Highway Safety Improvement Program (HSIP) Evaluation Guide

1. Planning

Communication (Feedback Loop)

2. Implementation

3. Evaluation

	Number of "before" years		
3	Number of "after" years	1	
4	Number of "before" comparison		193
5	Variance		0.001
6	Desired		0.1
7	Cumulative probability	=ABS(NORMSINV(B...))	1.64
8	Expected reduction [100*(1-CMF)]	User Input	20
11	...of "before" crashes in comparison group	=B4*B2	193
12	Estimated number of "after" crashes in comparison		193
13	Estimated number treatment group	...B11	193
14	Estimate of the variance without change	...*B5+B9/B11+B9 /B12)	616.249
15	Estimated index of effectiveness [CMF]	=B10/B13	0.8

FHWA Safety Program

May 2017

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

Foreword

Evaluation is a key component of the Highway Safety Improvement Program (HSIP) and should be standard business practice for state and local agency safety management programs. Through evaluation, we will know if our efforts are making a difference, or if we should pursue a different course of action.

This HSIP Evaluation Guide is a major revision to the 1981 Highway Safety Evaluation Procedural Guide (FHWA-TS-81-219) and serves as a companion guide to the HSIP Manual (FHWA-SA-09-029). The purpose of the HSIP Evaluation Guide is to support States' HSIP evaluation efforts and suggest practical ways to incorporate regular evaluations into HSIP management practices. HSIP managers, project managers, data analysts, and researchers can use this guide to recognize the importance of HSIP evaluation; understand the different levels of evaluation and associated methods and data requirements; enhance HSIP evaluation practices; and overcome challenges related to evaluation.

Employing more consistent and reliable evaluation methods will support future HSIP decisions, optimize return on investment of safety funding, and increase the effectiveness of projects and programs. Evaluation ensures continued success of the safety management program and demonstrates contributions to long-term safety outcomes and annual safety performance targets.

Michael S. Griffith, Acting Director
FHWA Office of Safety Programs

Notice

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document. This report does not constitute a standard, specification, or regulation.

The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

Quality Assurance Statement

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

TECHNICAL DOCUMENTATION PAGE

1. Report No. FHWA-SA-17-039	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Highway Safety Improvement Program (HSIP) Evaluation Guide	5. Report Date May 2017	
	6. Performing Organization Code	
7.Author(s) Frank Gross	8. Performing Organization Report No.	
9. Performing Organization Name and Address VHB 940 Main Campus Drive, Suite 500 Raleigh, North Carolina 27606	10. Work Unit No.	
	11. Contract or Grant No. DTFH61-16-D-00005	
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Safety 1200 New Jersey Ave., SE Washington, DC 20590	13. Type of Report and Period Final Report, June 2016 – May 2017	
	14. Sponsoring Agency Code FHWA	

15. Supplementary Notes
The Contracting Officer's Task Manager (COTM) was Karen Scurry.

16. Abstract
The Highway Safety Improvement Program (HSIP) comprises three components: planning, implementation, and evaluation. This guide focuses on the evaluation component, which provides a critical feedback loop to update processes and inform future decisions. The target audience includes safety program managers, project managers, data analysts, and researchers involved with project, countermeasure, or program evaluations, in particular HSIP evaluations. The purpose of the HSIP Evaluation Guide is to provide direction and resources to agencies that want to begin or enhance HSIP evaluation efforts. Readers will be prepared to track critical project details in support of HSIP evaluation, perform three different levels of evaluation (i.e., project, countermeasure, and program level), and use the results of HSIP evaluations to improve decisions and processes. Readers will also understand the methods, data requirements, and considerations associated with the different levels of evaluation. Examples from various States highlight noteworthy practices and the benefits of evaluation. Employing more consistent and reliable evaluation methods will support safety decision making, positively affect processes intended to reduce crashes, optimize return on investment of safety funding, and increase the effectiveness of projects and programs.

17. Key Words: Highway, Safety, Improvement, Program, HSIP, Evaluation, Data, Analysis	18. Distribution Statement No restrictions.		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 116	22. Price

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

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

APPROXIMATE CONVERSIONS FROM SI UNITS

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

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

A	Suspected serious injury on KABCO scale
AADT	Annual average daily traffic
B	Suspected minor injury on KABCO scale
C	Possible injury on KABCO scale
CMF	Crash modification factor
CRASH	Crash Reduction Analysis System Hub
DOT	Department of Transportation
EB	Empirical Bayes
FARS	Fatality Analysis Reporting System
FHWA	Federal Highway Administration
HSIP	Highway Safety Improvement Program
K	Fatal injury on KABCO scale
KTC	Kentucky Transportation Center
KYTC	Kentucky Transportation Cabinet
LTAP	Local Technical Assistance Program
MOE	Measure of effectiveness
O	No apparent injury on KABCO scale
RTM	Regression-to-the-mean
Safety PM	Safety Performance Management Measures
SHSP	Strategic Highway Safety Plan
SPF	Safety performance function
TOPS	Traffic Operations and Safety
PDO	Property damage only
PIES	Post-Implementation Evaluation System
PPRM	Program Project and Resource Management
VMT	Vehicle miles traveled

EXECUTIVE SUMMARY

The Highway Safety Improvement Program (HSIP) is a core Federal-aid highway program with the purpose to achieve a significant reduction in fatalities and serious injuries on all public roads through the implementation of highway safety improvement projects. As described in FHWA's HSIP Manual, the HSIP includes three components: planning, implementation, and evaluation. While planning and implementing projects are important steps to addressing existing and future safety opportunities, evaluating these efforts is critical to understanding the return on investment and improving the efficiency and effectiveness of future decisions.

The HSIP requires States to establish an evaluation process to analyze and assess results achieved by highway safety improvement projects. [23 CFR 924.13(a)(1)] While States are not required to report evaluation results for individual project locations to FHWA, the evaluation of individual projects supports higher levels of evaluations (e.g., countermeasure and program). The following are specific benefits to evaluating the safety effectiveness of individual projects, countermeasures (i.e., groups of similar projects), and programs.

- **Understand the return on investments**: Each year, transportation agencies invest nearly $4 billion on HSIP projects with the intent to reduce fatalities and serious injuries on all public roads. HSIP evaluations can help to demonstrate the value of these expenditures in terms of lives saved and serious injuries prevented. By demonstrating the value of past investments, there is an opportunity to justify the need for and appropriate level of future HSIP funding.

- **Identify and address potential opportunities**: Not all safety improvement projects result in a safety performance benefit. HSIP evaluation can help to identify investments that did not perform as intended. If an agency identifies a project that is not meeting safety performance expectations based on the evaluation results, then there is an opportunity to address the situation as appropriate for the location (e.g., remove the countermeasure or install supplemental countermeasures).

- **Inform future decisions**: With competing demands and limited funds, there is a need to prioritize efforts and justify decisions. Evaluations can help to develop or refine estimates of effectiveness used to prioritize projects and manage programs. For example, if certain programs or countermeasures are consistently effective (i.e., reduce the expected frequency and severity of crashes), then agencies may choose to continue those programs and implement similar countermeasures at additional locations.

- **Improve processes**: Evaluation can help to assess the HSIP process from start to finish, identifying opportunities to improve planning, implementation, evaluation, and documentation processes and procedures. For example, as part of countermeasure or program evaluation, it is useful to examine crashes that occur during construction to identify work zone configurations or construction practices that enhance safety.

- **Demonstrate accountability**: There is an ever-increasing demand for accountability at all levels of government. HSIP evaluation can help agencies to measure progress toward achieving their long-term safety goals and annual safety performance targets.

- **Meet federal requirements**: 23 CFR Part 924 requires each State to develop, implement, and evaluate on an annual basis a HSIP that has the objective to significantly reduce fatalities and serious injuries resulting from crashes on all public roads.

The HSIP Evaluation Guide provides agencies with the knowledge, tools, and insights to begin or enhance HSIP evaluation efforts. For those new to evaluation, the guide introduces the process and provides tips and templates to prepare for HSIP evaluation. It also describes the considerations, methods, and data requirements for HSIP evaluation, providing instructions and examples for application. For those familiar with HSIP evaluation, the guide can help to enhance current practices. Specifically, it identifies common challenges and opportunities to overcome those challenges. It provides several examples from States with noteworthy HSIP evaluation practices, which other States may consider adopting or modifying for their own use. Finally, the guide explains how to use the results from project, countermeasure, and program evaluations to inform future planning and implementation decisions and enhance current processes.

In summary, HSIP evaluation is critical to understanding the return on investment and the effectiveness of past decisions. This guide can help agencies prepare for and conduct HSIP evaluations, and use the results of HSIP evaluations to inform future decisions and improve future investments.

CHAPTER 1: INTRODUCTION

The Highway Safety Improvement Program (HSIP) is a core Federal-aid highway program with the purpose to achieve a significant reduction in fatalities and serious injuries on all public roads through the implementation of highway safety improvement projects. The HSIP, like other Federal-aid highway programs, is a federally funded, State administered program. The Federal Highway Administration (FHWA) establishes the HSIP requirements via 23 CFR Part 924, and the States develop and administer a program to best meet their needs.

> **Chapter 1 at a Glance**
> 1.1 Objective
> 1.2 Audience
> 1.3 Background
> 1.4 Organization and Use
> 1.5 Resources

The HSIP requires a data-driven, strategic approach to improving highway safety on all public roads that focuses on performance. This is accomplished through the planning, implementation, and evaluation of highway safety improvement projects, as shown in Figure 1.[1]

Planning

Problem Identification: the process of collecting, managing, and analyzing crash and other data needed to identify highway safety problems and opportunities.

Countermeasure Identification: the process of identifying factors that contribute to crashes, and developing countermeasures to address or mitigate the underlying factors.

Project Prioritization: the process of ranking individual countermeasures and projects to develop a portfolio of safety improvements.

Implementation

Schedule and implement safety improvement projects: the process of identifying funding sources, allocating resources, programming projects, and developing evaluation plans.

Evaluation

Determine the effectiveness of highway safety improvements: the process of performing project, countermeasure, and program evaluation to improve future decision-making through the HSIP and Strategic Highway Safety Plan (SHSP) planning processes.

Figure 1. Chart. HSIP process.

HSIP evaluation provides a critical feedback loop to inform decisions in future planning and implementation efforts. This HSIP Evaluation Guide is a major revision to the 1981 Highway Safety Evaluation Procedural Guide and serves as a companion document to the 2010 HSIP Manual.[1,2] This guide elaborates on the material in the HSIP Manual, providing agencies with the knowledge, tools, and insights to begin or enhance HSIP evaluation efforts. The remainder of this chapter identifies the objective and target audience of the guide, provides a general background on HSIP evaluation, including the value to agencies, and describes the organization and use of the guide.

HSIP evaluation provides a critical feedback loop to inform decisions in future planning and implementation efforts.

1.1 OBJECTIVE

The objective of this guide is to support States' HSIP evaluation efforts and suggest practical ways to incorporate regular evaluations into HSIP management practices. This guide can help readers to:

- Understand the HSIP evaluation component.

- Understand the importance of HSIP evaluation.

- Identify appropriate HSIP evaluation procedures.

- Identify the data requirements to employ HSIP evaluation procedures.

- Identify opportunities to enhance current HSIP evaluation practices.

- Overcome challenges related to HSIP evaluation.

1.2 AUDIENCE

The primary audience for this guide is safety program managers, project managers, data analysts, and researchers involved with project, countermeasure, or program evaluations. These users may be responsible for:

- Tracking HSIP projects.

- Developing annual HSIP reports.

- Evaluating projects, countermeasures, subprograms, or the HSIP program as a whole.

This guide will help to answer questions such as:

- What is the appropriate level of evaluation?

- What are the state-of-the-art HSIP evaluation procedures?

- What data are required for the different HSIP evaluation procedures?

- What tools and resources are available to support HSIP evaluation procedures?

A secondary audience for this guide is project managers (e.g., regional, district, and local staff) and internal support staff (e.g., IT staff). Project managers may be responsible for managing project implementation, including project tracking. IT staff may be responsible for establishing, maintaining, and providing access to project management resources and HSIP evaluation tools.

1.3 BACKGROUND

Before delving into the details of HSIP evaluation, it is important to understand the basic terminology associated with HSIP evaluation, the importance of HSIP evaluation, and the general timeline for evaluation.

What is HSIP Evaluation?

In general, evaluations include the analysis of data for a specific performance measure to understand the effectiveness of a given effort. HSIP evaluation comprises multiple levels of efforts and various performance measures. This guide describes performance measures related to safety, economics, and efficiency, as well as the process of analyzing these performance measures to assess the value of highway safety improvement projects and programs and to inform future decisions.

- **Safety** performance measures focus on the change in crash frequency and severity as well as the corresponding rates per measure of exposure.

- **Economic** performance measures focus on the cost-effectiveness (i.e., cost to change crash frequency and severity) and return on investment (i.e., benefit-cost ratio).

- **Efficiency** performance measures focus on project management activities comparing actual to planned values such as level of implementation, budget, and schedule.

This guide describes three levels of HSIP evaluation: project, countermeasure, and program.

> The following are examples of safety, economic, and efficiency performance measures related to shoulder rumble strips.
>
> **Safety** performance measures may include the change in total, fatal and serious injury, and run-off-road crashes over a defined period.
>
> **Economic** performance measures may include the benefit-cost ratio of a given rumble strip project or a group of similar rumble strip projects.
>
> **Efficiency** performance measures may include comparisons of the actual miles, cost, and schedule of shoulder rumble strip projects compared to the planned miles, cost, and schedule of projects for a given year.

Project Evaluation

Project evaluation is the evaluation of individual HSIP projects. Project evaluation measures the effectiveness of a safety improvement project by changes in the frequency and severity of crashes resulting from project implementation against the before condition. An example is the evaluation of the installation of a left-turn lane at a signalized intersection three years after project completion. The following are general considerations related to project evaluation.

- Consider **tracking individual projects over time** to collect and maintain information required to conduct project evaluations, and subsequently countermeasure and program evaluations.

- Consider the **methodology and study period** for evaluating crash-based performance measures. There is generally a need for several years of crash data to perform before-after evaluations.

- Consider the potential to **evaluate non-crash-based performance measures** to assess the implementation process and intermediate effectiveness of a completed improvement project. Non-crash-based performance measures include project management factors (e.g., schedule and budget) and changes in non-crash safety measures (e.g., operating speed, driver compliance, driver response).

Countermeasure Evaluation

Countermeasure evaluation is the evaluation of groups of similar projects. Countermeasure evaluation measures the effectiveness of a group of similar projects by changes in the frequency and severity of crashes at the treated locations, often with the intention to develop a crash modification factor (CMF). An example is the evaluation of a group of left-turn lane installations at signalized intersections. The following are general considerations related to countermeasure evaluation.

> A **CMF** is a multiplicative factor that indicates the expected change in crashes associated with a countermeasure. A CMF of 1.0 indicates no expected change in crash frequency. A CMF less than 1.0 indicates an expected reduction in crashes. A CMF greater than 1.0 indicates an expected increase in crashes.

- Consider the need to **link individual projects with specific countermeasures**. This should occur at the project tracking level to facilitate future countermeasure evaluations.

- Consider the **consistency in projects** (e.g., countermeasures and site characteristics) when combining multiple project locations for countermeasure evaluation. Different combinations of countermeasures and site characteristics along with variations in vehicles and driver behavior can result in differences in countermeasure effectiveness.

- Consider the **appropriate evaluation method**. If the intent is to develop agency-specific CMFs for use in future decision-making, then study designs should account for potential sources of bias in estimating countermeasure effectiveness.

- Consider the **sample size required** to obtain reliable results. While larger sample sizes generally provide more statistically reliable results, there is a need to balance the desired reliability with the resources required to collect and analyze the data.

Program Evaluation

Program evaluation is the evaluation of the overall HSIP program or subprograms. Program evaluation measures the effectiveness of a program by changes in the frequency, severity, and rate of crashes at the system level. An example is the evaluation of an intersection safety program (i.e., all HSIP projects targeting intersection safety). The following are general considerations related to program evaluation.

- Consider the need to **link individual projects with specific programs** and subprograms. This should occur at the project tracking level to facilitate future program and subprogram evaluations. For example, a shoulder rumble strip project may rollup under a State's roadway departure program.

- Consider the potential to **assess non-crash-based performance measures**. Project and program management evaluations can assess differences between planned and actual resource expenditures and the productivity of implementing highway safety projects and programs.

- Consider the opportunity to **use the results** from project, countermeasure, and program evaluations to enhance HSIP processes and inform the evaluation and update of the State Strategic Highway Safety Plan (SHSP).

Importance of HSIP Evaluation

While identifying candidate project locations, selecting appropriate countermeasures, and implementing projects are important steps to addressing existing and future safety problems and opportunities, evaluating these efforts is critical to understanding the return on investment and improving the efficiency and effectiveness of future decisions. The following is a summary of the various benefits related to evaluating the safety effectiveness of individual projects, countermeasures (i.e., groups of similar projects), and programs.

Understand the Return on Investments

HSIP evaluation is critical to understanding the return on safety investments. Each year, transportation agencies invest nearly $4 billion on HSIP projects with the intent to reduce fatalities and serious injuries on all public roads. There is a need to justify these expenditures to those responsible for funding and managing HSIP projects and programs. HSIP evaluations can help to demonstrate the value of these expenditures in terms of lives saved and serious injuries prevented as well as other measures such as the cost to save a life or prevent a serious injury. By demonstrating the value of past investments, there is an opportunity to justify the need for an appropriate level of future HSIP funding.

Identify and Address Potential Opportunities

Not all safety improvement projects result in a safety performance benefit. HSIP evaluation can help to identify investments that did not perform as intended and opportunities to increase the benefits returned from the projects. If an agency identifies a project that is not meeting safety performance expectations based on the evaluation results, then there is an opportunity to address the situation as appropriate for the location (e.g., remove the countermeasure or install supplemental countermeasures).

Inform Future Decisions

With competing demands and limited funds, there is a need to prioritize efforts and justify decisions. Quantitative evaluations can help to develop or refine estimates of effectiveness used to prioritize projects and manage programs. For example, if certain programs or countermeasures are consistently effective (i.e., reduce the expected frequency and severity of crashes), then agencies may choose to continue those programs and implement similar countermeasures at additional locations.

> **Quantitative evaluations** focus on the safety effectiveness of projects, countermeasures, and programs (e.g., change in crash frequency and severity).
>
> **Qualitative evaluations** focus on the strengths, limitations, and opportunities to improve processes and procedures.

Improve Processes

While there is a tendency to focus on the safety effectiveness of projects and programs during evaluation efforts, there is also an opportunity to identify opportunities related to the underlying processes and procedures. Both quantitative and qualitative evaluations can help to assess the HSIP process from start to finish, identifying opportunities to improve planning, implementation, evaluation, and documentation processes and decisions. The following are specific examples of process improvements related to HSIP planning and implementation.

- There is **value in comparing the safety performance after implementation to both the before and interim periods** to determine if there is a difference in results. If projects continually demonstrate a larger benefit from the before period to the after period compared to the interim period to the after period, then this suggests potential site selection bias and the agency may need to enhance their network screening practices.

> **Interim period** is the time between site selection and project implementation.

- There is **value in evaluating safety performance during construction**. While evaluations often exclude the construction or implementation period, it is useful to examine crashes during construction separately as part of countermeasure or program evaluation. This can help to identify work zone configurations or construction practices that enhance safety.

- There is **value in evaluating changes over time after implementation** of a given project. If an agency identifies an initial increase in crashes followed by a long-term reduction in crashes, there may be opportunities to improve future implementation practices (e.g., educate road users on navigating a roundabout). This is an example where project evaluation at one location can help to inform the rollout of similar countermeasures at other locations in the future.

Demonstrate Accountability

There is an ever-increasing demand for accountability at all levels of government. HSIP evaluation can help agencies to measure progress toward achieving their long-term safety goals and annual safety performance targets.

Meet Federal Requirements

Program and project effectiveness evaluations have been a key part of highway safety improvement programs since the 1966 Highway Safety Act. The following are current (as of the date of publication) requirements related to HSIP evaluation.

- 23 CFR Part 924.5(a) requires each State to develop, implement, and evaluate on an annual basis a HSIP that has the objective to significantly reduce fatalities and serious injuries resulting from crashes on all public roads.

- 23 CFR Part 924.13(a)(1) requires each State's HSIP evaluation process to include a process to analyze and assess the results achieved by the program of highway safety improvement projects in terms of contributions to improved safety outcomes and the attainment of safety performance targets established as per 23 U.S.C. 150.

- 23 CFR Part 924.13(a)(2) requires each State's HSIP evaluation process to include an evaluation of the SHSP as part of the regularly recurring update process to 1) confirm the validity of the emphasis areas and strategies based on analysis of current safety data, and 2) identify issues related to the SHSP's process, implementation, and progress that should be considered during each subsequent SHSP update.

- 23 CFR Part 924.13(b) requires each State to use the HSIP evaluation results for 1) updating safety data used in the planning process, 2) setting priorities for highway safety improvement projects, 3) assessing the overall effectiveness of the HSIP, and 4) reporting.

In summary, HSIP evaluation can help States manage and implement projects, countermeasures, and programs more efficiently and effectively, resulting in decision-making that has a higher potential to return on safety investments in the future. Further, national tools are becoming increasingly available to help agencies select countermeasures and implement performance-based decision-making practices. These data-driven safety analysis tools rely on a predictive

approach to estimate the safety performance under given geometric and operational conditions; however, the safety effects of geometric and operational conditions may be different for a specific State (or city, county, or region) than those presented in national tools. Agencies may improve their future decision-making by evaluating the effectiveness of projects and countermeasures in their jurisdiction as opposed to relying on national estimates.

1.4 GUIDE ORGANIZATION AND USE

The remainder of the guide includes seven chapters, four appendixes, and a glossary to provide agencies with the knowledge, tools, and insights to establish or enhance HSIP evaluation efforts. In general, each chapter builds on the last. Specifically, readers will learn how to prepare for HSIP evaluation, track and maintain relevant data for individual projects, analyze project-level data to evaluate projects, aggregate individual projects to evaluate countermeasures, aggregate the results of countermeasure evaluations to evaluate overall program effectiveness, and use the results of HSIP evaluations to inform future decisions. In some cases, there are multiple potential methods available for a given evaluation level. The methods differ in terms of data requirements and reliability of results. Analysts should select the most appropriate methods based on the goals, objectives, and resources associated with the evaluation. The following is a brief overview of each chapter and how readers can use the guide to establish or enhance various aspects of an HSIP evaluation program.

- **Chapter 2: Preparing for HSIP Evaluation** describes general considerations to prepare for HSIP evaluation. Readers will identify potential challenges and learn about opportunities to overcome those challenges based on examples presented from other States. Readers will also learn about the HSIP evaluation timeline, how to develop an evaluation plan, and how to use the results from HSIP evaluations.

- **Chapter 3: Project Tracking** describes procedures to monitor and track individual projects, which provides the foundation for any level of evaluation. Readers can use the examples and templates to determine the relevant project data for various levels of HSIP evaluation and to establish or enhance project tracking procedures.

- **Chapter 4: Project Evaluation** focuses on project-level evaluation. Readers will find appropriate measures of effectiveness, evaluation methods, and data requirements.

- **Chapter 5: Countermeasure Evaluation** focuses on countermeasure-level evaluation. Readers will find appropriate measures of effectiveness, evaluation methods, and data requirements.

- **Chapter 6: Program Evaluation** focuses on program-level evaluation. Readers will find appropriate measures of effectiveness, evaluation methods, and data requirements.

- **Chapter 7: Using HSIP Evaluation Results** describes the use of HSIP evaluation results, highlighting the feedback loop within the HSIP process. Readers will find opportunities to use the results from HSIP project, countermeasure, and program evaluations to improve HSIP processes and future decisions.

- **Chapter 8: Closing** provides a summary. Readers can use chapter 8 to review the salient points and critical elements related to HSIP evaluation.

- **Appendix A: Project Tracking** provides information to develop a project tracking template.

- **Appendix B: Evaluation Templates** provides templates to conduct evaluations.

- **Appendix C: Sample Size Templates** provides templates to estimate sample size requirements for evaluations.

- **Appendix D: Additional Resources** provides a summary of related resources and training for additional information on evaluation.

1.5 BACKGROUND RESOURCES

The following resources provide additional background information on the HSIP and HSIP evaluation.

- 2010 HSIP Manual provides an overview of the HSIP and offers practitioners with a review of standards, new and emerging technologies, and noteworthy practices for each step in the HSIP process.[1]

- The Art of Appropriate Evaluation: A Guide for Highway Safety Program Managers provides an overview of the traffic safety evaluation process, specifically for highway safety program managers.[3]

CHAPTER 2: PREPARING FOR HSIP EVALUATION

There are many considerations to prepare for and conduct successful evaluations. First, agencies should consider the evaluation timeline and develop an evaluation plan. HSIP staff and management support, as well as staffing availability and expertise are essential to establish and sustain a successful HSIP evaluation practice. Agencies should also consider the appropriate scale and scope of evaluations and provide related guidance and tools. Further, agencies should consider how they will use the results of HSIP evaluations to support future decisions. The following is a detailed description of various considerations to help prepare for HSIP evaluation, including potential challenges and opportunities to overcome these challenges.

> **Chapter 2 at a Glance**
> 2.1 Evaluation Timeline
> 2.2 HSIP Staff and Management Support
> 2.3 Staffing Availability and Expertise
> 2.4 Right-Sizing Evaluations
> 2.5 Guidance and Tools
> 2.6 Application of Evaluation Results
> 2.7 Summary
> 2.8 Resources

2.1 EVALUATION TIMELINE

While evaluation occurs after implementation of a given project, countermeasure, or program, the evaluation process should begin during the planning stage and end with a feedback loop to inform future decisions.

The Challenge

Agencies should consider evaluation during the planning and implementation stages so they are prepared to conduct evaluations later in the process. For example, as agencies plan projects, they can begin collecting preconstruction data and documenting important preconstruction details for future evaluations. If an agency does not plan for evaluation, then they may be limited in terms of the data available for later analysis. Further, evaluation is not the end of the process, and it is important to use the results to inform future decisions.

The Opportunity

Preparing for evaluation early in the HSIP process can help to save time and enhance the reliability of evaluation results in the long-term. Figure 2 identifies opportunities to document details and prepare for evaluation throughout the HSIP process.

Consider evaluation needs during the planning and implementation stages.

