of facility locations include signalized and unsignalized intersections or midblock locations along the street segments. Step 3 of this Guide provides information on how to determine desired geographic scale(s) for exposure.

Task 3: Once the location-specific details of the model are identified, the analyst needs to determine the variable of interest, i.e. the model output or outcome whose variation is being examined. For example, in the context of non-motorized direct demand models, the analyst might be interested in obtaining annual average pedestrian volume and peak hour bicycle traffic volume. In statistical terms, the variable of interest is named as the dependent variable of the model.

Phase B: Data Preparation

Objective: The main objective of Phase B is to prepare the data needed for model development.

Primary Tasks: Phase B includes five primary tasks:

- Task 1: Identify and compile data needed for the dependent variable of the model.
- Task 2: Process the data and perform data quality checks for the dependent variable.
- Task 3: Identify and compile candidate explanatory (i.e. independent) variables of the model.
- Task 4: Process the data and perform data quality checks for the independent variables.
- Task 5: Combine the datasets for dependent and independent variables.

Processing the Phase:

Task 1: The first task in this phase is to compile data needed for the dependent variable of the model. As aforementioned in the Site Counts part of Step 6 of this Guide, site counts serve as the common source of data for exposure estimation especially when the desired facility coverage is limited and data collection for all facilities within a large network or region is cost-prohibitive. Therefore, site counts are the main ingredient used for creating the dependent variable of the model.

Sampling: In the ideal conditions, it is desired to have site counts at all facility locations across the study area, but this is not feasible given the limited amount of resources (i.e. budget, time, equipment and manpower constraints). Therefore, the analyst needs to develop a sampling strategy to select sites from which data will be collected. There are two different types of sampling technique for data collection: non-probabilistic sampling technique and probabilistic sampling technique.

- Non-probabilistic sampling techniques are the most commonly used techniques due to their low cost, easy-to-implement methodology. In this technique, the analyst identifies some subjective criteria (e.g. convenience, engineering judgment, local knowledge, quota, etc.) and collects data based on those criteria. This technique does not allow the analyst to control for sampling error, and usually a small number of samples are collected. If the analyst needs to draw conclusions about the entire population, this technique is not recommended since the sites collected might not be representative.

- Probabilistic sampling techniques are based on statistical approaches, involves random selection process, and therefore allows the analyst to compute a sample size and to draw conclusions about the entire population. In general, this technique requires more resources because of the increased sample size needed to obtain a representative sample. There are various different methodologies that can be selected in applying probabilistic sampling techniques including simple random sampling, strategic sampling, cluster analysis and multi-stage random sampling.

Table 17 is adapted from Greene-Roesel et al. 2010 and provides summary information on each of these sampling techniques.

Table 17. Sampling Techniques

Sampling Technique	Methodology	Definition	Example	Advantage	Disadvantage
Non-probabilistic method	Convenience	Obtaining a sample of people or units that are most convenient to study.	Selecting intersections with available collision data	Low Cost; Easy method of sample design.	No representative sample; Not recommended for descriptive or casual studies.
	Judgement	Selecting a sample based on individual judgment about the desirable characteristics required of the sampling units.	Selecting signalized intersections because of experience or intuition that they have higher pedestrian flow.	Low cost; Allow to draw some conclusions about the characteristics of the selected sample.	Does not allow drawing general conclusions about the entire population.
	Quota	It is similar to the judgment sample, but requires that the various subgroups in a population are represented.	Making sure to select some signalized and some unsignalized intersections in a sample.	Low cost; Allow to draw some conclusions about the characteristics of the selected sample.	Does not allow drawing general conclusions about the entire population, or sample subgroups.

Sampling Technique	Methodology	Definition	Example	Advantage	Disadvantage
	Snowball	Additional survey respondents are obtained from information provided by the initial sample of respondents.	Used when surveying individuals about their behaviors (e.g. how much they walk in specific areas)	Some characteristics about the target population can be known	Requires a lot of time and resources; Used only for surveys.
Probabilistic	Simple random	A sampling procedure that ensures each element in the population will have an equal chance of being included in the sample	When there are enough resources; to inquire about the characteristics of the entire population	Simple; Conclusions about the population can be drawn.	Subgroups within the target population may not be represented in the sample; Larger samples are necessary.
	Systematic random	Samples are randomly selected from a list in order, but not everyone has an equal chance of being selected.			The sample may not be representative because of the ordering of the original list.
	Stratified	Sub-samples are drawn within different strata. Each stratum is composed of samples with similar characteristics (e.g. taking into account similarity of intersection characteristics – signalized or non-signalized.	When representation of all subgroups within a particular sample is necessary.	More efficient sample (variance differs between the strata); Small sampling error between strata; Smaller samples.	May be difficult to determine characteristics of individuals to appropriate classify them in specific strata.

Sampling Technique	Methodology	Definition	Example	Advantage	Disadvantage
	Cluster	Entire groups, not individuals, are selected to participate in the data collection; Simple random sampling is applied to the representative “clusters” to select the clusters in which all members will participate.	When the population is too big or when there is a lack of information about individual sampling units (e.g. all vehicle occupants in the United States)	Efficient for large numbers. Do not need to identify all units. Smaller samples; Less expensive relative to the population size.	Sample may not be as representative as desired; Error may be greater than with other techniques; Pilot studies may be necessary to identify the clusters.
	Multi Stage Random	Stratification techniques within the clusters used to refine and improve the sample. Examples of this kind of sampling: National Safety Belt Survey.			Like cluster sampling but more representative within clusters.

Source: (adapted from) Greene-Roesel et al. 2010

Sampling size: Similar to the selection of sampling technique, there are various considerations in determining the sample size. When selecting the sample size, it is recommended that the analyst first identifies what is available, and if the data can be obtained by adjusting/combining readily available data or modifying the existing data collection system. This will help to develop an effective and practical approach in sampling and determination of sample size needed. In addition, it is of utmost importance to continuously evaluate the study objectives to effectively use the existing resources.

The following provides resource examples in determining sample size, which might be helpful in making decisions about the sampling size and technique (adapted from Greene-Roesel et al. 2010):

- <u>Evaluate change over time:</u> The analyst might be interested in understanding the change in pedestrian traffic volume at a particular facility location over time. For example, if the only focus of a study is to conduct a before and after evaluation of one particular intersection in the region, then there is no need to draw information about the general population. In such cases, non-probabilistic sampling techniques (e.g. professional judgement) are commonly used. The sampling focus should be given to collect representative data at that particular intersection taking into account potential biases regarding the time of data collection (e.g. seasonal changes).

- Evaluate risk related to infrastructure type: The analyst might be interested in comparing the pedestrian safety between different facility locations, such as signalized intersections versus unsignalized intersections in a city. This will require one to collect two random samples to determine the pedestrian exposure at signalized intersections and pedestrian exposure at unsignalized intersections, respectively. Simple random sampling technique might be appropriate and are easy to apply assuming that pedestrian exposure will be similar across similar sites (i.e. minimal variance across the selected sample) and the complete list of targeted intersections are available. In this case, the following formulae can be used to compute an approximate value of each sample size (Garder 2004):

$$n = \frac{z^2 \times CV^2}{ME^2}$$

  where n is the sample size, z is the z-value determined based on the desired level of confidence, CV is the coefficient of variation, and ME is the margin of error.

  For example, assuming a 95 percent confidence interval (z=1.96), with low variation (e.g., 10 percent), and with acceptable margin of error (e.g., 5 percent), the minimum sample size is computed as 16. For the example described above, this will yield to about 32 intersections with 16 of them signalized and 16 of them unsignalized. It is also important to reexamine the coefficient of variation during data collection, and re-compute the sample size as needed.

- Sampling exposure in geographic area: The analyst might be interested in determining exposure across an area for example to assess pedestrian risk in a city. In this case, a probabilistic sampling technique is needed since the analyst wants to draw information across the entire area and in need of representative sample of pedestrian volume at different facility types and locations across the area.

  The analyst can use different probabilistic approaches. For instance, the analyst can apply stratified sampling by choosing stratification variables and their corresponding categories. Let's assume the analyst identified intersection type (with 2 categories - signalized versus unsignalized) and geographic area (with four categories - CBD, urban, suburban, and rural). Then, based on the above assumptions, the minimum sample size needed can be computed as 128 (16x2x4). This number may also need to be proportionally adjusted based on the shares of each stratum in the region.

  While stratified sampling is an effective method in obtaining observations with different levels for the variables used, the analyst may need to increase the number of strata for better representation, which will eventually increase the sample size needed. In that case, the analyst might consider other probabilistic sampling techniques. For example, cluster sampling can be used by classifying all the intersections into different clusters with similar characteristics. Cluster analysis helps control sample size while adding more variables, but might not be as representative as stratified sampling. Finally, the analyst can choose to combine clustering and stratification to obtain more representative samples within the clusters as in the case of multi-stage random sampling.

Task 2: Once the site count data are sampled and compiled, the second task of this phase includes processing the data (e.g. factor adjustment) and conducting data quality checks (for a representative sample) needed. Site counts can be conducted for various different durations and at different times of the year. While some studies have used the count data directly (i.e., for the specific collection period) as their dependent variable of the models, other studies have processed and expanded the count data to longer time periods. Direct demand models typically require high-quality volume count information that might be supplemented/validated with travel surveys (such as to account for demographics and trip generators). The analysts are referred to Site Counts part of Step 6 of this Guide, which provides detailed information on site counts and key considerations in processing and obtaining a representative sample of site counts to be used as the dependent variable of the model. This process completes the preparation of the model's dependent variable.

Task 3: Next, the analyst needs to focus on the preparation of explanatory variables of the model. Explanatory variables represent the cause or reason for the outcome. In statistical terms, explanatory variables are also named as the independent variables of the model whose relationships with the dependent variable are being examined. At this stage, the analyst needs to first identify candidate independent variables to be considered in the model, and then compile the data needed for the identified variables. It is likely that the analyst might have an initial (desired) set of independent variables based on the local knowledge, professional judgement, data availability, practicality in usage, etc. However, it is also likely that the analyst might not have any preference or knowledge on the candidate independent variables to be considered in the model. It is important to always keep in mind the goals of developing the model when selecting variables. The final explanatory variables of the model should be composed of variables that are intuitive, logical and relevant to the action items in the decision making process.

Table 18 provides an overview of the key significant variables used across the studies (based on the extensive literature review conducted by Munira and Sener 2017). The analyst is recommended to review the variables in Table 18 before making a final selection of candidate independent variables. The table provides information on frequency (i.e., use in the model) and impact (i.e., direction of the variable).

The model variables showed some differences based on the mode (i.e., pedestrian model versus bicycle model) and the analysis method. While choosing model variables, it is important to consider the context specific nature of explanatory variables of the direct demand models. As indicated by Munira and Sener (2017), "choice of independent variables and their magnitude and direction of impact on non-motorized activity largely depend on community, people, and location". For example, while the availability of sidewalks and land use characteristics might be more influential in motivating walking trips, cycling trips might be more likely to be influenced by various factors across spatial areas beyond the trip origin (Munira and Sener 2017; Winters et al. 2010).

Table 18. Key Explanatory Variables of Pedestrian and Bicyclist Direct Demand Models

Category	Variable	Pedestrian		Bicycle	
		Frequency	Impact	Frequency	Impact
Demographic	Population density	◐	+	○	+
	Total population	○	+	○	+
	% of non-white residents	○	+	○	+
	% of black residents	○	-	○	-
	% residents with a college education	◐	+	◐	+
	% residents younger than 5 and older than 65 years			○	+
Socioeconomic	Household income	○	-	●	+/-
	Total employment	○	+	◐	+
	Employment density	○	+/-	○	+
Network/ interaction with vehicle traffic	Number of lanes	○	+	○	+/-
	Speed limit			○	-
	Arterial street (of count location)	◐	+	○	+
	% major arterials	○	-		
	Collector street (of count location)	◐	+		
	Presence of four-way intersection	○	+		
Bicycle- or pedestrian-specific infrastructure	Presence of bike lane	○	+	○	+
	Presence of sidewalk	○	+		
	Footway pavement width	○	+		
	On-street bicycle facility length			●	+
	Presence of a cycle track			○	+
	Bicycle-trail access			○	+
	Bike lane or curb lane width			○	+
	Separated path			○	+
	Presence of bicycle markings on any approach			○	+
Transit facilities	Number of bus/transit stops	●	+	◐	+
	Presence of subway station	○	+	○	+
	Bus frequency	○	+		
	Accessibility to an underground station	○	+		
Major generators	Distance from the central business district/downtown	◐	-	○	-

Category	Variable	Pedestrian		Bicycle	
		Frequency	Impact	Frequency	Impact
	Proximity to a university campus	○	+	○	+
	Number of schools	○	+	○	+
Weather and environmental	Precipitation	○	-	◐	-
	Temperature	○	-	○	+
	Very warm temperature (max. temperature >32°C)	○	-		
Land use	Residential land use	◐	+/-	○	-
	Land-use mix (area of retail, office, and commercial space per housing unit)	○	+	●	+
	Retail area	◐	+/-	○	+
	Office space area	○	+		
	Industrial area	○	-	○	-
	Cultural and entertainment space area	○	+		
	Job accessibility	○	+		
	Dwell count	○	+		
	Commercial space	◐	+	◐	+
	Maximum/mean slope	○	-	○	-
	Traffic signal-controlled intersection	○	+		
	Patch richness density	○	+		
	Single-family residential areas	○	-		
	Average visibility within the street network	○	+		
	Tourist and downtown area	○	+		
	Job accessibility			○	+
	Centrality			○	+
	Low-density residential space			○	+
	Institutional space			○	+
	Presence of three approaches			○	-
	Presence of parking entrance			○	-

Legend: ○ = to a small extent(1,2); ◐ = to a moderate extent (3,4); ● = to a great extent (>=5)

Source: Based on the literature review of 22 studies conducted by Munira and Sener 2017.

Task 4: Upon identification and compilation of independent variables, the next task includes processing the data and conducting data quality checks needed. Several different specifications and alternative functional forms of independent variables might need to be considered to identify the best data fit during the development of direct demand models. For example, while some variables may need to be considered in categorical or binary forms, some other might need to be used as continuous variables (e.g. development of income categories versus income as a continuous variable). Similarly, some independent variables might work best if they are transformed into other scales (e.g. natural logarithmic scale), which might be helpful for an easier interpretation of model variables as well as better data fit. In addition, all independent variables should be carefully assessed and statistical tests should be performed to ensure the database compiled for independent variables is logical and free of error (e.g. identification of missing values and outliers). This process completes the preparation of the model's candidate independent variables.

Task 5: The final task of the data preparation phase is to combine datasets prepared for dependent and independent variables to obtain one final dataset to be used in model development.

Phase C: Model Development

Objective: The main objective of Phase C is to develop the model based on the data prepared.

Primary Tasks: Phase C includes four key tasks:

- Task 1: Identify statistical method to be adopted for model development.
- Task 2: Perform statistical checks to identify independent variables to be tested in the model.
- Task 3: Estimate the model.
- Task 4: Perform model validation.

Processing the Phase:
Task 1: The first task in this phase is to identify the statistical method to describe the relationship between dependent variable and independent variables. A wide variety of methods have been used in predicting non-motorized activity using direct-demand models. Linear regression, Poisson regression, and negative binomial regression models are among the most commonly used statistical methods used in direct demand models for bicycle and pedestrian exposure estimation. In order to select the best model to the data, the analyst needs to examine the nature of the data. For example, Poisson distribution assumes that the mean and variance are the same; however, we often found that count data exhibits over-dispersion with a variance greater than the mean. In that case, it has been shown by many studies that negative binomial provides a better data fit.

Task 2: In this second task, the analyst needs to screen the pre-identified model variables and the relationships between variables. The following describes some of the key considerations in evaluating the variables.

- Compute correlations across variables, and screen all variables for high and low correlation values.
- Identify independent variables that are relatively highly correlated with the dependent variable.
- Check for multicollinearity to avoid including highly correlated independent variables simultaneously in the same model.

Task 3: Once the statistical method is identified and statistical relationships across variables are initially screened, the analysts can start working on the model estimation. In this task, the analyst may need to conduct several model iterations and different combinations of explanatory variables to find the best model structure. It is recommended that the final model is selected based on both statistical and intuitive considerations (e.g. statistical significance, goodness of fit, insights obtained from the literature, practicality, and engineering judgment). The model performance should be evaluated through statistical checks, such as by overall goodness of fit, residual plots, etc. A 5 percent level of significance is recommended in general to include variables in the final model; however, the analyst is also recommended to check logical and intuitive variables that might not be very highly statistically significant to reduce any bias.

Task 4: It is important to validate the model, which is defined as the application of the models and comparison of the results to observed data that was not used to estimate the model . It is required that the observed data used for model validation are not the same data used for model estimation. Model sensitivity tests will also be useful to determine if the model results are reasonable and sensitive to the changes in explanatory variables.

The analyst is recommended to check the process from the selection of statistical method to the identification and examination of variables included in the model until a good model performance is obtained. As/if needed, the model should be re-specified. This process helps develop more robust models with good fit and intuitive explanatory variables that would be useful for both evaluating risk and informing safety policy and investment decisions.

The completion of this task concludes the development of the model. The exposure output obtained from the model can then be used in risk analysis. The direct demand models can also be used in predicting volumes at locations where the count data are not available, extending the study to an areawide level in the application process.

Examples of Direct Demand Models

Table 19 provides an overview of example direct demand models from the recent literature. The table provides information on the coverage, data collection scale, analysis methods, and significant explanatory variables of the final estimated models. Next, we provide example studies that have developed and applied direct demand models for exposure estimation in non-motorized safety analysis.