1. Planning

- Develop an analysis and evaluation plan.
- Document the method for selecting the project, funding sources, and high-level project details such as crash history, traffic volume, and underlying crash contributing factors in the before period.
- Timeline: Preconstruction

Communication (Feedback Loop)

- Communicate evaluation results to program managers, project managers, and decision-makers to help inform future decisions.
- Timeline: Post-evaluation

2. Implementation

- Refine the analysis and evaluation plan.
- Document project details including countermeasures, specific treated locations, implementation dates, and final costs.
- Timeline: During construction

3. Evaluation

- Document crashes and traffic volume in the after period.
- Evaluate the safety effectiveness of projects, countermeasures, and programs.
- Timeline: Post-construction

Figure 2. Chart. Illustration of HSIP process and evaluation timeline.

An evaluation plan is important to provide guidance and direction to the evaluation process. Agencies can develop an evaluation plan at any stage of the HSIP process, but there is a benefit to developing the plan during the planning stage. By considering evaluation during the planning and implementation stages, agencies will be prepared for a proactive rather than reactive evaluation process. This allows the agency to collect additional details as needed before and during implementation rather than relying on data that is generally available.

An evaluation plan should address the following questions:[1,4]

- What will you measure?
- How will you measure it (and when)?
- How will you analyze the data?
- How will you use the results?

What Will You Measure?

Identify the measure(s) of effectiveness (MOE), which should relate directly to the objective of the project, countermeasure, or program. The objective of the HSIP is to reduce fatalities and serious injuries on all public roads, so HSIP evaluations should capture at least these MOEs. Applications for HSIP funding and results from safety studies can help to identify other target crash types and project objectives.

How Will You Measure It (And When)?

Determine data requirements and how to gather the information needed to make the measurement. Considerations include where and when to collect the data, how much data to collect, and what procedures to use to collect the data. It is useful to develop systematic procedures and templates to improve consistency in data collected by multiple people. If MOEs include non-crash-based measures (e.g., changes in driver behavior or speed), then it is critical to plan and collect these measures before implementation because these data elements are not often readily available in databases.

How Will You Analyze the Data?

Determine the appropriate evaluation method based on the question at hand. If the objective is to simply determine if safety improved at treated locations, then the analyst may select a less rigorous method. If the objective is to develop agency-specific CMFs for use in future countermeasure selection and economic analysis, then agencies should employ more statistically rigorous methods. Another consideration related to this question is, who will conduct the analysis? It is important to determine if the skills exist in-house or if it will be necessary to contract with a private consultant or local college or university.

How Will You Use the Results?

Evaluation is not the end of the process. While evaluation is the third component of the HSIP, there is a need to communicate evaluation results back to program managers, project managers, and decision-makers to help inform future decisions. The evaluation plan should identify the intended reporting mechanism, stakeholders to receive the results, and who is responsible for disseminating the results. Examples of reporting mechanisms are progress reports, briefings, and meetings. Stakeholders may include safety, planning, design, and operations staff at various levels from analysts to executives. In decentralized States, there is also a need to communicate results to the districts and regions.

2.2 HSIP STAFF AND MANAGEMENT SUPPORT

HSIP evaluation requires dedicated staff time and funding, which requires support from HSIP staff and management. HSIP staff and management should recognize the value of conducting HSIP evaluations and using the results to inform decisions. Management should then allocate appropriate staff time and funding to support evaluation.

The Challenge

HSIP staff may not seek additional funds and staff to support HSIP evaluation if they do not recognize the value added. Further, management may hesitate to shift or allocate funds and staff to HSIP evaluation if they do not recognize the value of conducting HSIP evaluations and using the results to inform decisions.

The Opportunity

There is an opportunity to describe the value of HSIP evaluation to HSIP staff and management. In general, HSIP evaluation results demonstrate the return on investments. For example, evaluation results can demonstrate whether an agency should continue efforts that are cost-beneficial or perhaps they might indicate the agency should strengthen or address efforts that are not providing the expected return on investment.

To describe the value of HSIP evaluation to current management, request a meeting to explain the State's HSIP process and the value of the evaluation component. Be prepared to discuss specific needs (e.g., staff time, funding, and tools) required to establish and maintain an HSIP evaluation program. The following is an example to help explain the value of HSIP evaluation.

The North Carolina Department of Transportation (DOT) Traffic Safety Unit conducts safety evaluations for all completed safety projects, select countermeasures, and the safety program as a whole. For the Spot Safety program, NCDOT calculated a 14:1 benefit-cost ratio across 600 projects within the program. By institutionalizing HSIP evaluation, safety staff have been successful in retaining safety project funding and even obtaining increases in funding. The safety evaluations have also impacted other areas of NCDOT operations, leading to increased confidence in investments.

It is easy to understand the urge to spend more on implementation and less on evaluation; more projects provide more opportunity to improve safety and address known contributing factors. The problem with this approach is that agencies may not implement the most effective strategies if they do not perform evaluations to understand the effects. For example, consider the decision to install more overhead flashing beacons at two-way stop-controlled intersections. Research shows that installing flashing beacons at two-way stop-controlled intersections can reduce angle crashes in rural areas; however, this strategy may increase angle crashes in urban areas.[5] Without the appropriate evaluation, agencies may continue to install flashing beacons at urban stop-controlled intersections, when another strategy may increase the return on investment.

Once there is support from HSIP staff and management, it is helpful to develop, document, and adopt an HSIP evaluation policy to establish the frequency and quality of HSIP evaluations. This will help to ensure the long-term success and continuation of HSIP evaluation, particularly as there is turnover in HSIP staff and management. There is also an opportunity to integrate the data requirements for evaluation into the State's Safety Data Business Plan as a means of institutionalizing the process.

2.3 STAFFING AVAILABILITY AND EXPERTISE

Agencies should consider the availability and expertise of in-house staff to complete HSIP evaluations or determine if it would be more appropriate to seek additional support from outside resources.

The Challenge

Even with HSIP staff and management support, there are still limitations with respect to the time and budget to conduct HSIP evaluations. There are also potential limitations related to in-house staff availability and technical expertise.

The Opportunity

With respect to in-house staff availability and budget allocation, a program manager may be faced with the following decision: how much time and money should be allocated to each component of the HSIP? During the planning stage of the HSIP, States identify sites with promise and develop projects to address the target crash types and contributing factors.

The North Carolina DOT estimates an average time of 16 to 24 hours per individual project evaluation. This includes a simple before-after analysis and comparison of target crashes using collision diagrams. This estimate is based on evaluations of more than 400 locations among five different staff members.

For implementation, States construct projects to address target crash types and contributing factors identified during the planning component. For evaluation, States estimate the crash frequency and severity reduction of individual projects, countermeasures, and the HSIP. To ensure adequate resources for HSIP evaluation, this should be one of the first considerations in allocating the HSIP budget. First, agencies should know the time and cost required to screen the network, evaluate the HSIP program as a whole, and prepare the annual HSIP report. These are relatively fixed annual expenses that do not change substantially from year to year in each State. Next, agencies should estimate the average time and cost to evaluate an individual HSIP project. For each proposed HSIP project, the agency should include average evaluation costs in the total estimated project cost to cover a subsequent evaluation. The manager can then balance the budget and determine the number of projects for implementation.

A related opportunity is to build relationships with key stakeholders and rely on universities, local technical assistance programs (LTAPs), and consultants when there is a need for additional support. This can help to overcome challenges related to limited availability of staff and lack of in-house expertise. The following are several examples of agency partnerships to support HSIP evaluation.

The Florida DOT collaborates with universities on HSIP evaluations, relying on the LTAP to assist with contractual arrangements.

The Kentucky Transportation Cabinet (KYTC) collaborates with the Kentucky Transportation Center (KTC) at the University of Kentucky to support HSIP efforts. While KYTC staff administer the accounting, designing, letting, and construction consulting on current HSIP Projects, they have limited in-house capacity and capability to develop CMFs and delve deeper into specific statistical analyses. KYTC established an agreement with KTC that works like an on-call contract. KTC provides additional staff and statistical expertise to support the HSIP evaluation process.

The Wisconsin Traffic Operations and Safety (TOPS) Laboratory, at the University of Wisconsin, supports the Wisconsin DOT by evaluating HSIP projects. In the past, the Wisconsin DOT Division of Transportation Investment Management provided the TOPS Laboratory with a list of completed HSIP projects. The TOPS Laboratory then performed statistically rigorous before-after evaluations and economic analysis to estimate the safety effectiveness and benefit-cost ratio of individual projects. The TOPS Laboratory also developed a process to extract appropriate crashes (by location, type and year) from the Wisconsin DOT crash database based on the project locations and project start and completion dates. Now, the Wisconsin DOT is developing in-house expertise to perform evaluations.

2.4 RIGHT-SIZING EVALUATIONS

Evaluation does not have to be overly complicated or expensive. There is a need to "right-size" the evaluation to meet the needs within a given budget.

The Challenge

If HSIP staff or management perceive evaluation as a time-consuming and expensive process, then they may avoid HSIP evaluations altogether. Further, if HSIP staff assume all evaluations require the use of the same methods, then they may spend more time than needed on a few select projects when they could have used the evaluation budget to include more evaluations with less time-consuming methods.

The Opportunity

There is an opportunity to match the type and reliability of the procedures with the needs of the specific evaluation. For example, if the goal is to estimate the individual project effects, then it may be sufficient to use a simple procedure. However, the effects of a single project are

unreliable for estimating the expected safety effects of similar future projects. If the goal is to estimate the expected safety effectiveness of a countermeasure, then there is a need to employ more reliable safety evaluation procedures and data from several similar projects.

The Montana DOT established HSIP evaluation guidelines that define the scope of HSIP evaluations. Specifically, Montana has elected to evaluate the HSIP based on groups of similar projects on an annual basis. The evaluations focus on projects with a cost exceeding $100,000. As the evaluation program matures, they plan to consider evaluation of lower cost projects (< $100,000).

2.5 GUIDANCE AND TOOLS

There is a benefit for States to establish HSIP guidance and provide appropriate tools to support HSIP evaluation.

The Challenge

Turnover in agency staff presents a challenge to the consistency of existing programs and procedures, including HSIP evaluation. Further, it is difficult to perform HSIP evaluations efficiently without the proper tools.

The Opportunity

There is an opportunity to develop State-specific HSIP guidance to maintain consistency in existing programs and procedures. The guidance will serve to document the State-specific processes and procedures for each component of the HSIP, including evaluation. When developing State-specific guidance, agencies should keep it simple and focus on the application of procedures.

The Montana DOT established an HSIP evaluation process with guidelines in response to a program assessment. Montana's HSIP Evaluation Guidelines describe evaluation methodologies, potential sources of bias, and data requirements. The following project groups guide the evaluation:

- *Geometric improvements at a specific location (e.g., curve realignment or shoulder widening).*
- *Slope flattening or elimination of roadside hazards.*
- *Signing, striping and delineation including the installation of warning flashers.*
- *Installation of guardrail.*
- *Intersection improvements. This may require further grouping depending on the types of projects completed in the evaluation year (e.g., separate the evaluation of geometric improvements such as turn lanes from traffic signals and roundabouts).*
- *Installation of rumble strips (centerline and shoulder).*

> *While the evaluation focuses on groups of projects, analysts evaluate crash data for each of the locations individually and then prepare a summary based on the project grouping. The guidelines provide a project evaluation template, which is summarized in a memorandum for distributing evaluation results to the Chief Engineer, the Preconstruction Engineer, the Traffic and Safety Engineer, the Traffic Operations Engineer, the Traffic Design Engineer, the Traffic Safety Engineer, personnel in the Safety Engineering Section, and others in the Department who may have an interest in highway traffic safety or the results of the HSIP projects.*

The following are links to examples of HSIP manuals and guidelines from other States that include specific information on HSIP evaluation:

- [California Local Road Safety Manual](6)
- [Pennsylvania Pub 638 (District Highway Safety Guidance Manual)](7)
- [Utah HSIP Manual](8)
- [Virginia HSIP Guidelines](9)

Another opportunity to improve the consistency and efficiency of HSIP evaluations is to provide project tracking, evaluation, and reporting tools. Agencies are more frequently using web-based project tracking and data visualization software. These tools allow staff and partners from districts, regions, and local agencies (i.e., non-central agency partners) to transmit or upload information for project or program-wide evaluation in a timely and consistent manner. Chapters 3 through 6 provide specific examples of web-based software applications to support project tracking, evaluation, and reporting.

2.6 APPLICATION OF EVALUATION RESULTS

There is an opportunity for agencies to apply lessons learned from evaluations. For example, an agency may learn about the safety performance of strategies in their State-specific context and opportunities to reduce crash potential in future designs and operations. There is also an opportunity to expand evaluations beyond the typical analysis of project effectiveness to include process evaluations.

The Challenge

Some may consider analysis and evaluation as the final step in the process with results indicating the success or failure of past efforts. This may discourage agencies from conducting analysis and evaluations if they are worried about the potential for unexpected results (i.e., efforts were not as promising as they had planned) or less than optimal projects or programs in reducing crashes. Further, solely focusing on the safety effectiveness of projects and programs misses the opportunity to identify challenges related to the underlying processes and procedures that, if addressed, could benefit other safety-related activities.

The Opportunity

There is an opportunity to learn from the results of any evaluation, and even unexpected results are useful to inform future decisions. Specifically, evaluation results identify future opportunities for investment of similar strategies and opportunities for improvement. For example, if an agency identifies a project that is not meeting safety performance expectations based on the evaluation results, then there is an opportunity to address the situation as appropriate for the location (e.g., remove the countermeasure or install supplemental countermeasures).

> *There is an opportunity to learn from the results of any evaluation; even unexpected results inform future decisions.*

> *Wisconsin DOT evaluated 19 HSIP projects completed in fiscal year 2006. The results indicated a reduction in crashes at 15 of the 19 locations, and an increase in crashes at four locations. The agency was able to further investigate the locations with increases in crashes based on the results of the safety evaluation. Further, the report noted that while the results of the crash and benefit-cost analyses indicated positive results (i.e., crash reductions) at nearly all project locations, the cost of reducing these crashes was much higher at some locations than others. As a result, there is an opportunity to develop a database of evaluation results to use as a future reference for comparing the expected safety and cost-effectiveness of contemplated countermeasures in Wisconsin.*

Evaluation should also focus on more than project effectiveness. By evaluating the HSIP process from start to finish, there is an opportunity to improve planning, implementation, evaluation, and documentation processes and decisions.

> *The Alaska Department of Transportation & Public Facilities evaluates all HSIP projects statewide. Headquarters staff compares the difference between planned and actual benefit-cost ratios. They use the information to adjust effectiveness estimates for future planning and identify types of projects that tend to overrun costs and that do not produce proposed benefit-cost ratios.*

> *The Pennsylvania DOT designates certain sections of roadway as "highway safety corridors". Pennsylvania regulations (67 Pa. Code § 214.2) define highway safety corridors as "the portion of a highway determined by a traffic study to be targeted for the application of signs, increased levels of enforcement, and increased penalties specifically for the purpose of eliminating or reducing unsafe driving behaviors that are known to result in crashes and fatalities." Once designated a highway safety corridor, engineers evaluate the safety performance of the corridor, focusing on targeted crashes. For example, if enforcement targets aggressive or unbelted driving, then the analysis focuses on these factors in the evaluation. Staff from the Department's engineering districts then meet with the respective law enforcement officers to share the evaluation results and determine if they need to make program adjustments.*

2.7 SUMMARY OF PREPARING FOR HSIP EVALUATION

The following is a summary of general HSIP evaluation considerations, highlighting tips and tricks to prepare for successful evaluation, including the following:

- Promote the value of evaluation and share success stories. For example, explain how other States are using evaluation results to justify funding and support safety projects, especially when particular countermeasures receive pushback from the public or local businesses (e.g., rumble strips or non-traversable medians).

- Build relationships with stakeholders. There is an opportunity to rely on universities, LTAPs, and consultants when in-house resources are not available.

- Provide State-specific guidance, tools, and resources. When developing State-specific guidance, keep it simple and focus on the application of methods.

- Share information outside of HSIP staff. The results of evaluation (project, countermeasure, and program level) are useful to inform future decisions. These decisions go beyond the safety program. By sharing this information, there is an opportunity to generate support and additional funding for continued evaluations.

2.8 GENERAL HSIP EVALUATION RESOURCES

The following resources provide additional information on HSIP evaluation challenges and opportunities to overcome those challenges.

- 2010 HSIP Manual provides an overview of the HSIP and offers practitioners with a review of standards, new and emerging technologies, and noteworthy practices for each step in the HSIP process.[1]

- Developing an Effective Evaluation Plan provides guidance for developing a living "Evaluation Plan," helping program administrators identify answers to three questions about their program: "What?", "How?", and "Why it matters?".[4]

- The following are examples of State-specific HSIP manuals and guidelines:
    o California Local Road Safety Manual[6]
    o Pennsylvania Pub 638 (District Highway Safety Guidance Manual)[7]
    o Utah HSIP Manual[8]
    o Virginia HSIP Guidelines[9]

- HSIP Assessment Toolbox provides support for agencies to conduct an assessment of their HSIP. This toolbox provides strategies, methods, and best-practices for agencies to consider incorporating into their program.[10]

- HSIP National Scan Tour Report provides a summary of notable practices in the areas of HSIP administration, planning, implementation, and evaluation.[11]

- HSIP Noteworthy Practice Series provides examples from around the United States of best-practices for various aspects of the HSIP process. It provides a series of case studies with noteworthy examples for other agencies to consider incorporating into their programs.[12]

- HSIP Self-Assessment Tool provides a question-based method for managers to perform a self-evaluation of an agency's HSIP, which helps to identify and address opportunities to improve current programs and processes.[13]

CHAPTER 3: PROJECT TRACKING

Project tracking provides the foundation for all HSIP evaluations. With detailed project information, agencies can readily conduct project-level evaluations (chapter 4), countermeasure evaluations (chapter 5), or program-level evaluations (chapter 6). This section describes general project types and practices to monitor and track individual projects, including the tracking timeline, relevant project data, use of templates, and opportunities to engage stakeholders to support the process.

Chapter 3 at a Glance
3.1 Project Types
3.2 Project Tracking Timeline
3.3 Project Data
3.4 Project Tracking Templates
3.5 Stakeholder Support
3.6 Summary
3.7 Resources

Project tracking provides the foundation for all HSIP evaluations.

3.1 PROJECT TYPES

The HSIP is a systematic (i.e., repeatable) safety management process to identify locations and implement projects with the goal of reducing fatalities and serious injuries. In general, there are two approaches to implementing safety projects: crash-based and systemic.

Crash-based refers to the selection and treatment of sites based on site-specific crash frequency and severity.

Systemic refers to the selection and treatment of sites based on site-specific geometric and operational attributes known to increase crash potential.

The crash-based approach focuses on selecting and treating sites based on site-specific crashes. In the crash-based approach, analysts first identify sites based on site-specific, crash-based performance measures. For example, agencies may conduct crash-based network screening to identify candidate locations for safety projects with the highest frequency of crashes, highest potential for safety improvement, or highest crash severity change. Diagnostic analyses serve to hone in on what actions and behaviors are leading to crashes at each individual site identified during network screening. Based on the site-specific collision patterns and crash contributing factors, agencies develop and implement appropriate countermeasures to mitigate the contributing factors at each site. For example, an agency may identify a specific rural, two-way stop-controlled intersection for further review through network screening based on a history of right-angle crashes and the potential for safety improvement. After a detailed diagnosis of the intersection, the agency identified limited intersection sight distance and speeding as factors contributing to the relatively high number of angle crashes. To target these crash contributing factors, the agency proposed several potential strategies (e.g., improving intersection sight distance, converting the two-way stop to an all-way stop, and converting the two-way stop to a roundabout), and then determined the all-way stop would be the most appropriate and cost-effective option for this location.

The systemic approach focuses on selecting and treating sites based on site-specific geometric and operational attributes known to increase crash potential. The first step in the systemic approach is to select focus crash type(s), facility types, and contributing factors. Using crash or severity potential as a guide, the next step is to identify sites with these specific geometric and operational characteristics as candidate locations for potential safety improvement. Given the list of potential contributing factors for the focus crash type(s), an agency can develop targeted countermeasures to address or mitigate the specific contributing factors at the specific locations across the focus facility type. For example, consider a scenario where an agency identified head-on crashes as a focus crash type based on the number of fatal and serious injury crashes. They noted these crashes were most prevalent on rural, two-lane roads and selected this as the focus facility type for head-on crashes. The agency reviewed the data for all head-on crashes on rural, two-lane roads and determined that common roadway features (potential contributing factors) include narrow cross-section, numerous horizontal curves, and high-risk passing. Alternatives to address the underlying contributing factors from an infrastructure perspective might include installing centerline rumble strips, widening the cross-section, or adding a passing lane. The agency deemed the latter two options as not effective in returning benefit for the investment for wide-scale deployment. As such, they selected centerline rumble strips as an appropriate measure to address the underlying contributing factors.

The crash-based and systemic approaches are complementary and support a comprehensive approach to safety management. The primary difference is the way in which analysts identify opportunities and develop projects in the planning component of the HSIP. These differences should be considered in the tracking and evaluation of crash-based and systemic projects.

The crash-based and systemic approaches are complementary and support a comprehensive approach to safety management.

3.2 PROJECT TRACKING TIMELINE

Project tracking follows the project development process from planning and programming through implementation and operations. Figure 3 provides an illustration of the project development process, indicating relevant points for project tracking.

Project tracking should begin with the planning and programming stage as analysts identify locations for potential improvement. By starting project tracking early in the project development process, agencies can take advantage of early diagnosis efforts, documenting site conditions and crash history prior to any improvements. This will expedite future evaluations.

Planning and Environmental
- Document the project location
- Identify the approach used to identify potential projects
- Document crash contributing factors, target crashes, crash history, and related programs

Right-of-Way and Design
- Document ROW and design costs
- Document site conditions, traffic volume, and project details (e.g., type and number of countermeasures)

Maintenance and Operations
- Document maintenance costs
- Document after period data (crashes and traffic volume)
- Evaluate the safety effectiveness of projects, countermeasures, and programs

Programming
- Determine if programmed projects are still needed
- Document the preliminary cost estimate
- Document the funding source(s) and amount(s)

Construction
- Document construction dates
- Document actual project costs
- Document change orders and modify project details
- Take photos of completed work
- Evaluate on-time and on-budget delivery

Figure 3. Chart. Project tracking in relation to the project development process.

Project tracking is not exclusive to HSIP evaluation efforts, and there are opportunities to piggyback on existing project delivery or project management systems. Several agencies have systems in place to help manage projects from start to finish, and some have customized their system to facilitate evaluations.

Virginia uses Tableau to track all safety projects, including HSIP projects, bike and pedestrian safety program projects, and highway-rail grade crossing safety program projects. Tableau is an enterprise business intelligence and data visualization software. The full software allows the district project managers to populate and update the project database. District project managers use the software to track project schedules and budgets, which is especially useful when the district engineers meet with the State Safety Engineer to discuss on-time and on-budget delivery. The central office staff use the software to track and evaluate projects as well as to track and follow up with slow-responders.

If your agency does not have a project management system in place, or it does not have the desired capabilities for HSIP evaluation, there is an opportunity to start from scratch. In this case, it may be necessary to start small and build to something more sophisticated over time. A simple spreadsheet can serve as an effective mechanism for documenting project details and tracking performance over time.

Alaska developed spreadsheet templates to collect project-level details and evaluation results for completed projects. The central office provides the templates to the regions. The region engineers then enter the project details as well as before and after crash data for each project and submit the completed worksheets to the central office annually. Refer to Appendix A in Alaska's HSIP Handbook for HSIP project tracking and evaluation templates.[14]

3.3 PROJECT DATA

Depending on the level of HSIP evaluation, analysts will need certain data (i.e., required data) while other data will help to enhance the evaluation (i.e., desired data).

Required Data

Required data include project location, construction dates, countermeasure details, project cost and crash data. A brief description of each follows, with an expanded list and description of individual data elements to consider for project tracking located in Appendix A.

Project Location

Agencies should specify the project location using a reference system for linking to other databases such as crash, roadway, and traffic data. Some agencies use X-Y coordinates for geospatial locating. Others use a route-milepost reference system.

Improvements at a single location: identifying the location of improvements at a single location is relatively straightforward. For example, an agency may identify the location of an intersection project by the X-Y coordinates of the intersection as well as the major road and cross street names. For a curve improvement project, the agency may identify the location by the route number and beginning and ending milepost of the curve.

Improvements at multiple locations: identifying the location of multiple improvements within a single project is more complicated. This is often the case for systemic projects and other projects that include similar improvements at multiple locations along a corridor or within a jurisdiction. For example, an agency may undertake a project to install retroreflective backplates on all signal heads within a given city. While the project may be completed under a single contract, it could include dozens of intersections. For this example, it would be important for the agency to identify the location of each intersection included in the project and preferably the number of retroreflective backplates installed at each location. As another example, an agency may let multiple contracts to enhance curve signing along multiple routes in multiple jurisdictions. For this example, it would be important for the agency to identify the jurisdiction, route, and beginning and ending milepost for each curve included in each project, and preferably the number of enhanced signs within each curve. It would also be useful for the agency to indicate that each of the projects supported a systemic effort to enhance curve signing. In general, to facilitate tracking for projects with multiple locations, the tracking mechanism should allow for the identification of specific locations and treatment details at the multiple treated locations.

Construction Dates

Agencies should specify the dates of construction as precise as possible, including the beginning of construction, end of construction, and date on which the facility was open to the public. These dates are critical to defining the periods before, during, and after construction. At a minimum, the dates should include the year, and preferably the month as well.

Countermeasure Type and Details

Agencies should define the specific countermeasure(s) with sufficient detail to facilitate the desired level of evaluation. Refer to FHWA's [HSIP Reporting Guidance](#) for a list of HSIP project categories and subcategories.[15] This list represents the minimum level of detail required for annual HSIP reporting, and hence for project tracking. For countermeasure evaluations, analysts need to group similar projects, which may require additional details such as the number of treated intersection approaches or type of signal improvements rather than a general countermeasure description such as "intersection improvement" or "traffic signal improvement." It is also important to know if the project represents a single countermeasure or multiple countermeasures. A lack of details can limit the usefulness of the project data and evaluation results. For example, consider a countermeasure listed as "modify skew angle." Without further details, it is difficult to understand the extent of improvements and the applicability of results. With additional information such as the number of improved approaches and the change in skew (e.g., change from 15 degrees to 0 degrees or change from 45 degrees to 0 degrees), the results are more useful for future decisions.

Project Cost

Agencies should enter the proposed project costs during the planning and programming stage, and then enter the actual project costs after construction. It is useful to document costs separately for preliminary engineering, right-of-way, construction, and maintenance.

Crash Data

Agencies should enter the relevant crash data before (and possibly after) implementation. During the planning stage, there is often an assessment of crash history to identify locations for improvement or to diagnose the site-specific safety opportunities. The project tracking database should retain this information for future evaluations. Relevant crash data may include the count of crashes by type and severity for each year of the before period. It is useful to track both total and target crashes (i.e., the crashes targeted by the proposed project) separately.