Table 19. Examples of Recent Direct-Demand Models to Estimate Pedestrian and Bicycle Volumes

Author (Date)	Coverage	Data Collection Scale	Analysis Methods	Significant Explanatory Variables (Buffer Size)		Model Performance and Validation
				Pedestrian	Bicyclist	
Hankey et al. (2017)	Blacksburg, VA	Pedestrian and bicyclist counts at 101 locations on different street and	Stepwise linear regression model	Sidewalk length; off-street trail length; household income; residential addresses count	Household income; centrality; population density; on-street facility	Bicycle Model: Adj-R^2=0.52 Pedestrian Model: Adj-R^2=0.71

Author (Date)	Coverage	Data Collection Scale	Analysis Methods	Significant Explanatory Variables (Buffer Size)		Model Performance and Validation
				Pedestrian	Bicyclist	
		trail segments		in buffer; population density; bus stop count in buffer	length; major roads length	Validated by goodness of fit, internal validation, and a Monte Carlo–based 20% holdout analysis
Hankey and Lindsey (2016)	Minneapolis, MN	Pedestrian and bicyclist counts at 471 locations on different street and trail segments	Stepwise linear regression model	Major roads (200 m); off-street trails (3000 m); transit stops (400 m); retail areas (100 m); industrial areas (1250 m); open space areas (100 m); job accessibility; population density (750 m)	Off-street trails (200 m); on-street facilities (100 m); retail areas (100 m); industrial areas (1250 m); open space areas (200 m); job accessibility; population density (1250 m); precipitation; temperature	Bicycle Model: Adj-R^2=0.58 Pedestrian Model: Adj-R^2=0.53 Internal validation and Monte Carlo–based 10% holdout analysis
Fagnant and Kockelman (2016)	Seattle, WA	Bicycle counts at 251 intersections	Negative binomial and Poisson models	Not reported	Employment density; bicycle-trail access; bridges; number of lanes; curb-lane width; bike-lane width; separated paths; speed limit; residential areas; morning period count; League of American Bicyclists gold (bicycle-friendly community listings)	Not reported

Author (Date)	Coverage	Data Collection Scale	Analysis Methods	Significant Explanatory Variables (Buffer Size)		Model Performance and Validation
				Pedestrian	Bicyclist	
Tabeshian and Kattan (2014)	Calgary, Canada	Pedestrian and bicycle counts at 34 intersections located on major arterials (excluding downtown)	Multiple linear and Poisson models	Number of bus stops (0.1 mi); street length (0.5 mi); total bus-km of bus routes (0.75 mi); total number of dwell count (0.5 mi); hectares of commercial space (0.25 mi); number of schools (0.5 mi); pathway length (0.25 mi)	Hectares of commercial space (0.10 mi); hectares of low-density residential space (0.10 mi); number of bus stops (0.25 mi); hectares of institutional space (0.50 mi); number of street lanes reaching intersection	Multiple linear regression model Bicycle Model: Adj-R^2=0.90 Pedestrian Model: Adj-R^2=0.92 Validation based on prediction models of 18 intersections in southwest Calgary
Strauss et al. (2013)	Island of Montreal, Quebec, Canada	Bicycle activity counts at 647 signalized intersections	Bayesian model	Not reported	Number of employment (400 m); presence of schools (400 m); presence of subway stations (800 m); land-use mix (800 m); length of bicycle facilities (800 m); commercial land-use area (50 m); presence of three approaches	Not reported

Author (Date)	Coverage	Data Collection Scale	Analysis Methods	Significant Explanatory Variables (Buffer Size)		Model Performance and Validation
				Pedestrian	Bicyclist	
Schneider et al. (2012)	San Francisco, CA	Count of pedestrians who crossed each leg of the 50 intersections	Log-linear model	Number of households (0.25 mi); total employment (0.25 mi); intersection is in a high-activity zone; maximum slope on any intersection approach leg; intersection is within 0.25 mi of a university campus; intersection is controlled by a traffic signal		Adj-R^2 values between 0.78 and 0.83 Validated against 2002 pedestrian volume at other 49 four-way intersections
Hankey et al. (2012)	Minneapolis, MN	Pedestrian and bicyclist counts at 259 locations, midblock portion of each street or sidewalk segment	Ordinary least squares and negative binomial models	% of non-white residents; % residents with a college education; distance from the central business district (CBD); distance from nearest body of water; recorded	% of non-white residents; % residents with a college education; median household income; measure of mixing of land uses; distance from the CBD;	Bicycle Model: Adj-R^2=0.38 (OLS); Cox–Snell R^2=0.48 (NB) Pedestrian Model: Adj-R^2=0.30 (OLS); Cox–Snell R^2=0.42 (NB)

Author (Date)	Coverage	Data Collection Scale	Analysis Methods	Significant Explanatory Variables (Buffer Size)		Model Performance and Validation
				Pedestrian	Bicyclist	
				precipitation; principal arterial street (of count location); arterial street (of count location); and collector street (of count location)	recorded precipitation; off-street trail (of count location); arterial street (of count location) and year	Validates models based on predicted non-motorized traffic at 85 locations (46 new and 39 previously sampled locations)

Source: Adapted from Munira and Sener (2017)

Detailed Example

This section provides an example study with details to help analysts follow the process described above. The example is based on a study conducted by Schneider et al. (2012), and presents a typical example for direct demand model development for non-motorized exposure estimation.

Development of Pedestrian Intersection Volume Model in San Francisco, California

Phase A – Study Identification: Schneider et al. (2012) developed a pedestrian intersection volume model in San Francisco, California, focusing on annual pedestrian crossing intersection volume.

Phase B – Data Preparation: The authors identified 50 intersections to collect a sample of counts for the San Francisco pedestrian intersection volume model. The intersections were selected to represent the range or urban characteristics across the city. The data collection was conducted at different time periods (2009 and 2010). The authors aimed at increasing the geographic representation of locations across the study area by collecting data at various locations, e.g. high-crash locations, regional count locations, locations near planned or completed projects, locations near key transit hubs, etc. Next, the authors applied automated counter, temporal, and weather adjustment factors to extrapolate an annual pedestrian volume estimate from the two-hour counts at the 50 study intersections. The logarithm of the annual pedestrian crossing volume constituted the dependent variable. For independent variables, they considered 16 explanatory variables (e.g. total number of households within 0.25 mile of the intersection without a car, ratio of population to jobs within 0.25 mile of the intersection, intersection in a high-activity zone, etc.). The authors examined descriptive statistics for all variables (i.e. mean, standard deviation, minimum and maximum).

Phase C – Model Development: Next, the authors developed a log-linear regression model to identify the relationship between annual pedestrian volume estimate and various different explanatory variables including land use, transportation system, local environment, and socioeconomic characteristics near each sampled intersection. After conducting various model runs, the authors identified 12 potential models of annual volumes of pedestrian intersection crossings. The variables with high level of correlation, without precise estimates, and with counterintuitive relationships with pedestrian volume were excluded from the model. The 12 potential models were indicated to have good fits (adjusted R2-

values between .78 and .83) and were significantly better than a model based only on a constant with no independent variables (F-values between 28.4 and 34.4). The final recommended model was selected based on a combination of good overall fit and intuitive, logical and practical explanatory variables. The model performance was evaluated by spatially reviewing the difference between predicted and observed counts. Sensitivity tests were conducted, and the model was validated against pedestrian volumes collected at 49 four-way intersections (different than the intersections used in the model estimation) in 2002, which showed that the model ranked intersections similarly to the previous counts in overall volume. The authors indicated the model included only six significant factors because of the relatively small sample of intersections, and highlighted the importance of various other variables that might need to be considered in future studies.

The model was then used to evaluate pedestrian crossing risk at each intersection based on the exposure measure of the number of pedestrian crashes per 10 million crossings.

Additional Examples

This section provides additional examples that are briefly described to present different thought process during the development of direct demand models for non-motorized safety analysis.

- Molino et al. (2009; 2012) developed a log-linear regression model (with Poisson distribution) to estimate pedestrian counts at signalized intersections in Washington, D.C. While 15-min pedestrian counts served as the dependent variable, the independent variables of the model included land use variables (e.g., commercial, residential) and characteristics of the day (e.g., day of the week, time of the day). Using the parameter estimates of the model and follow-up adjustment procedures, a total number of miles traveled were estimated "…by multiplying the total number of pedestrians by the mean width of all the sampled signalized intersections." This result was then used as an exposure measure in pedestrian crash rate computation.

- Using a count database of 954 observations and 471 locations, Hankey and Lindsey (2016) employed a stepwise linear regression model that allowed for varying spatial scale of independent variables including land use and transportation network variables. Relying on the modeled values of bicycle traffic from this work, Wang et al. (2016) then estimated peak-hour bicycle traffic volumes for many segments in Minneapolis. The model results were then converted to bicycling volume for intersections and used for computing bicycle crash rates by intersections and segments.

- Strauss et al. (2013; 2014) used a relatively improved version of modeling to estimate non-motorized demand. First, Strauss et al. (2013) developed a bivariate Bayesian Poisson model to simultaneously estimate cyclists' injury occurrence and bicycle activity at 647 signalized intersections on the island of Montreal, Quebec, Canada. In a follow-up study, Strauss et al. (2014) applied their Bayesian modeling methodology as part of a multimodal approach aimed at examining the safety at intersections for both non-motorized and motorized traffic. After model calibration, the study compared injury and risk between modes and intersections by using the "expected number of injuries (obtained from the models) per million cyclists, pedestrians or motor-vehicle occupants per year" as the expected risk.

Travel Surveys

This section of the guide provides an overview of the most commonly used travel surveys in estimating exposure including the ACS, NHTS, and regional household travel surveys. This section includes information on the general purpose of the survey and applicability in estimating exposure for bicyclists and pedestrians, limitations and benefits, and data availability.

American Community Survey (ACS)

The ACS is an important tool for tracking non-motorized (bicycling and walking) travel patterns. This national ongoing survey of a sample of U.S. households conducted by the U.S. Census Bureau gathers a wide variety of information in addition to their primary travel mode from home to work.[2] The ACS does not have trip information for non-commute trips (whereas NHTS does but only conducted once a decade). Thus, the ACS can be used to estimate non-motorized exposure expressed as the daily person commute trips by walking or bicycling per the specified areawide geography.

Benefits and Limitations

In many cases, non-motorized trips are secondary modes of travel to the longest distance mode (driving or transit). However, it provides one of the most robust sources of information on non-motorized commuting by bicycle and walking in smaller spatial units like Census block groups. Table 20 summarizes the strengths and limitations of the ACS in estimating non-motorized trips.

All survey and census estimates incorporate errors to a certain extent. Sampling error is the quality measure that can adversely affect any survey result. Sampling error usually occurs when data are based on a sample of a population rather than the full population. For every ACS estimate, margins of error are provided and can be easily converted into confidence intervals.[3] For different calculations associated with the ACS, it is important to consider sampling error.

Table 20. Strengths and Limitations of the ACS in Estimating Pedestrian and Bicyclist Travel

Strengths	Limitations
• Deliver useful, relevant data, similar to data from previous census long forms, and updated every year instead of every 10 years. • ACS is more accurate than the decennial census long form sample. • ACS item allocation rates are lower, and non-sampling error is reduced.	• Only home-to-work commute trips. • Does not capture trips by children or most trips by older adults. • Requires sound statistical knowledge to understand and use multi-year estimates. • Relatively large confidence intervals associated with ACS data for smaller geographic areas and subgroups of the population.

[2] McKenzie, B. *Modes Less Traveled—Bicycling and Walking to Work in the United States: 2008–2012 American Community Survey Reports*. 2014.
[3] U.S. Census Bureau, *A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know*, October 2008.

Data Availability

The ACS provides estimates for different levels: a) 1-year estimates, b) 3-year estimates, and c) 5-year estimates. A key U.S. Census document[4] lists the distinguishing features of 1-year, 3-year, and 5-year estimates. It is important to note that using 3-year or 5-year ACS is beneficial due to large sample size relative to 1-year estimates, thus reducing margins of error of estimates for small subpopulations. For analyzing areas with larger population (e.g., states, congressional districts), 1-year ACS is beneficial.[5]

For spatial units with smaller populations, the ACS samples may have insufficient numbers of households to provide reliable single-year estimates. For these spatial units, multiple years (3 or 5) worth of data will be merged together to create more reliable estimates. The multi-year estimates have advantages of statistical reliability for less populated areas and small population subgroups. The level of precision improves considerably with multi-year estimates.

National Household Travel Survey (NHTS)

FHWA conducts periodically the NHTS to gather detailed information on the travel behavior of the American public. The survey collects a wide array data from respondents, including household characteristics, demographics of each member in the households, vehicle details, and trip attributes (mode used, trip length, trip time, trip purpose). These data are stored in separate files: household, person, vehicle, and travel day (i.e., travel diary).

The basic concept of using travel surveys to calculate the amount travel for an area is based on developing estimated statistics developed from the survey sample and expanding those estimates to the population by using weights (for example, see Schneider et al.[6]). Analysts can use the NHTS to estimate the following exposure measures for pedestrian and bicyclist travel:

- Total estimated annual trips
- Total estimated annual miles traveled
- Total estimated annual hours traveled

Benefits and Limitations

The 2009 NHTS national sample estimates are statistically valid down to the state level. However, if additional add-on samples were purchases by a particular state or MPO, then estimates in those areas may be valid at a smaller geography depending on the methodology used by the analyst. Keep in mind that the NHTS documentation warns that standard errors or margins or error should generally be used when looking at estimates at geographies smaller than the national level.

While providing a rich national sample, the NHTS sample sizes might have sparse coverage at small geographic scales. Transferability of the NHTS results to small geographic area (e.g., census tracts) is limited to estimates of average weekday household person trips, vehicle trips, person-miles traveled, and VMT. Though these estimates could serve as exposure measures for non-motorized travel risk, they are not the best choice since specific mode of travel is not offered or is vehicle-based.

---

[4] U.S. Census Bureau, *A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know*, October 2008.

[5] *NCHRP Report 588: A Guidebook for Using American Community Survey Data for Transportation Planning*. Transportation Research Board, Washington DC, 2007.

[6] Schneider, R.J., J. Vargo, and A. Sanatizadeh. Comparison of US Metropolitan Region Pedestrian and Bicyclist Fatality Rates, *Accident Analysis and Prevention*, Forthcoming, 2017.

Unfortunately, since the NHTS is not conducted on a more frequent and regular basis, it cannot be used to directly track short-term trends in non-motorized travel exposure. It can, however, be used in sketch planning or travel demand modeling efforts to estimate or predict non-motorized exposure based on more current census demographic information. For example, the generalized daily person trips per person by mode, generated from the NHTS, can be used to estimate the total non-motorized trips produced in subsequent years by using current ACS population estimates. These trip rates need to be updated periodically, and if possible, supplemented by more localized travel data to better reflect local nuances and unique characteristics of the transportation infrastructure and traveling public.

In terms of sketch planning, it is possible to apply generalized person trip rates that were produced to represent the statewide population to local areas. This is done by multiplying the person type trip rate by the total number of corresponding population within the local area. This operation has statistical drawbacks since the generalized person trips rates were produced by using a statewide sample, which is likely to not be statistically representative of the local area population. It is always best to use local data for local purposes; however, the NHTS provides an opportunity to estimate local exposure when local data do not exist.

Data Availability

A NHTS was conducted for 1969, 1977, 1983, 1990, 1995, 2001, and 2009. Most recently, a 2017 NHTS that began in March 2016 was released in early 2018. It is comprised of a 26,000 national household sample representing all U.S. States and the District and Columbia along with an additional 103,112 add-on sample households. Additional add-on samples are made available to the states and regional/MPOs for purchase. These add-on samples provide the opportunity to populate different exposure measures at a finer geographic level and to develop more robust safety analyses. The 2017 NHTS add-on sample sizes for the state DOTs and MPOs are listed in Table 21.

Table 21. 2017 NHTS State DOT & MPO Add-on Household Sample Sizes

Study Area	Sample Size
National	26,000
Arizona DOT	2,444
California DOT	24,000
Des Moines Area MPO	1,200
Georgia DOT	8,000
Indian Nations Council of Governments	1,000
Iowa Northland Regional Council of Governments	1,200
Maryland DOT	1,000
New York State DOT	15,851
North Carolina DOT	8,000
South Carolina DOT	6,500
Wisconsin DOT	11,000
Texas DOT	20,000
North Central Texas Council of Governments	2,917
TOTAL	**129,112**

Source: NHTS Task C: Sample Design, Dec. 31, 2015, page 5

The NHTS data can be used to compute or statistically model several different exposure measure estimates (e.g., population, miles traveled, number of trips) nationally, by census region/division, state, and urban/rural area types, depending on the survey year. It is possible to produce these same measures at smaller census geographies like Core Based Statistical Area (CBSA) if the particular location(s) participated in the add-on program and were specified in the sampling design.

Regional Household Travel Survey

A regional household travel survey is typically conducted by an MPO to develop a regional travel demand model. The frequency of these surveys varies from city-to-city, with some planning agencies conducting household travel surveys every eight-to-ten years or longer. Just like the NHTS, regional household travel surveys collect data from respondents on the household characteristics, demographics of each member in the households, vehicle details, and trip attributes via a travel diary. Figure 19 depicts the relationship between the four separate data components. Exposure measures (e.g., miles traveled or number of trips) can be estimated for household and person types and expanded to the population to provide statistically valid areawide estimates.

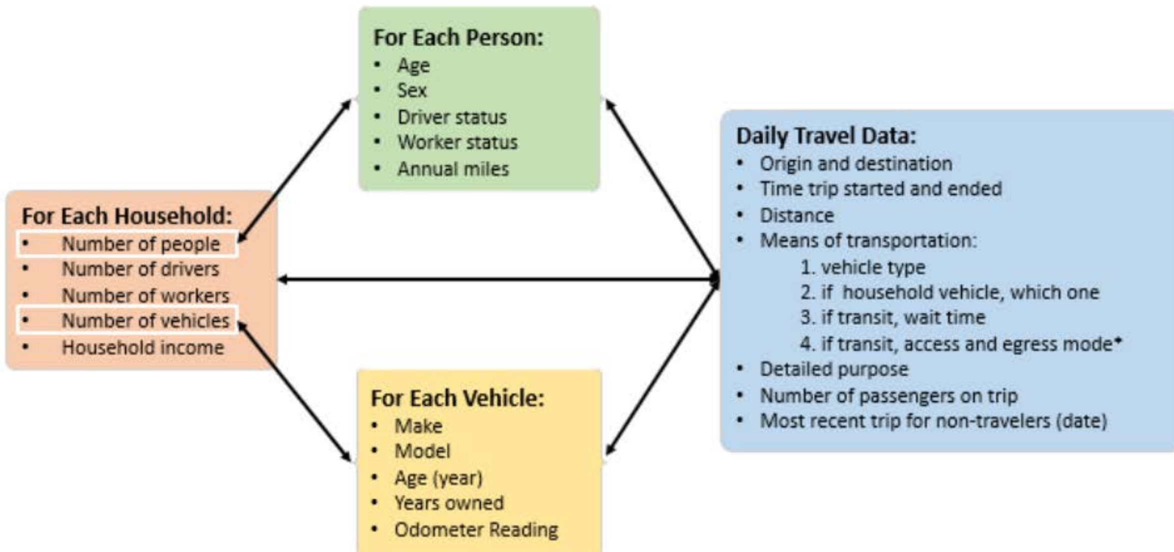


Figure 19. Household Survey Data Relationships

Source: 2017 NHTS Data Users Guide, https://nhts.ornl.gov/assets/2017UsersGuide.pdf

Benefits and Limitations

Household travel surveys can be used to measure the population proportion, distance traveled, duration traveled, and number of trips by a specific mode for the survey region. Survey respondents typically fill out a travel diary indicating origins and destinations with the start and end times of trips along with the mode that was used. Since the survey represents only a stratified sample of the population, weights must be applied to expand the survey sample so that it represents the entire population of the study area (see Figure 20). Survey weights indicate how many households each survey observation represents of the total population of households – these weights are typically provided along with the survey data.

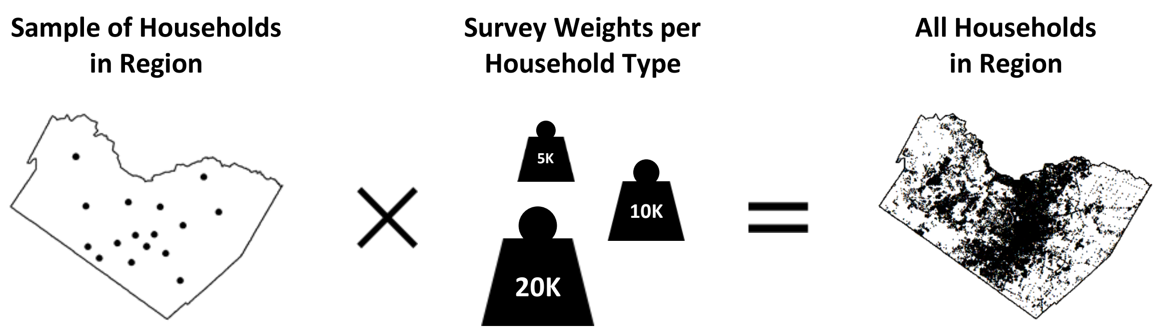


Figure 20. Expanding Household Samples to Represent All Households in a Region

For example, the regional household travel survey for Austin, Texas could be used to estimate the total amount of time traveled by walking for the five-county region. The survey sample is comprised of 3,000 households and 8,100 persons and can be expanded to represent the population of the study area by applying the survey weights. To do so, the total duration of trips by mode must be enumerated per household type, as defined in the survey stratification. The totals then must be multiplied by their corresponding survey weight to equal the total daily duration of trips by mode for the entire study area. The result is an estimated total of 189,256 daily walk trips with an average trip duration of 16 minutes equaling approximately 50,437 hours of walking per weekday.

The main limitations of region household travel surveys include their high cost and expertise required to process and analyze the survey data. The data may not be publically available due to survey respondent privacy concerns.

Areawide Non-Motorized Exposure Tool

The Areawide Non-Motorized Exposure Tool described here makes it easier for practitioners to obtain and summarize nationwide travel survey data to estimate pedestrian and bicyclist exposure to risk at statewide and MPO area scales. The first part of the tool titled *Statewide Exposure Estimates* fills that gap between years when the NHTS is conducted by using the more current ACS data to estimate non-motorized exposure at the state-level. The second part of the tool titled *MPO Area Exposure Estimates* also uses NHTS and ACS data to estimate non-motorize exposure but for individual MPOs throughout the nation. Both parts of the tool produce annual non-motorized exposure estimates by mode for years 2009 to 2016 in terms of trips, miles of travel, and hours of travel for their respective areawide geography. The results are offered in tabular form along with graphics like the examples shown in Figure 21. The following sections describe the tool's capabilities, as well as instructions on how to use the tool.

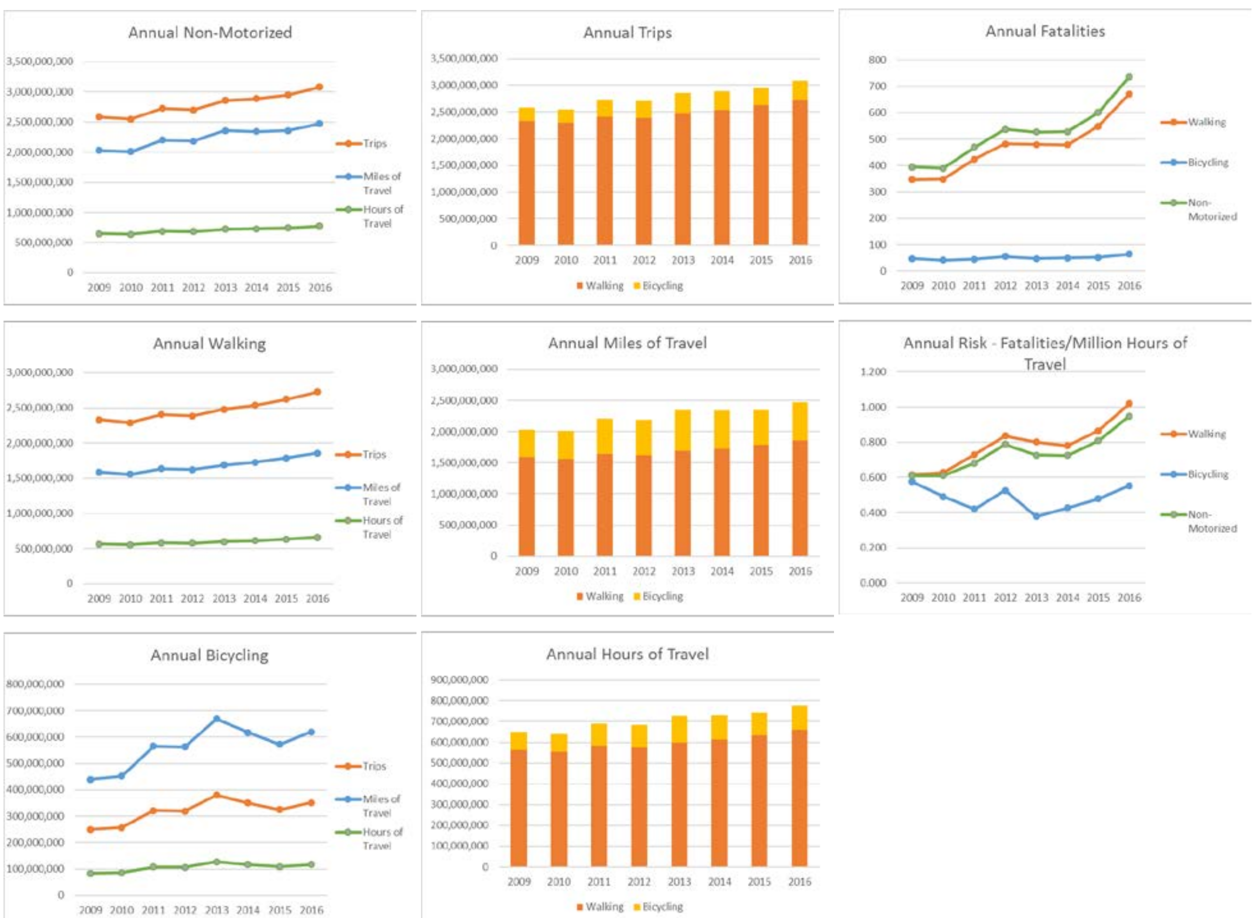


Figure 21. Annual Non-Motorized Exposure Estimates, Non-Motorized Fatalities, and Risk

Statewide Exposure Estimates
The FHWA's Safety Performance Management (Safety PM) Final Rule currently requires each State DOT to report the number of non-motorized fatalities and serious injuries (without considering exposure). To understand the relationship between these crashes and non-motorized risk, exposure is desirable to help measure the magnitude of bicyclist and pedestrian vulnerability. However, users should note that, at this time, the Safety PM Final Rule does not require non-motorized exposure to be reported or considered.

The *Statewide Exposure Estimates* component offers a method for practitioners to estimate statewide non-motorized exposure in order to calculate non-motorized risk. The tool provides the following exposure measure estimates for both bike and walk travel modes per state for the individual years 2009-2016:

- Total estimated annual trips
- Total estimated annual miles traveled
- Total estimated annual hours traveled

The estimates are based on a combination of the 2009 NHTS and the U.S. Census Bureau's ACS data for each respective year. The 2009 NHTS total annualized trips per state are adjusted to better represent the selected year for analysis by using the more current ACS population and daily commute trip estimates (tables B01003 and B08301, respectively). The adjustment factors account for change in both population and the number of commute trips per mode over time. The population adjustment factor ($AF_{pop,i}$) is based on the 2009 ACS population estimate since the NHTS data represent 2009. It can be written as:

$$AF_{pop,i} = \frac{POP_i}{POP_{2009}}$$

Where:
$AF_{pop,i}$ = Population adjustment factor in i^{th} year (i = 2009 to 2016) for state
POP_i = ACS population estimate in i^{th} year for state
POP_{2009} = ACS population estimate in 2009 for state

In order to expand daily person commute biking and walking trips, the relationship between commute and total trips is required. The commute trip adjustment factor is based on the 2009 NHTS annualized person trips per person by mode (bike and walk) and the annualized ACS daily persons commuting by mode (bike or walk) for the selected year (i.e., i^{th} year) for analysis. The equation is as follows:

$$AF_{CT} = \frac{PT_{2009}}{PC_{2009} \times 365}$$

Where:
AF_{CT} = Commute trip adjustment factor by mode for state
PT_{2009} = NHTS annualized person trips by mode in 2009 for state
PC_{2009} = ACS daily persons commuting by mode in 2009 for state

The adjustment factors (AF) are applied to the selected year ACS commute trips per person by mode to provide estimated annual person trips:

$$PT_i = PC_i \times AF_{pop,i} \times AF_{CT} \times 365$$

Where:
PT_i = Estimated annual person trips by mode (biking or walking) in i^{th} year for state
PC_i = ACS daily persons commuting by mode in i^{th} year for state
$AF_{pop,i}$ = Population adjustment factor in i^{th} year for state
$AF_{CT,i}$ = Commute trip adjustment factor by mode for state

Finally, to calculate the estimated total annual miles and hours traveled, the 2009 NHTS average trip durations and trip lengths per state were then applied to the total trips to estimate the amount of hours and miles traveled annually per mode for each state.

$$HT_i = PT_i \times TD_{2009}$$
$$MT_i = PT_i \times TL_{2009}$$

Where:
HT_i = Estimated annual hours traveled by mode (biking or walking) in i^{th} year for state
PT_i = Estimated annual person trips by mode in i^{th} year for state
TD_{2009} = 2009 NHTS average trip duration (in hours) by mode for state
MT_i = Estimated annual miles traveled by mode in i^{th} year for state
TL_{2009} = 2009 NHTS average trip length (in miles) by mode for state

Data sources for the above variables are as follows:

Variable	**Data Source**
POP_i & POP_{2009}	ACS 1-year estimate, table B01003 - Total Population
PC_i & PC_{2009}	ACS 1-year estimate, table B08301 - Means Of Transportation To Work
PT_{2009}	2009 NHTS
TD_{2009}	2009 NHTS
TL_{2009}	2009 NHTS

This method assumes that the average trip durations and lengths remain constant between year 2009 and 2016 due to the lack of more current data. However, the tool does provide the user the option to input their own values if available. In addition, the tool should be updated with the newly published 2017 NHTS data to produce the 2017 estimates based on current travel behavior data.

The NHTSA Fatality Analysis Reporting System (FARS) person data were used to calculated the total annual non-motorized fatalities per state from 2009 to 2016. The totals are provided in the spreadsheet tool along with total annual risk per state that is based on the total annual non-motorized fatalities per million hours of travel. The total annual non-motorized fatalities are defined as individuals classified as a bicyclist or pedestrian that sustained a fatal injury in a motor-vehicle crash. As of May 2018, the 2016 FARS data were incomplete. The data can be found online: https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars.

The interface of the *Statewide Exposure Estimates* component is shown in Figure 22.

Statewide Exposure Estimates

State: New York

❶ Select **State** of interest

❷ Select the source (Default or User Input) of the required inputs. For the User Input option, values are required in the cell below.

		Walking							
		2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Daily Persons Commuting		574,322	542,579	575,553	568,540	574,861	576,752	583,151	577,983
Commute-to-Total Trips Adjustment Factor		25.49	25.49	25.49	25.49	25.49	25.49	25.49	25.49
Population Adjustment Factor		1.00	0.99	1.00	1.00	1.01	1.01	1.01	1.01
Estimated Annual Pedestrian Trips		5,343,405,740	4,997,592,893	5,354,858,779	5,289,610,879	5,401,904,720	5,419,674,236	5,479,804,926	5,431,241,806
Average Trip Length (Miles)	Source:	Default	Default	Default	Default	Default	Default	Default	Default
	Default Value:	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76
	User Input Value:				1				
Estimated Annual Pedestrian Miles of Travel		4,060,988,362	3,798,170,599	4,069,692,672	4,020,104,268	4,105,447,587	4,118,952,419	4,164,651,744	4,127,743,772
Average Trip Duration (Minutes)	Source:	Default	Default	Default	Default	Default	Default	Default	Default
	Default Value:	14.82	14.82	14.82	14.82	14.82	14.82	14.82	14.82
	User Input Value:								
Estimated Annual Pedestrian Hours of Travel		1,319,821,218	1,234,405,445	1,322,650,118	1,306,533,887	1,334,270,466	1,338,659,536	1,353,511,817	1,341,516,726
Fatalities		290	288	273	287	293	262	295	300
Fatalities/Million Hours of Travel		0.220	0.233	0.206	0.220	0.220	0.196	0.218	0.224

		Bicycling							
		2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Daily Persons Commuting		39,185	41,232	44,418	53,119	62,021	58,198	61,618	66,595
Commute-to-Total Trips Adjustment Factor		11.54	11.54	11.54	11.54	11.54	11.54	11.54	11.54
Population Adjustment Factor		1.00	0.99	1.00	1.00	1.01	1.01	1.01	1.01
Estimated Annual Bicyclist Trips		165,051,139	171,936,574	187,093,058	223,742,540	263,851,041	247,587,154	262,136,590	283,309,847
Average Trip Length (Miles)	Source:	Default	Default	Default	Default	Default	Default	Default	Default
	Default Value:	1.93	1.93	1.93	1.93	1.93	1.93	1.93	1.93
	User Input Value:								
Estimated Annual Bicyclist Miles of Travel		318,548,697	331,837,588	361,089,602	431,823,102	509,232,508	477,843,207	505,923,618	546,788,006
Average Trip Duration (Minutes)	Source:	Default	Default	Default	Default	Default	Default	Default	Default
	Default Value:	20.81	20.81	20.81	20.81	20.81	20.81	20.81	20.81
	User Input Value:								
Estimated Annual Bicyclist Hours of Travel		57,245,237	59,633,335	64,890,109	77,601,371	91,512,336	85,871,478	90,917,707	98,261,299
Fatalities		28	36	57	42	36	46	36	36
Fatalities/Million Hours of Travel		0.489	0.604	0.878	0.541	0.393	0.536	0.396	0.366

	Non-Motorized							
	2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Estimated Annual Non-Motorized Trips	5,508,456,878	5,169,529,467	5,541,951,837	5,513,353,419	5,665,755,761	5,667,261,390	5,741,941,515	5,714,551,653
Estimated Annual Non-Motorized Miles of Travel	4,379,537,059.48	4,130,008,186.64	4,430,782,273.63	4,451,927,370.05	4,614,680,095.64	4,596,795,626.23	4,670,575,361.39	4,674,531,778.15
Estimated Annual Non-Motorized Hours of Travel	1,377,066,454.24	1,294,038,779.66	1,387,540,227.31	1,384,135,258.03	1,425,782,801.77	1,424,531,014.16	1,444,429,523.79	1,439,778,024.83
Non-Motorized Fatalities	318	324	330	329	329	308	331	336
Non-Motorized Fatalities/Million Hours of Travel	0.231	0.250	0.238	0.238	0.231	0.216	0.229	0.233

Figure 22. Interface of the Statewide Exposure Estimates Tool Component

Step-by-Step Instructions
1: Open the spreadsheet containing the tools.
2: Read the “Introduction” page for an overview of the available tools.
3: Click on the “Statewide Exposure Estimates” tab to select the tool.
4: Select the state of interest from the drop down menu titled “State” (also marked by ❶ in the spreadsheet tool).
5: Select the source (default or user input) of the required inputs, “Average Trip Length (Miles)” and “Average Trip Duration (minutes)”, from the dropdown menus highlighted in green for each year (also marked by ❷ in the spreadsheet tool).
6: For those required inputs where the default option is chosen, no further action is required.
7: For those required inputs where the user input option is chosen, enter the desired user input value into the appropriate cell (no color highlights) below the source dropdown menu.