Desired Data

Additional information can help to facilitate more detailed evaluations and link projects with specific efforts. Desired data elements include funding source and amount, relation to SHSP emphasis areas and safety programs, information from pre-construction safety analyses and photos. A brief description of each desired elements follows.

Funding Source and Amount

Agencies should indicate the funding source and amount. States currently report the cost with respect to HSIP funding; however, it is useful to document the project costs by amount and funding source if joint-funded. It is also useful to track project details for non-safety projects. This can facilitate more robust countermeasure evaluations by pooling HSIP and non-HSIP projects. The following are potential categories to designate funding source:

New York uses their Post-Implementation Evaluation System (PIES) to track project details for both HSIP and non-HSIP funded projects. PIES includes data for capital improvements constructed as part of major highway projects as well as low cost improvements (e.g., pavement markings, signs, and signal improvements) implemented through maintenance efforts.

Tennessee DOT uses their Program Project and Resource Management (PPRM) database to track all projects, providing a one-stop-shop for users to manage and evaluate projects.

- Highway Safety Improvement Program (23 U.S.C. 148).
- High Risk Rural Roads Special Rule.
- Penalty Transfer Funds (23 U.S.C. 154 and 164).
- Other Federal-aid Funds (Surface Transportation Block Grant Program, National Highway Performance Program).
- State and Local Funds.

Relation to SHSP Emphasis Areas and Safety Programs

Agencies should indicate the emphasis area(s) and safety programs related to the project. Relating projects to SHSP emphasis areas will help with program evaluation and in updating the State SHSP. Relating projects to specific safety programs (e.g., roadway departure, intersection, pedestrian, bicycle) allows for individual program evaluation and can help to justify future funding of these efforts. Projects may relate to more than one emphasis area. For example, a project at an intersection may relate to both intersection and pedestrian emphasis areas.

Information from Pre-Construction Safety Performance Analyses

Agencies should identify the type of approach employed to identify the project (i.e., crash-based or systemic). This information is also useful for identifying the target crash type based on the pre-construction safety diagnosis.

Photos

Agencies should include photos before and after implementation to document the specific conditions in the before and after periods. This can help to understand the extent of the improvements. There is an opportunity to use resources such as Google Earth™ or other similar applications to verify pre- and post-implementation conditions; however, there is no guarantee the timeframe or resolution of these resources will meet the evaluation needs.

North Carolina uses photos to document conditions before and after implementation, noting that this helps to document the 'end of construction' or 'substantial completion' rather than waiting for project closeout when there are backlogs in inspections due to limited staff.

Data Collection Guidelines

Agencies should develop tracking mechanisms to capture at least the required data elements. It is also useful to document and explain the data elements in State-specific HSIP guidelines or manuals. Refer to the following section for further discussion of project tracking templates and refer to Appendix A for a detailed list of data elements to consider for project tracking.

California's Local Roadway Safety Manual provides details related to the type of project data needed for HSIP project evaluation.[6] At a minimum, the manual recommends that local agencies collect the following project data to prepare for evaluation of local HSIP projects: project location, countermeasure type, date of installation, three to five years of crashes before and after implementation, and duration and severity of the crash summaries before and after implementation. The manual notes that an effective project tracking system facilitates project evaluation, which, in turn, informs the practitioner on the effectiveness of past improvements and provides quantitative data to help justify the value of continuing and expanding the local agency's safety program in the future.

3.4 PROJECT TRACKING TEMPLATES

A project tracking template may help to improve consistency among data entries. As applicable, the template could include dropdown menus with predefined categories. Predefined categories ensure consistency among data entries. One challenge with predefined categories is the lack of flexibility if there is a need to deviate from the standardized list.

> *The Alaska Department of Transportation & Public Facilities developed a standard project tracking template and established a list of improvement type codes based on eligible HSIP project types. The improvement type codes help to improve consistency in project reporting and allow analysts to filter by the project type of interest when grouping similar projects for countermeasure or program evaluation. Table 1 shows a modified version of Alaska's project tracking template, including an explanation of the expected data entry for improvement type. Refer to Appendix A in Alaska's HSIP Handbook for the complete HSIP project tracking template and additional details.[14]*

Table 1. Sample of Alaska project tracking template with hypothetical example.

Region	Location	FHWA Road Functional Classification	Improvement Type[1]	Total Project Cost	Before and Interim Crash Data	After Crash Data	B/C Ratio
Southcoast	Seward Highway (MP X to Y)	Interstate	Skid-resistant surface	$125,000	15	3	7.97

1. Improvement type: base entry on descriptions from page A-11 of the Alaska HSIP Handbook. If project includes multiple improvement types, list the predominant category. Report project categories related to railway-highway grade crossing safety improvements separately using the form under Tab 130 Eff.

3.5 STAKEHOLDER SUPPORT

A common challenge related to HSIP evaluation is limited time and resources. Many agencies do not have a dedicated staff focused solely on project evaluations. Instead, evaluations become one of many responsibilities for HSIP staff. To minimize the burden on any individual, there is an opportunity to enlist district and regional staff as well as local partners to support project tracking and reporting. With proper training and tools, these stakeholders can help to define project limits (begin and end mileposts), identify relevant construction dates (begin, end, and open to public), and provide detailed project information (project cost and type of improvement).

To track project-level details, agencies should strengthen communication between project managers and construction managers and between the central office and any district or regional offices. In decentralized States, it may be necessary to establish one point of contact in the central office to provide templates and answer questions. It may also be useful to identify a

project manager in each local, district, and regional office as the point of contact for HSIP efforts, including evaluation. It is also useful to establish a SharePoint site or central repository for information sharing and templates.

> *In Alaska, the central office developed spreadsheet templates for project tracking and evaluation. Data for tracking the effectiveness of HSIP projects comes from the regional offices. Specifically, the regions prepare an HSIP Project Evaluation Spreadsheet to compile project-level details for each completed project with three years of after data. The spreadsheet provides an overview of project performance, including the benefit-cost ratio based on both the construction and maintenance costs over the analysis period. The spreadsheet annualizes crash data for comparison of before and after periods, which is particularly useful if the before and after periods are different duration. The regional offices are responsible for collecting post-project crash data, entering project data in the spreadsheet, and submitting the spreadsheets to headquarters. Headquarters aggregates the individual project data into a master spreadsheet to evaluate the effectiveness of countermeasures and the entire HSIP program, track the frequency of implemented countermeasures, and provide an historical listing of completed projects. The benefits of Alaska's project tracking and evaluation templates reach beyond a streamlined reporting system, allowing the regions to maintain control over the projects from start to finish, and allowing the central office to focus on other aspects of the HSIP. The process has also resulted in improved funding allocation, countermeasure identification, project evaluation, and HSIP reporting. Refer to [Alaska's HSIP Handbook](#) for additional discussion of the process and the HSIP project tracking and evaluation templates.[14]*

> *The Florida DOT uses a central repository called the Crash Reduction Analysis System Hub (CRASH) to track safety projects. Districts submit information through work orders via a SharePoint site to populate a safety improvement database. This facilitates project, countermeasure, and program evaluation described in chapters 4, 5, and 6, respectively. Refer to [Florida's CRASH Portal](#) for details.*

3.6 SUMMARY OF PROJECT TRACKING

The following is a summary of project tracking, highlighting tips and tricks to prepare for successful evaluation:

- Establish a project-tracking database. This critical element feeds evaluations at the project, countermeasure, and program levels. It also helps to improve consistency in reporting. There are opportunities to start small or piggyback on existing systems.

- Establish a list of improvement type codes. This may improve consistency in project reporting and help to identify the specific type of project. This is useful when grouping similar projects for countermeasure or program evaluation.

- Track and confirm projects with photos to document the type of countermeasures.

- Tie projects to the specific project type (i.e., crash-based or systemic), funding sources, and SHSP emphasis areas. This helps with program evaluation. This will also help in updating the State SHSP. The underlying methodology and project-level indicators should be embedded in the State's HSIP manual or guidelines.

- Establish a SharePoint site or other web-based portal to serve as a central repository for HSIP project information and resources. This can facilitate communication and data sharing, particularly in a decentralized agency.

- Enlist district/region staff to support the project tracking and reporting process. With proper training and tools, the district/region staff can help to identify implementation dates, define the beginning and end of project limits, and submit project information.

3.7 PROJECT TRACKING RESOURCES

The following resources provide additional information on project tracking:

- Alaska's HSIP Handbook provides HSIP project tracking and evaluation templates.[14]

- California's Local Roadway Safety Manual provides details related to the type of project data needed for HSIP project evaluation.[6]

- FHWA's HSIP Reporting Guidance provides a list of HSIP project categories and subcategories, representing the minimum level of detail required for annual HSIP reporting, and hence for project tracking.[15]

- Florida's CRASH Portal provides an example of a SharePoint site for information sharing and templates.

CHAPTER 4: PROJECT EVALUATION

The HSIP requires States to establish an evaluation process to analyze and assess results achieved by highway safety improvement projects. [23 CFR 924.13(a)(1)] While FHWA does not require the States to report evaluation results for individual project locations, the evaluation of individual projects helps to determine if projects achieved the intended results. Project evaluations also support higher levels of evaluations (e.g., countermeasure and program). This section describes considerations and practices to evaluate individual projects, including measures of effectiveness, potential sources of bias, evaluation methods, sample size and study periods, data requirements, and accessibility and information sharing.

Chapter 4 at a Glance
4.1 Measures of Effectiveness
4.2 Potential Sources of Bias
4.3 Evaluation Methods
4.4 Sample Size and Study Periods
4.5 Data Requirements
4.6 Accessibility/Information Sharing
4.7 Summary
4.8 Resources

Individual project evaluations support countermeasure and program evaluations

4.1 MEASURES OF EFFECTIVENESS

MOEs for individual projects typically include localized safety impacts (e.g., site-specific change in crashes, injuries, and fatalities) and economic measures (economic effectiveness and benefit-cost ratio). Regarding these crash-based performance measures, agencies should evaluate target and correctable crashes in addition to total, injury, and property damage only (PDO) crashes. If a project targets specific contributing factors or crashes, then it is informative to evaluate the change in target crashes as well as the change in crashes by severity and the change in total crashes. For example, if an agency installs cable median barrier to address cross-median and head-on crashes, then target crashes would include cross-median and head-on crashes. In this case, it may be useful to evaluate the cable median barrier project with respect to total, target, fatal plus injury, and PDO crashes because the safety effects of the cable median barrier may be different for each crash type.

It is informative to evaluate the change in target crashes as well as the change in total crashes.

North Carolina DOT conducts individual project evaluations, using collision diagrams and focusing on changes in target crashes to evaluate individual project effectiveness. The North Carolina DOT noted the importance of focusing on the change in target crashes rather than total crashes. Specifically, they noted cases where total crashes increased or remained unchanged while target crashes

decreased. Including target crashes in the evaluation helps to understand if the project achieved the initial objective (i.e., to address a specific crash type or crash contributing factor). If a project does not address the target crashes, then alternative or supplemental countermeasures may be required. If the project achieved a reduction in target crashes, but total crashes increased or remained the same, then other countermeasures may be required to address the other crash types.

Figure 4 and Figure 5 show an example of individual project evaluation from North Carolina using collision diagrams. Figure 4 shows the collision diagram for the before period. Figure 5 shows the collision diagram for the after period. In this example, the target crashes included right-angle and turning-related crashes as indicated by red circles in the collision diagrams; non-target crashes are indicated by black circles. A note indicates that one run-off-road crash is included in the target crashes because a near angle crash resulted in the vehicle swerving and running off the road. To address the target crashes, the North Carolina DOT converted the intersection from two-way stop control to all-way stop control. Before the conversion, there were 21 total crashes and 19 target crashes. After the conversion, there were 8 total crashes and 4 target crashes, resulting in a 62 percent reduction in total crashes and a 79 percent reduction in target crashes.
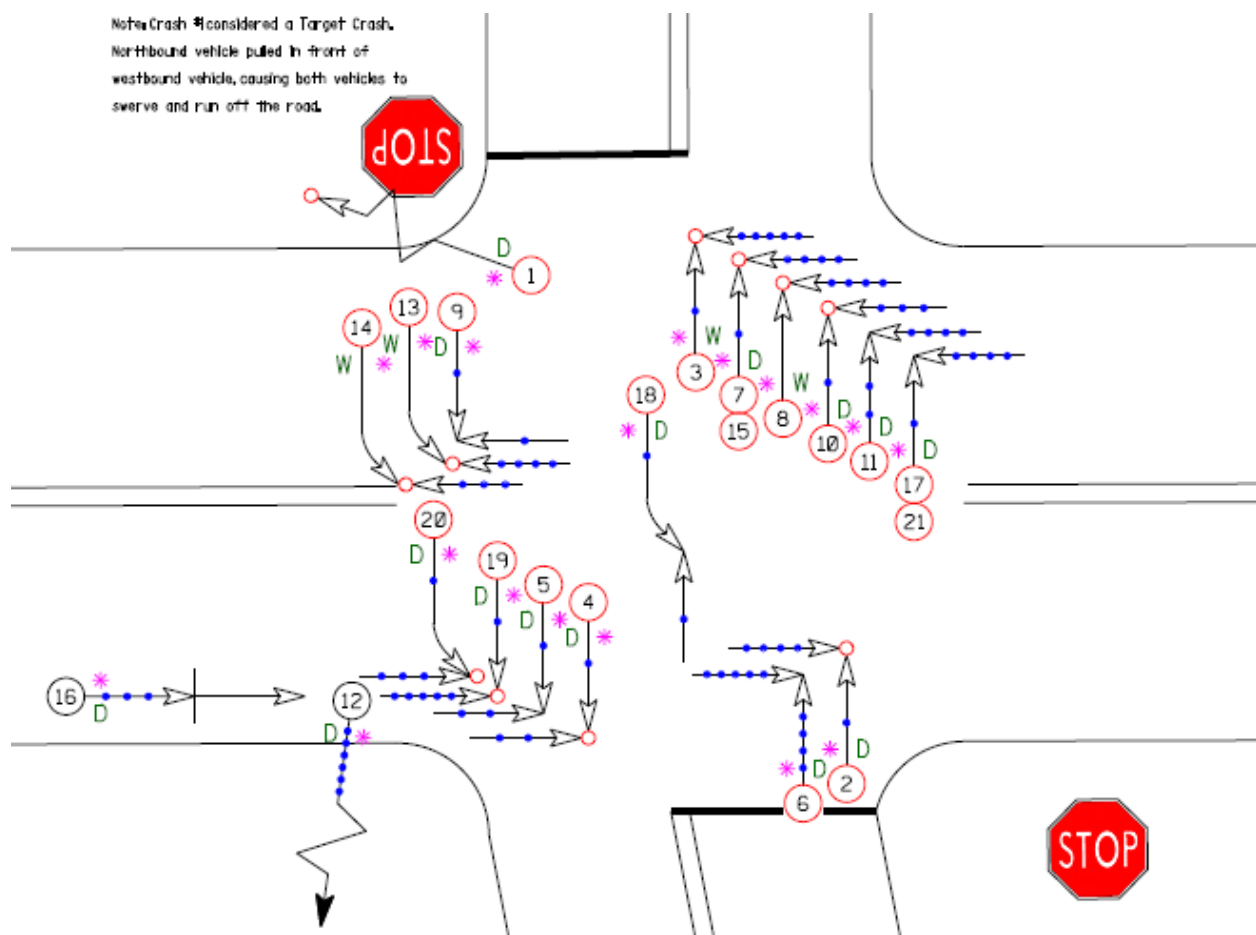
Figure 4. Illustration. Collision diagram before conversion of two-way to all-way stop control.
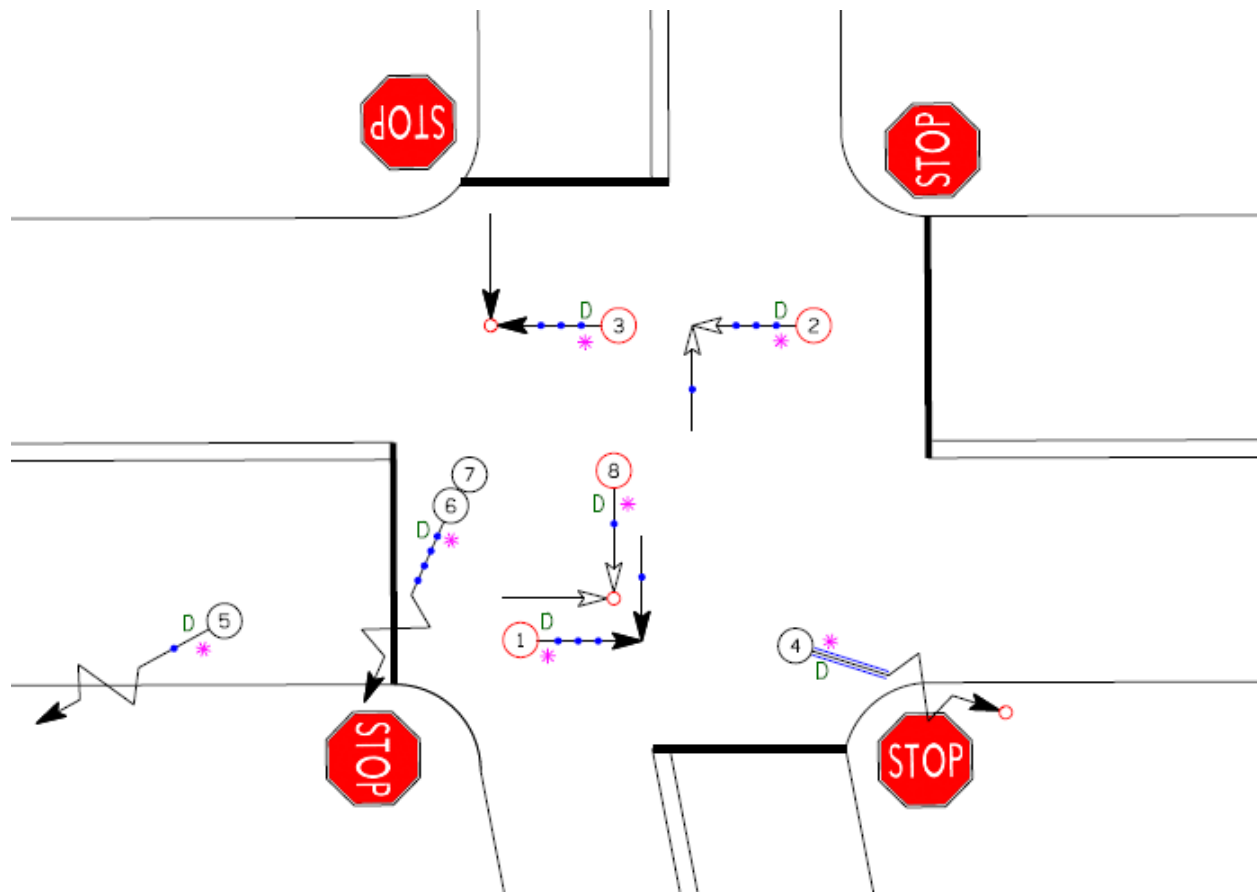
Figure 5. Illustration. Collision diagram after conversion of two-way to all-way stop control.

Wisconsin DOT defines target crash types for each individual project. Target crashes include the primary crash type or types that the safety improvements were intended to address or mitigate. The process of identifying target crashes relies on information from the HSIP project files. For example, a project including the construction of exclusive left-turn lanes may target left-turn and rear-end crashes at an intersection. In this case, the target crashes would be left-turn and rear-end crashes on the approaches with newly-installed left-turn lanes. The Wisconsin DOT also considers the potential increase in non-target crashes. For example, the installation of a traffic signal may target right-angle and turning crashes, but they also consider the potential increase in rear-end crashes separately as a non-target crash type.

The benefit-cost ratio is another potential performance measure. While many agencies use the benefit-cost ratio, it is susceptible to influence from a few severe crashes. For example, if there were two fatal crashes in the before period and no fatal crashes in the after period, then the observed reduction of two fatal crashes can have a substantial impact on the estimated benefit since the value of a fatal crash is on the order of $5 to $9 million while the value of an injury crash is on the order of $50,000 to $250,000 depending on the severity and the value of a PDO crash is on the order of $10,000. As such, the cost of one fatal crash is equivalent to the cost of

approximately 20 or more injury crashes, or 500 or more PDO crashes. If agencies are concerned with the influence from a few severe crashes, then they may consider additional or other performance measures for crash-based project evaluations or consider using weighted crash costs.

4.2 POTENTIAL SOURCES OF BIAS

To accurately estimate safety effectiveness, agencies should consider the approach used to select the project location and the potential for site selection bias and changes over time such as changes in traffic volume and other temporal trends. Site selection bias occurs when agencies select sites with high crash counts for improvement, not at random. When this is the case, there is potential for regression-to-the-mean (RTM). As shown in Figure 6, RTM describes the situation when periods with relatively high crash frequencies are followed by periods with relatively low crash frequencies simply due to the variability of crashes, not necessarily the project in question. RTM also implies that periods with relatively low crash frequencies are likely to be followed by periods with relatively high crash frequencies. If an agency selects sites based on high short-term average crash history, then crashes at those sites may be lower in the following years due to RTM, even if the agency does not treat those sites.
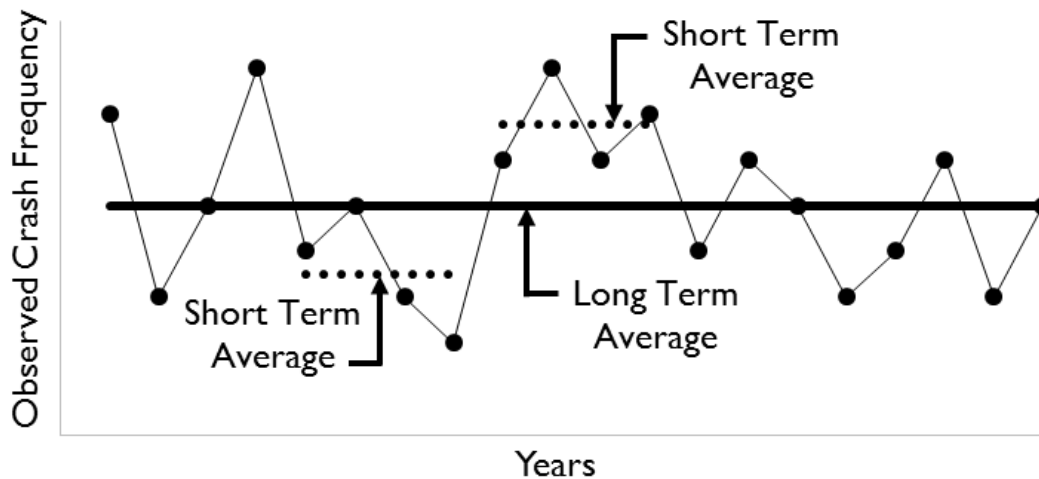
Figure 6. Chart. Illustration of RTM comparing short- and long-term averages.

If the analyst does not account for potential sources of bias, then the evaluation results may incorrectly overestimate or underestimate the safety effectiveness of the project or countermeasure. Refer to Reliability of Safety Management Methods: Safety Effectiveness Evaluation for further discussion of potential sources of bias, the potential impacts on safety effectiveness evaluations, and opportunities to address these issues.[16] In discussing the various evaluation methods, this guide indicates the ability of each method to address potential sources of bias.

4.3 EVALUATION METHODS

There are several techniques to evaluate individual projects, and the appropriate technique depends on the objective of the evaluation and the MOE. It is often useful to employ multiple techniques to evaluate individual projects. For individual crash-based project evaluations, the simple before-after study, before-after study with traffic volume correction, and the before-after with shift of proportions are appropriate. It is important to recognize that small sample size is a common limitation for evaluations of individual projects. Chapter 5 describes opportunities to combine data from multiple projects to produce a more reliable estimate of countermeasure effectiveness. Chapter 5 also describes advanced before-after methods to account for potential sources of bias such as RTM, the nonlinear relationship between crashes and traffic volume, and other changes over time. While the advanced before-after methods generally provide more reliable estimates than simple before-after studies, they also require additional data and more time to conduct the analysis, which may not be viable for individual project evaluations.

Simple Before-After Study

A simple before-after study is a basic comparison of crashes before and after implementation of a particular treatment. The safety effect of a countermeasure is assessed by directly comparing the crash frequency in the after period with the crash frequency in the before period. The simple before-after study design does not account for possible bias due to RTM and does not account for temporal effects or trends such as changes in traffic volume, changes in driver behavior, and changes in crash reporting. While the simple before-after method is common for individual project evaluations, analysts should consider the results from a single site with caution because they may not represent the general countermeasure effect. For example, a simple before-after study that includes an economic recession could overestimate the countermeasure effect, while a similar study during an economic recovery could underestimate the countermeasure effect because there is no account for changes such as traffic volume and driver behavior. The before-after method with traffic volume correction is more appropriate if there are changes in traffic volume during the study period. Again, chapter 5 describes opportunities to combine data from multiple projects to produce a more reliable estimate of countermeasure effectiveness.

Simple before-after studies do not
account for RTM or changes over time.

Table 2 presents sample data for a simple before-after study of an individual project. In this example, the study period is seven years, including three years before implementation, three years after implementation, and excluding the implementation year. The simple before-after method is relatively straightforward, and includes a comparison of crashes before and after implementation for different crash categories depending on the intent of the analysis. Again, the

analyst may assess the change in total crashes, target crashes, and other crash types and severities. For this example, the sample data indicate a 6.7 percent increase in total crashes (100*(1 − 16/15)), a 60 percent reduction in fatal and serious injury crashes, a 55 percent reduction in target crashes, and a 75 percent reduction in target fatal and serious injury crashes. If traffic volumes changed during the study period, then the percent change in the various crash categories would reflect the effect of changes in traffic volume in addition to the effect of the project. The before-after method with traffic volume correction is more appropriate to account for changes in traffic volume as described in the next section.

Table 2. Sample data for simple before-after study.

Crash Category	Crashes Before Implementation (3-year period)	Crashes During Implementation Year(s)	Crashes After Implementation (3-year period)	Percent Change
Total (all types and severities combined)	15	Excluded	16	6.7% increase
Fatal and serious injury (all types combined)	5	Excluded	2	60% reduction
Target (all severities of focus crash types of countermeasure)	11	Excluded	5	55% reduction
Target fatal and serious injury (focus crash types of countermeasure)	4	Excluded	1	75% reduction

Before-After Study with Traffic Volume Correction

A before-after study with traffic volume correction is a variation of the simple before-after study that accounts for changes in traffic volume over time. For example, one option to account for changes in traffic volume during the study period is to multiply the before period crashes by the ratio of average traffic volume after to average traffic volume before implementation. A more reliable option is to use the ratio of predicted crashes after to the predicted crashes before implementation based on a calibrated safety performance function (SPF). The before-after with traffic volume correction is still a rather simplistic method, but is more reliable than the simple before-after method when traffic volume is available.

> **Safety performance functions (SPFs)** are equations used to predict the average crash frequency at a location as a function of traffic volume and in some cases roadway or intersection characteristics (e.g., number of lanes, traffic control, or median type).

The traffic volume correction method is more appropriate than a simple before–after study to account for changes in traffic volume.

Shift of Proportions

For project-level evaluations, it is often useful to evaluate the shift in the proportions of crashes by type or severity level when a countermeasure targets specific crashes. For example, if an agency replaces a two-way stop-controlled intersection with a roundabout to address fatal and serious injury crashes, then it would be useful to know if the proportion of fatal and serious injury crashes decreases after the conversion. This method is particularly useful, and generally more appropriate than the simple before-after method, when there is a suspected change in traffic volume over time but traffic volume is not available to perform the before-after method with traffic volume correction.

> ### *The shift of proportions method is useful to account for changes in traffic volume when traffic volume data are not available.*

For this method, the analyst compares the proportion of target crashes to total crashes before and after implementation. The target crashes may include specific crash types or severities. The Wilcoxon signed rank test helps to determine statistical significance. Refer to Chapter 9 of the Highway Safety Manual for details on how to test the statistical significance of the shift in proportions.[5]

> *The KYTC employed the shift of proportions method to determine the safety effectiveness of projects such as cable median barrier and high-friction surface treatment. These countermeasures target specific crash types. For cable median barrier projects, the target crash type was cross-median crashes. For high-friction surface treatment, the target crash type was wet-weather lane departure crashes. After comparing the shift in proportions, KYTC used the Wilcoxon signed rank test to determine the statistical significance. For the shift in proportion of cross-median crashes, the test indicated a statistically significant reduction at the 99 percent confidence level. For the shift in proportion of wet-weather lane departure crashes, the test indicated a statistically significant reduction at the 95 percent confidence level.*

Table 3 presents sample data for a before-after study of the shift in proportions. In this example, the study period is seven years, including three years before implementation, three years after implementation, and excluding the implementation year. The before-after shift in proportions method is relatively straightforward, and includes a comparison of target crashes to total crashes before and after implementation. The target crashes will depend on the intent of the analysis. For this example, assume an agency converted a two-way stop-controlled intersection to a roundabout, targeting fatal and serious injury crashes. In this case, the target crashes are fatal and serious injury crashes. In Table 3, the proportion of target to total crashes is 0.67 before implementation and 0.33 after implementation. The result is a difference of -0.33 from the before to the after period, indicating a 50 percent reduction (100*(1 − 0.33/0.67)) in the proportion of target crashes.

Table 3. Sample data for before-after shift in proportions study.

Total Crashes Before (3-years)	Target Crashes Before (3-years)	Proportion of Target to Total Crashes Before	Crashes During Implementation Year(s)	Total Crashes After (3-years)	Target Crashes After (3-years)	Proportion of Target to Total Crashes After
18	12	0.67	Excluded	9	3	0.33

4.4 SAMPLE SIZE AND STUDY PERIODS

For project evaluations, the sample size is typically a single project. Chapters 5 and 6 discuss the combination of multiple projects to evaluate countermeasure effectiveness (i.e., develop CMFs) and evaluate programs, respectively.

For project-level evaluations, the total study period is typically seven to eleven years. While a longer study period generally provides a larger sample of crashes for analysis, it also increases the chances for other changes over time such as physical or operational improvements, driver behavior, vehicle fleet, or natural degradation (e.g., reduced surface friction, reduced sign reflectivity, or reduced pavement marking reflectivity). As such, agencies should balance the length of the study period with the potential for other changes over time.

It is preferred to use the same duration for the before and after period. For example, with a seven-year study period, it is common to include three years before implementation and three years after implementation, excluding data for the implementation year. It is possible to use different durations for the before and after period, but it becomes necessary to normalize the analysis by comparing crashes per year rather than total crashes before and after. For example, if an analyst has three years of data before implementation but only two years of data after implementation, then it is necessary to divide the crashes before and after implementation by three and two, respectively, to perform a simple before-after analysis. The shift in proportions method is not affected by different durations of the before and after periods, but it is preferable to include at least three years before and three years after implementation to increase the number of crashes in the sample.

Project evaluation should include 12-month increments to avoid seasonal impacts. For example, if the analyst includes 2.5 years before implementation (April through September) and 2.5 years after implementation (November through April), then there are more winter months represented in the after period, which could bias the results. It is also common to use full calendar years (e.g., 2015, 2016) as opposed to 12-month periods spanning multiple years (e.g., May 2015 – April 2016) for ease of assembling data.

The before period must end prior to implementation and the after period begins after implementation. There is also an interim period (i.e., time between network screening and implementation). It may take several years to program a capital project, and other maintenance activities or safety efforts may resolve the underlying crash contributing factors in the time between network screening and project implementation. An agency can review the safety performance of a site during the interim period to determine the continued need for a project. This relates to network screening and requires continual monitoring of sites from year to year (i.e., include sites selected from a given network screening in subsequent network screening). From an evaluation perspective, there is potential to compare safety performance after implementation to both the before and interim periods to see if there is a difference in the safety effects. If projects continually demonstrate a larger benefit from the before to the after period compared to the interim to after period, then this suggests potential bias due to RTM and the agency may need to enhance their network screening practices. Specifically, there may be potential to employ network screening methods that account for RTM.

> *Project evaluation can help to identify opportunities to enhance the HSIP planning and implementation processes.*

While analysts often exclude the construction or implementation period from project evaluations, it is useful to examine crashes during construction separately. This can be done as part of countermeasure or program evaluation to identify work zone configurations or construction practices that enhance safety.

While analysts often include a buffer period after implementation to allow drivers to adjust to the change in conditions, it may also be of interest to evaluate changes over time after implementation. For example, the analyst could compare the before period to the first year after implementation and then to the next two to three years after implementation to determine if there are short-term countermeasure effects that change over time. If an agency identifies an initial increase in crashes followed by a long-term reduction in crashes, there may be opportunities to educate road users to expedite the learning process. For example, some drivers have less experience with roundabouts and therefore require time to gain familiarity with them to drive confidently. While roundabouts physically eliminate crossing-path movements, which are typically more serious, there is potential to increase other crash types (rear-end and sideswipe) that tend result in PDO or minor injury. This may be more pronounced during the first few months after installation as drivers adjust to the new traffic pattern. This is an example where project evaluation at one location can help to inform the rollout of similar countermeasures at other locations.

4.5 DATA REQUIREMENTS

For project evaluation, agencies should have basic countermeasure and crash data, and possibly traffic data for the project site(s) of interest. For both before-after and shift of proportions, agencies should also have project-level details. Specifically, they should identify the specific type of countermeasure, countermeasure location(s), and implementation date. Refer to chapter 3 for details on project tracking. The following are specific data requirements for before-after and shift of proportions evaluations.

Before-After Evaluations

Table 4 provides a template for assembling the following data to facilitate before-after evaluations. These data include:

- Three to five years of crash data before and after implementation of the countermeasure.

- Crash data include details for the target crash type(s) of interest.

- Corresponding years of traffic data (if using the before-after with traffic volume correction).

- Duration of before and after periods (when duration of before and after periods differ).

Refer to section 4.2 for further discussion of before-after methods for evaluating these data.

Table 4. Template to assemble data for project-level before-after evaluations.

Site ID	Before Period Crash Frequency	Before Period Traffic Volume (vehicles/day)	Before Period Duration (years)	After Period Crash Frequency	After Period Traffic Volume (vehicles/day)	After Period Duration (years)
Example	12	5,000	3	5	5,500	2

Shift of Proportions

Table 5 provides a template for assembling the following data to evaluate shifts in crash proportions. These data include:

- Three to five years of crash data before and after implementation of the countermeasure.

- Crash data include details for total crashes as well as the target crash type(s) of interest.

Refer to section 4.2 for further discussion of the shift of proportions method for evaluating these data.

Table 5. Template to assemble data for shift of proportions evaluations.

Site ID	Total Crashes Before	Target Crashes Before	Total Crashes After	Target Crashes After	Proportion of Target to Total Before	Proportion of Target to Total After	Percent Change in Proportion
Example	18	12	9	3	0.67	0.33	50% reduction

4.6 ACCESSIBILITY AND INFORMATION SHARING

As noted in chapter 3, it is useful to establish a SharePoint site or other web-based portal to facilitate communication and data sharing, particularly in a decentralized agency. A SharePoint site or web-based portal serves as a central repository for templates and information sharing. Again, templates and web-based evaluation tools help to expedite the evaluation process and improve consistency among evaluations. Information sharing is a two-way flow and mutually beneficial to the central office and other regional, district, or local offices. The central office receives information required for Federal reporting, project tracking, and evaluations. Subsequently, the central office can share the results of project evaluations back to the regional, district, and local offices to help inform future decisions.

> *It is useful to establish a SharePoint site or other web-based portal to facilitate communication and data sharing.*

The Wisconsin DOT evaluated 19 individual HSIP projects completed in fiscal year 2006 and produced a project evaluation report to share the results. The project level evaluations included both before-after analysis and benefit-cost analysis based on total and target crashes. Figure 7 shows an example project evaluation summary from the Wisconsin report. The example is representative of the other project evaluation summaries, including basic project information (e.g., location, traffic control, traffic volume, project cost, and study period), summary of safety opportunities, description of countermeasure(s), and project evaluation results (i.e., change in crashes and benefit-cost analysis). In addition to the one-page project evaluation summaries, the report includes the before-after analysis worksheet and benefit-cost analysis worksheet for each project. Refer to Wisconsin's HSIP Project Evaluation Report for further information and additional examples.[17]

Highway Safety Improvement Program Project Evaluation Report

Project Information

County	City/Town	Project Type	Major Road
Milwaukee	Franklin	Intersection	USH 45/STH 36 (Loomis Rd.)
Minor Road	Region	FOS ID	Traffic Control
CTH H (Ryan Rd.)	SE	22401570	Two Way Stop
Traffic Volume	Project Cost	Before Period	After Period
18100	$140,000.00	5 Years	2.2 Years

Safety Opportunity and General Information

1. The intersection of USH 45/STH 36 at CTH H is located in Milwaukee County.

2. The intersection was originally constructed with a skew of 49 degrees. This factor, combined with a heavy WB to SB turning movement, resulted in increased numbers of angle crashes, specifically to a crash rate of 1.03 in 2004.

3. Target crashes were angle and rear end.

Countermeasures

1. Traffic signals were installed in 2006.

Results Summary

1. There were 19 crashes (3.8 crashes per year) before the safety treatment and 3 crashes (1.4 crashes per year) after the implementation of the safety treatment. The number of crashes per year was reduced after implementation.

2. The Empirical Bayes analysis for total crashes indicates that approximately 8 crashes were expected at the site without safety treatment while 3 crashes were observed after safety treatment during the 2.2-year after period. This difference is a reduction of approximately 5 crashes (2.3 crashes per year). The reduction is not statistically significant.

3. The benefit-cost ratios using EB estimates and before-after differences for target crashes were 7.72 and 7.91 respectively, both showing that the project was economically worthwhile.
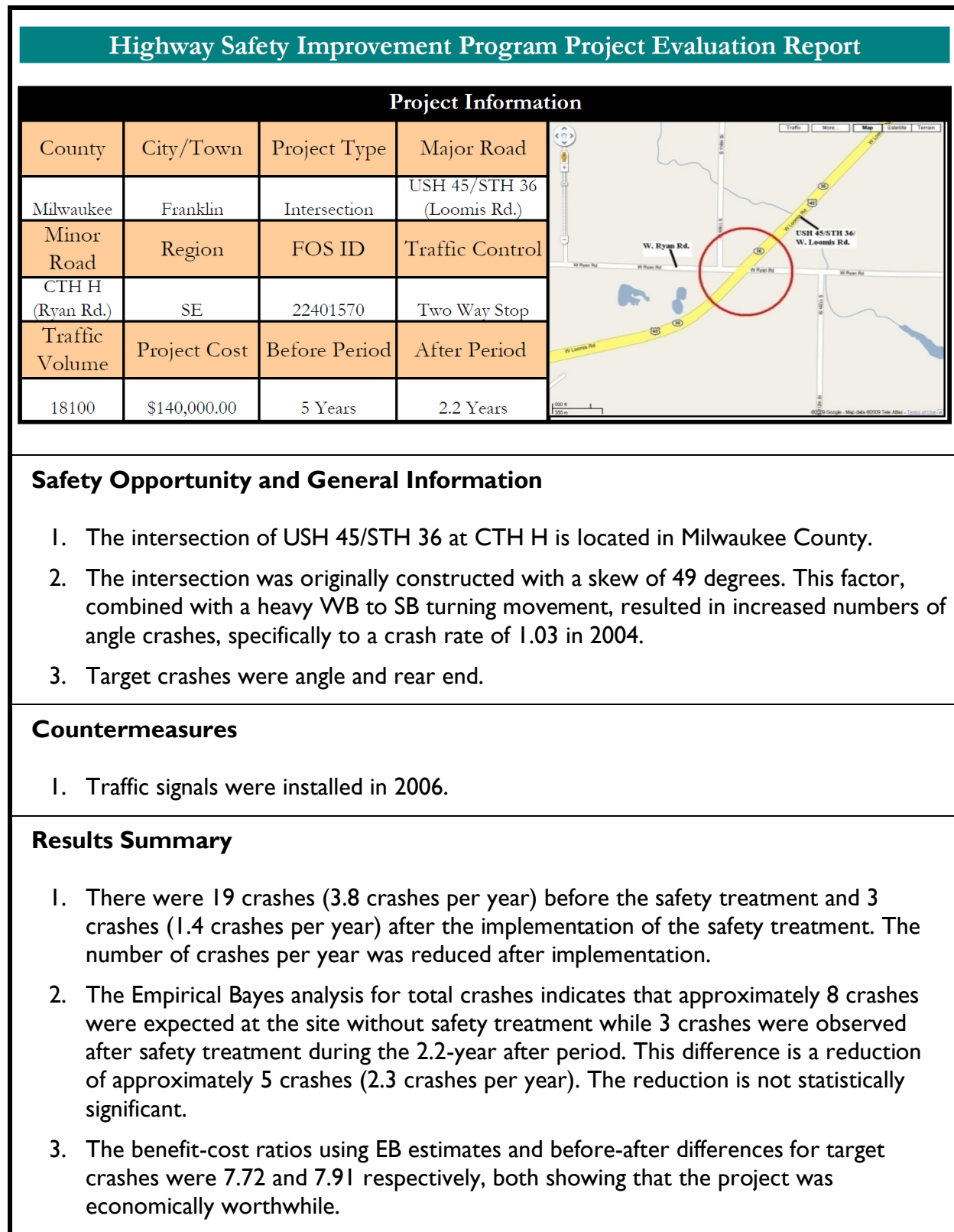
Figure 7. Screenshot. Wisconsin summary of individual project evaluation.[17]

The Colorado DOT produces a report each year, presenting the results of individual HSIP project evaluations. Figure 8 shows the first page of a sample project evaluation, summarizing general project information, underlying contributing factors, specific countermeasures, target crashes, and results. Refer to [Colorado's HSIP project evaluation report](#) for further information and additional examples.[18]

Before-After Safety Analysis   **CDOT Project #: 16313**

Project Name:	Colfax Avenue (US 40) / Youngfield Street	
Project Description:	Upgrade signal	

CDOT Region: 6	Project Def: 16313	County: Jefferson (Lakewood)
Location: US 40	Mile Points: 289.38	Length: N/A
Schedule:	Work Start Date: Approx. 2009	Completion Date: Approx. 2010

Problem Description

As described in the HSIP application for this project, the 3-year crash history showed a higher than expected number of rear-end and broadside type crashes. The cause of these crashes was assumed to be old span wire mounted signals that were subject to wind damage and visibility problems on gusty days.

Improvement Description

In 2009/2010, the intersection was realigned to improve turns and add a protected-permissive southbound left-turn lane. The span wire was replaced with mast arms. The cost of construction was $622,904.

The HSIP application anticipated that this improvement would impact four crash types: rear-end, approach turn, broadside, and pedestrian crashes. The anticipated crash reduction was 20% for these crash types. The expected benefit/cost ratio was 1.18.

Summary and Findings

The before-after analysis showed safety improvements. For this intersection, there were 36 total crashes during the 4-year period before the upgrades (2004 – 2007). In the 4 years after construction (2011 – 2014), the number of crashes decreased to 19. While traffic volumes decreased slightly, the crash rate was still reduced. In addition, the number of injuries also diminished.

The signal and geometry upgrade resulted in a decrease in the number and severity of rear-end and broadside crashes. The actual benefit-cost ratio for this project shows that benefits outweigh costs by a ratio of 3.33 to one, showing that the improvement was justified.

Figure 8. Screenshot. Colorado sample project evaluation report.

The North Carolina DOT publishes individual project evaluations on the web, including a description of the project location, project background, summary of improvements, and results and discussion of simple before-after analysis for both total and target crashes. The web-based project evaluation documents allow others within the State to access the evaluation results to inform future decisions and to demonstrate the benefits of past projects when justifying proposed projects to the public. The following is a link to the project evaluation report from the example in section 4.1 for the conversion from two-way to all-way stop control.[19] Based on a simple before-after analysis, the project evaluation report indicates the project resulted in a 62 percent reduction in total crashes and a 79 percent reduction in target crashes, despite a 58 percent increase in traffic volume. Figure 9 shows a screenshot of North Carolina's website for project evaluation reports. The website allows users to sort and filter by project category (e.g., Traffic Signal Revisions), subcategory (e.g., add signal head, backplates, coordination), division, county, analysis type (e.g., intersection, section), location type (e.g., two-lane undivided, three-leg), and geometry (e.g., two-lane, four-lane at two-lane). As shown in Figure 9, the website indicates the number of related projects in parentheses following each project category (e.g., 40 projects for converting to all-way stop control).

Completed Safety Project Evaluations							
Name	Category	SubCategory	Division	County	Analysis Type	Location Type	Geometry
▷ Category : Advance Activated Warning Sign (17)							
▷ Category : All Way Stop (40)							
▷ Category : Bridge Fencing (2)							
▷ Category : Bridge Removal/Replacement (2)							
▷ Category : Clear Zone Improvements (1)							
▷ Category : Closed Loop Signal System (3)							
▷ Category : Congestion Detection System (1)							
▷ Category : Curve Superelevation Improvements (12)							
▷ Category : Curve Wedging and Resurfacing (7)							
▷ Category : Deceleration Lane for Ramp (1)							
▷ Category : Directional Crossover (65)							
▷ Category : Drainage Improvements (3)							

Figure 9. Screenshot. North Carolina website for project evaluation reports.

4.7 SUMMARY OF INDIVIDUAL PROJECT EVALUATION

The following is a summary of highlights, tips, and tricks for successful evaluation of individual projects:

- Set-up spreadsheet tools to facilitate the evaluation process and develop templates for documentation of project evaluations. Standardized spreadsheets and templates help to improve efficiency and consistency for project evaluation and subsequent reporting.

- Evaluate target and correctable crashes in addition to total, fatal and injury, and PDO crashes. If strategies target specific contributing factors or crashes, then an evaluation of total crashes may not indicate the actual effectiveness. It is more informative to know the impact on target crashes as well as different severity levels.

- Normalize data by year. When there are differential before and after periods, it is necessary to normalize the analysis by comparing crashes per year rather than total crashes before and after.

- Use Google Earth or refer to pre- and post-implementation photos to verify site improvements. It is critical to understand the extent of the improvements.

- Consider the interim crash data period. It may take several years to program a project, and the underlying contributing factors may be resolved by other means in the time it takes to program a formal project. An agency can review the safety performance of a site during the interim period (i.e., time between network screening and construction) to determine if a project is still justified. There is also the potential to compare the safety performance after implementation to both the before and interim periods to see if there is a difference in the safety effects.

4.8 PROJECT EVALUATION RESOURCES

The following resources provide additional information on project evaluation:

- A Guide to Developing Quality Crash Modification Factors explains numerous methodologies for evaluating the safety effectiveness of a project and helps the reader to select an appropriate evaluation method based on the objectives and available data.[20]

- Colorado's HSIP Project Evaluation Report provides further information and additional examples related to project evaluation reporting.[18]

- Manual of Transportation Engineering Studies is an updated and expanded version of the Manual of Traffic Engineering Studies, 4th Edition. The primary focus of this manual is on "how to conduct" transportation engineering studies in the field, particularly for conducting non-crash-based evaluations.[21]

- [North Carolina's Website for Project Evaluation Reports](#) provides an example of a web-based repository for sharing project evaluation results.

- [Reliability of Safety Management Methods: Safety Effectiveness Evaluation](#) provides an overview of methods for conducting observational before-after studies, including the associated strengths and limitations such as RTM and the potential impacts on safety effectiveness evaluations.[16]

- [Wisconsin's HSIP Project Evaluation Report](#) provides further information and additional examples related to project evaluation reporting.[17]

CHAPTER 5: COUNTERMEASURE EVALUATION

While it is important to evaluate individual projects, results from the evaluation of a single project may not be reliable for estimating the general safety effectiveness of a countermeasure or group of similar projects. Instead, it is more reliable to aggregate data from multiple similar projects to estimate the safety effectiveness of a countermeasure. This section describes considerations and practices to aggregate individual project data to perform countermeasure evaluations, including measures of effectiveness, evaluation methods, sample size and study periods, data requirements, documentation, and evaluation of systemic strategies.

> **Chapter 5 at a Glance**
>
> 5.1 Measures of Effectiveness
> 5.2 Evaluation Methods
> 5.3 Sample Size and Study Periods
> 5.4 Data Requirements
> 5.5 Documentation
> 5.6 Evaluating Systemic Strategies
> 5.7 Disaggregate Analysis
> 5.8 Summary
> 5.9 Resources

5.1 MEASURES OF EFFECTIVENESS

MOEs for countermeasure evaluation are similar to those for individual project evaluation (e.g., change in crashes, injuries, and fatalities, economic effectiveness, and benefit-cost ratio). As described in project evaluation, it is appropriate to focus on target crashes in addition to total crashes.

The agency's desired use of countermeasure effectiveness estimates (i.e., CMFs) in countermeasure selection, economic analysis, and project prioritization should guide the selection of target crashes for developing countermeasure effectiveness estimates. For example, if an agency generally focuses on the potential benefit of a countermeasure with respect to fatal and serious injury crashes, then it is useful to develop safety effectiveness estimates based on fatal and serious injury crashes. If an agency generally focuses on the net benefits of potential projects, considering all crashes, then it is useful to develop safety effectiveness estimates by crash severity (e.g., fatal, injury, and PDO). Some agencies combine fatal and injury crashes into a single category, which increases the sample size for evaluation and reduces the number of categories for future analysis. In some cases, it is appropriate to create multiple categories for injury crashes (e.g., fatal and serious injury separately from minor injury). This is appropriate when there is reason to believe the countermeasure has differential effects by injury level (e.g., a median barrier may increase PDO and minor injury crashes while reducing fatal and serious injury crashes). The ability to develop countermeasure effectiveness estimates by severity depends on the sample size and quality of crash data to distinguish between severity levels.

> *The desired use of CMFs should guide the selection of target crashes for developing countermeasure effectiveness estimates.*

Severity is a common means to stratify and estimate net benefits because there are few discrete injury categories. It is more difficult to estimate the effectiveness of countermeasures for each crash type because there are more potential categories (e.g., right-angle, rear-end, sideswipe, run-off-road, head-on, etc.). In addition, there are correlations between crash type and crash severity (e.g., right-angle crashes tend to be more severe and rear-end crashes tend to be less severe). Although not as common, agencies may develop safety effectiveness estimates by crash type. Each agency should consider how they intend to use the countermeasure effectiveness estimates in future decisions, and establish MOEs accordingly.

5.2 EVALUATION METHODS

There are several available methods to evaluate countermeasures, and the appropriate method depends on the objective of the evaluation and the MOE. For crash-based evaluations, the Highway Safety Manual lists three basic study designs:[5]

- Experimental before-after studies.

- Observational cross-sectional studies.

- Observational before-after studies.

In experimental studies, agencies randomly select sites for treatment and control, and then administer the countermeasure to the treatment group, leaving the control group untreated. The control group serves as a baseline to control for other factors and estimate what would have happened to the treatment sites had the agency not administered the treatment. Experimental studies are not commonly used in highway safety.

In observational studies, agencies do not randomly select sites for treatment. Instead, evaluations are limited to sites already selected for treatment based on other reasons including safety concerns. Observational studies are more common than experimental studies in highway safety because agencies do not use random selection to identify sites for treatment. Hence, this guide focuses on observational studies.

There are two broad classifications of observational studies: cross-sectional studies and before-after studies. In either case, the objective is to compare the safety performance of a group of sites with the countermeasure of interest to what would have been the safety performance for the same group of sites without the countermeasure. The primary difference between cross-sectional and before-after studies is the process of estimating what would have occurred in the treatment group had the agency not implemented the treatment.

In **cross-sectional studies**, there is not necessarily a physical change or countermeasure during the study period. Instead, analysts compare the safety performance of a group of sites with the feature of interest to a group of sites without the feature of interest. Cross-sectional studies use the group of sites without the feature of interest to estimate what would have been the safety performance without the countermeasure. For example, to estimate the safety effectiveness of shoulder rumble strips, an agency might use a control group comprised of sites with paved shoulders and no shoulder rumble strips to estimate what would have been the safety performance of the treatment group (i.e., a group of sites with paved shoulders and shoulder rumble strips) without the countermeasure.

In **before-after studies**, some change occurs during the study period, and analysts compare the safety performance of the treatment group over time. Before-after studies use information from the before period for the treatment group (i.e., sites with the countermeasure of interest) to estimate what would have been the safety performance for the same group without the countermeasure. There is also the potential to use information from a comparison or reference group (i.e., sites without the countermeasure) to adjust for other changes over time that affect safety performance. For example, the use of a comparison or reference group can help to account for the impact of an economic recession (e.g., changes in traffic volume and driver behavior).

Since HSIP projects typically consist of specific countermeasures with clearly defined before and after periods, this guide focuses on before-after evaluation methods. Table 6 provides an overview of five before-after methods. In general, the Empirical Bayes (EB) before-after method is one of the more reliable methods for developing quality CMFs because it can properly account for RTM, changes in traffic volume, and other changes over time. The sections following Table 6 provide additional discussion of each method, including strengths, limitations, and applicability (i.e., when the method may be an acceptable alternative to the EB before-after method). Refer to A Guide to Developing Quality CMFs and Recommended Protocols for Developing CMFs for additional information and equations related to before-after and cross-sectional studies as they apply to developing CMFs.[20,22] Refer to Appendix B for templates to conduct various before-after evaluations.