MPO Area Exposure Estimates

The *MPO Area Exposure Estimates* component offers a method for practitioners to estimate MPO-wide non-motorized exposure for calculating non-motorized risk. The tool provides the following exposure measure estimates for both bike and walk travel modes per MPO for the individual years 2009-2016:

- Total estimated annual trips
- Total estimated annual miles traveled
- Total estimated annual hours traveled

Non-motorized exposure estimates at the MPO level are derived from a combination of ACS and the 2009 NHTS. The Census data offers estimates at relatively small geographies that can be interpolated to the MPO level. The 2009 NHTS data provides information on travel behavior for a sample of the travel public from around the nation and can be used to calculate average person trip rate, average trip length, and average trip duration per mode.

Total person trips by bike and walk can be estimated with a generalized 2009 person trip rate per mode applied to the total population of the year of interest and then annualized (365 days). The product is then adjusted to account for any change in the mode-specific commuting population between 2009 and the year of interest. However, this adjustment does not capture any change in non-motorized recreational travel that may be induced from communities investing in bicycle and pedestrian infrastructure. Total person trips are then be applied to the average trip length and average trip duration to equal total miles and total hours traveled per mode, respectively. It also important to note that any error in the 2009 NHTS estimates of walking or bicycling is carried through to the subsequent years.

The MPO-level estimated annual person trips by mode equation is as follows:

$$PT_i = (PTR_{2009} \times POP_i \times 365)\frac{PC_i}{PC_{2009}}$$

Where,
PT_i = Estimated annual person trips by mode (biking or walking) in i^{th} year for MPO
PTR_{2009} = 2009 NHTS average daily person trip rate by mode for CBSA peer group
POP_i = ACS 5-year population estimate in i^{th} year for MPO

PC_i = ACS 5-year daily persons commuting by mode estimate in i^th year for MPO
PC_{2009} = ACS daily persons commuting by mode 2009 for MPO

To calculate the estimated total annual miles and hours traveled, the 2009 NHTS average trip durations and trip lengths per MPO are applied to the total trips to estimate the amount of hours and miles traveled annually per mode for each MPO.

$$HT_i = PT_i \times TD_{2009}$$
$$MT_i = PT_i \times TL_{2009}$$

Where,
HT_i = Estimated annual hours traveled by mode (biking or walking) in i^th year for MPO
PT_i = Estimated annual person trips by mode in i^th year for MPO
TD_{2009} = 2009 NHTS average daily person trip duration (in hours) by mode for CBSA peer group
MT_i = Estimated annual miles traveled by mode in i^th year for MPO
TL_{2009} = 2009 NHTS average daily person trip length (in miles) by mode for CBSA peer group

Data sources for the above variables are as follows:

Variable	Data Source
PTR_{2009}	2009 NHTS
POP_i	ACS 5-year estimate, table B01003 - Total Population
PC_i & PC_{2009}	ACS 5-year estimate, table B08301 - Means Of Transportation To Work
TD_{2009}	2009 NHTS
TL_{2009}	2009 NHTS

Several caveats do apply to this method. The end user should keep in mind that the MPO-level estimates for their area are based on an estimated average for their CBSA (Core Based Statistical Area) peer group. Also, like the statewide tool, the MPO-level method assumes that the average trip durations and lengths remain constant between year 2009 and 2016 due to the lack of more current data. However, the tool does provide the user the option to input their own values if available.

The NHTSA FARS person data were used to calculate the total annual non-motorized fatalities per MPO from 2009 to 2016. The crashes were plotted in a GIS based on the coordinated provided and spatially joined the underlying MPO layer. The totals along with total annual risk per MPO that is based on the total annual non-motorized fatalities per million hours of travel are provided in the spreadsheet tool. The total annual non-motorized fatalities are defined as individuals classified as a bicyclist or pedestrian that sustained a fatal injury in a motor-vehicle crash. As of May 2018, the 2016 FARS data were incomplete. The data can be found online: https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars.

The interface of the *MPO Area Exposure Estimates* component is shown in Figure 23.

MPO Area Exposure Estimates		
State: Minnesota	❶ Select **State** of interest	❸ Select the source (Default or User Input) of the required inputs. For the User Input option, values are required in the cell below.
MPO: Metropolitan Council	❷ Select **MPO** of interest	

		Walking								
		2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	
Person Trip Rate	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	0.47995	0.47995	0.47995	0.47995	0.47995	0.47995	0.47995	0.47995	
	User Input Value:									
MPO Population Estimate	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	2,825,790	2,844,011	2,870,005	2,898,571	2,927,805	2,959,604	2,991,363	3,018,255	
	User Input Value:									
Population Adjustment Factor	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	1.00000	1.00699	0.97189	0.96451	0.96615	0.99991	1.01908	1.04069	
	User Input Value:									
Estimated Annual Pedestrian Trips		495,031,193	501,704,633	488,645,373	489,759,351	495,542,319	518,427,275	534,037,123	550,261,550	
Average Trip Length (Miles)	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	0.71634	0.71634	0.71634	0.71634	0.71634	0.71634	0.71634	0.71634	
	User Input Value:									
Estimated Annual Pedestrian Miles of Travel		354,610,413	359,390,862	350,035,997	350,833,984	354,976,553	371,369,952	382,551,903	394,174,101	
Average Trip Duration (Minutes)	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	15.95757	15.95757	15.95757	15.95757	15.95757	15.95757	15.95757	15.95757	
	User Input Value:									
Estimated Annual Pedestrian Hours of Travel		131,658,224	133,433,088	129,959,854	130,256,128	131,794,162	137,880,633	142,032,219	146,347,259	
Fatalities		26	17	25	22	17	10	19	32	
Fatalities/Million Hours of Travel		0.197	0.127	0.192	0.169	0.129	0.073	0.134	0.219	

		Bicycling								
		2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	
Person Trip Rate	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	0.05439	0.05439	0.05439	0.05439	0.05439	0.05439	0.05439	0.05439	
	User Input Value:									
MPO Population Estimate	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	2,825,790	2,844,011	2,870,005	2,898,571	2,927,805	2,959,604	2,991,363	3,018,255	
	User Input Value:									
Population Adjustment Factor	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	1.00000	1.00682	1.04289	1.12132	1.08677	1.15036	1.29125	1.31320	
	User Input Value:									
Estimated Annual Bicyclist Trips		56,102,133	56,848,709	59,424,083	64,528,489	63,171,015	67,593,590	76,686,582	78,691,003	
Average Trip Length (Miles)	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	3.07657	3.07657	3.07657	3.07657	3.07657	3.07657	3.07657	3.07657	
	User Input Value:									
Estimated Annual Bicyclist Miles of Travel		172,602,331	174,899,227	182,822,553	198,526,634	194,350,265	207,956,644	235,931,900	242,098,648	
Average Trip Duration (Minutes)	Source:	Default	Default	Default	Default	Default	Default	Default	Default	❸
	Default Value:	22.69772	22.69772	22.69772	22.69772	22.69772	22.69772	22.69772	22.69772	
	User Input Value:									
Estimated Annual Bicyclist Hours of Travel		21,223,177	21,505,604	22,479,856	24,410,829	23,897,303	25,570,343	29,010,179	29,768,442	
Fatalities		4	5	3	4	3	2	4	1	
Fatalities/Million Hours of Travel		0.188	0.232	0.133	0.164	0.126	0.078	0.138	0.034	

	Non-Motorized							
	2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Estimated Annual Non-Motorized Trips	551,133,326	558,553,341	548,069,455	554,287,840	558,713,334	586,020,865	610,723,705	628,952,552
Estimated Annual Non-Motorized Miles of Travel	527,212,743.90	534,290,088.67	532,858,550.69	549,360,618.35	549,326,817.95	579,326,595.13	618,483,803.29	636,272,748.53
Estimated Annual Non-Motorized Hours of Travel	152,881,401.13	154,938,691.69	152,439,710.05	154,666,956.47	155,691,465.24	163,450,975.46	171,042,398.83	176,115,701.41
Non-Motorized Fatalities	30	22	28	26	20	12	23	33
Non-Motorized Fatalities/Million Hours of Travel	0.196	0.142	0.184	0.168	0.128	0.073	0.134	0.187

Figure 23. Interface of the MPO Area Exposure Estimates Tool Component

Step-by-Step Instructions

1: Open the spreadsheet containing the tools.
2: Read the "Introduction" page for an overview of the available tools.
3: Click on the "MPO Area Exposure Estimates" tab to select the tool.
4: Select the state of interest from the drop down menu titled "State" (also marked by ❶ in the spreadsheet tool).
5: Select the state of interest from the drop down menu titled "State" (also marked by ❷ in the spreadsheet tool).
6: Select the source (default or user input) of the required inputs, "Person Trip Rate", "MPO Population Estimate", "Population Adjustment Factor", "Average Trip Length (Miles)", and "Average Trip Duration (minutes)", from the dropdown menus highlighted in green for each year (also marked by ❸ in the spreadsheet tool).
7: For those required inputs where the default option is chosen, no further action is required.
8: For those required inputs where the user input option is chosen, enter the desired user input value into the appropriate cell (no color highlights) below the source dropdown menu.

Census Data for MPOs

The Census does not offer data specific to MPO geographies; therefore, tract-level ACS Census data are used to provide the finest resolution to areal interpolate the population and commuter population for the MPOs. Only ACS 5-year estimates are available at the tract level; therefore, the estimates represent a given year within the five-year period as opposed to any individual year. 1- and 3-year estimates are unavailable due to inadequate ACS sample sizes at small geographies (i.e., tracts and counties). Figure 24 offers a visual comparison example of the Census Core-base Statistical Areas (CBSA), MPO and tract geographies for Memphis, TN.

Variables that require ACS data:

POP_i = MPO population in ith year (derived through areal interpolation of tract-level ACS data)
PC_i = MPO commuter population in ith year (derived through areal interpolation of tract-level ACS data)
PC_{2009} = MPO commuter population in 2009 (derived through areal interpolation of tract-level ACS data)

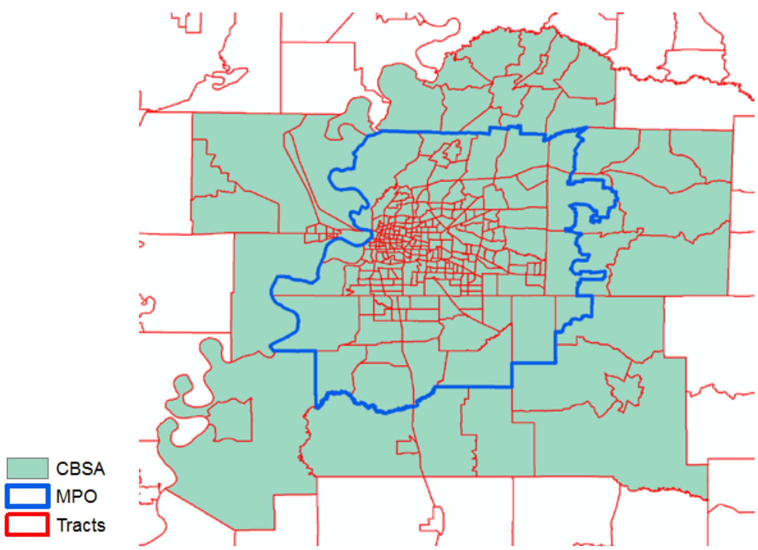


Figure 24. Geography Comparison

Developing Person Travel Estimates from the 2009 NHTS

CBSA Peer Grouping Methodology: The 2009 NHTS survey data represent only a sample of the traveling public from both rural and urban areas. A portion of the 2009 NHTS data are labeled as being located within Census Core-base Statistical Areas (CBSA) that represent metropolitan areas. The CBSA geography is the smallest geography in the 2009 NHTS data and the only way to locate a portion of the survey sample.

The survey sample data are grouped by their CBSA as indicated in the original 2009 NHTS data to represent metropolitan areas around the nation. The CBSA metropolitan areas serve as proxies for MPOs in terms of developing of travel estimates from the 2009 NHTS; however, sample sizes vary between CBSAs and are possibly not statistically representative of the local populations. CBSAs are then grouped together based on 2009 ACS 1-year estimates for bicycle and walk commute percentages to increase samples sizes. Tables 22 (bike) and 23 (walk) list the ACS commute percentage ranges for the initial CBSA peer groupings along with their corresponding 2009 NHTS trip sample and the generalized travel estimates.

Generalized Travel Estimates Applied to MPOs: The 2009 NHTS survey data are used to generate an average person trip rate, average trip length (miles), and average trip duration (minutes) for bicycling and walking per CBSA grouping. These generalized travel estimates are applied to the MPOs that possess similar bicycle and walk commute percentages as their peer 2009 NHTS CBSAs. The MPOs area assigned a CBSAs peer grouping for every year between 2009 and 2016 based on the annual release of ACS 5-year estimates for bicycling and walking commute percentages. Refer to the Appendix for a list of the MPOs with their corresponding ACS population and commuter information along with their CBSA bike and walk grouping assignments.

In developing the generalized travel estimates, the NHTS survey weights are not applied because:

- Estimates are for CBSAs with unrepresentative samples
- Unrepresentative samples do not include all segments of the population
- NHTS weights are to replicate the national population (meaning, each person is weighted to be represented within the national population)

Table 22. Bike - CBSA Peer Groupings

CBSA Name	2009 ACS Bike Commute Percentage	Quintile Grouping	2009 NHTS Bike Trips	Average Person Trip Rate	Average Person Trip Length (miles)	Average Person Trip Duration (minutes)
Memphis, TN-MS-AR	0.02%	1	531	0.00598	2.43	18.29
Nashville-Davidson--Murfreesboro--Franklin, TN	0.09%					
Charlotte-Gastonia-Concord, NC-SC	0.11%					
Birmingham-Hoover, AL	0.12%					
Dallas-Fort Worth-Arlington, TX	0.13%					
Cincinnati-Middletown, OH-KY-IN	0.18%					
San Antonio, TX	0.18%					
Oklahoma City, OK	0.20%					
Atlanta-Sandy Springs-Marietta, GA	0.20%					
Kansas City, MO-KS	0.21%					
Cleveland-Elyria-Mentor, OH	0.22%					
Pittsburgh, PA	0.24%	2	609	0.00779	2.47	21.15
Hartford-West Hartford-East Hartford, CT	0.24%					
Louisville-Jefferson County, KY-IN	0.26%					
Houston-Sugar Land-Baytown, TX	0.27%					
Riverside-San Bernardino-Ontario, CA	0.27%					
Providence-New Bedford-Fall River, RI-MA	0.28%					
St. Louis, MO-IL	0.30%					
Richmond, VA	0.31%					
Indianapolis-Carmel, IN	0.32%					
Detroit-Warren-Livonia, MI	0.32%					
Baltimore-Towson, MD	0.33%	3	662	0.00719	2.77	21.45
Las Vegas-Paradise, NV	0.34%					
Buffalo-Niagara Falls, NY	0.35%					
Raleigh-Cary, NC	0.36%					
New York-Northern New Jersey-Long Island, NY-NJ-PA	0.40%					
Virginia Beach-Norfolk-Newport News, VA-NC	0.41%					
Columbus, OH	0.42%					
Milwaukee-Waukesha-West Allis, WI	0.43%					
Orlando-Kissimmee, FL	0.45%					
Rochester, NY	0.50%					
Chicago-Naperville-Joliet, IL-IN-WI	0.57%	4	1,406	0.00970	2.82	22.07
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.57%					
Miami-Fort Lauderdale-Pompano Beach, FL	0.61%					
San Diego-Carlsbad-San Marcos, CA	0.62%					
Jacksonville, FL	0.64%					
Tampa-St. Petersburg-Clearwater, FL	0.70%					
Denver-Aurora-Broomfield, CO	0.72%					
Austin-Round Rock, TX	0.72%					
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.73%					
Minneapolis-St. Paul-Bloomington, MN-WI	0.86%					
Los Angeles-Long Beach-Santa Ana, CA	0.86%	5	1,430	0.01227	3.08	22.70
Salt Lake City, UT	0.87%					
Phoenix-Mesa-Scottsdale, AZ	0.91%					
Seattle-Tacoma-Bellevue, WA	0.92%					
New Orleans-Metairie-Kenner, LA	0.96%					
Boston-Cambridge-Quincy, MA-NH	1.03%					
San Jose-Sunnyvale-Santa Clara, CA	1.43%					
San Francisco-Oakland-Fremont, CA	1.54%					
Sacramento--Arden-Arcade--Roseville, CA	1.62%					
Portland-Vancouver-Beaverton, OR-WA	2.13%					

Table 23. Walk - CBSA Peer Groupings

CBSA Name	2009 ACS Walk Commute Percentage	Quintile Grouping	2009 NHTS Walk Trips	Average Person Trip Rate	Average Person Trip Length (miles)	Average Person Trip Duration (minutes)
Orlando-Kissimmee, FL	0.97%	1	10,692	0.07503	0.70	14.35
Nashville-Davidson--Murfreesboro--Franklin, TN	1.10%					
Birmingham-Hoover, AL	1.30%					
Memphis, TN-MS-AR	1.33%					
Richmond, VA	1.34%					
Dallas-Fort Worth-Arlington, TX	1.40%					
Atlanta-Sandy Springs-Marietta, GA	1.41%					
Tampa-St. Petersburg-Clearwater, FL	1.43%					
Kansas City, MO-KS	1.48%					
Raleigh-Cary, NC	1.51%					
Houston-Sugar Land-Baytown, TX	1.55%					
Indianapolis-Carmel, IN	1.57%	2	6,023	0.09039	0.67	14.55
Jacksonville, FL	1.58%					
Charlotte-Gastonia-Concord, NC-SC	1.58%					
St. Louis, MO-IL	1.64%					
Detroit-Warren-Livonia, MI	1.65%					
Louisville-Jefferson County, KY-IN	1.66%					
Oklahoma City, OK	1.66%					
Austin-Round Rock, TX	1.77%					
Miami-Fort Lauderdale-Pompano Beach, FL	1.77%					
Las Vegas-Paradise, NV	1.78%					
Phoenix-Mesa-Scottsdale, AZ	1.80%	3	7,854	0.08853	0.73	15.93
Sacramento--Arden-Arcade--Roseville, CA	1.84%					
San Antonio, TX	2.02%					
Riverside-San Bernardino-Ontario, CA	2.03%					
San Jose-Sunnyvale-Santa Clara, CA	2.13%					
Columbus, OH	2.14%					
Denver-Aurora-Broomfield, CO	2.15%					
Cincinnati-Middletown, OH-KY-IN	2.16%					
Hartford-West Hartford-East Hartford, CT	2.19%					
Cleveland-Elyria-Mentor, OH	2.25%					
Minneapolis-St. Paul-Bloomington, MN-WI	2.26%	4	13,552	0.10889	0.72	15.96
Salt Lake City, UT	2.27%					
Virginia Beach-Norfolk-Newport News, VA-NC	2.40%					
New Orleans-Metairie-Kenner, LA	2.59%					
Los Angeles-Long Beach-Santa Ana, CA	2.63%					
Providence-New Bedford-Fall River, RI-MA	2.79%					
San Diego-Carlsbad-San Marcos, CA	2.80%					
Baltimore-Towson, MD	2.85%					
Milwaukee-Waukesha-West Allis, WI	2.88%					
Buffalo-Niagara Falls, NY	3.10%					
Chicago-Naperville-Joliet, IL-IN-WI	3.17%	5	14,211	0.14297	0.68	14.50
Portland-Vancouver-Beaverton, OR-WA	3.17%					
Washington-Arlington-Alexandria, DC-VA-MD-WV	3.21%					
Rochester, NY	3.37%					
Seattle-Tacoma-Bellevue, WA	3.57%					
Pittsburgh, PA	3.71%					
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	3.75%					
San Francisco-Oakland-Fremont, CA	4.40%					
Boston-Cambridge-Quincy, MA-NH	5.12%					
New York-Northern New Jersey-Long Island, NY-NJ-PA	6.28%					

Manual Data Extraction

The following section offers details on how to the manually extract data from the ACS and NHTS sources. These data sources offer a variety of additional demographic and travel behavior information that may be of value to safety analysis projects or outreach efforts.

ACS Data Extraction

ACS provides pedestrian and bicycle commuting estimates in geodatabase and CSV formats.[7] Data are available for different measures related to pedestrian and bicycle commuting. Users can download data based on their requirements. To get estimate the number of workers that commute to work by walking or bicycling, four major variables can be used (see Table 24).

Table 24. ACS Data Attributes for Pedestrian and Bicycle Commute Estimates from Table B08301

Census Code	Variable Name
B08301e18	Means of Transportation to Work: Bicycle: Workers 16 years and over -- (Estimate)
B08301m18	Means of Transportation to Work: Bicycle: Workers 16 years and over -- (Margin of Error)
B08301e19	Means of Transportation to Work: Walk: Workers 16 years and over -- (Estimate)
B08301m19	Means of Transportation to Work: Walk: Workers 16 years and over -- (Margin of Error)

To create maps, users can use ESRI ArcGIS software to join TIGER/Line (Topologically Integrated Geographic Encoding and Referencing) shapefile with the ACS data tables.

Example Problem: Determine recent bicycle commuting estimates for Census Tracts in Texas

With the following steps, users can determine bicycle-commuting estimates for census tracts in Texas. It is important to note that manual data extraction requires expertise in ArcGIS tool.

1: Download 2011-2015 ACS Data from U.S. Census[8] (see Figure 25).

[7] American FactFinder https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

[8] TIGER/Line® with Selected Demographic and Economic Data https://www.census.gov/geo/maps-data/data/tiger-data.html Accessed on Sept 29, 2017.

Figure 25. ACS Data for Texas

2: By using ArcMap 10.3, users can join ACS shapefile with ACS commuting data provided in the geodatabase. Shapefile's variable 'GEOID' is required to be joined with 'GEOID' variable in ACS commuting data (see Figure 26).

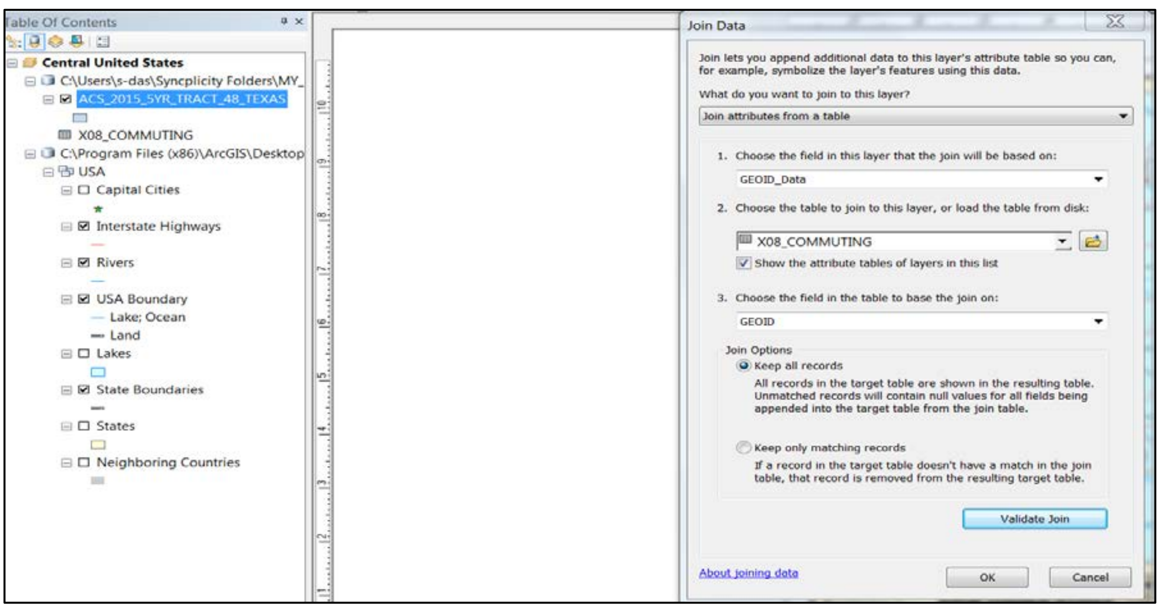


Figure 26. Joining of ACS Commuting Data with ACS Shapefile Data

3: Export data to text file format from the joined ACS shapefile attribute table.

4: To generate a choropleth map (see Figure 27), users need to follow the following steps:

- Right click on the generated shapefile to select properties
- Under the 'Symbology' tab, select Quantiles and Graduated Maps
- Select variable 'B08301e18' from the drop-down lists in values.

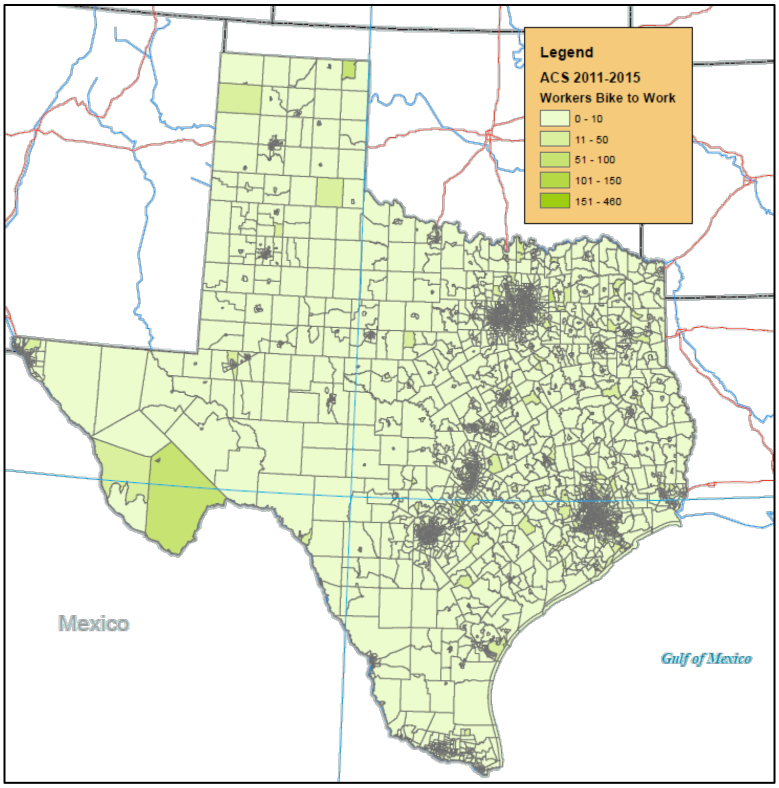


Figure 27. Map for Bicycle Commuting in Texas Census Tracts

NHTS Data Extraction

The following section highlights NHTS online tools that provide relevant pedestrian and bicyclist travel summaries at the state and CBSA geographies based on the NHTS data files as presented in Table 25.

Since these data are collected via a survey, and not a census, they must be weighted according to ACS demographic information to represent the entire population and produce valid estimates. Survey weights for households and persons are available for all useable households in the NHTS databases.

The 2009 and 2017 NHTS national sample datasets (i.e., adjusted for oversampling due to add-ons) offer the ability to estimate non-motorized travel exposure by person trips at the state level. Figure 28 shows the online analysis tool with the total annual person trips by mode per state from the 2009 NHTS. The tool provides choropleth maps of 2009 NHTS person trips per state with drill-down capability to state statistics on trips, mode, and purpose. For example, Colorado produced approximately 7 billion person trips in 2009 with 9.7% (approximately 675 million) being walking trips.

Table 25. Structure of 2009 and 2007 NHTS Data Files

Data Files	Information Included	Record Level	ID Variables	Weight Variables
Household	Data unique to a household. Example questions from interview sections: • Number of vehicles • Person Data • Type of Residence • Location of Home • Household Income • Education	One record per household unit	HOUSEID	WTHHFIN
Person	Data determined once for each completed person interview. Example questions from interview sections: • Age • Driver Status • Race & Ethnicity • Travel to Work • Miles driven • Education	One record per person	HOUSEID PERSONID	WTPERFIN
Vehicle	Data relating to each of the household's vehicles. Example questions from interview section: • Vehicle Data • Type of Residence • Verified Vehicle Data • Annualized Vehicle Miles • Household Income	One record per vehicle, if present	HOUSEID VEHID	WTHHFIN
Travel Day Trip	Data about each trip the person made on the household's randomly-assigned travel day. Example questions from interview section • Person Data • Travel Day Data	One record per travel day person trip	HOUSEID PERSONID TDTRPNUM	WTTRDFIN

Sources: 2009 NHTS Users Guide V2, page 6-2 and 2017 NHTS Users Guide, page 53

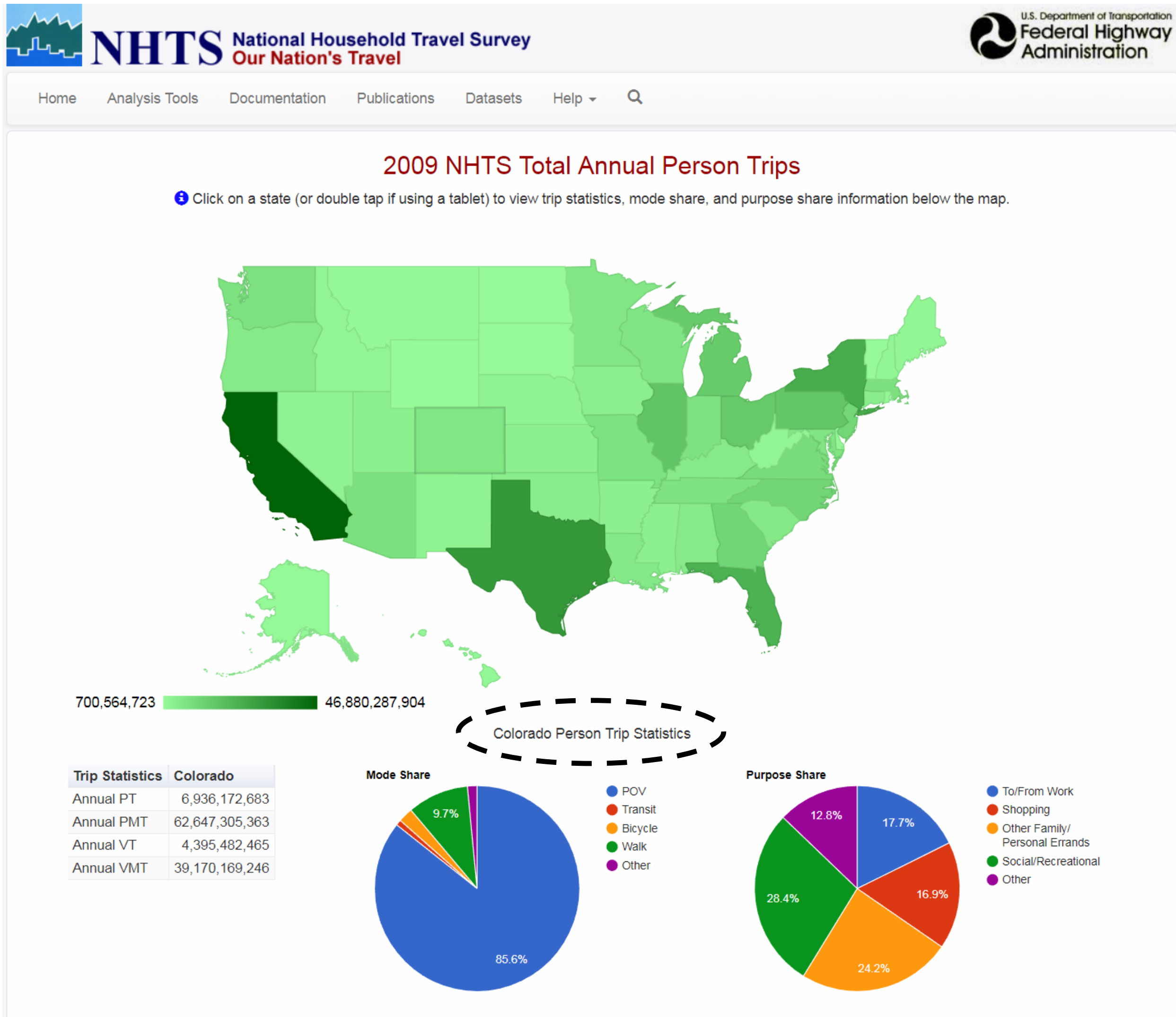


Figure 28. State-Level Annual Person Trips in NHTS Online Analysis Tool
Source: http://nhts.ornl.gov/tools/pt.shtml

Another online analysis tool based on the 2009 and 2017 NHTS national sample datasets is the table designer shown in Figure 29 and Figure 30. The table designer tools allow users to build customized data tabulations quickly and easily. The tabulation outputs are offered in either HTML, Excel, or CSV formats.