> ### *The EB before-after method is one of the more reliable methods for developing quality CMFs.*

Table 6. Overview of before-after methods for countermeasure evaluations.

Method	Accounts for RTM	Accounts for Changes in Traffic Volume	Accounts for Nonlinear Relationship between Crashes and Traffic Volume	Accounts for Other Changes Over Time
Simple				
Simple with linear traffic volume correction		●		
Simple with non-linear traffic volume correction		●	●	
Comparison group		●		●
Empirical Bayes	●	●	●	●

Simple Before-After Study

A simple before-after study is a basic comparison of crashes before and after implementation. The safety effect of a countermeasure is assessed by directly comparing the crash frequency in the after period with the crash frequency in the before period. A simple before-after study is generally not appropriate for developing quality CMFs because it does not account for possible bias due to RTM and does not account for temporal effects or trends such as changes in traffic volume, changes in driver behavior, and changes in crash reporting. A simple before-after study may be appropriate if the analyst has reason to believe there is limited or no potential for RTM and there are no other changes over time that affect safety other than the treatment of interest. There may be limited RTM in cases where 1) crash frequency is not considered in selecting a site for safety treatment, 2) the safety evaluation is strictly related to a change implemented for operational reasons, or 3) a blanket treatment is applied to all sites of a given type. In practice, except for blanket treatments, it is difficult to confirm that there is no RTM, and only a truly random selection of sites for treatment will ensure there is no selection bias.

> *A simple before-after study is generally not appropriate for developing quality CMFs.*

Before-After Study with Traffic Volume Correction

A before-after study with traffic volume correction is a variation of the simple before-after study that accounts for changes in traffic volume over time. For example, comparing the crash rates (i.e., crashes per some measure of exposure such as vehicle miles traveled) before and after implementation rather than the crash counts helps to account for changes in traffic volume.

The traffic volume correction may be a linear or nonlinear trend. The use of crash rates implicitly assumes the relationship between crash frequency and traffic volume is linear; however, many studies have shown the relationship between crash frequency and traffic volume is nonlinear. Further, the use of crash rates may not account for the annual variation in traffic volume within the before and after periods. SPFs are more reliable to account for changes in traffic volume because they reflect the nonlinear relationship between crash frequency and traffic volume.

> *Nonlinear traffic volume correction methods such as SPFs*
> *may increase statistical reliability when compared to*
> *linear traffic volume correction methods.*

Figure 10 illustrates the difference between a linear and nonlinear trend to define the relationship between crash frequency and traffic volume. Hypothetically, if the traffic volume increases from 5,000 to 10,000 vehicles per day, the nonlinear trend from Figure 10 predicts a 25 percent increase in crashes (i.e., 9 crashes at 5,000 vehicles per day versus 12 crashes at 10,000 vehicles per day). Using the linear trend, this same increase in traffic volume is associated with a 50 percent increase in predicted crashes. Nonlinear traffic volume correction methods such as SPFs may increase statistical reliability when compared to linear traffic volume correction methods such as crash rates.
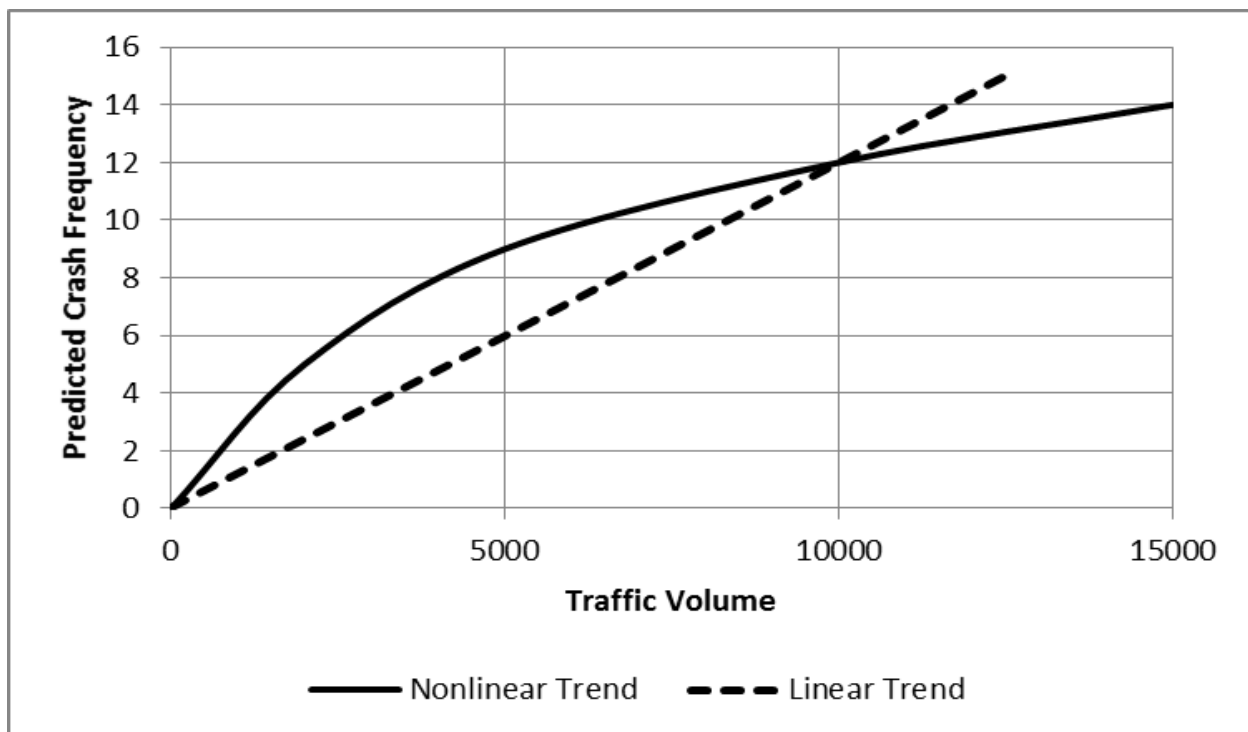

Figure 10. Graph. Relationships between crash frequency and traffic volume.

A before-after study with traffic volume correction is generally not appropriate for developing quality CMFs because it does not account for possible bias due to RTM, and does not account for temporal effects or trends such as changes in driver behavior and changes in crash reporting. A before-after study with traffic volume correction may be appropriate if the analyst has reason to believe there is limited or no potential for RTM and there are no changes in driver behavior or crash reporting over time. There may be limited RTM in cases where 1) crash frequency is not considered in selecting a site for safety treatment, 2) the safety evaluation is strictly related to a change implemented for operational reasons, or 3) a blanket treatment is applied to all sites of a given type. In practice, except for blanket treatments, it is difficult to confirm that there is no RTM, and only a truly random selection of sites for treatment will ensure there is no selection bias.

Before-After Study with Comparison Group

The before-after study with comparison group incorporates information from an untreated group of sites to account for temporal effects and changes in traffic volume. One way to apply this method is to use the comparison group to calculate a *comparison ratio,* which is the ratio of observed crash frequency in the after period to that in the before period. The observed crash frequency in the before period at the treated sites is multiplied by this comparison ratio to estimate the number of crashes at the treated sites in the after period had the countermeasure not been implemented. The estimated crashes at the treated sites in the after period, had the countermeasure not been implemented, is then compared with the observed crashes at the treated sites in the after period to determine the countermeasure effect.

This approach assumes that the trends in the crash counts in the treatment and comparison groups are similar. Hauer proposes a test to determine if the trends in the two groups are indeed similar, using a sequence of sample odds ratios.[20,23] Analysts typically select comparison sites from the same jurisdiction as the treated sites to increase the likelihood that comparison sites will have similar trends as the treated sites.

Another possible approach for applying this method is to develop or calibrate SPFs using data from the comparison group. Using SPFs, the comparison ratio is the ratio of the predicted crash frequency in the after period to the predicted crash frequency in the before period. By using an SPF, this approach accounts for changes in traffic volume from the before to the after period, and the nonlinear relationship between crash frequency and traffic volume.

The comparison group method does not account for RTM unless treatment and comparison sites are also matched on the basis of the observed crash frequency in the before period. Specifically, the analyst would need to match a control site to each treated site based on the annual crashes in the before period. There are difficulties to matching on the basis of crash occurrence.[24] In addition, the necessary assumption that the comparison group is unaffected by

the treatment is difficult to test and can be an unreasonable assumption in some situations. For example, in an evaluation of red light running cameras, the use of nearby untreated signalized intersections as a comparison group may not be appropriate because the treatment may affect driver behavior at those intersections as well, particularly if drivers are unaware of the location of the cameras.

The comparison group method may be a viable approach to developing CMFs if the analyst has reason to believe there is limited or no potential for RTM. There may be limited RTM in cases where 1) crash frequency is not considered in selecting a site for safety treatment, 2) the safety evaluation is strictly related to a change implemented for operational reasons, or 3) a blanket treatment is applied to all sites of a given type. In practice, except for blanket treatments, it is difficult to confirm that there is no RTM, and only a truly random selection of sites for treatment will ensure there is no selection bias.

> *The comparison group method may be a viable approach to developing CMFs if there is limited or no potential for RTM.*

Empirical-Bayes Before-After Study

The EB before-after method is one of the more reliable methods for developing CMFs because it can properly account for bias due to RTM, changes in traffic volume, and temporal effects.[16,23] The intent of the EB method is to estimate the expected number of crashes that would have occurred had there been no change, and compare that with the number of observed crashes after implementation. The following steps describe how to estimate the expected number of crashes that would have occurred had there been no change:

1. **Identify Reference Group**: Identify a group of sites without the countermeasure, but similar to the treated sites in terms of the factors contributing to crashes, including traffic volume and other site characteristics.

2. **Develop or Calibrate SPFs**: Using data from the reference sites, estimate or calibrate an SPF relating crashes to independent variables such as traffic volume and other site characteristics. As discussed in the following steps, the EB method incorporates information from SPFs to predict crashes based on traffic volume and site characteristics. By selecting a reference group that is similar to the treatment group in terms of the factors contributing to crashes, analysts can reduce the possible bias due to confounding factors.

3. **Estimate Predicted Crashes**: Use the calibrated SPFs and traffic volume data to estimate the predicted number of crashes for each year in the before and after periods at each treated site.

4. **Estimate Ratio of Predicted Crashes**: Using the results of step 3, compute the ratio of total predicted crashes after implementation to total predicted crashes before implementation.

5. **Estimate Expected Crashes Before Implementation**: Using the EB method, compute the expected crashes in the before period at each treated site as the weighted sum of observed crashes before implementation and predicted crashes before implementation from step 3.

6. **Estimate Expected Crashes After Implementation**: For each treated site, estimate the expected crashes after implementation as the product of the expected crashes before implementation (step 5) and the ratio of predicted crashes (step 4). This is the expected number of crashes that would have occurred had there been no change. In addition, estimate the variance of this expected number of crashes.

CMFs are developed by comparing the expected number of crashes that would have occurred had there been no change to the observed crashes with the countermeasure of interest.[20,23] Refer to [A Guide for Developing Quality CMFs](#) for more details.[20]

5.3 SAMPLE SIZE AND STUDY PERIODS

For countermeasure evaluations, the study period considerations are similar to project evaluations. The following is a brief summary (refer to chapter 4 for details):

- **Study period**: 7 to 11 years with 3 to 5 years before and after implementation.

- **Duration of before and after periods**: It is possible to use different durations for the before and after period, but it becomes necessary to normalize the analysis based on the duration before and after. It may be of interest to evaluate multiple after periods to determine if there are differential effects over time.

- **Increments**: 12-month periods to avoid seasonal bias. It is common to use full calendar years for ease of assembling data.

- **Implementation period**: Exclude the implementation period from the analysis to estimate the safety effectiveness of the countermeasure, but consider for safety effects of work zone configuration and related construction or implementation practices.

To increase the statistical reliability of countermeasure evaluations, there is a need to include multiple similar projects in the analysis rather than a single project. While analysts can estimate the effect of a countermeasure based on only a few sites, this will typically result in a large standard error and lower confidence in the estimate. The primary sample size consideration for countermeasure evaluation is to balance the size of the sample needed for reliable results

against the time and resources available to compile and analyze the data. The following four variables impact whether a sample is sufficiently large.

1. Size of the treatment group, in terms of the number of sites and the number of crashes in the before period. More sites and more crashes are preferred to fewer sites and fewer crashes.

2. Relative duration of the before and after periods. Longer study periods are preferred to shorter study periods, although there is potential for other changes over time when the before or after periods are longer than five years. As such, three to five years is typically preferred for both the before and after period.

3. Hypothesized countermeasure effect. The purpose of countermeasure evaluation is to estimate the safety effectiveness of the countermeasure; however, agencies should assume the effect size for the purpose of planning evaluations. This is an educated guess of the potential countermeasure effectiveness based on the results of previous studies. Countermeasures that result in smaller effects require larger sample sizes to detect the effect at the desired level of significance.

4. Size of the comparison group in terms of the number of sites and number of crashes in the before and after periods. More sites and more crashes are preferred to fewer sites and fewer crashes.

Balance the sample size against the time and resources available to compile and analyze the data.

At the time of this guide, there is no formal method for determining required sample sizes for the EB before-after method. Hauer provides a method for estimating the required sample size for a before-after study with comparison group.[23] The method is based on the desired significance level (e.g., 0.05 or 0.10) and the expected change in safety (e.g., 10 percent reduction in crashes). Analysts can use Hauer's method to approximate the sample size required for an EB before-after study, recognizing the sample size estimates will be conservative because the EB approach reduces uncertainty in the estimate of expected crashes.

There are two different perspectives for sample size considerations:

1. Case A: What sample size is required to detect an expected change at a given level of significance?

2. Case B: What change can be detected at a given level of significance with a given sample size?

The following sections provide a brief overview of the process and sample size considerations associated with each perspective followed by an example. Refer to appendix C for further details, templates, and examples.

Case A: What sample size is required to detect an expected change at a given level of significance?

Agencies should use the following four-step process to determine sample size.

1. **Determine desired significance level**: The first step is to determine the desired level of significance (α). Ideally, the level of significance would be close to zero (i.e., confidence level close to 100 percent). Practically, this is difficult, if not impossible to achieve. In practice, the desired level of significance is typically 0.05 (95 percent confidence), but in some cases analysts accept significance levels of 0.10 (90 percent confidence) or 0.15 (85 percent confidence) due to limitations in data availability.

2. **Estimate expected level of effect**: The second step is to estimate the expected level of effect (i.e., approximate value of the CMF). This is an educated guess of the potential countermeasure effectiveness based on the results of previous studies. For example, you might expect higher cost countermeasures to have a greater impact on crashes than lower cost countermeasures. Refer to the CMF Clearinghouse for examples of similar countermeasures and related effectiveness by crash type and severity.

3. **Determine required number of crashes (sample size)**: The third step is to determine the required number of crashes for the desired level of significance and estimated level of effect. Table 7 presents the number of crashes required for select levels of effect and common levels of significance. For example, if the expected CMF is 0.80 and the desired level of significance is 0.10 (90 percent confidence), then the required number of crashes is 193. These estimates assume the number of comparison sites is equal to the number of treated sites and the duration of the before and after periods are equal. Hence, you would need 193 crashes in the before and after periods for both the treatment and comparison groups for this example. Do not use linear interpolation or extrapolation to estimate sample sizes for other levels of significance or levels of effect from the numbers in Table 7 because the trends are nonlinear. For scenarios not listed in the table, refer to the spreadsheet template in appendix C.

Table 7. Sample size requirements (number of crashes) by level of effect and desired level of significance.

Expected Level of Effect (CMF)	0.05 Level of Significance (95% Confidence)	0.10 Level of Significance (90% Confidence)
0.90	1858	1155
0.80	279	193
0.70	95	67
0.60	41	29

4. **Determine number of site-years**: At this point, the analyst knows the required sample size in terms of the number of crashes. Now, it is important to estimate the number of sites and years of data to obtain the required number of crashes. The fourth step is to identify the required number of site-years based on a rough estimate of the average crashes at sites similar to the treated sites as well as the number of sites available for analysis. For segments, express the average crashes as crashes per mile-year. For intersections, express the average crashes as crashes per intersection-year. Mile-years is the product of the number of miles and number of years. For example, two years of data for three miles of roadway equals six mile-years.

Case A Example: What sample size is required to detect a 20 percent change at a 0.05 level of significance?

In this example, assume the analyst would like to estimate the sample size required to evaluate a given signalized intersection countermeasure. The analyst assumes the countermeasure will reduce crashes by at least 20 percent (CMF = 0.80) based on information from the CMF Clearinghouse and anecdotal evidence from similar strategies. The analyst would like to detect the change in crashes at the 0.05 level of significance (95 percent confidence). This example uses the four-step process from Case A:

1. **Determine desired significance level**: 0.05 (95 percent confidence).

2. **Estimate expected level of effect**: 20 percent reduction (CMF = 0.80).

3. **Determine required number of crashes (sample size)**: 279 crashes in the before and after periods for both the treatment and comparison groups (from Table 7).

4. **Determine number of site-years**: Based on a rough estimate from similar sites, the analyst assumes an average of 3.0 crashes per year for each signalized intersection. With 279 crashes required, there is a need for 93 site-years (279 crashes / 3.0 crashes per site-year) in the before and after periods for both the treatment and comparison groups. Given three years of before and three years of after data, there would be a need

for 31 treatment and 31 comparison sites. Given five years of before and five years of after data, there would be a need for 19 treatment and 19 comparison sites. As noted in chapter 4, there is a need to balance the duration of the study period with the potential to introduce other changes over time unrelated to the countermeasure of interest. In general, five years before and after implementation is an upper limit for the duration of the study period. If there are not enough sites to achieve the required sample size with five years of data before and after implementation, then it may be necessary to extend the study period or use a different level of significance (e.g., $\alpha = 0.10$ or 0.15). While a smaller sample is required to detect the same effect at a significance level of 0.10 compared to 0.05, the smaller sample size will increase the confidence interval associated with the estimate.

Case B: What change can be detected at a given level of significance with a given sample size?

In this case, the analyst has a sample of sites (e.g., 10 intersections converted from two-way stop-control to roundabouts). The question is whether the analysis is likely to produce a statistically significant effect. This question involves the following two-step process:

1. **Determine number of crashes available for analysis**: The first step is to determine the number of crashes available for analysis from the sample of treated sites.

2. **Determine level of effect based on desired significance and number of crashes**: The second step is to determine the detectable level of effect based on a desired level of significance and the available number of crashes. Use Table 7 to compare the available sample to the cells within the table to identify the relative level of effect for a desired level of confidence. For a more precise answer, refer to the spreadsheet template in appendix C.

Case B Example: What magnitude of effect can be detected at 0.10 level of significance with 70 crashes per year, 3 years of before data, and 3 years of after data?

In this example, assume the analyst would like to determine the detectable effect at a 0.10 level of significance (90 percent confidence) with a sample of 70 crashes per year, three years of before data, and three years of after data. The available data are for the treatment group. This example uses the two-step process from Case B:

1. **Determine number of crashes available for analysis**: 210 crashes before and 210 crashes after.

2. **Determine level of effect based on desired significance and number of crashes**: From Table 7, use the column for a significance level of 0.10 to determine the minimum detectable effect. The minimum sample to detect a 20 percent effect (CMF = 0.80) is 193 crashes. As such, the analyst could detect effects of 20 percent or greater at 0.10 level of significance with the given sample. The analyst would need to collect a similar sample for a comparison group because the available data is for the treatment group only.

Additional Sample Size and Study Period Considerations

In estimating required sample sizes and establishing study periods, agencies should balance the study period duration with the potential for other changes over time. As such, the duration of the study period is generally 10 years or less and it is important to account for changes other than the countermeasure of interest during the given period.

For some countermeasures, it may be difficult to collect the required sample size. For example, pedestrian and bicycle countermeasures may involve relatively few pedestrian and bicycle crashes per intersection or per mile. Similarly, systemic improvements often target crashes over a wider area, and some locations within that area may experience few or no crashes. Increasing the duration of the study period is one option to increase the number of crashes for analysis, but there is the potential to introduce other changes over time as discussed previously. When other options are unavailable, an alternative is to accept a lower level of significance (e.g., 0.15 or 0.20) as opposed to the typical value of 0.05 or 0.10 as an interim step. If there is reason to not accept a lower level of significance, then the analyst could consider extending the geographic area of the study to include more sites.

Extending the geographic study area has potential benefits and drawbacks. The obvious benefit is the increased sample size. Another benefit is the applicability of the results. If a study includes data from multiple jurisdictions or States, then the results will apply beyond a single geographic area. A potential drawback is that countermeasures can have different effects under different conditions, and including multiple geographic areas can introduce variability in the results. This can lead to increased standard errors and less certainty in the countermeasure effectiveness. To account for this issue, agencies should determine the presence of differential effects across the geographic areas. This is accomplished through disaggregate analysis as discussed at the end of this chapter.

Geographic diversity and extended applicability are important to national countermeasure effectiveness studies as is the care taken in selecting similar locations for analysis. When developing agency-specific CMFs, national applicability is less of a priority; however, agencies should still consider the applicability of results across the jurisdiction. For example, in large States such as California or New York, there is a need to consider whether the estimated effect of the countermeasure applies statewide.

5.4 DATA REQUIREMENTS

For countermeasure evaluation, agencies should have basic project, crash, and traffic data for sites with the countermeasure of interest. In addition, agencies may also need similar information from a suitable comparison or reference group, depending on the evaluation method. Table 8 summarizes the data requirements for observational before-after methods in countermeasure evaluation, and the following is a description of each data element:

- **Countermeasure Details**: Identify the countermeasure(s) for evaluation, including the specific type of countermeasure, treated locations, and implementation date.

- **Crash Data**: Summarize the crashes before and after implementation for each site included in the analysis. It is often useful to evaluate countermeasures with respect to total crashes as well as specific crash types (e.g., run-off-road) and crash severities (e.g., fatal and injury).

- **Traffic Volume Data**: Summarize the traffic volume before and after implementation for each site included in the analysis. It is desirable to obtain at least one traffic volume estimate in the before period and one in the after period for each site. For years where traffic volumes are not available, consider estimating the value based on linear interpolation.

- **Reference Group**: Identify a group of sites without the countermeasure, but similar to the treated sites in terms of the factors contributing to crashes, including traffic volume and other site characteristics.

- **Comparison Group**: Identify a group of sites without the countermeasure, but nearby the treated sites to account for temporal factors contributing to crashes such as changes in crash reporting, weather, and driver populations.

- **SPFs**: Calibrate an existing SPF or develop a new SPF using data from the reference sites.

Table 8. Data requirements for observational before-after methods.

Method	Countermeasure Details	Crash Data	Traffic Volume Data	Reference or Comparison Group	SPF
Simple Before-After	●	●			
Before-After with Linear Traffic Volume Correction	●	●	●		
Before-After with Non-Linear Traffic Volume Correction	●	●	●		●
Before-After with Comparison Group	●	●		●	
Empirical Bayes Before-After	●	●	●	●	●

5.5 DOCUMENTATION

Countermeasure evaluations commonly result in CMFs, and it is critical to document the CMF as well as the underlying assumptions and conditions for use in future decision-making. The intent of this section is twofold: 1) identify critical factors related to documentation of countermeasure evaluations, and 2) help analysts to understand the factors that affect the quality of countermeasure evaluations.

> *It is critical to document the CMF as well as the underlying assumptions and conditions for use in future decision-making.*

Countermeasure evaluations range in terms of statistical rigor, leading to results of varying quality and reliability. The CMF Clearinghouse includes criteria such as study design, sample size, standard error, source of the data, and other potential biases to evaluate the quality of a CMF. Refer to the [Recommended Protocols for Developing CMFs](#) for details about these factors.[22] The following is a brief overview of each factor.

Study Design

Agencies should document the method employed in the evaluation. High-quality evaluations use a statistically rigorous study design with reference group or randomized experiment and control group. An example is a thorough EB before-after study. The simple before-after method is a low-quality study design.

Sample Size

Agencies should document the required and available sample size in terms of the number of sites (e.g., intersections, miles, etc.), number of years, and number of crashes. It is useful to report these details separately for the before and after period, total and target crashes, and the treatment and reference or comparison groups. High-quality before-after evaluations typically include at least 20 to 30 treated sites, a suitable reference or comparison group, three to five years before and after implementation, and the minimum required crashes.

Standard Error

Agencies should document the standard error of the CMF. High-quality evaluations result in small standard errors relative to the countermeasure effect. Even if the countermeasure effect is not statistically significant at the 90 or 95 percent confidence level, a small standard error increases the reliability of the results when compared to a large standard error at the same confidence level.

Potential Bias

Agencies should document the potential sources of bias and how the study addresses each. High-quality evaluations control for all sources of known potential bias. The following is a list of potential sources of bias related to before-after evaluations.

- Regression-to-the-mean.
- Changes or uncertainty in traffic volumes.
- Other safety and operational investments.
- Changes in crash reporting.
- Among-State differences if using multiple States.
- Suitability of comparison or reference groups.

Refer to the [Recommended Protocols for Developing CMFs](#) for discussion of how to address these sources of bias.[22]

Data Source

Agencies should document the sources and geographic representation of data used in the evaluation. High-quality evaluations include data from multiple locations with consistent results. Increasing the geographic diversity of an evaluation, without considering the differential effects among different locations, does not improve study quality. In fact, increasing the geographic diversity of sites can increase variance in the results, which leads to a lower quality rating. Similarly, including data from only one jurisdiction with limited sites has limited applicability beyond the sites included in the analysis.

5.6 EVALUATING SYSTEMIC STRATEGIES

In general, the methods and data requirements to evaluate systemic countermeasures are identical to those for evaluating countermeasures implemented through the crash-based approach. If it is possible to define the project details (e.g., systemic countermeasure, treated locations, and implementation date) and collect the required crash and traffic volume data before and after implementation for each improved site, then the analyst may proceed with the countermeasure evaluation methods described in section 5.2. However, the nature of the systemic approach can lead to potential evaluation challenges. For example, if there are many treated sites with no recent crashes, or there is a lack of site-specific crash data, or the sample size is limited, then consider the following alternatives to evaluate systemic countermeasures.

The nature of the systemic approach can lead to potential evaluation challenges.