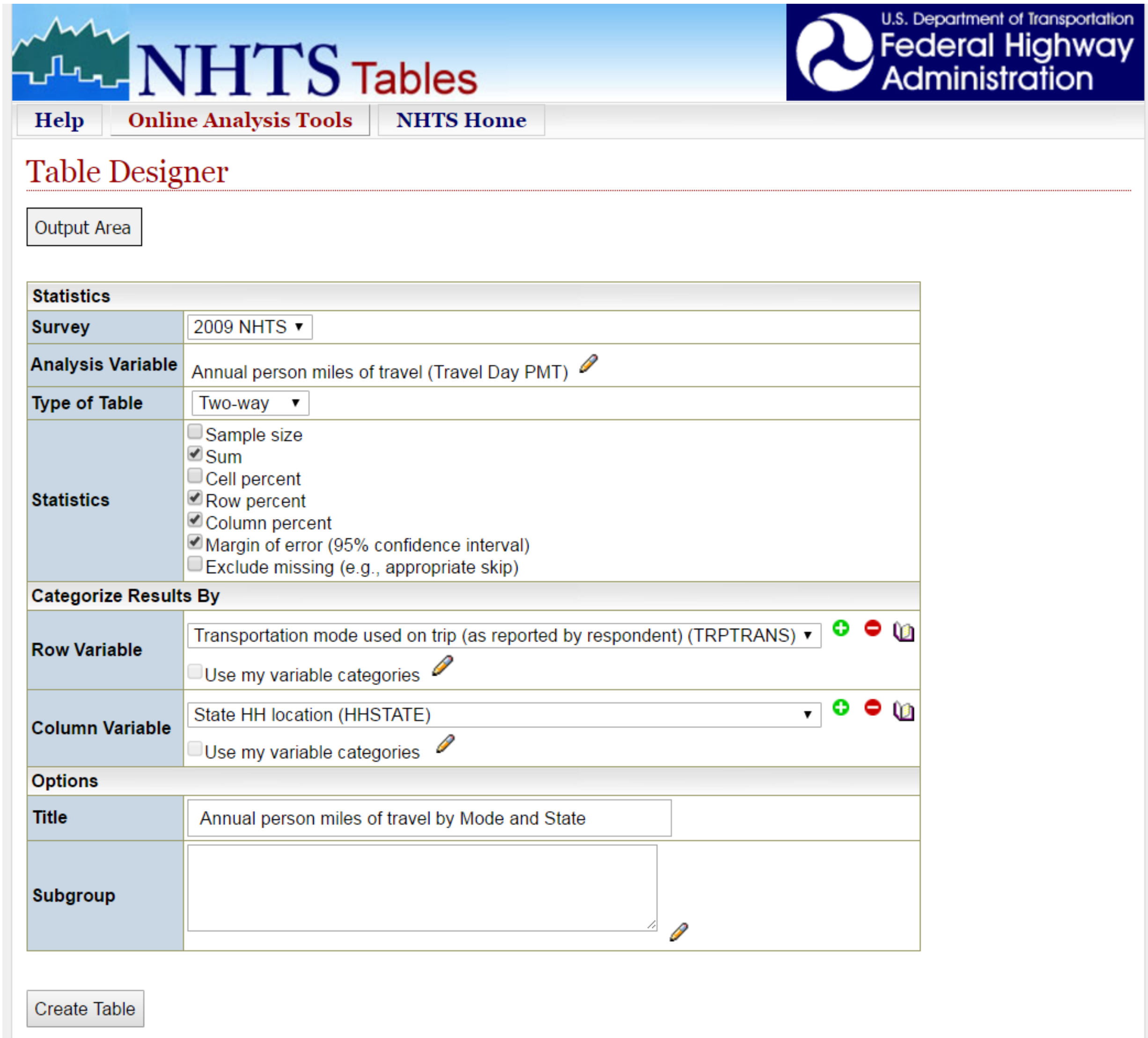


Figure 29. Table Designer Custom Query Tool for the 2009 NHTS

Source: http://nhts.ornl.gov/tables09/Default.aspx

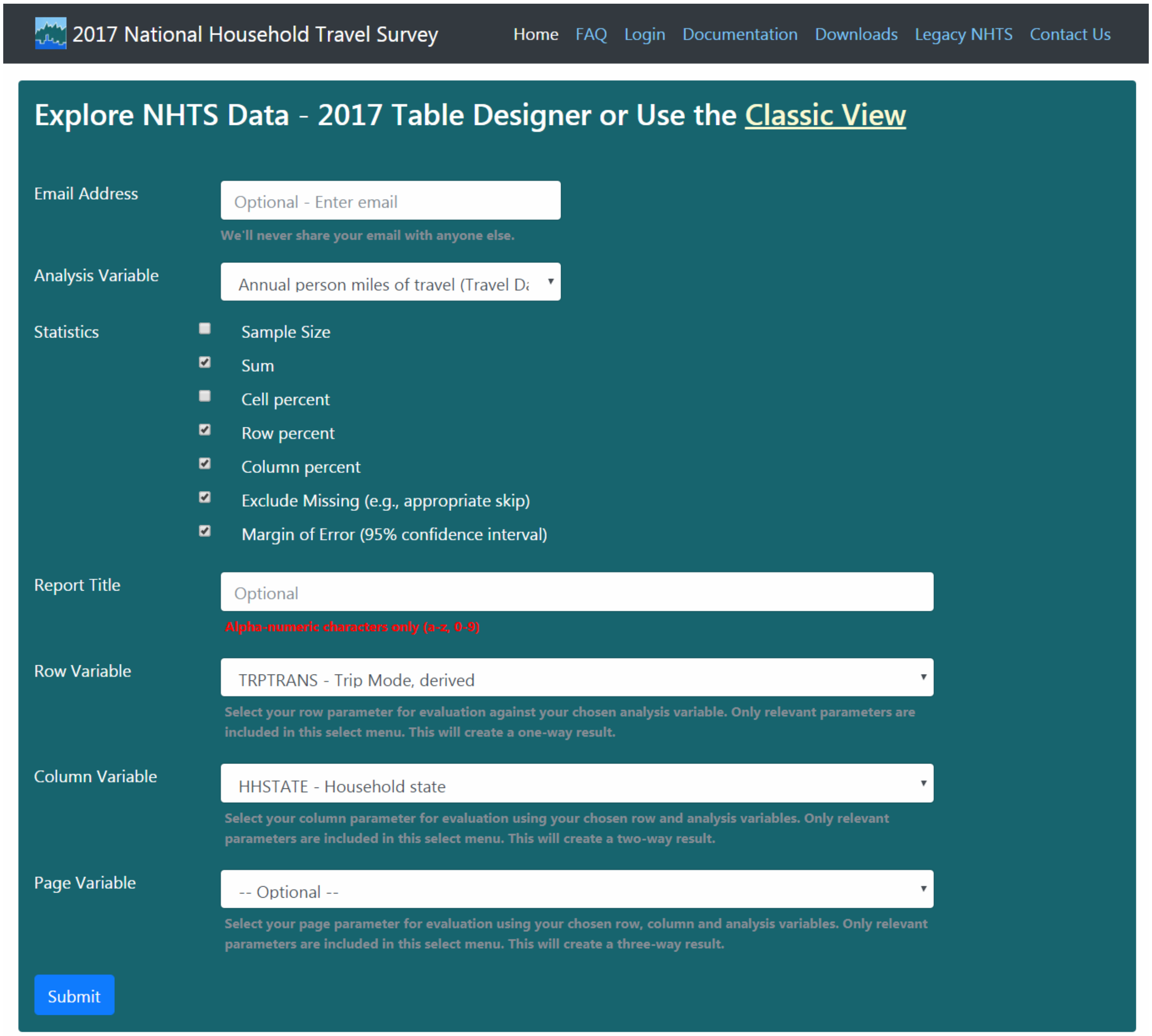


Figure 30. Table Designer Custom Query Tool for the 2017 NHTS
Source: https://nhts.ornl.gov/

The NHTS attributes can also be tabulated for various census-based geographies, such as state, CBSA, and rural/urban (http://nhts.ornl.gov/2009/pub/UsersGuideClaritas.pdf). For example, the exposure measure of annual person miles of travel can be calculated by mode per state with these tools. Table 26 shows that in 2017 Arizona generated approximately 114 billion person miles of travel, of which 0.5% (approximately 585 million) were by walking.

Table 26. 2017 NHTS Annual Person Miles of Travel (millions) by Mode and State

Trip Mode, derived	AK	AL	AR	AZ	CA	WA	WI	WV	WY	All
Walk	47	111	113	585	3,891	520	401	63	33	33,651
Bicycle	28	1	31	294	1,161	231	176	.	26	8,499
Car	1,200	55,222	10,030	34,206	205,354	34,944	27,258	9,731	2,764	1,532,612
SUV	1,065	11,654	6,514	15,881	80,072	19,160	15,739	3,873	1,222	817,847
Van	288	3,115	3,920	6,926	22,955	3,902	7,210	1,566	238	260,856
Pickup truck	1,970	5,660	8,062	10,781	31,378	6,117	7,700	2,115	2,407	386,559
Golf cart / Segway	.	0	0	21	12	.	1	.	.	461
Motorcycle / Moped	.	.	41	91	1,307	481	257	142	.	9,676
RV (motor home, ATV, snowmobile)	67	103	.	18	131	6	17	12	2	4,502
School bus	34	1,045	1,200	795	1,262	838	722	396	167	38,682
Public or commuter bus	35	88	.	491	6,258	2,131	505	58	29	37,827
Paratransit / Dial-a-ride	.	54	.	44	411	33	59	.	.	2,724
Private / Charter / Tour / Shuttle bus	267	308	.	523	2,543	979	431	.	48	19,607
City-to-city bus (Greyhound, Megabus)	.	.	60	0	165	.	109	.	.	2,563
Amtrak / Commuter rail	.	.	.	9	3,482	84	734	.	.	34,213
Subway / elevated / light rail / street car	47	.	.	405	4,470	403	35	.	.	31,694
Taxi / limo (including Uber / Lyft)	3	45	5	169	3,419	101	76	.	3	15,292
Rental car (Including Zipcar / Car2Go)	140	201	.	1,189	2,799	544	512	.	.	19,385
Airplane	2,562	644	.	41,568	105,459	24,234	10,404	.	1,156	667,211
Boat / ferry / water taxi	71	.	.	0	279	338	12	.	.	1,909
Something Else	168	774	0	306	1,796	118	1,177	151	50	43,844
All	7,989	79,025	29,977	114,303	478,606	95,163	73,537	18,106	8,145	3,969,614

Source: Federal Highway Administration, 2017 NHTS

STEP 7. COMPILE OTHER REQUIRED DATA

Step 7 in the scalable risk assessment process consists of compiling other data besides exposure that is required based upon the risk definition selected in Step 3. The three possible risk definitions are:

A) Observed crash rate
B) Expected crashes
C) Additional risk indicators

Detailed instructions for compiling other required data for these three risk definitions is beyond the scope of this guide. There is extensive guidance and examples in several other reports, manuals, and guides. Therefore, the following sections provide summary information and pointers to these other guidance documents.

Observed Crash Rate

To calculate observed crash rate, reported pedestrian and bicyclist crash data are compiled from existing state and local crash databases. The exact procedures for obtaining and compiling crash data vary from state to state (as well as the crash data attributes). Therefore, this section provides an overview and points to other published resources and guide. In particular, these FHWA documents are relevant for compiling crash data:

- *Highway Safety Improvement Program (HSIP) Manual*, FHWA-SA-09-029, January 2010.
- *Guidebook on Identification of High Pedestrian Crash Locations*, publication pending, June 2018.

Each agency that provides crash data will typically provide documentation and data dictionaries that describe crash database attributes. Typically, a crash database contains three major components:

- Crash-level data sets contain information about the entire crash, such as crash location, crash date, total fatalities in the crash, and light level.
- Vehicle- or unit-level data sets contain information about each vehicle (or unit) in the crash, such as vehicle type and harmful events. Pedestrians and bicyclists (pedalcyclists) are included as non-motorized vehicles.
- Person-level data sets contain information about all people in crashes, such as age and belt usage. The data set includes one record for each person involved in the crash.

If observed crashes are being used to quantify risk, the possibility of unreported crashes should be considered as a potential bias. In some cases, safety analysts will supplement official crash databases with other sources of data, such as that from emergency medical services, hospital outcomes, and public health databases. Considering these other sources may help to provide a more comprehensive database of pedestrian and bicyclist crashes.

Expected Crashes

Expected pedestrian and bicyclist crashes can be estimated by:

- Using HSM procedures to estimate point (i.e., intersection) or segment level predicted crashes, then using Empirical Bayes procedures to blend observed crashes and predicted crashes to estimate expected crashes at a specific location.
- Developing a crash prediction model at aggregate or disaggregate level using observed crashes and other causal factors, then using Empirical Bayes procedures to estimate expected crashes.

In some situations, a safety performance function is not available within the HSM and the blending can use locally developed safety performance functions with the observed crashes to estimate expected crashes.

In either case, exposure is considered an important factor in estimating expected crashes. Therefore, the exposure values developed in Step 6 will be used in this step to estimate expected crashes.

Safety analysts estimate expected crashes to overcome several issues associated with observed crashes. Observed crashes (especially pedestrians and bicyclists) can be a rare occurrence, and the actual observed number of crashes at specific locations may not accurately represent the risk to pedestrians and bicyclists.

The HSM has developed safety performance functions that are used to calculate predicted crashes. Then, Empirical Bayes procedures are used to estimate expected crashes (which is a weighted average of observed crashes and predicted crashes). However, at the time of this writing, the HSM procedures for pedestrian and bicyclist crashes are still being refined and are not comprehensive (e.g., they do not address crashes on rural roads). NCHRP Project 17-84 was initiated in early 2017 to improve guidance for pedestrian and bicyclist crash prediction in future editions of the HSM.

Several efforts have developed crash prediction models aside from those in the HSM. The development of crash prediction models is outside the scope of this Guide, but the following list includes examples of crash model development for interested readers:

- Turner, S., Wood, G., Hughes, T., Singh, R. Safety Performance Functions for Bicycle Crashes in New Zealand and Australia. Transportation Research Record 2236, pp. 66–73, 2011.
- Pulugurthaa, S., and Sambhara, V. Pedestrian Crash Estimation Models for Signalized Intersections. Accident Analysis and Prevention, Vol. 43, pp. 439–446, 2011.
- Nordback, K., Marshall, W., and Janson, B. Bicyclist safety performance functions for a U.S. city. Accident Analysis and Prevention, Vol. 65, pp. 114– 122, 2014.
- Alluri, P., Haleem, K., Gan, A., Lavasani, M., and Saha, D. Comprehensive Study to Reduce Pedestrian Crashes in Florida. Florida Department of Transportation, Grant: BDK80 977-32, 2015.
- Amoh-Gyimah, R., Saberi, M., Sarvi, M. Macroscopic modeling of pedestrian and bicycle crashes: A cross-comparison of estimation methods. Accident Analysis & Prevention, Vol. 93, pp. 147-159, 2016.
- Thomas, L., Lan, B., Sanders, R., Frackelton, A., Gardner, S., and Hintze, M. Changing the Future? Development and Application of Pedestrian Safety Performance Functions to Prioritize Locations in Seattle, WA. Transportation Research Board 96th Annual Meeting Compendium Papers, Washington D.C., 2017.

Additional Risk Indicators

In this definition of risk, analysts develop and compile additional risk indicators that have been defined in Step 3. The actual risk indicators may vary depending upon the location and facilities being analyzed, and are identified as part of a systemic safety evaluation process (or similar process). FHWA provides several resources for systemic safety at https://safety.fhwa.dot.gov/systemic/. In particular, three documents are relevant:

- Systemic Safety Project Selection Tool, FHWA-SA-13-019, July 2013

- Systemic Safety Project Selection Tool Supplemental Case Studies, FHWA-SA-17-002, December 2016
- Thomas et al. Systemic Pedestrian Safety Analysis, NCHRP Project 17-73, anticipated 2018.

The case studies in Report FHWA-SA-17-002 include an example of systemic analysis for pedestrian and bicyclist crashes, and this example was included earlier in this guide.

Attributes from roadway inventory and traffic count databases are often the starting point for identifying risk factors. For example, FHWA recommends the following list of potential risk factors for consideration in systemic safety analyses.

Roadway and Intersection Features

- Number of lanes
- Lane width
- Shoulder surface width/type
- Median width/type
- Horizontal curvature, delineation, or advance warning
- Horizontal curve and tangent speed differential
- Roadside or edge hazard rating
- Driveway density
- Presence of shoulder or centerline rumble strips
- Presence of lighting
- Presence of on-street parking
- Intersection skew angle
- Intersection traffic control device
- Number of signal heads vs. number of lanes
- Presence of backplates
- Presence of advanced warning signs
- Intersection located in/near horizontal curve
- Presence of left-turn or right-turn lanes
- Left-turn phasing
- Allowance of right-turn-on-red
- Overhead versus pedestal mounted signal heads
- Pedestrian crosswalk presence, crossing distance, signal head type

Traffic Volume

- Average daily traffic volumes
- Average daily entering vehicles

Other Features

- Posted speed limit or operating speed
- Presence of nearby railroad crossing
- Presence of automated enforcement
- Adjacent land use type, such as schools, commercial, or alcohol-sales establishments
- Location and presence of bus stops

Source: https://safety.fhwa.dot.gov/systemic/pdf/FHWA_SystemicApproach_PotentialRiskFactors.pdf

STEP 8. CALCULATE RISK VALUES

Step 8 in the scalable risk assessment process is to calculate risk values based on the outputs from previous steps. That is, Step 6 provides exposure estimates and Step 7 provides observed crashes, expected crashes, or additional risk indicators that are then used to calculate final risk values at the geographic scale chosen in previous steps.

Case studies are provided in this chapter to tie together the eight steps described in this guide. The case studies are based on actual examples of risk assessment for pedestrians and bicyclists.

Case Study: Pedestrian Risk Assessment in Michigan

The Michigan DOT partnered with the University of Michigan Transportation Research Institute to develop a risk assessment tool (http://www.cmisst.org/pedbike-risk-exposure/) for pedestrian crashes for all 83 counties throughout the state. Michigan DOT's goal was to create a risk score, based on mapping crashes and the risk characteristics, for a defined area or network for the entire state of Michigan. We now present a fictional case example based on this project.

This case example focuses on pedestrian risk assessment to identify corridors in Detroit Michigan in need of countermeasures. Often the characteristics that make walking safe (or unsafe) persist over space. For example, along busy roads, land use features like business districts or the lack of lighting are often consistent over space. Due to this spatial continuity, transportation engineers often would like to improve the facilities in an entire corridor, not just one location.

Table 27. Eight Steps for the Scale Risk Assessment Methodology as Applied to this Case

Steps	Explanation
Step 1: Define Use(s) of Risk Values	Network screening, Area Based
Step 2: Select Geographic Scale	Areawide->Network->Corridor
Step 3: Select Risk Definition	Definition 2: Expected Crashes
Step 4: Select Exposure Measure	Trips made
Step 5: Select Exposure Estimation Method	Demand Estimation ->Pedestrian Trip generation and flow models
Step 6: Estimate Exposure	Estimate binomial and logistic regressions
Step 7: Compile Other Required Data	
Step 8: Calculate Risk Values	

Step 1: Determine use(s) of risk values. Network screening, Area Based.
The MDOT engineers were interested in estimating pedestrian risks to identify corridors in need of countermeasures.

Step 2: Select Geographic Scale. Areawide->Network->Corridor
MDOT project team was interested analyzing risk at the corridor level.

Step 3: Select risk definition. Expected Crashes
The definition of risk combined various risk indicators to estimate the expected number of pedestrian-vehicle crashes.

Step 4: Select exposure measure. Trips made
The project team measured exposure in trips per day in a Pedestrian Analysis Zone (PAZ), which is a 400m x 400m unit of analysis. These units are aggregated up to the level of the corridor.

Step 5: Select analytic method to estimate exposure.
Demand Estimation ->Pedestrian Trip generation and flow models

The analytic method used a statewide travel survey, land-use data, and household characteristics to generate pedestrian trips at the PAZ level.

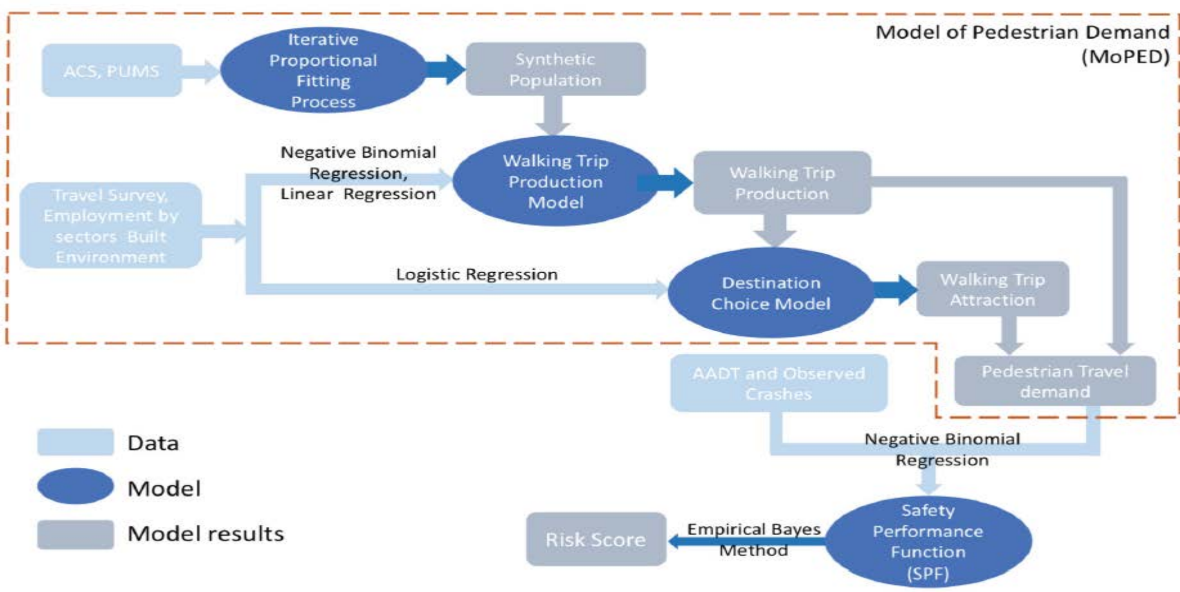


Figure 31. Step 5 Demand Estimation Using Pedestrian Trip Generation and Flow Models

STEP 6. Use analytic method to estimate selected exposure measure
The project team used the Michigan household travel survey (MTC III) to fit our trip production and destination choice models. We divided trips into five categories, namely home-based other (HBOther), home-based shopping (HBShopping), home-based school (HBSchool) and non-home-based other (NHBO) and non-home-based work (NHBW) and run the regression separately. We highlight the results for home based other (HBO) trips. For home-based trips, we estimate the number of trips per day at the household level using a negative binomial regression of the of the form,

Number of HB walking trips = f(number of households+household characteristics + built environment).

Below are representative regression results for the “home-based other” trip purpose.

Variable	Coef	Std err	P > |z|	[95.0% Conf. Int.]	
Constant	-0.7348	0.034	0.000	-0.801	-0.669
HHSIZE_1	-0.6565	0.071	0.000	-0.795	-0.518
HHSIZE_2	-0.1999	0.051	0.000	-0.299	-0.101
HHSIZE_3_or_more	-0.1217	0.073	0.0094	-0.021	0.264
HHVEH_0	0.8794	0.095	0.000	0.693	1.065
HHVEH_1	-0.2959	0.055	0.000	-0.404	-0.188
HHVEH_2	-0.5610	0.053	0.000	-0.664	-0.458
HHVEH_3_or_more	-0.7573	0.067	0.000	-0.888	-0.626
HHCHILD_0	-0.4110	0.051	0.000	-0.512	-0.310
HHCHILD_1_or_more	-0.3238	0.059	0.000	-0.439	-0.208
HHWORKER_0	-0.2714	0.040	0.000	-0.349	-0.194
HHWORKER_1_or_more	-0.4635	0.039	0.000	-0.540	-0.387
PIE	0.0272	0.002	0.000	0.022	0.032
Sample size	12062				
Log-Likelihood	-7457.5				

Figure 32. Pedestrian Trip Generation Model Results for home-based other trips

After completing all of the steps in the methodology, we obtained pedestrian exposure estimates. For more detailed information about the methodology, see (Cai et. al., 2018). Figure 33 shows the results for Wayne county Michigan where Detroit is located.

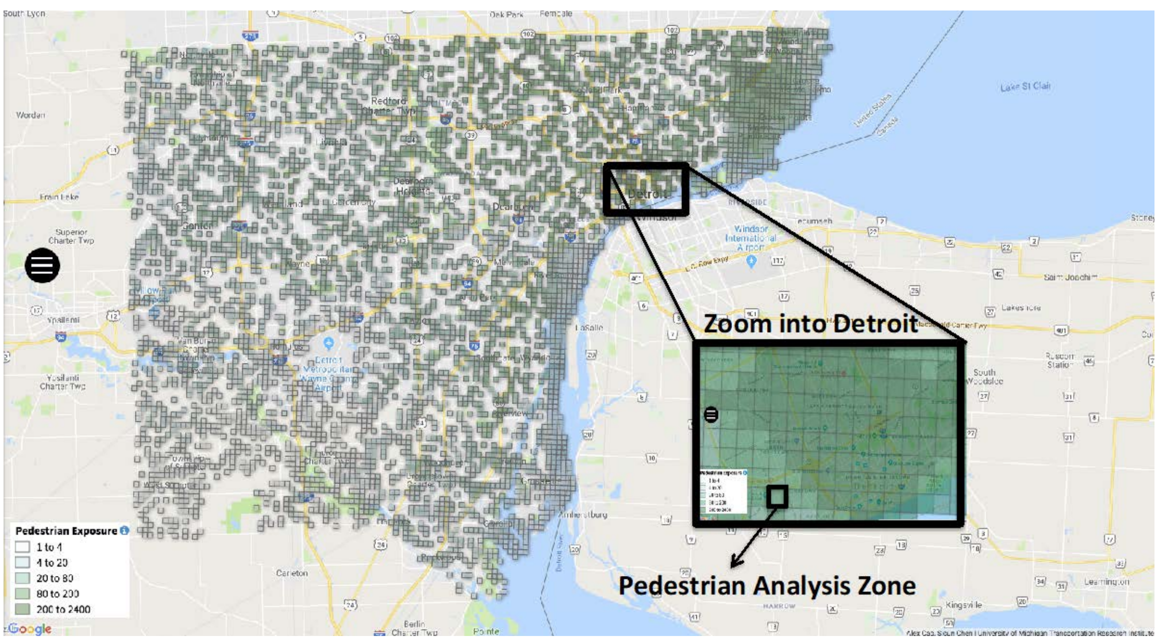


Figure 33. Daily pedestrian trips made per PAZ for Wayne county Michigan

STEP 7. Compile other required data.

The approach required many other data sources to calculate the risk values. The schematic below shows the various data sources used in the risk model.

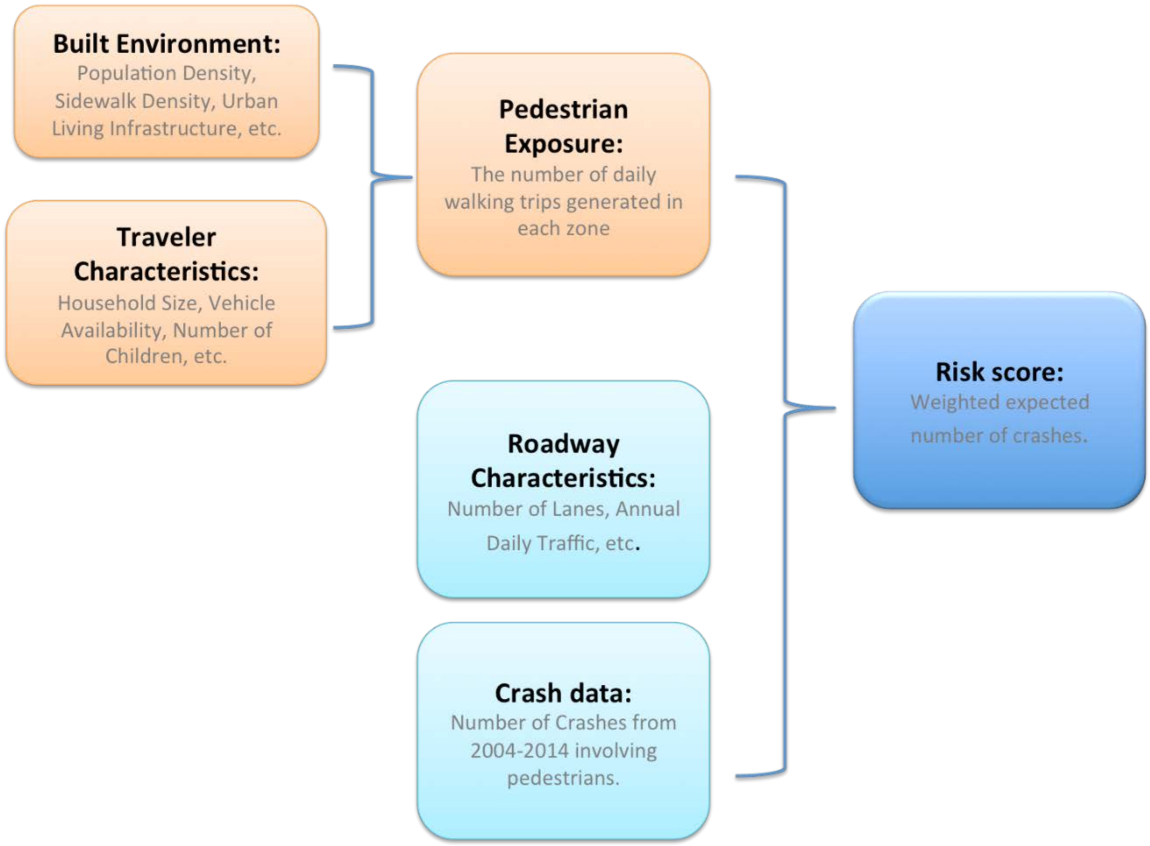


Figure 34. Step 7: Compile other required data

STEP 8. Calculate risk values

We used the Empirical Bayes framework from the HSM to create customized safety performance functions (SPFs) for both bicyclist and pedestrians (Cai et al., 2018).

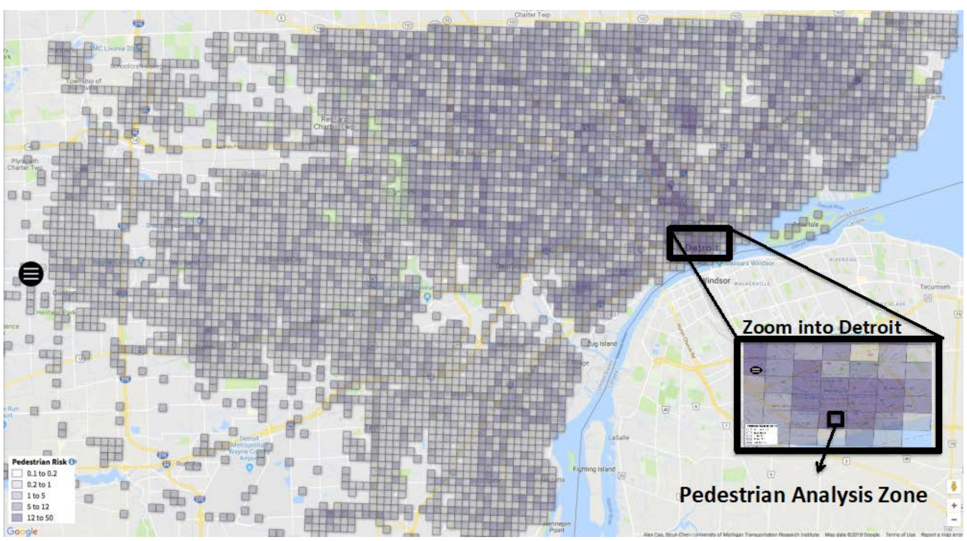


Figure 35. Risk measured as the expected number of crashes per PAZ

The MDOT engineers also wanted to analyze corridors. Thus we applied the Level-of-Service-of-Safety (LOSS) metric from the highway safety manual. LOSS divides the risk-scores of candidate areas into 4 categories based on its standard deviation from the average risk score (Kononov et al., 2003; Kononov et al., 2015). The areas in the highest quantile are the most dangerous for pedestrians. Figure 36 shows that LOSS map for pedestrian risk. In order to calculate the risk values of the corridors, we add together the risk score of each PAZ in the Gratiot corridor to arrive at a cumulative risk score of 50. The corresponding cumulative risk fir the Woodward corridor was 91.

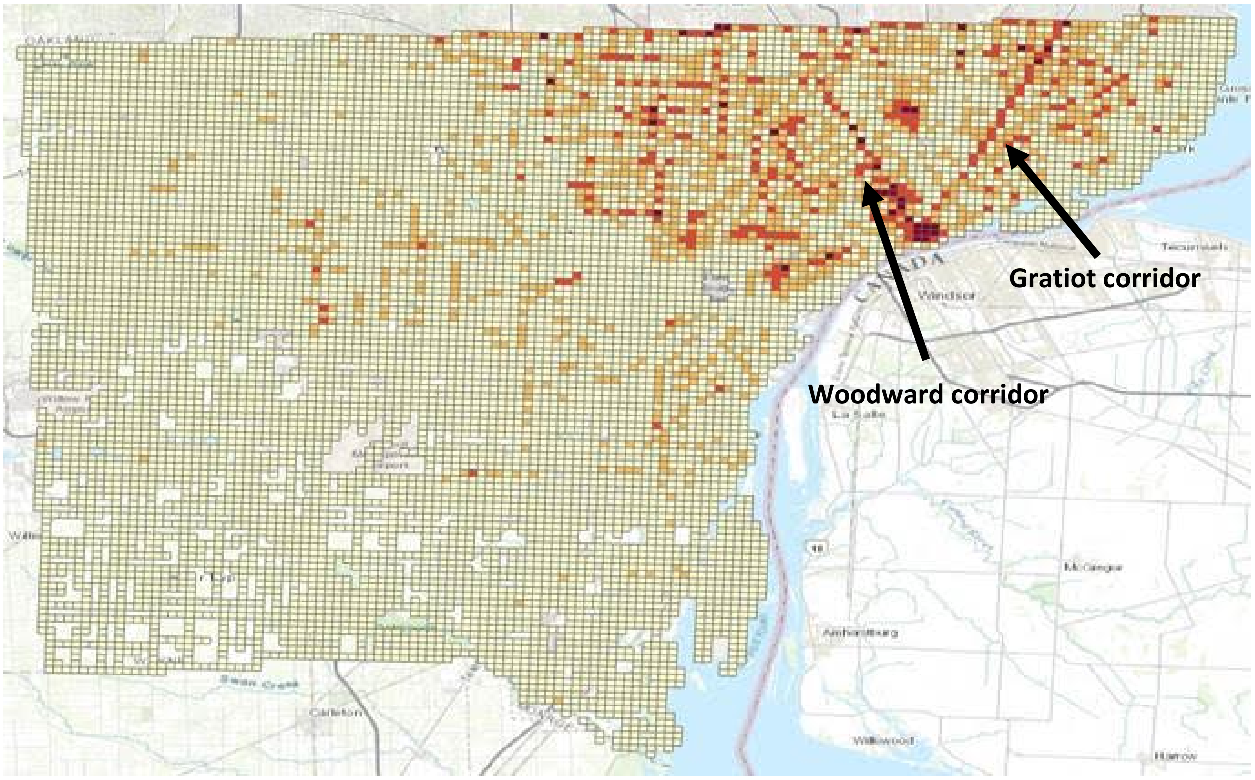


Figure 36. Level-of-Service-of-Safety (LOSS) map derived from the risk values

Lessons Learned

- Pedestrian Trip generation and flow models require significant technical capabilities.
- Methods from the Highway Safety Manual can be modified and used for non-motorized risk assessment.
- While these results are promising, extensive work is required to the validate exposure and risk models in order to integrate them into MDOT's business processes.

Case Study Example for Arizona DOT Bicycle Safety Action Plan

The scalable risk assessment methods were applied in Arizona as part of the Arizona DOT Bicycle Safety Action Plan as part of a statewide review of segments and intersection with a high priority for bicycle safety improvement projects.

Steps	**Explanation**
Step 1: Define Use(s) of Risk Values	C. Network Screening - Facility-Based
Step 2: Select Geographic Scale	Facility Specific - Segments and Points (Intersections)
Step 3: Select Risk Definition	Definition 3: Additional Risk Indicators were used because bike count information is not feasible for the statewide network and the exposure estimation is not applicable to a statewide network
Step 4: Select Exposure Measure	Not applicable for Risk Indicators
Step 5: Select Exposure Estimation Method	Not applicable for Risk Indicators
Step 6: Estimate Exposure	Not applicable for Risk Indicators
Step 7: Compile Other Required Data	
Step 8: Calculate Risk Values	

The study team applied a network planning analysis approach to identify priority corridor locations and countermeasures to provide safety improvements for bicyclists. Emphasis was placed on providing safe conditions for bicycle travel all along a corridor (segment) and within the bicycle travel network. To apply this network analysis approach to the 2018 Bicycle Safety Action Plan Update for the Arizona State Highway System, high-crash intersections and segments and high-crash potential segments were grouped into Priority Locations. A Priority Location may consist of one or more high-crash segments, intersection, or high-crash potential segments. The high crash potential segments were identified through a risk assessment methodology. These Priority Locations comprise 94% of the high-crash segments, 100% of the high-crash intersections, and 74% of the high-crash potential segments.

The approach included an initial review of high bicycle crash locations on the Arizona State Highway System. These locations were identified using GIS and subsequently verified by visual inspection. The locations are separated into highway segments and intersections/interchanges. A high-crash intersection/interchange and segment location includes at least three bicycle crashes within the five-year period. In addition, bicycle count data were included where available for the intersection/interchange and segment location. The count data were from the recent Arizona DOT efforts to develop a bicyclist and pedestrian count strategy plan for the State Highway System. The purpose of the counts is to provide insight into the bicycle exposure on these selected high-crash locations.

A key element of improving bicycle safety in Arizona is to proactively identify locations where bicycle improvements are needed, leading to projects to address the need. This section introduces a risk assessment methodology that can assist ADOT in identifying state highway segments and intersections where investment can help to lower the risk of bicycle crashes. The proposed methodology is similar to the process used in the 2016 *ADOT Pedestrian Safety Action Plan*. The assessment methodology represents an approach through which high-probability segments can be identified and addressed before bicyclist/motor-vehicle crashes occur.

Methodology

The methodology considers factors that are frequently identified as contributing factors or environmental/facility conditions that are common to bicycle crashes on the SHS. These factors are associated with the roadway facilities' existing conditions that relate to the absence of sufficient bicycle accommodation and bicycle demand as data is available. Bicycle demand can be estimated based on the facilities' proximity to specific land uses such as institutional areas that include schools, colleges, or universities, or being part of a known popular cycling route or corridor. Strava is a tool that can be used as a tool to help identify the popularity of cycling routes and corridors, although the Strava app data may be used more focused by recreational bicyclists.