System-Level Evaluation

As part of the systemic approach, agencies may select focus crash type(s), facility types, and contributing factors to target with systemic improvements. As such, it is appropriate to employ a system-level approach to the evaluation, focusing on the change in target crashes for the targeted facility types. This may include comparing the level of implementation by year (e.g., cost and number of treated miles

Utah evaluates systemic projects differently than crash-density-based projects because of the fundamental differences. Specifically, Utah evaluates systemic projects by comparing the focus crash types in a large defined area before and after improvement. While this does not produce a CMF, it provides a general sense of the effectiveness of the systemic improvements.

or sites) with the target crashes for the focus facility types. For example, if an agency installs shoulder rumble strips to target run-off-road crashes on all rural, two-lane roads with a minimum shoulder width of five feet, then a system-level evaluation of shoulder rumble strips would focus on run-off-road crashes on all treated rural, two-lane roads with a minimum shoulder width of five feet rather than evaluating individual sites or corridors. For this approach, the agency should track the number and cost of improvements by year and compare to the annual number of target crashes on the focus facility type. This is similar to a program evaluation, but focused on a specific crash and facility type. Refer to chapter 6 for examples of charts and tables used in program evaluations.

Shift of Proportions

The shift of proportions is a useful supplement to any evaluation method, whether evaluating a single project, developing a CMF, or evaluating a program. As described in section 4.2, it is useful to evaluate the shift in the proportions of crashes by type or severity level when a countermeasure targets specific crashes. As a reminder, the analyst compares the proportion of target crashes to total crashes before and after implementation, using the Wilcoxon signed rank test to determine statistical significance. Refer to section 4.2 for details and an example related to the shift of proportions.

The KTC evaluated the safety effectiveness of systemic applications of cable median barriers, rumble strips, and high-friction surface treatments using the shift in proportions. Analysts employed the Wilcoxon signed rank test to determine statistical significance. The results indicated statistically significant reductions in target crashes for each systemic improvement. Subsequently, the KTC employed the EB before-after method to confirm the results. The CMFs from the EB before-after method confirmed statistically significant reductions in the target crashes for each countermeasure.

Alternative Analysis Methods

If the sample size is limited (i.e., few observed crashes), or if the average crash frequency per site is relatively small (i.e., low sample mean), then it may be necessary to rely on more sophisticated methods (e.g., Full Bayes before-after) to evaluate the impact of these countermeasures. It is important to note that data limitations still exist and methods are not a substitute for sufficient quality data.

5.7 DISAGGREGATE ANALYSIS

A disaggregate analysis is a detailed investigation of a countermeasure to determine conditions under which the countermeasure may be more or less effective. For example, would a curve warning sign have the same effect under the two conditions shown below in Figure 11? The answer is likely "no" because drivers may not expect a curve along the road shown in the photo on the left, particularly after driving several miles without encountering a curve. In this case, the curve warning sign may be helpful to identify the change in roadway conditions. A curve warning sign may be less effective in areas with numerous curves, such as the one shown by the photo on the right, because drivers should expect curves after driving several miles through other curves.

Source: Frank Gross   Source: Frank Gross

Figure 11. Photo. Example of opportunities for a disaggregate analysis.

One advantage of a disaggregate analysis is that it helps to clarify the applicability of the CMF. Rather than estimating a single CMF to represent the effect of a countermeasure in all situations, the analyst can identify differential effects by characteristics such as area type (e.g., rural or urban), geometry (e.g., number of lanes, lane width, median width, and degree of curve), traffic operations (e.g., traffic volume and posted speed), geographic area (e.g., State, region, county, or city if multiple jurisdictions), and crash history (e.g., number, type, and severity of crashes).

A disaggregate analysis can help to clarify the applicability of a CMF.

Another advantage of a disaggregate analysis is that it helps to prioritize locations for improvement based on the potential effectiveness of the countermeasure. If there are ten potential project locations for installing a countermeasure, and only budget enough to treat five locations, then the analyst could use the results of a disaggregate analysis to identify locations that are likely to benefit the most. If there is only one CMF available from an aggregate analysis, then the analyst assumes that the countermeasure will have the same effect at all locations.

Another advantage of a disaggregate analysis is that it helps to test the sensitivity of the effects of a countermeasure. If a disaggregate analysis indicates the CMF is similar under various conditions, then the analyst can be more confident that the countermeasure will have a similar effect at other locations in the future. On the other hand, if the CMF varies substantially from one location to the next, then the analyst should carefully consider the applicability and potential differential effects of the contemplated countermeasure in the future.

The primary disadvantage of a disaggregate analysis is reduced sample size. By definition, a disaggregate analysis is based on a subset of the data used to compute the aggregate CMF. As such, the sample size is smaller in each disaggregate group, and sometimes too small to produce reliable results. As the sample size decreases, the variance increases. As a result, disaggregate analyses are often associated with larger confidence intervals (i.e., less confidence in the results). Conversely, the disaggregate analysis can lead to smaller standard errors if there is substantial variation in countermeasure effectiveness between different site conditions. For example, if the countermeasure effectiveness differs between urban and rural areas, then an aggregate CMF based on all sites may have a larger standard error than the respective disaggregate CMFs for urban and rural areas.

Another disadvantage of disaggregate analysis is the difficulty in accounting for the effect of other differences among sites that affect safety performance (e.g., non-infrastructure countermeasures). For example, there may be unaccounted differences in educational and enforcement activities among sites.

5.8 SUMMARY OF COUNTERMEASURE EVALUATION

The following is a summary of highlights, tips, and tricks for successful countermeasure evaluation:

- Use more reliable methods that properly account for potential bias due to RTM, changes in traffic volume, the nonlinear relationship between crash frequency and traffic volume, and general temporal effects.

  - The EB before-after method is one of the more reliable methods for developing quality CMFs because it accounts for potential bias due to RTM, changes in traffic volume, the nonlinear relationship between crash frequency and traffic volume, and general temporal effects.

  - The before-after with comparison group method may serve as a viable alternative to the EB method when RTM is not an issue.

  - The before-after method with traffic volume correction is generally not a preferred method for developing quality CMFs because it does not account for possible bias due to RTM or temporal effects.

  - The simple before-after method may overestimate or underestimate the safety effect of a treatment because it does not account for potential bias due to RTM, changes in traffic volume, and general temporal effects. As such, it is generally not a preferred method for developing quality CMFs.

- Consider the appropriate performance measures for the given evaluation. As described in project evaluation, it is appropriate to focus on target crashes in addition to total crashes for the given context.

- Consider the appropriate study period for the given evaluation. It may be necessary to extend the study period for systemic evaluations or evaluations of countermeasures that target rare crash types (e.g., pedestrian crashes). When extending study periods, consider changes other than the countermeasure of interest during the given period.

- Consider the appropriate geographic area for the given evaluation. While extending the geographic area can help to increase sample size, it can also introduce variability in the results. Use disaggregate analysis to determine the presence of differential effects across the geographic areas.

- Evaluate systemic projects with a high-level approach. It may be sufficient to compare the level of implementation (e.g., number of treated miles or sites) with the target crashes for the focus facility types.

5.9 COUNTERMEASURE EVALUATION RESOURCES

The following resources provide additional information on countermeasure evaluation:

- [A Guide to Developing Quality Crash Modification Factors](#) explains numerous methodologies for developing countermeasure safety effectiveness estimates and helps the reader to select an appropriate evaluation method based on the objectives and available data.[20]

- [FHWA's CMF Clearinghouse](#) provides countermeasure safety effectiveness estimates by crash type and severity, helping users to estimate the expected level of effect (i.e., approximate value of the CMF) for similar countermeasures.

- [Recommended Protocols for Developing Crash Modification Factors](#) provides details on potential sources of bias, opportunities to address sources of bias, and important information to document in evaluation reports when developing CMFs.[22]

- [Systemic Safety Project Selection Tool](#) provides guidance for a crash potential-based approach to identifying and addressing systemic safety opportunities. It provides guidance for identifying and implementing systemic countermeasures and methods for assessing their effectiveness.[25]

- [The Art of Appropriate Evaluation: A Guide for Highway Safety Program Managers](#) provides an overview of the traffic safety evaluation process, helping highway safety program managers to identify the best evaluation method and choose a well-qualified professional evaluator.[3]

CHAPTER 6: PROGRAM EVALUATION

A highway safety program is a portfolio of projects implemented to achieve a common goal. The HSIP is a program of highway safety infrastructure improvements with the common goal to reduce fatalities and serious injuries on all public roads. Within the HSIP, there may be subprograms focused on emphasis areas such as fatalities and serious injuries related to roadway departure, intersections, pedestrians, bicycles, speed, local roads, or rural roads. Other programs could be defined by the approach used to identify, diagnose, and treat the locations (i.e., crash-based or systemic).

There are two primary types of program evaluation: crash-based and activity-based. Crash-based evaluations focus on the progress in meeting the safety goals of a program (e.g., a reduction in crashes, serious injuries, and fatalities). Activity-based evaluations focus on the process and actions within a program (e.g., project management, resource allocation, implementation, etc.). While the ultimate goal of the HSIP is to achieve a significant reduction in fatalities and serious injuries on all public roads, it is also helpful to review the process and incremental progress that can impact the overall success of the HSIP.

The following sections present performance measures and techniques for crash-based and activity-based program evaluations as well as the data requirements and link to the SHSP.

6.1 CRASH-BASED PERFORMANCE MEASURES

Typical crash-based performance measures include the frequency and change in crashes, injuries, and fatalities over time and the corresponding rates per measure of exposure such as vehicle miles traveled (VMT) or population. The following are five mandatory performance measures for the HSIP based on the five-year rolling averages [23 CFR 490.207]:

1. Number of fatalities.
2. Rate of fatalities per 100 million VMT.
3. Number of serious injuries.
4. Rate of serious injuries per 100 million VMT.
5. Number of non-motorized fatalities and non-motorized serious injuries.

Refer to the [Safety Performance Management Measures (Safety PM) Final Rule](#) for further information. Figure 12 provides an example of program evaluation based on the five-year rolling average number of fatalities.
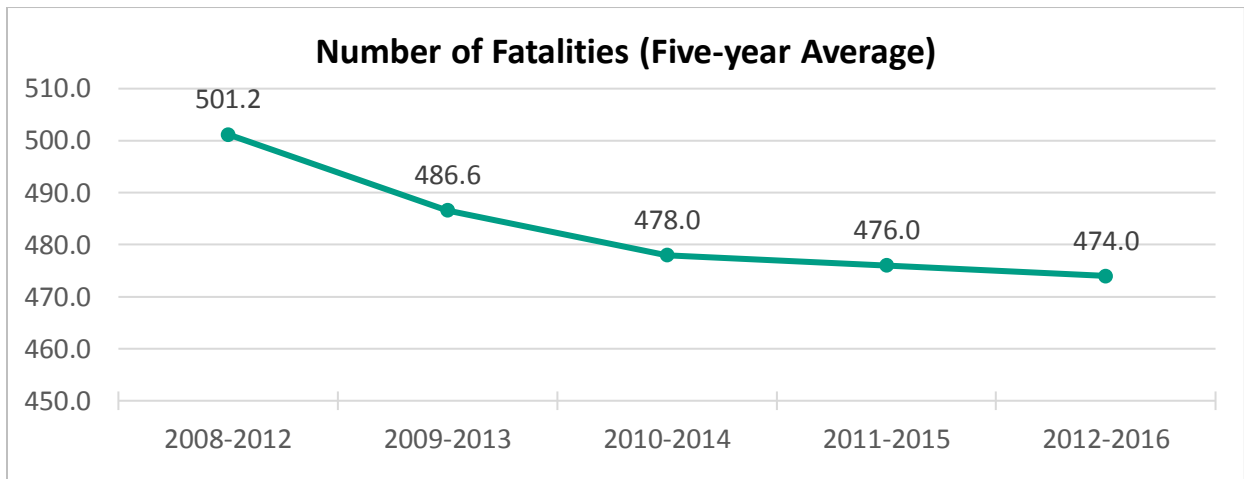
Number of Fatalities (Five-year Average)

	2008-2012	2009-2013	2010-2014	2011-2015	2012-2016
	501.2	486.6	478.0	476.0	474.0

Figure 12. Graph. Example of tracking five-year average number of fatalities.

Beyond the mandatory performance measures, it is useful to evaluate other performance measures such as the change in total crashes, target crashes, and different injury severities, particularly when evaluating subprograms. The following is a list of potential performance measures to consider for crash-based program evaluations, followed by a discussion of each:

- Lives saved (or crashes or injuries prevented).
- Net economic safety benefits and benefit-cost ratio.
- Number of projects with reduction in target crashes.
- Difference in effectiveness on target and total crashes.
- Percent change in crashes versus absolute number of crashes.
- Effectiveness of older versus newer projects.
- Effectiveness of HSIP-funded projects versus non-HSIP projects.
- Effectiveness of projects by region.

Lives Saved (or Injuries Prevented)

As the goal of the HSIP is to achieve a significant reduction in fatalities and serious injuries on all public roads, agencies should track the lives saved and injuries prevented, by implementing highway safety improvement projects. This can help to evaluate the effectiveness of the overall program, subprograms (e.g., intersections, curves, roadway departure), and emphasis areas within an SHSP. To evaluate subprograms or emphasis areas, agencies aggregate crash data based on the associated characteristics (e.g., intersection-related, curve-related, run-off-road

crash type). Lives saved (and injuries prevented) can help to inform decision-makers of the return on past investments and make a case for the amount of funding needed to continue to reduce fatalities and serious injuries.

> *The Texas Traffic Safety Task Force developed a report to summarize potential future lives saved for various safety countermeasures based on the results of past evaluations. The report indicates that Texas could save more than 540 lives each year for an average annual investment of $540 million. Including the additional benefits of injuries and crashes prevented, this investment could return up to $12 for every $1 spent in five years, and many of the safety improvements will continue saving lives and preventing injuries and crashes for up to 20 years. Table 9 provides a summary of the highway safety engineering improvements included in the report. Refer to the Solutions for Saving Lives on Texas Roads for additional information on individual engineering and behavioral measures.*[26]

Table 9. Summary of potential lives saved from highway safety engineering improvements.

Highway Safety Improvement	Potential Lives Saved Over Service Life	Investment (millions)	Potential Cumulative Benefit (billions)	Potential Return per $1 Spent
Rumble Strips	850 – 900	$360	$4.3	$12
Urban Intersections	50 – 100	$313	$1.3	$4
High-Friction Surface Treatments	25 – 50	$100	$0.2	$2
Median Barriers	400 – 700	$590	$2.3	$4
Modernize Bridge Rail	10 – 50	$70	$0.2	$2
Widen Narrow Highways and Bridges	200 – 300	$636	$1.3	$2
Traffic Management Systems	500 – 900	$300	$17.1	$57

Net Economic Safety Benefits and Benefit-Cost Ratio

It is important to understand the net benefits, considering potential increases in certain crash types or severities along with the reductions. If an agency focuses only on the change in fatalities and serious injuries, there is the potential to overlook the added benefits from reductions in minor injuries and PDO crashes. Similarly, certain countermeasures may increase minor injury and PDO crashes (e.g., median barrier). Converting changes in crashes to dollar values by crash type and severity, agencies can estimate the net economic safety benefit (or disbenefit). Using the net economic safety benefit, agencies can then compute a program-level benefit-cost ratio by dividing the net present value benefit by the present value cost of the program. Refer to the HSIP Manual for details on conducting benefit-cost analysis.[1] Refer to the National Safety Council or the Crash Cost Estimates by Maximum Police-Reported Injury Severity within Selected Crash Geometries for crash cost information by severity.[27,28]

California estimates a statewide benefit-cost ratio for HSIP projects on the State Highway System. California measures the effectiveness of the State HSIP through a three-step process: 1) compare three years of crash data before and after the implementation of safety improvements at individual project locations, 2) aggregate individual project costs and benefits to estimate the overall HSIP cost and benefit, and 3) compare the overall HSIP cost and benefit to estimate the statewide HSIP benefit-cost ratio. In 2015, California evaluated 78 projects on the State Highway System with a total implementation cost of $119.4 million. They estimated the annual benefit, in terms of reductions in crash frequency and severity, at $78.3 million, with a present value benefit of $1,565.2 million assuming a 20-year project life. The overall benefit-cost ratio for these 78 projects was 13.1 to 1. In 2016, they evaluated 81 projects on the State Highway System with a total implementation cost of $144.9 million, a present value benefit of $1,725 million, and a benefit-cost ratio of 11.9 to 1.

North Carolina uses a similar approach to California, evaluating individual HSIP projects and then aggregating the results to estimate the overall benefit-cost ratio of the HSIP. Based on the 2015 and 2016 North Carolina HSIP annual reports, the average benefit-cost ratio for HSIP projects is approximately 14 to 1.

Number of Projects with Reduction in Target Crashes

In some cases, a given project may not perform as intended. It is useful to determine the number and percentage of projects that achieved the intended reduction in target crashes and those that did not to inform future decisions related to similar projects. This is a simple tally of individual HSIP projects, comparing the percentage of projects associated with reductions in target crashes to the percentage of projects with no change or increases in target crashes. For projects that achieved the intended reduction in target crashes, identify common characteristics to bring forward for future projects. For projects that did not achieve intended goals, agencies should identify lessons learned to inform future projects.

The Wisconsin DOT evaluated 19 individual HSIP projects. The individual results indicated crash reductions at nearly all of the project locations; however, the cost of reducing these crashes was much higher at some locations than others. As such, the analysts recommended that Wisconsin DOT create a database of such evaluations to use as a future reference for comparing the safety and cost effectiveness of HSIP projects in Wisconsin.

North Carolina has evaluated hundreds of individual safety improvement projects. As they complete multiple evaluations for a particular type of project, they are able to conduct large scale studies using data from locations across the State. Refer to North Carolina's All-Way Stop-Control Evaluation for an example. They share this information to project and program managers, including the regional offices, to provide more objective and definite information regarding the expected safety benefits of similar future projects. If many of the individual project evaluations indicate a lower than expected effectiveness, then there is an opportunity to consider alternative countermeasures in future projects.

Difference in Effectiveness on Target and Total Crashes

It is useful to compare the effect of projects and programs with respect to total and target crashes. Similar to targeted projects, focused safety programs should have a larger effect on target crashes compared to total crashes. Agencies can employ any of the previous performance measures, focusing on target crashes as opposed to total crashes.

> *The Wisconsin DOT evaluated 19 HSIP projects as part of an overall HSIP program evaluation. The evaluation focused on the change in crashes by crash type and severity as well as the benefit-cost ratio. In most cases, the benefit-cost ratio includes only target crashes. Target crashes include the primary crash type or types that the safety improvements were intended to address or mitigate. This helps to understand the return on investment based on the focus of the countermeasure. This also helps to avoid the influence of changes in crashes unrelated to the project. The drawback of using only targeted crashes is the potential to ignore increases in certain crash types. For example, a traffic signal may target right-angle crashes, but there is the potential to increase rear-end crashes. In these cases, Wisconsin includes other related crashes such as rear-end crashes in the economic analysis.*

Percent Change in Crashes versus Absolute Number of Crashes

Agencies should consider the magnitude of safety effects in terms of the absolute number of crashes, injuries, and fatalities. Focusing on the percent change in crashes can be deceiving. For example, a 50 percent reduction in a given crash type may appear to be a substantial improvement; however, if the percent reduction only represents a handful of crashes, then this is a relatively small change relative to the statewide crashes. For program-level evaluations, it is generally appropriate to focus on the change in the number of crashes by type and severity rather than the percent change.

Effectiveness of Older versus Newer Projects

Agencies should consider the differences in project and program effectiveness over time. In general, this involves a comparison of the average project or program effectiveness for different periods of time (e.g., projects or programs from different fiscal years). This can help to understand the efficiency of the HSIP planning component (network screening, diagnosis, and countermeasure selection), and is particularly useful if processes change over time. For example, if an agency implements changes to enhance the process for identifying locations for improvement, and this results in more effective projects, then this is an indication of more effective decision-making. These types of program evaluations can help to assess the effectiveness of past changes in decision-making processes and to justify requests for future enhancements to the process.

Effectiveness of HSIP-funded Projects versus Non-HSIP Projects

There are more opportunities for safety performance investment than what can be addressed by HSIP projects alone. As such, there is an opportunity for agencies to incorporate safety improvements in other infrastructure projects. As agencies implement safety improvements, it is useful to track the effectiveness of projects and programs implemented through different funding sources and processes. In general, this involves a comparison of the average project or program effectiveness by funding source (e.g., HSIP versus non-HSIP safety projects).

Evaluating effectiveness by funding source can help to identify factors that increase the efficiency and effectiveness of projects. For example, safety projects developed and implemented through the HSIP may result in larger safety benefits than projects implemented through other funding sources. If an agency can link enhanced safety planning practices (e.g., network screening, diagnosis, countermeasure selection) to more effective safety projects, then this can help to demonstrate the need to incorporate HSIP-type analyses in other (non-HSIP) project planning practices.

Evaluating effectiveness by funding source can also help to refine estimates of program and countermeasure effectiveness, which can be used in future planning and evaluation efforts. For example, if countermeasures implemented through HSIP projects are generally more effective than similar countermeasures implemented through other funding sources, then an agency could use different CMF values for future planning purposes based on the funding source.

Effectiveness of Projects by Region

Similar to understanding the effects of projects and programs over time and by funding source, there is an opportunity to evaluate projects by region. Again, this involves a comparison of the average project or program effectiveness by subsets of data, in this case by district or region. For districts or regions with above-average results, there is an opportunity to follow-up to identify factors that increase the efficiency and effectiveness of safety projects. For districts or regions with less than expected results, there is an opportunity to follow-up to identify challenges and determine opportunities to overcome those challenges. This may be particularly useful in decentralized States where each district or region is responsible for developing and implementing safety projects.

The appropriate performance measure or combination of measures depends on the underlying objective of the evaluation. The previous discussions indicate the types of decisions an agency could support using evaluation results based on the various performance measures.

The appropriate MOE depends on the underlying objective of the evaluation, and multiple MOEs may be appropriate.

North Carolina employs multiple performance measures to evaluate program success. These include number of projects with crash reduction, number of projects with statistically significant crash reduction, differences between the effect related to target versus total crashes, percent change versus absolute change in the number of crashes, and older versus newer projects.

6.2 ACTIVITY-BASED PERFORMANCE MEASURES

While the ultimate goal of a safety program is to reduce crashes and the related injuries and fatalities, it can take several years to realize (and observe) these results. As such, there is an opportunity to employ interim performance measures to track activity and potential safety surrogates. The following are examples of activity-based performance measures:

- Number of projects implemented.

- Timeliness of project implementation.

- Comparison of estimated project cost versus actual project cost.

- Proportion of program funds allocated.

Activity-based evaluations can help to identify and address opportunities to improve performance. For example, if the State identifies challenges related to countermeasure implementation, then there may be a need to modify implementation procedures or consider alternative countermeasures. Some States such as Colorado and New York produce and distribute performance reports to encourage timely delivery of projects and safety-related activities.

Colorado has created a system for monitoring safety programs and reporting statewide and regional-level information and financial data to highlight areas with opportunities for improvement and areas of excellence. Colorado DOT noted the districts appreciate the feedback and embrace the reporting as a means for statewide accountability.

In New York, the regions are responsible for implementing safety projects and the central office is primarily responsible for conducting evaluations and tracking statewide statistics and progress as part of the HSIP. Figure 13 shows a sample statewide quarterly progress report from the New York State Safety Program Management Bureau. The report presents the vision and goals of the safety program as well as the seven-year trend in fatalities and fatality rate. The report also presents the progress in obligating HSIP funds and letting projects by region and fiscal year.

SAFETY PROGRAM MANAGEMENT BUREAU

Quarterly Report for period ending XX/XX/XXXX

NEW YORK STATE OF OPPORTUNITY. | Department of Transportation

Report Date: XXXX

Vision

New York's safety community will continue to work to ensure that its customers - those who live, work and travel in New York State - have a safe, efficient, balanced and environmentally sound transportation system, and that safety is appropriately considered in all education, enforcement, engineering and emergency medical services activities in New York State in order to reduce fatal and injury crashes.

Safety Goals
- Decrease total fatalities XX percent from the XXXX-XXXX calendar base year average of XX to XX by XXXX. Decrease fatalities/100 million VMT XX percent from XX in XXXX to XX by XXXX.
- Complete required number of Highway Safety Investigations (XX% of identified PILS each year).
- Construct Centerline Audible Roadway Delineators (CARDS) on XX% of the eligible miles by the end of XXXX.
- Install Pedestrian Countdown Timers at XX% of the eligible intersections by the end of XXXX.
- Obligate XX% of the Annual HSIP funds.
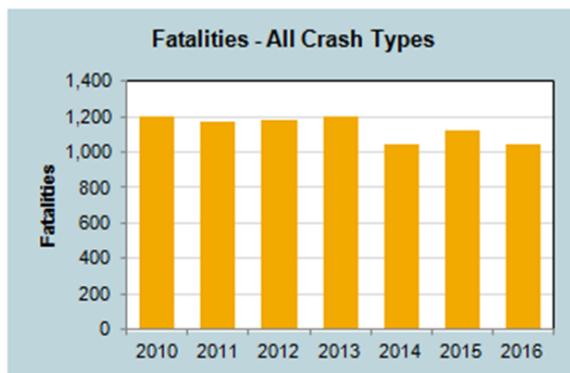
Source: 2009-2015 FARS; 2016 – Preliminary ITSMR
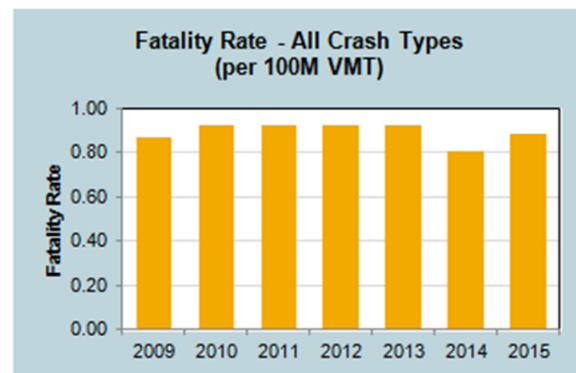
Source: 2009-2015 FARS

Total HSIP Obligations by Federal Fiscal Year (FFY)								
Region	FFY XX		FFY XX		FFY XX		FFY XX	
	Obligations	# of projects let	Obligations	# of projects let	Obligations	# of projects let	Obligations	# of projects let to date
1	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
2	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
3	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
4	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
5	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
6	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
7	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
8	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
9	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
10	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
11	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
MO	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X
Total	$X.XX	X	$X.XX	X	$X.XX	X	$X.XX	X

Source: PSS: Construction Contract Letting Report by Fund Use

Safety Program Management Bureau Quarterly Report   Page 1

Figure 13. Screenshot. New York sample program quarterly report.