Application of the methodology occurred through a **GIS-based screening** that utilized available statewide GIS data to identify and screen potential SHS locations where bicycle facilities should be considered, consistent with an established set of risk criteria. *Note that interstates were excluded from the screening as the intent of this is application was to identify and direct resources to where they will be the most effective.*

Table 28. Summarizes the Factors and Scoring for the Risk Assessment Process

Factor	Score
Operating Environment/Width of Roadway	
6-Lane Highway	6
4- or 5-Lane Undivided Highway	3
2- or 3-Lane Undivided Highway	2
2- or 3- or 4-Lane Divided Highway	1
Posted Travel Speed	
50 mph or greater	6
35-45 mph	4
25-30 mph	2
20 mph or less	0
Paved Effective Shoulder Width/Wide Curb Lane	
0-4 feet	6
4-8 feet	0
Bicyclist Exposure to Vehicles	
>7,500 ADT	6
2,500-7,500 ADT	3
<2,500 ADT	0
Designated U.S. Bicycle Route (USBR) 90*	
Yes	3
No	0
Environment Type	
Urban	6
Rural	3

*The USBR is not so much as risk factor, but is used to gain higher priority for improvements with the designation.

Screening Results

A scale was developed based on the distribution of the overall scores assigned to the SHS. The scale is defined in Table 29. A total of 31 higher-risk locations were identified and are shown in Figure 37.

Table 29. Bicyclist/Motor-Vehicle Risk Assessment Levels

SCALE	RISK LEVEL
>20	Higher Risk
14-19	Medium Risk
<13	Lower Risk

Lessons Learned

- Although there was a desire to incorporate an exposure measure into the risk assessment, it was not appropriate at the statewide level for this project. Bicycle count data would have been needed at enough locations to estimate for the entire network and that is not feasible at this time.
- The use of population data and/or national census results were evaluated but determined not to be an appropriate application for exposure on a statewide basis.
- The use of population density as a surrogate factor related to exposure within the risk assessment is being evaluated.

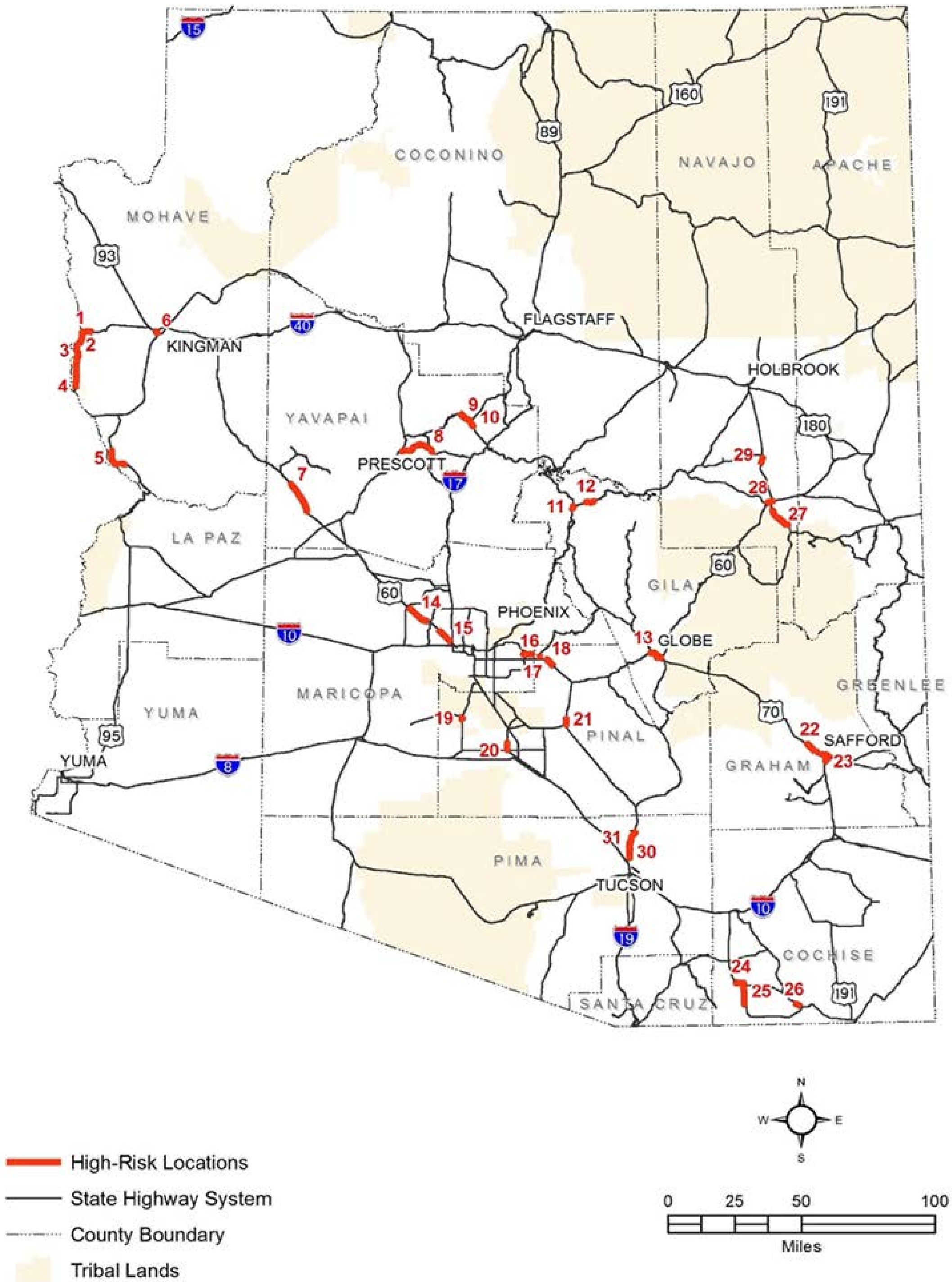


Figure 37. Higher-Risk Locations for Bicyclists on the Arizona State Highway System

BIBLIOGRAPHY

Abdelghany, A., Adbelghany, K., Mahmassani, H.S., and Al-Zharani, A. (2012). Dynamic Simulation Assignment Model for Pedestrian Movements in Crowded Networks. Transportation Research Record 2316. pp. 95–105.

Alliance for Biking & Walking (2016). Bicycling and Walking in the United States: 2016 Benchmarking Report.

Aoun, A., Bjornstad, J., DuBose, B., Mitman, M., Pelon, M., and Fehr & Peers. (2015). Bicycle and Pedestrian Forecasting Tools: State of the Practice. DTFHGI-11-H-00024. Federal Highway Administration, Washington, DC.

Cai, Yiming, Li, Xiao, T. Fabusuyi, Molnar, L. & Hampshire, R.C. (2018). A Statewide Pedestrian Crash Risk Assessment for Michigan: An application of the Model of Pedestrian Demand (MoPED). Transportation Research Board Annual Meeting.

Cambridge Systematics, Inc. and CH2M Hill, Inc. (December 2016). Systemic Safety Project Selection Tool Supplemental Case Studies. Report FHWA-SA-17-002.

Castiglione, J., Bradley, M. and John G. (2014). Activity-Based Travel Demand Models: A Primer. Strategic Highway Research Program (SHRP2) Report S2-C46-RR-1. Transportation Research Board, Washington, DC.

Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J. (2016). Representing pedestrian activity in travel demand models: Framework and application. Journal of transport geography, 52, 111-122.

Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J. (2016). Development of destination choice models for pedestrian travel. Transportation Research Part A: Policy and Practice, 94, 255-265.

Clifton, K.J., Burnier, C.V., Schneider, R., Huang, S., Kang, M.W. (2008). Pedestrian Demand Model for Evaluating Pedestrian Risk Exposure. Report prepared for the Maryland State Highway Administration, Office of Traffic and Safety.

Clifton, K.J., Davies, G., Allen, W.G., and Radford, N. (2004). Pedestrian Flow Modeling for Prototypical Maryland Cities. Technical report prepared for the Maryland Department of Transportation Division of Highway Safety Programs Hanover, MD, November 2004.

Clifton, K.J., Singleton, P., Muhs, C.D., Schneider, R.J., Lagerwey P. (2013). Improving the Representation of the Pedestrian Environment in Travel Demand Models – Phase I. Technical report prepared by the Oregon Transportation Research and Education Consortium.

DKS Associates, University of California, Irvine, University of California Santa Barbara, and Utah State University. (2007). Assessment of Local Models and Tools for Analyzing Smart-Growth Strategies. Final Report, Prepared for the California Department of Transportation.

Fagnant, D. J., and Kockelman, K. (2016). A Direct-Demand Model for Bicycle Counts: The Impacts of Level of Service and Other Factors. Environment and Planning B: Planning and Design, Vol. 43, No. 1, pp. 93–107.

Federal Highway Administration. (2016). FHWA New Non-Motorized Travel Analysis Toolkit. Available at http://catsip.berkeley.edu/news/fhwa-new-non-motorized-travel-analysis-toolkit.

Fitzpatrick, K., Park, E.S. (July 2010). Safety Effectiveness of the HAWK Pedestrian Crossing Treatment. Report FHWA-HRT-10-042.

Greene-Roesel, R., Diogenes, M.C., Ragland, D.R. (2007). Estimating Pedestrian Accident Exposure: Protocol Report. (In Estimating Pedestrian Accident Exposure: Final Report, California PATH Research Report UCB-ITS-PRR-2010-32, Final Report for Task Orders 5211/6211).

Hankey, S., and Lindsey, G. (2016). Facility-Demand Models of Peak Period Pedestrian and Bicycle Traffic: Comparison of Fully Specified and Reduced-Form Models. Transportation Research Record, No. 2586, pp. 48–58.

Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B., and Xu, Z. (2012). Estimating Use of Nonmotorized Infrastructure: Models of Bicycle and Pedestrian Traffic in Minneapolis, MN. Landscape and Urban Planning, Vol. 107, No. 3, pp. 307–316.

Hankey, S., Lu, T., Mondschein, A., and Buehler, R. (2017). Merging Traffic Monitoring and Direct-Demand Modeling to Assess Spatial Patterns of Annual Average Daily Bicycle and Pedestrian Traffic. Transportation Research Board 2017 Annual Meeting.

Hong, J., Shankar, V.N. and Venkataraman, N. (2016). A Spatially Autoregressive and Heteroskedastic Space-time Pedestrian Exposure Modeling Framework with Spatial Lags and Endogenous Network Topologies. Analytic Methods in Accident Research. Vol. 10, pp. 26–46.

Hood, J. Sall, E., Charlton, B. (2011). A GPS-based Bicycle Route Choice Model for San Francisco, California. Transportation Letters: International Journal of Transportation Research. Vol. 3, No. 1., pp. 63-75.

Kononov, J., & Allery, B. (2003). Level of service of safety: Conceptual blueprint and analytical framework. Transportation Research Record: Journal of the Transportation Research Board, (1840), 57-66.

Kononov, J., Durso, C., Lyon, C., & Allery, B. (2015). Level of service of safety revisited. Transportation Research Record: Journal of the Transportation Research Board, (2514), 10-20.

Kuzmyak, J. R., Walters, J., Bradley, M., and Kockelman, K.M. (2014). Estimating Bicycling and Walking for Planning and Project Development: A Guidebook. NCHRP 770. Transportation Research Board, Washington, DC.

Lagerwey, P., Hintze, M., Elliott, J., Toole, J., and Schneider, R. (2015). NCHRP Report 803: Pedestrian and Bicycle Transportation Along Existing Roads—ActiveTrans Priority Tool Guidebook.

Lassarre, S., Papadimitriou, E., Yannis, G. and Golias, J. (2007). Measuring Accident Risk Exposure for Pedestrians in Different Micro-Environments. Accident Analysis and Prevention. Vol. 39, No. 6, pp. 1226–1238.

Lee, K. and Sener, I.N. (2017). Emerging Data Mining for Pedestrian and Bicyclist Monitoring: A Literature Review Report. Technical Report, prepared for the Safety through Disruption (Safe-D) National University Transportation Center. Texas A&M Transportation Institute, September 2017.

McDaniel, S., Lowry, M. and Dixon, M. (2014). Using Origin-Destination Centrality to Estimate Directional Bicycle Volumes. Transportation Research Record 2430. pp. 12–19.

Molino, J.A., Kennedy, J.F., Inge, P.J., Bertola, M. A., Beuse, P A., Fowler, N.L., Emo, A.K., Do, A. (2012). A Distance-Based Method to Estimate Annual Pedestrian and Bicyclist Exposure in an Urban Environment. Report No. FHWA-HRT-11-043.

Molino, J.A., Kennedy, J.F., Johnson, P.L., Beuse, P.A., Emo, A.K., Do, A. (2009). Pedestrian and Bicyclist Exposure to Risk: Methodology for Estimation in an Urban Environment. Transportation Research Record 2140. pp. 145-156.

Munira, S. and Sener, I.N. (2017). Use of the Direct-Demand Modeling in Estimating Nonmotorized Activity: A Meta-Analysis. Technical Report, prepared for the Safety through Disruption (Safe-D) National University Transportation Center. Texas A&M Transportation Institute, October 2017.

National Highway Traffic Safety Administration. (February 2017). Traffic Safety Facts: 2015 Data, Pedestrians. NHTSA Report DOT-HS-812-375.

Papadimitriou, E., Yannis, G., and Golias, J. (2012). Analysis of Pedestrian Exposure to Risk in Relation to Crossing Behavior. Transportation Research Record 2299. pp. 79–90.

Proulx, F. and Pozdnoukhov, A. (2017). Bicycle Traffic Volume Estimation using Geographically Weighted Data Fusion. Technical Paper.

Radwan, E., Abou-Senna, H., Mohamed, A., Navarro, A., Minaei, N., Wu, J., and Gonzalez, L. (2016). Assessment of Sidewalk/Bicycle-Lane Gaps with Safety and Developing Statewide Pedestrian Crash Rates. Prepared for the Florida Department of Transportation. Final Report, Contract No. BDV24-977-07.

Raford, N. and Ragland, D. (2004). Space Syntax: Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety. Transportation Research Record 1878. pp. 66-74.

Raford, N. and Ragland, D.R. (2006). Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: Case of Boston, Massachusetts. TRB 2006 Annual Meeting Compendium of Papers.

Salon, D. (2016). Estimating Pedestrian and Cyclist Activity at the Neighborhood Scale. Journal of Transport Geography. Vol. 55, pp. 11–21.

Schmiedeskamp, P. and Zhao, W., (2016). Estimating Daily Bicycle Counts in Seattle, Washington, from Seasonal and Weather Factors. Transportation Research Record 2593. pp. 94–102.

Schneider, R. and Stefanich, J. (2015). Wisconsin Pedestrian and Bicycle Crash Analysis: 2011-2013. Prepared for the Wisconsin Department of Transportation (WisDOT), Bureau of Transportation Safety (BOTS). Final Draft.

Schneider, R.J., Henry, T., Mitman, M.F., Stonehill, L., Koehler, J. (2012). Development and Application of a Pedestrian Volume Model in San Francisco, California. Transportation Research Record 2299. pp. 65–78.

Seattle DOT. (September 2016). City of Seattle Bicycle and Pedestrian Safety Analysis.

Sener, I.N., Eluru, N. and Bhat, C.R. (2009). An Analysis of Bicycle Route Choice Preferences in Texas, U.S. Transportation, 36(5): 511-539.

Sener, I.N., Ferdous, N., Bhat, C.R., and Reeder, P.R. (2009). Tour-based Model Development for TxDOT: Evaluation and Transition Steps. Technical Report, prepared for the Texas Department of Transportation (TxDOT), October 2009.

Southern California Association of Governments (SCAG) (2016). SCAG Regional Travel Demand Model and 2012 Model Validation. March 2016.

Strauss, J., Miranda-Moreno, L.F., and Morency, P. (2013). Cyclist Activity and Injury Risk Analysis at Signalized Intersections: A Bayesian modelling approach. Accident Analysis and Prevention. Vol. 59, pp. 9–17.

Strauss, J., Miranda-Moreno, L.F., and Morency, P. (2014). Multimodal Injury Risk Analysis of Road Users at Signalized and Non-Signalized Intersections. Accident Analysis and Prevention. Vol. 71, pp. 201–209.

Tabeshian, M., and Kattan, L. (2014). Modeling Nonmotorized Travel Demand at Intersections in Calgary, Canada: Use of Traffic Counts and Geographic. Transportation Research Record, No. 2430, pp. 38–46.

Thomas, L., Lan, B., Sanders, R.L., Frackelton, A., Gardner, S., Hintze, M. (2017). Changing the Future? Development and Application of Pedestrian Safety Performance Functions to Prioritize Locations in Seattle, WA. TRB 2018 Annual Meeting Compendium of Papers.

Travel Forecasting Resources. http://tfresource.org/Travel_Forecasting_Resource.

Turner, S., Sener, I., Martin, M., Das, S., Shipp, E., Hampshire R., Fitzpatrick, K., Molnar, L., Wijesundera, R., Colety, M., Robinson, S. (January 2017). Synthesis of Methods for Estimating Pedestrian and Bicyclist Exposure to Risk at Areawide Levels and on Specific Transportation Facilities. Report FHWA-SA-17-041.

Wang, J., Lindsey, G., and Hankey, S. (2016). Exposure to Risk and the Built Environment, an Empirical Study of Bicycle Crashes in Minneapolis. TRB 2017 Annual Meeting Compendium of Papers.

Winters, M., Brauer, M., Setton, E. M., and Teschke, K. (2010). Built Environment Influences on Healthy Transportation Choices: Bicycling Versus Driving. Journal of Urban health, Vol. 87, No. 6, pp. 969-993.

Zimmermann, M., Mai, T., and Frejinger, E. (2017). Bike Route Choice Modeling Using GPS Data without Choice Sets of Paths. Transportation Research Part C: Emerging Technologies, Vol. 75, pp. 183–196.