In addition to annual HSIP and statewide progress reports, the central office develops regional reports to track overall progress toward the regional work plans and other organizational goals such as:

- *Number of highway safety investigations completed.*
- *Percent obligation of HSIP funds.*
- *Implementation of focused safety programs (e.g., installation rates of centerline rumble strips and pedestrian countdown timers).*

The intent of the regional reports is to improve performance management and encourage regions through positive feedback. If a region is lagging, then the central office works with the region to identify and resolve the challenges. Additionally, the central office provides one-page dashboards summarizing fatalities, serious injuries, and progress toward SHSP emphasis area performance measures. Given the regular progress reports, the central and regional offices have a better understanding of how activities and projects contribute to the statewide safety goals and outcomes.

Once sufficient time (based on the study design and intent) has elapsed, there is an opportunity to overlay the crash-based and activity-based performance measures. For example, if an agency implements a centerline rumble strip program to address head-on crashes on rural, two-lane, undivided roads, then it may be useful to compare the number of miles of centerline rumble strips installed over time with the number of head-on crashes on rural, two-lane, undivided roads over the same time period. Figure 14 provides an example of such a comparison.
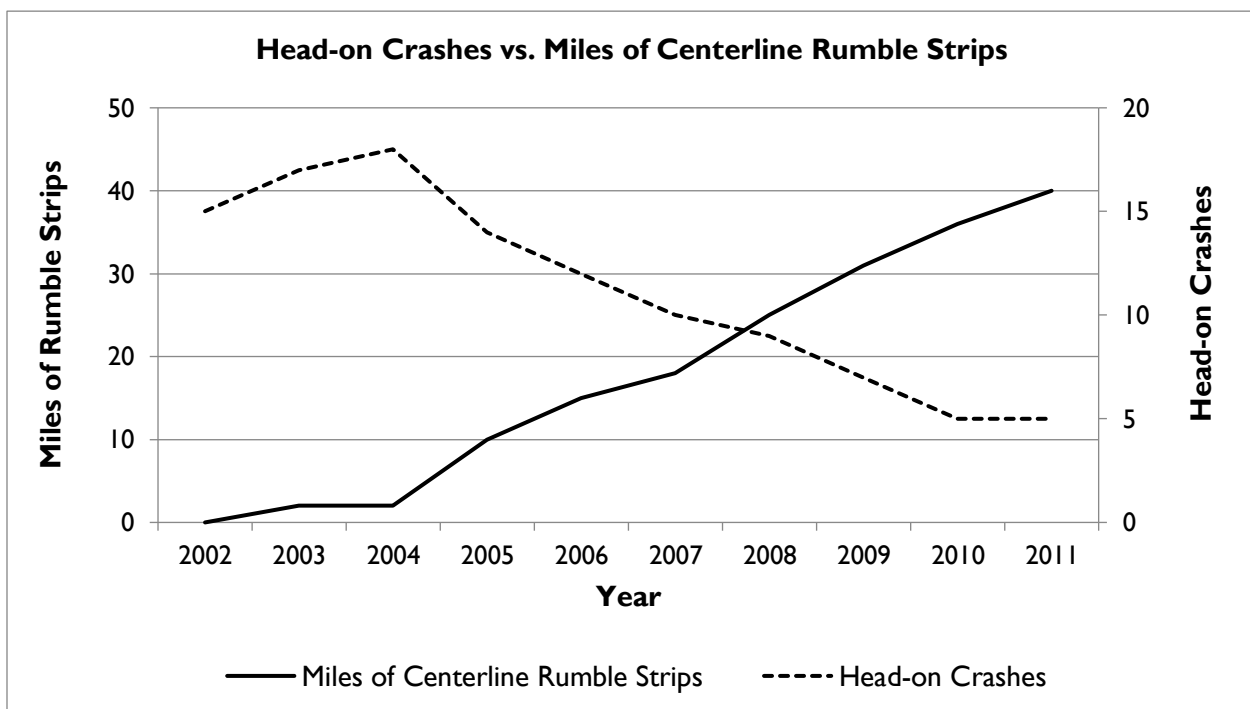

Figure 14. Graph. Example of tracking implementation versus target crashes.

6.3 DATA REQUIREMENTS

For program evaluation, the data requirements are similar to project evaluation: basic project, crash, and traffic data for sites with the countermeasure of interest. In addition, there is a need to link specific projects to programs and subprograms. For example, consider a scenario where a State implements several projects with HSIP funds to address roadway departure crashes (e.g., shoulder rumble strip projects, high friction surface treatment projects, and enhanced curve delineation projects). The State could evaluate the combined effect of these projects as part of a roadway departure program. The State could also combine these projects with all other HSIP-funded projects to evaluate the overall effect of HSIP projects. To conduct this type of evaluation, the State would need some type of identifier to know the project included HSIP dollars and the specific subprograms related to each project. Agencies should use the following information to support program-level evaluations (refer to chapter 3 for details):

- **Project-level effectiveness**: results from the project-level evaluation.
- **Countermeasure type and details**: specific countermeasure(s), treated locations, and implementation dates.
- **Project cost**: proposed and actual project costs.
- **Crash data**: crash frequency by type and severity by year.
- **Exposure data**: vehicle-miles traveled by year.
- **Funding source**: project costs by amount and funding source if joint-funded.
- **Relation to SHSP emphasis areas and safety programs**: indicator linking to emphasis area(s) and safety programs related to the project.

6.4 LINK TO STATE STRATEGIC HIGHWAY SAFETY PLANS

The SHSP is a data-driven, multi-year plan that establishes statewide goals, objectives, and key emphasis areas to reduce fatalities and serious injuries. Each State has developed a SHSP, and 23 CFR Part 924.9(a)(3)(i) requires States to update their SHSP no later than five years from the previous approved version. States are required to evaluate their HSIP as part of the SHSP update (23 USC 148(c)(1)(C) and 23 CFR Part 924.13(a)(2)).

HSIP projects relate directly to the SHSP, contributing to SHSP goals and targets. In fact, 23 CFR Part 924.5(b) requires States to use HSIP funds for highway safety improvement projects that are consistent with the State's SHSP. The FHWA considers highway safety improvement projects consistent with a State's SHSP if they logically flow from identified SHSP emphasis areas and strategies. As such, HSIP projects are one implementation component of the State SHSP, primarily the implementation of infrastructure strategies.

As States update their SHSPs, they should consider how past HSIP projects have contributed to progress toward the SHSP goals and objectives. Understanding the effectiveness of HSIP

projects supports the update of infrastructure-related emphasis areas and strategies as well as the related safety targets and performance measures. For example, if the agency did not meet an SHSP goal or if the safety performance of a strategy is not performing as expected, then the SHSP update may include new strategies or revisions to the old strategy based on lessons learned. Recall the importance of linking HSIP projects to specific safety programs. In this case, there is a need to relate HSIP projects to SHSP emphasis areas.

The Louisiana Department of Transportation and Development developed an online tool to allow users to generate data reports and dashboards for various safety performance measures. Users can customize the reports and dashboards to focus on specific geographic areas and specific emphasis areas. This is particularly useful for monitoring progress related to SHSP emphasis areas. Figure 15 shows an example dashboard from the Louisiana online tool, indicating the number of severe injury crashes by parish from 2006 through 2015 where the contributing factor was intersection related. Users can customize the reporting period, crash severity level, and contributing factor or click on the charts to filter the results and obtain more details on crashes by parish, day of week, time of day, month of year, highway type, contributing factor, and manner of collision. The following is a link to [Louisiana's Data Reports and Dashboards](#).
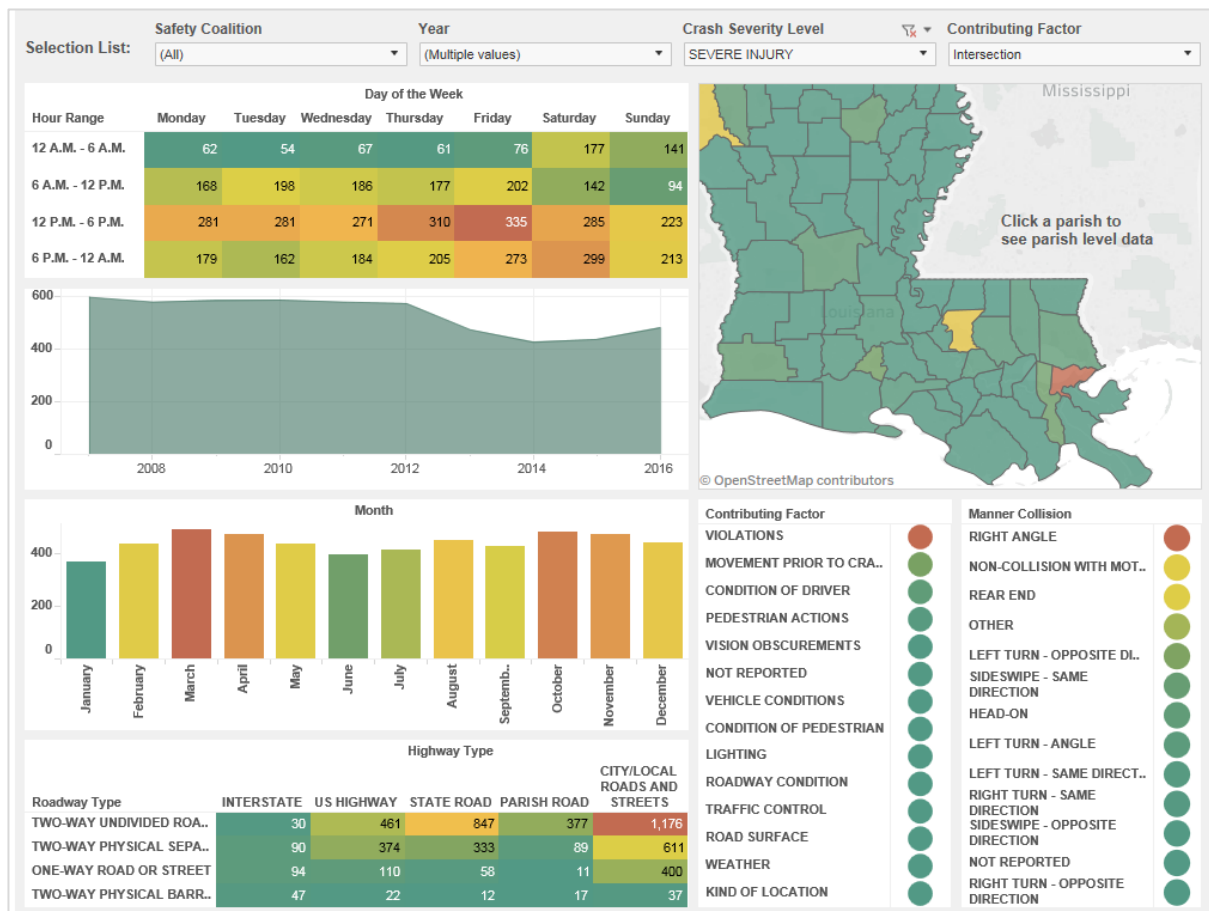

Figure 15. Screenshot. Example of Louisiana SHSP crash dashboard.

> *In New York, the central office provides quarterly reports to the regions. The quarterly reports include progress toward implementation of the HSIP (e.g., number of activities completed and percent of HSIP funds obligated) and one-page dashboards. The dashboards summarize fatalities, serious injuries, and progress toward SHSP emphasis area performance measures.*

6.5 SUMMARY OF PROGRAM EVALUATION

The following is a summary of highlights, tips, and tricks for successful program evaluation:

- Select the appropriate MOE of combination of MOEs based on the underlying objective of the evaluation and intended use of the results.

- Employ crash-based program evaluations to assess the progress toward achieving long-term safety goals and annual safety performance targets (e.g., a reduction in fatalities and serious injuries).

- Employ activity-based evaluations to review processes and measure incremental progress within a program (e.g., project management, resource allocation, implementation, etc.).

- Collect data for individual HSIP projects and perform project-level evaluations to prepare for program level evaluation. By link specific projects to programs and subprograms, it becomes easier to aggregate the results from individual project evaluations to estimate program benefits.

- Consider how HSIP projects have contributed to progress toward the SHSP goals and objectives. Understanding the effectiveness of HSIP projects supports the update of infrastructure-related emphasis areas and strategies as well as the related safety targets and performance measures.

6.6 PROGRAM EVALUATION RESOURCES

The following resources provide additional information on program evaluation:

- 2010 HSIP Manual provides an overview of the HSIP and offers practitioners with a review of standards, new and emerging technologies, and noteworthy practices for each step in the HSIP process. It also provides details on conducting benefit-cost analysis.[1]

- Crash Cost Estimates by Maximum Police-Reported Injury Severity within Selected Crash Geometries provides crash cost information by severity.[28]

- Developing an Effective Evaluation Report provides guidance for producing a final evaluation report, including a summary of how the program was monitored and evaluated as well as the review findings and possible improvements to be implemented into the program.[29]

- **FHWA's Safety Performance Management Measures Final Rule** provides additional information on the five mandatory safety performance measures for the HSIP.

- **HSIP Assessment Toolbox** provides support for agencies to conduct an assessment of their HSIP, including strategies, methods, and best-practices for agencies to consider incorporating into their program.[10]

- **HSIP National Scan Tour Report** provides a summary of notable practices in the areas of HSIP administration, planning, implementation, and evaluation.[11]

- **HSIP Noteworthy Practice Series** provides examples from around the United States of best-practices for various aspects of the HSIP process. It provides a series of case studies with noteworthy examples for other agencies to consider incorporating into their programs.[12]

- **HSIP Self-Assessment Tool** provides a question-based method for managers to perform a self-evaluation of an agency's HSIP. This tool will help agencies track progress, improve strategy development, identify potential areas of improvement for their current program, and ultimately improve their process.[13]

- **Louisiana's Data Reports and Dashboards** provides an example of an online tool to generate data reports and dashboards for various safety performance measures.

- **National Safety Council** provides crash cost information by severity.[27]

- **North Carolina's All-Way Stop-Control Evaluation** provides an example of reporting the results of a program evaluation.

- **Texas's Solutions for Saving Lives on Texas Roads** provides an example of estimating potential future lives saved for various safety countermeasures based on the results of past evaluations.[26]

- **The Art of Appropriate Evaluation: A Guide for Highway Safety Program Managers** provides an overview of the traffic safety evaluation process, helping highway safety program managers to identify the appropriate evaluation method and choose a well-qualified professional evaluator.[3]

CHAPTER 7: USING HSIP EVALUATION RESULTS

Evaluation is the third component of the HSIP process, but this is not the end of the process. While much of evaluation focuses on the effectiveness of past efforts, evaluation provides a critical feedback loop to improve processes and future decisions. The following sections discuss key elements to using the results from HSIP project, countermeasure, and program evaluations to enhance future efforts.

> **Chapter 7 at a Glance**
>
> 7.1 Demonstrating the Value of Investments and Evaluation
>
> 7.2 Communicating Evaluation Results
>
> 7.3 Updating the SHSP
>
> 7.4 Automated Evaluation and Reporting
>
> 7.5 Summary
>
> 7.6 Resources

7.1 DEMONSTRATING THE VALUE OF INVESTMENTS AND EVALUATION

Evaluation results are useful to inform future decisions at the project, countermeasure, and program level. These decisions go beyond the safety program. By sharing evaluation results, there is an opportunity to generate support and additional funding for safety projects as well as continued support and funding to perform evaluations.

To generate interest in evaluations, there is an opportunity to share success stories from other States. For example, an agency could explain how other States are using evaluation results to justify funding and projects such as median barrier, especially when receiving public or political pushback.

The North Carolina DOT used the lives saved approach to demonstrate the value of median barriers. This came at a time when there was considerable pushback from the public and politicians regarding the installation of median barriers because it restricted access.

The Texas DOT reported lives saved estimates as part of its State Bond Program. Specifically, Texas produced a report on the effectiveness of several common engineering and behavioral measures. The report details the potential implementation level and cost as well as the benefits. The report indicates benefits in multiple forms, but all are simple and relatable to the public and decision-makers.[26]

7.2 COMMUNICATING HSIP EVALUATION RESULTS

Communication is a very important element of HSIP evaluation. Early in the process, communication is important to prepare for evaluations. For example, there is an opportunity for project managers to track and communicate project-level details during project development and implementation. Once evaluations are complete, it is important for the HSIP manager to communicate the evaluation results to stakeholders. This requires effective

communication between the safety program and other offices (e.g., planning, design, operations). In decentralized States, there is also a need for effective communication between the central office and the locals, districts, and regions.

One way to communicate results is through progress reports. There is an opportunity to develop progress reports to track performance measures statewide and by local, district, and regional level. In doing so, it is necessary to ensure the results are understandable to program and project managers. Agencies can share progress reports with various groups to highlight successes and challenges. This conveys a sense of ownership and promotes accountability to the local, district, and regional transportation managers.

The New York State DOT initiated progress reporting by providing regions with statewide statistics. This helped to generate interest and the regions requested a breakdown of the numbers by region to see how they were performing. Now, the central office develops and distributes regional reports, demonstrating the progress at the regional level.

Colorado created a system for monitoring safety programs for the purposes of statewide reporting and accountability. These program reports include statewide and regional-level information and financial data to highlight areas with opportunities for improvement and areas of excellence. Figure 16 is an example of Colorado's statewide progress report for the HSIP benefit-cost ratio. The goal is to achieve an average benefit-cost ratio of at least 2.0 for HSIP projects advertised in 2015.

Lead Metric:
Highway Safety Improvement Program (HSIP) Benefit/Cost

Metric description: Achieve an average benefit/cost ratio of at least 2.0 for HSIP projects advertised for CY15.

Why it matters: A high benefit to cost ratio ensures that the program is run in the most cost effective manner and that locations with a high degree of severe crashes are being targeted in order to achieve reductions in fatalities and serious injuries

Additional Lead Metric:
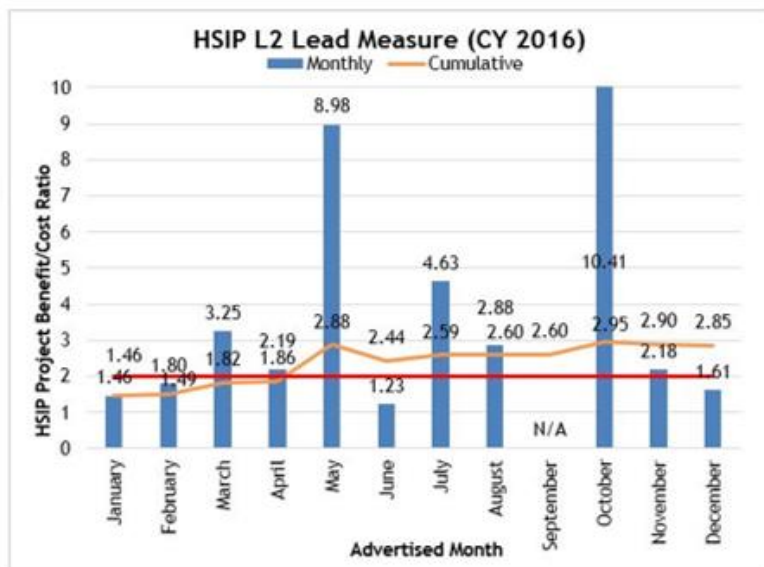**HSIP Expenditure Performance Index (XPI)**

Figure 16. Screenshot. Colorado sample program-level report.

In addition to providing progress reports, it is useful to meet regularly with program managers and project managers at the local, district, and regional level to explain the results, communicate needs, identify opportunities for improvement, and address specific concerns. Meetings are a good time to

The Florida DOT noted that regular communication—quarterly meetings with FDOT executives and monthly meetings with FDOT districts—emphasizes the importance of safety and evaluation to project managers.

review progress toward activity-based performance measures (e.g., are projects on time and under budget) and ultimately review progress toward crash-based performance measures (e.g., are projects improving safety performance in terms of the number or severity of crashes).

7.3 UPDATING THE SHSP

States should use evaluation results to inform updates to their SHSP. There is an opportunity to use crash-based evaluation results to confirm or modify SHSP emphasis areas and strategies. Specifically, the State may identify countermeasures or programs that are particularly effective in reducing specific crash types as well as countermeasures or programs that did not meet intended goals. These findings could influence the selection of strategies for the SHSP update.

There is an opportunity to use activity-based evaluation results to identify and address opportunities to improve performance. For example, if the State identifies challenges related to the implementation of strategies from the SHSP, then there may be a need to modify the SHSP implementation plan or consider alternative strategies in the SHSP update.

7.4 AUTOMATED EVALUATION AND REPORTING

Some agencies have developed automated evaluation approaches that rely on underlying project-level data to summarize progress by various performance measures. There is also an opportunity to consider the use of business intelligence software (e.g., Tableau) to facilitate progress reporting and generate visual representations of results (tables and dashboards). Standardized spreadsheets, templates, and automated approaches help to improve efficiency and consistency for evaluation and reporting.

The Louisiana Department of Transportation and Development developed an online tool to allow users to generate data reports and dashboards for various safety performance measures. Users can customize the reports and dashboards to focus on specific geographic areas and crash types. The following is a link to Louisiana's Data Reports and Dashboards.

The Florida DOT uses their CRASH software to automate parts of the evaluation process. CRASH includes a historical crash database (updated annually), a database of countermeasure effectiveness (i.e., CMFs) based on statewide projects, and a user database to maintain user access and permissions. District engineers can use the system to develop benefit-cost ratios and to generate other crash-based statistics during the project development process. They can also populate the safety improvement database with work orders and update information as they program and complete projects. Three years after project completion, the CRASH software computes the project effectiveness and combines similar projects to generate estimates of countermeasure effectiveness. This ultimately rolls up to support annual HSIP reporting and program evaluation. The following is a link to [Florida's CRASH Portal](#).

The Missouri DOT developed an online reporting tool similar to Louisiana to facilitate evaluation and reporting. Missouri provides a high-level [dashboard](#) on their website, indicating the annual fatalities to date and the percent not wearing a seatbelt. From their [crash statistics website](#), users can select from a dropdown menu of standard report templates and then select the year and location of interest. Report templates include crash summaries by city, county, region, statewide. Reports are also available for fatalities involving bicycles, pedestrians, motorcycles, signalized intersections, unsignalized intersections, and other road user and behavioral factors.

The New York State DOT also developed an automated evaluation process: PIES. PIES links 1,500 capital projects and 4,700 safety studies to crash and roadway data. The system retrieves construction start and end dates, project limits, and other data from program and construction management databases. Users can search the database by region and by year. The system generates estimates of countermeasure effectiveness (i.e., CMFs) for low-volume and high-volume sites along with the significance level for any given project type. PIES supports project, countermeasure, and program level evaluations. At the program level, the central office develops quarterly performance management reports, which contain both automated and manual components.

The Virginia DOT uses [Tableau](#)—an enterprise business intelligence and data visualization software. The full software allows districts to populate and update the project database through project managers. The central office can track safety and non-safety projects through this database as well as track and follow up with districts to address challenges and concerns related to project management (e.g., schedule and budget). The districts can access the database and use filters to find projects and track project schedules and budgets. This functionality is especially useful when district and State safety engineers meet monthly to discuss on-time and on-budget delivery. The Virginia DOT noted that Tableau is particularly useful for generating visual representations of performance measures and progress (tables and dashboards).

7.5 SUMMARY OF USING HSIP EVALUATION RESULTS

The following is a summary of automated evaluation and reporting approaches, highlighting tips and tricks for using HSIP evaluation results.

- Support improved communication between HSIP staff and district or region offices as well as other non-safety State transportation agency program offices (e.g. planning, project development, operations, and maintenance).

- Develop progress reports to track performance measures statewide and by district or region. The reports should present information clearly and concisely so that the results are understandable to the public and managers.

- Share progress reports with districts and regions to highlight successes and identify opportunities for improvement. This conveys a sense of ownership to the districts and regions and helps with accountability.

- Meet regularly with decision-makers and stakeholders (e.g., local and district/regional staff) to share results, communicate needs, identify opportunities for improvement, and address specific topics of common interest. Meetings are a good time to review progress and performance measures (e.g., are projects on time, under budget, improving safety performance).

- Consider the use of business intelligence software to facilitate progress reporting and generate visual representations of results (tables and dashboards).

- Consider the use of project tracking and data visualization software. These tools allow non-central staff to easily transmit or upload information for project or program-wide evaluation in a timely manner.

7.6 HSIP EVALUATION RESULTS RESOURCES

The following resources provide additional information on using HSIP evaluation results:

- [Developing an Effective Evaluation Report](#) provides guidance for producing a final evaluation report.[29]

- [Florida's CRASH Portal](#) provides an example of a software that automates parts of the evaluation process and supports annual HSIP reporting and evaluation.

- [Louisiana's Data Reports and Dashboards](#) provides an example of an online tool to generate data reports and dashboards for various safety performance measures.

- Missouri's Crash Statistics Website provides an example of an online tool to facilitate evaluation and reporting. Users select from a dropdown menu of standard report templates and then select the year and location of interest to generate crash summaries by various emphasis areas.

- Missouri's Dashboard provides an example of an online reporting tool, indicating the annual fatalities to date and the percent not wearing a seatbelt.

- Virginia's Tableau Application is an example of an enterprise business intelligence and data visualization software that allows users to generate visual representations of performance measures and progress (tables and dashboards).

CHAPTER 8: CLOSING

The HSIP comprises three components: planning, implementation, and evaluation. While planning and implementing projects are important steps to addressing existing and future safety opportunities, evaluating these efforts is critical to understanding the return on investment and improving the efficiency and effectiveness of future decisions. Beyond the requirements for States to conduct evaluation as part of the HSIP, the following are specific benefits to evaluating the safety effectiveness of individual HSIP projects, countermeasures (i.e., groups of similar HSIP projects), and the overall program:

- **Understand the return on investments**: Each year, transportation agencies invest nearly $4 billion on HSIP projects with the intent to reduce fatalities and serious injuries on all public roads. HSIP evaluations can help to demonstrate the value of these expenditures in terms of lives saved and serious injuries prevented. By demonstrating the value of past investments, there is an opportunity to justify the need for and appropriate level of future HSIP funding.

- **Identify and address potential opportunities**: Not all safety investments result in a safety performance benefit. HSIP evaluation can help to identify investments that did not perform as intended. If an agency identifies a project that is not meeting safety performance expectations based on the evaluation results, then there is an opportunity to address the situation as appropriate for the location (e.g., remove the countermeasure or install supplemental countermeasures).

- **Inform future decisions**: With competing demands and limited funds, there is a need to prioritize efforts and justify decisions. Evaluations can help to develop or refine estimates of effectiveness used to prioritize projects and manage programs. For example, if certain programs or countermeasures are consistently effective (i.e., reduce the expected frequency and severity of crashes), then agencies may choose to continue those programs and implement similar countermeasures at additional locations.

- **Improve processes**: Evaluation can help to assess the HSIP process from start to finish, identifying opportunities to improve planning, implementation, evaluation, and documentation processes and procedures. For example, as part of countermeasure or program evaluation, it is useful to examine crashes that occur during construction to identify work zone configurations or construction practices that support State safety performance goals.

- **Demonstrate accountability**: There is an ever increasing demand for accountability at all levels of government. HSIP evaluation can help agencies to measure progress toward achieving their long-term safety goals and annual safety performance targets.

- **Meet federal requirements**: 23 CFR Part 924 requires each State to develop, implement, and evaluate on an annual basis a HSIP that has the objective to significantly reduce fatalities and serious injuries resulting from crashes on all public roads.

This guide provided an introduction to HSIP evaluation and a description of the considerations, methods, and data requirements for various evaluation techniques. Each chapter focused on specific aspects of HSIP evaluation from project tracking and evaluation of individual projects to program evaluation and reporting. Examples highlighted successful practices from several States and the end of each chapter provided a summary of tips and tricks to prepare for and perform successful HSIP evaluation. The following is a brief review of each chapter.

Chapter 2 provided an overview of general considerations and opportunities to prepare for HSIP evaluation. To establish and sustain a successful HSIP evaluation practice, agencies should gain HSIP staff and management support, allocate funding and staffing, and develop guidance, tools, and resources. Further, agencies should determine how the agency will use the results of HSIP evaluations to support future decisions.

Chapter 3 described practices to monitor and track individual projects. Project tracking provides the foundation for project, countermeasure, and program evaluations. Project tracking follows the project development process from planning and programming through construction and operations. By tracking projects from the beginning of the project development process, agencies can take advantage of early diagnosis efforts to document site conditions and crash history prior to any improvements. There is also an opportunity to enlist district and regional staff as well as local partners to support project tracking. With proper training and tools, these stakeholders can help to define project limits (begin and end mileposts), identify relevant construction dates (begin, end, and open to public), and provide detailed project information (project cost and type of improvement).

Chapter 4 focused on project evaluations. MOEs for individual projects typically include localized safety performance impacts and economic measures, and consider target and correctable crashes in addition to total, injury, and PDO crashes. For project evaluations, the simple before-after, before-after with traffic volume correction, and before-after with shift of proportions are potential evaluation methods. While more advanced before-after methods generally provide more reliable estimates than simple before-after studies, they also require additional data and more time to conduct the analysis, which may not be necessary for individual project evaluations.

Chapter 5 focused on countermeasure evaluations. Countermeasure evaluations aggregate data from multiple similar projects to estimate the safety effectiveness of a countermeasure. MOEs for countermeasure evaluations are similar to those for individual projects, and again, it is appropriate to focus on target crashes in addition to total crashes. The EB before-after method

is one of the more reliable methods for evaluating countermeasures because it can properly account for bias due to RTM, changes in traffic volume, and temporal effects. It is important to understand that countermeasures may have different effects under different conditions. Disaggregate analyses can identify differential effects by characteristics such as area type, geometry, traffic operations, geographic area, and crash history.

For both project and countermeasure evaluations, the study period considerations are similar. In establishing study periods, agencies should balance the duration with the potential for other changes over time. As such, the duration of the study period is generally ten years or less (three to five years before and after implementation) and it is important to account for changes other than the countermeasure of interest during the given period. It is possible to use different durations for the before and after period, but it becomes necessary to normalize the analysis based on the duration before and after. It is common to use full calendar years for ease of assembling data, and it is necessary to use 12-month increments to avoid seasonal bias. It is also common to exclude the implementation period from the countermeasure evaluation, but it may be useful to analyze the safety performance during implementation to estimate the safety effects of work zone configuration and related construction or implementation practices.

Chapter 6 focused on program evaluations. There are two primary types of program evaluation: crash-based and activity-based. Crash-based evaluations focus on the progress in meeting the safety goals of a program. Activity-based evaluations focus on the process and actions within a program. In general, each State is required to track and report on five performance measures based on five-year rolling averages: 1) number of fatalities, 2) rate of fatalities per 100 million VMT, 3) number of serious injuries, 4) rate of serious injuries per 100 million VMT, and 5) number of non-motorized fatalities and non-motorized serious injuries [23 CFR 490.207, 490.213, and 924.15]. Other potential performance measures for crash-based program evaluations include estimates of lives saved, net economic safety benefits and benefit-cost ratio, number of projects with crash reduction or significant crash reduction, difference in effectiveness on target and total crashes, percent change in crashes versus absolute number of crashes, effectiveness of older versus newer projects, effectiveness of HSIP-funded projects versus other projects, and effectiveness of projects by region. Examples of activity-based performance measures include the number of projects implemented, the timeliness of project implementation, a comparison of estimated project cost versus actual project cost, and the proportion of program funds allocated. For program evaluation, the data requirements are similar to project evaluation with the additional need to link specific projects to programs and subprograms.

Evaluation provides a critical feedback loop to improve processes and future decisions. Evaluations allow agencies to refine future estimates of effectiveness used to plan and manage projects, countermeasures, and programs. Similarly, agencies can use evaluation results to identify and address opportunities for improvement in HSIP processes. As States update their

SHSPs, they should consider how past HSIP projects have contributed to progress toward the SHSP goals and objectives. An improved understanding of the effectiveness of HSIP projects allows States to update infrastructure-related emphasis areas and strategies as well as the related targets and performance measures. When projects, countermeasures, and programs are effective (i.e., reduce the expected frequency and severity of crashes), agencies may choose to continue implementing similar improvements at additional locations. When efforts do not deliver the intended results or present implementation challenges, there may be an opportunity to modify the implementation process or consider alternative strategies.

In summary, HSIP evaluation is critical to understanding the return on investment and the effectiveness of past decisions. This guide can help agencies prepare for and conduct HSIP evaluations, and use the results of HSIP evaluations to inform future decisions and improve future investments.

APPENDIX A: PROJECT TRACKING

Appendix A lists several data fields for tracking HSIP projects. These fields can serve as a starting point for an agency to develop a spreadsheet or database to track and monitor projects implemented through HSIP. These fields provide a method for organizing projects by location, countermeasure type, costs, project date, and safety performance. For text-based categories, agencies should make an effort to codify responses to ease filtering, sorting, and evaluation. While agencies can use the existing fields to track both segment and intersection projects, an agency may prefer to develop and maintain separate spreadsheets for segment-based projects and intersection-based projects. In general, agencies should adjust the fields to their needs. The following is a brief description of each data field.

Project Location and Information: data provided under this category will help users identify the location of the project. The template provides data entry fields for Road 1 and Road 2. Road 1 is the primary field for segment-related projects and it represents the major road for intersection-related projects. Road 2 will generally be blank for segment-related projects and it represents the minor road for intersection-related projects. Regarding the major and minor road terminology, the major roadway refers to the higher volume roadway, while the minor roadway refers to the lower volume roadway. For systemic projects and projects where improvements occur intermittently between the begin and end milepost, consider adding a column to indicate specific treated locations or create separate entries for each improvement location, using the project identification number to link multiple improvement locations from the same project:

- Project ID #: a unique project identification number to link various data to a given project.
- Facility Type: type of facility for the project (e.g., intersection, ramp, road segment, etc.).
- District: State agency's district in which the project occurred.
- County: county in which the project occurred.
- Municipality: city/township/municipality in which the project occurred.
- Latitude: latitude of the project.
- Longitude: longitude of the project.
- Road 1 Route Number: route number for the major road in the project.
- Road 1 Route Name: route name for the major road in the project.
- Road 1 Before AADT: annual average daily traffic volume for the major road in the project for the before period.
- Road 1 After AADT: annual average daily traffic volume for the major road in the project for the after period.
- Road 1 Begin Milepost: the starting milepost for the major road in the project.

- Road 1 End Milepost: the ending milepost for the major road in the project.
- Road 1 FHWA Functional Classification: the FHWA Functional Classification of the major road on the project.
- Road 2 Route Number: route number for the minor road in the project (leave blank for segment projects).
- Road 2 Route Name: route name for the minor road in the project (leave blank for segment projects).
- Road 2 Before AADT: annual average daily traffic volume for the minor road in the project for the before period (leave blank for segment projects).
- Road 2 After AADT: annual average daily traffic volume for the minor road in the project for the after period (leave blank for segment projects).
- Road 2 Begin Milepost: the starting milepost for the minor road in the project (leave blank for segment projects).
- Road 2 End Milepost: the ending milepost for the minor road in the project (leave blank for segment projects).
- Road 2 FHWA Functional Classification: the FHWA Functional Classification of the minor road on the project (leave blank for segment projects).

Project Dates: data provided under this category will provide users with relevant dates for the project. This will also help analysts identify the before and after periods for analysis. These data include:

- Notice to Proceed: the date on which the agency allows the contractor to proceed.
- Begin Construction Date: the date on which the contractor begins construction.
- End Construction Date: the date on which the contractor ends construction.
- Open to Traffic: the date on which the facility is open to live traffic.

Safety Focus: this section summarizes the safety focus of improvements included in the project:

- SHSP Emphasis Area: the emphasis area of the Strategic Highway Safety Plan the project falls under (e.g., roadway departure, bikes, pedestrians, head-on, etc.).
- Project Selection Criteria: the pre-construction safety analysis method used to select the project (crash frequency, crash rate, systemic, etc.).
- Countermeasure 1: the first countermeasure implemented in the project.
- Target Crash Type 1.1: the first crash type targeted by Countermeasure 1.
- Target Crash Type 1.2: the second crash type targeted by Countermeasure 1.

- Target Crash Type 1.3: the third crash type targeted by Countermeasure 1.
- Countermeasure 2: the second countermeasure implemented in the project.
- Target Crash Type 2.1: the first crash type targeted by Countermeasure 2.
- Target Crash Type 2.2: the second crash type targeted by Countermeasure 2.
- Target Crash Type 2.3: the third crash type targeted by Countermeasure 2.
- Countermeasure 3: the third countermeasure implemented in the project.
- Target Crash Type 3.1: the first crash type targeted by Countermeasure 3.
- Target Crash Type 3.2: the second crash type targeted by Countermeasure 3.
- Target Crash Type 3.3: the third crash type targeted by Countermeasure 3.

Project Costs: this section summarizes the cost and funding source(s) of the project:

- Estimated Cost—Planning: the estimated cost of the project in the planning stage.
- Estimated Cost—Programming: the estimated cost of the project in the programming stage.
- Bid Cost: the winning bid cost.
- Paid Cost: the final paid cost by the agency for the project.
- Funding Source 1: the first funding source of the project (e.g., HSIP, High Risk Rural Roads Program, Incentive Grants, etc.).
- Funding Amount 1: the funding amount provided by the first funding source.
- Funding Source 2: the second funding source of the project.
- Funding Amount 2: the funding amount provided by the second funding source.

Fatal Crashes: observed crashes in which the highest injury severity is a fatality (K on KABCO scale):

- Before Year 1: observed crashes in the first before year.
- Before Year 2: observed crashes in the second before year.
- Before Year 3: observed crashes in the third before year.
- After Year 1: observed crashes in the first after year.
- After Year 2: observed crashes in the second after year.
- After Year 3: observed crashes in the third after year.

Serious Injury Crashes: observed crashes in which the highest injury severity is a suspected serious injury (A on KABCO scale), subcategories are the same as Fatal Crashes.

Evident Injury Crashes: observed crashes in which the highest injury severity is an evident injury (B on KABCO scale), subcategories are the same as Fatal Crashes.

Possible Injury Crashes: observed crashes in which the highest injury severity is a possible injury (C on KABCO scale), subcategories are the same as Fatal Crashes.

Property Damage Only Crashes: observed crashes in which the highest injury severity is property damage only (O on KABCO scale), subcategories are the same as Fatal Crashes.

Target Crash Type 1 Crashes: observed crashes in which the crash type is the same as Target Crash Type 1, subcategories are the same as Fatal Crashes.

Target Crash Type 2 Crashes: observed crashes in which the crash type is the same as Target Crash Type 2, subcategories are the same as Fatal Crashes.

Target Crash Type 3 Crashes: observed crashes in which the crash type is the same as Target Crash Type 3, subcategories are the same as Fatal Crashes.

The above crash categories provide data through which an analyst may perform a project-level evaluation. Combining multiple projects, the analyst may perform countermeasure or program evaluations.

Project Evaluation: this section summarizes the evaluation results of the project:

- B/C Ratio: calculated benefit-cost ratio of the project.
- Countermeasure 1 CMF: the estimated CMF for countermeasure 1.
- Countermeasure 2 CMF: the estimated CMF for countermeasure 2.
- Countermeasure 3 CMF: the estimated CMF for countermeasure 3.

Media: this category provides the user with a repository for photos and a place to link to any online presence for the project:

- Photos: may provide photos directly in the respective cells or provide a hyperlink to a set of photos.
- Webpage: provide a link to a website if there is any online presence for the project.

APPENDIX B: EVALUATION TEMPLATES

Appendix B includes evaluation templates with detailed examples. These templates and examples serve as a resource to perform project and countermeasure evaluations. Refer to chapters 4 and 5 for detailed discussions of project and countermeasure evaluation methods. Analysts can use these templates for both spot and systemic safety improvement evaluations as well as intersection and segment evaluations.

SIMPLE BEFORE-AFTER METHOD

Table 10 presents sample data for an example project and Table 11 provides a spreadsheet template and completed example for estimating project or countermeasure effectiveness using the simple before-after method. As an example, consider a scenario where an agency installed left-turn lanes on both approaches of the major road at a two-way stop-controlled intersection along a rural, two-lane highway. The analyst would like to estimate the safety effectiveness and standard error for this project (i.e., installing left-turn lanes) using the simple before-after method. Using the spreadsheet template in Table 11, the green cells represent user inputs. The yellow cells represent the outputs, computed automatically based on the user inputs.

Table 10. Sample project data for simple before-after evaluation.

Treatment Site Data	Before	After
Total observed crashes	18	10
Duration (years)	3	2
Traffic volume (vehicles/day) from permanent traffic count station (AADT estimated from 365-day count)	7,500	8,300

Table 11. Spreadsheet for estimating project or countermeasure effectiveness using the simple before-after method.

Excel Row	Variable Inputs (Column A)	Excel Formula (Column B)	Example
1	Number of Observed Crashes "Before" in Treatment Group	User Input	18
2	Traffic Volume Before	User Input	7,500
3	Years Before	User Input	3
4	Number of Observed Crashes "After" in Treatment Group	User Input	10
5	Traffic Volume After	User Input	8,300
6	Years After	User Input	2
7	Number of Count Days to Estimate AADT Before	User Input	365
8	Number of Count Days to Estimate AADT After	User Input	365
9	Adjustment for Duration of Before and After Period [Years After/Years Before]	=B6/B3	0.67
10	Adjustment for Change in Traffic Volume [Traffic After/Traffic Before]	=B5/B2	1.11
11	Estimated Number of Crashes "After" in Treatment Group Without Change	=B1*B9*B10	13.28
12	Variance of Observed Crashes "After" in Treatment Group	=B4	10
13	Coefficient of Variation (v) Before	=1+(7.7/B7)+(1650/(B2)^0.82)	2.12
14	Coefficient of Variation (v) After	=1+(7.7/B8)+(1650/(B5)^0.82)	2.03
15	Variance of Adjustment for Change in Traffic Volume	=B10^2*(((B13/100)^2)+((B14/100)^2))	0.0011
16	Variance of Estimated Number of Crashes "After" in Treatment Group Without Change	=B9^2*(((B10^2)*B1)+((B1^2)*B15))	9.95
17	Estimate of Effectiveness	=(B4/B11)/(1+(B16/(B11^2)))	0.71
18	Variance of Estimate of Effectiveness	=(B17^2)*((B12/(B4^2))+(B16/(B11^2)))/ (1+(B16/(B11^2))^2)	0.08
19	Standard Error of Estimate of Effectiveness	=SQRT(B18)	0.28

BEFORE-AFTER WITH COMPARISON GROUP METHOD

Table 12 presents sample data for an example countermeasure and Table 13 provides a spreadsheet template and completed example for estimating project or countermeasure effectiveness using the before-after with comparison group method. As an example, consider a scenario where an agency installed advance warning flashers at 25 signalized intersections and identified a representative comparison group, including 25 similar signalized intersections without advance warning flashers. The analyst would like to estimate the safety effectiveness and standard error for this countermeasure (i.e., installing advance warning flashers) using the before-after with comparison group method. Using the spreadsheet template in Table 13, the green cells represent user inputs. The yellow cells represent the outputs, computed automatically based on the user inputs. This assumes the before and after periods are the same for the treatment group and comparison group.

Table 12. Sample data for before-after with comparison group evaluation.

Input Data	Before	After
Total observed crashes for treatment group	100	75
Total observed crashes for comparison group	84	80
Duration (years)	3	3

Table 13. Spreadsheet for estimating project or countermeasure effectiveness using the before-after with comparison group method.

Excel Row	Variable Inputs (Column A)	Excel Formula (Column B)	Example
1	Number of Observed Crashes "Before" in Treatment Group	User Input	100
2	Number of Observed Crashes "After" in Treatment Group	User Input	75
3	Number of Observed Crashes "Before" in Comparison Group	User Input	84
4	Number of Observed Crashes "After" in Comparison Group	User Input	80
5	Estimated Number of Crashes "After" in Treatment Group Without Change	=B1*(B4/B3)	95.2
6	Variance of Observed Crashes "After" in Treatment Group	=B2	75
7	Variance of Estimated Number of Crashes "After" in Treatment Group Without Change	=(B5^2)*((1/B1)+(1/B3)+(1/B4))	312.1
8	Estimate of Effectiveness (CMF)	=(B2/B5)/(1+(B7/(B5^2)))	0.76
9	Variance of CMF	=(B8^2)*((B6/(B2^2))+(B7/(B5^2)))/ (1+(B7/(B5^2))^2)	0.03
10	Standard Error of CMF	=SQRT(B9)	0.17

EMPIRICAL BAYES BEFORE-AFTER METHOD

Table 14 presents sample data for an example project and Table 15 provides a spreadsheet template for estimating project or countermeasure effectiveness using the EB before-after method. As an example, consider a scenario where an agency installed left-turn lanes on both approaches of the major road at a two-way stop-controlled intersection along a rural, two-lane highway. The analyst would like to estimate the safety effectiveness and standard error for this project (i.e., installing left-turn lanes) using the EB before-after method. Using the spreadsheet template in Table 15, the green cells represent user inputs. The yellow cells represent the outputs, computed automatically based on the user inputs. This assumes the analyst has developed or calibrated an SPF and dispersion parameter. It also assumes the annual crashes are "predicted" using an SPF and appropriate calibration factors. For this example, Figure 17 presents the uncalibrated SPF and Table 14 presents the calibration and dispersion parameters.

Table 14. Sample data for EB before-after evaluation.

Treatment Site Data	Before	After
Total observed crashes	18	10
Duration (years)	3	2
Traffic volume major road (vehicles/day)	7,000	7,700
Traffic volume minor road (vehicles/day)	500	600
Calibration Factors for SPF	0.69	1.08
Dispersion parameter of SPF (k)	0.24	0.24

$$SPF = \exp\left[-8.56 + 0.60 * \ln\left(AADT_{major\ road}\right) + 0.61 * \ln(AADT_{minor\ road})\right]$$

Figure 17. Equation. Sample SPF for EB before-after evaluation.

Table 15. Spreadsheet for estimating project or countermeasure effectiveness using the EB before-after method.

Excel Row	Variable Inputs (Column A)	Excel Formula (Column B)	Example
1	Number of Observed Crashes "Before" in Treatment Group	User Input	18
2	Number of Observed Crashes "After" in Treatment Group	User Input	10
3	Number of Predicted Crashes "Before" in Treatment Group	User Input: Sum of predicted annual crashes for each year in the before period	3.56
4	Number of Predicted Crashes "After" in Treatment Group	User Input: Sum of predicted annual crashes for each year in the after period	4.40
5	Dispersion parameter of SPF (k)	User Input	0.24
6	SPF Weight	=1/(1+B5*B3)	0.54
7	Expected Number of Crashes "Before" in Treatment Group	=B6*B3+(1-B6)*B1	10.22
8	Expected Number of Crashes "After" in Treatment Group Without Change	=B7*(B4/B3)	12.62
9	Variance of Observed Crashes "After" in Treatment Group	=B2	10
10	Variance of Expected Number of Crashes "After" in Treatment Group Without Change	=(B8^2)*((B4/B3)*(1-B6))	90.63
11	Estimate of Effectiveness (CMF)	=(B2/B8)/(1+(B10/(B8^2)))	0.51
12	Variance of CMF	=(B11^2)*((B9/(B2^2))+(B10/(B8^2)))/ (1+(B10/(B8^2))^2)	0.13
13	Standard Error of CMF	=SQRT(B12)	0.36

APPENDIX C: SAMPLE SIZE TEMPLATES

Appendix C includes sample size templates with detailed examples. The first template and example serve as a resource to estimate the required sample size to detect an expected change in safety performance at a given level of significance. The second template and example serve as a resource to estimate the minimum detectable effect based on the desired level of significance and a given sample size.

CASE A: WHAT SAMPLE SIZE IS REQUIRED TO DETECT AN EXPECTED CHANGE AT A GIVEN LEVEL OF SIGNIFICANCE?

Using the spreadsheet template in Table 16, the green cells represent user inputs. The yellow cells represent the outputs, computed automatically based on the user inputs. In this case, the sample size is the question at hand, so that will be the last input. Assuming the number of years before and after implementation is one, the sample size represents the total crashes required in a given period (before or after). Input the desired level of significance, the variance of the odds ratio, and the expected reduction. For the variance of the odds ratio, Hauer recommends examining the sensitivity assuming values from 0.001 to 0.01.[23] Refer to section 9.3 of Hauer (1997) for more information related to the variance of the odds ratio.[23]

With these variables input, enter a value for the number of "before" crashes per year in the treatment group. Adjust the value until the upper bound of the confidence interval is 1.0. If the upper bound is greater than 1.0 (not statistically significant), then increase the sample size. If the upper bound is less than 1.0 (statistically significant), then decrease the sample size until the upper bound is 0.9999. When the upper bound dips below 1.0, the associated sample size is the sample required to obtain significant results for the given level of effectiveness and significance. It is useful to start by changing the sample size by units of 100, then 10, and then 1 until the analyst identifies the threshold.

CASE A EXAMPLE: WHAT SAMPLE SIZE IS REQUIRED TO DETECT A 20 PERCENT LEVEL OF EFFECT (CMF = 0.80) AT A 0.10 LEVEL OF SIGNIFICANCE?

In this example, the analyst would like to estimate the required sample size to detect a 20 percent level of effect (CMF = 0.80) at a 0.10 level of significance. Using the spreadsheet template, input the desired level of significance (0.10 in this example), the variance of the odds ratio (0.001 in this example), and the expected reduction (20 percent in this example). For the sample size, start with 100 crashes per year for the treatment group. In this case, the sample size is too small because the upper bound of the confidence interval is greater than 1.0. Next, try 200 crashes per year; the sample size is too large because the upper bound of the confidence interval is less than 1.0. Now, work back from 200 crashes by increments of 5 or

10, and then work up from 190 crashes by increments of 1 until the upper bound bounces above and below 1.0 as the number of crashes changes from 192 to 193. In this example, the required number of crashes is 193. The analysis assumes the number of comparison sites is equal to the number of treated sites and the duration of the before and after periods are equal. As such, the required sample is 193 crashes in the before and after periods for both the treatment and comparison groups. If the before and after periods are different or if the number of comparison sites is not equal to the number of treatment sites, then the user can change the number of before and after years in cells B2 and B3 and enter a different number of "before" crashes per year for the comparison group in cell B4.

Table 16. Spreadsheet for estimating sample size requirements for observational before-after evaluations.

Excel Row	Variable Inputs (Column A)	Excel Formula (Column B)	Example
1	Number of "before" crashes per year in treatment group	User Input	193
2	Number of "before" years	1	1
3	Number of "after" years	1	1
4	Number of "before" crashes per year in comparison group	=B1	193
5	Variance of odds ratio	User Input	0.001
6	Desired level of significance (α)	User Input	0.1
7	Cumulative probability	=ABS(NORMSINV(B6/2))	1.64
8	Expected % reduction [100*(1-CMF)]	User Input	20
9	Number of "before" crashes in treatment group	=B1*B2	193
10	Estimated number of "after" crashes in treatment group	=B1*B3*(1-B8/100)	154.4
11	Number of "before" crashes in comparison group	=B4*B2	193
12	Estimated number of "after" crashes in comparison group	=B4*B3	193
13	Estimated number of "after" crashes in treatment group without change	=B9*B12/B11	193
14	Estimate of the variance of crashes "after" without change	=B9*(B12/B11)^2*(1+B9*B5+B9/B11+B9/B12)	616.249
15	Estimated index of effectiveness [CMF]	=B10/B13	0.8
16	Standard deviation of the estimated index of effectiveness [SE(CMF)]	=(B15^2(1/B10+B14/B13^2))^0.5	0.121
17	Lower bound of confidence interval	=B15-B7*B16	0.6003
18	Upper bound of confidence interval	= B15+B7*B16	0.9997

CASE B: WHAT CHANGE CAN BE DETECTED AT A GIVEN LEVEL OF SIGNIFICANCE WITH A GIVEN SAMPLE SIZE?

Using the spreadsheet template in Table 16, the green cells represent user inputs. The yellow cells represent the outputs, computed automatically based on the user inputs. In this case, the sample size is known and the question at hand is whether the analysis is likely to produce a statistically significant effect. Assuming the number of years before and after implementation is one, enter the sample size as the total crashes available in a given period (before or after). Input the desired level of significance and the variance of the odds ratio. For the variance of the odds ratio, Hauer recommends examining the sensitivity assuming values from 0.001 to 0.01.[23] Refer to section 9.3 of Hauer for more information related to the variance of the odds ratio.[23]

With these variables input, enter a value for the countermeasure effect (i.e., expected reduction). Adjust the value until the upper bound of the confidence interval is 1.0. If the upper bound is greater than 1.0 (not statistically significant), then increase the expected reduction. If the upper bound is less than 1.0 (statistically significant), then decrease the expected reduction until the upper bound is 0.9999. When the upper bound dips below 1.0, the expected reduction is the minimum detectable at the desired level of significance with the available sample size. It is useful to start by changing the expected reduction by units of 10, then 5, and then 1 until the analyst identifies the threshold.

CASE B EXAMPLE: WHAT MAGNITUDE OF EFFECT CAN BE DETECTED AT 0.10 LEVEL OF SIGNIFICANCE WITH 70 CRASHES PER YEAR, 3 YEARS OF BEFORE DATA, AND 3 YEARS OF AFTER DATA?

In this example, assume the analyst would like to determine the detectable effect at a 0.10 level of significance (90 percent confidence) with a sample of 70 crashes per year, three years of before data, and three years of after data. The available data are for the treatment group. Based on this information, there are 210 "before" crashes (3 years * 70 crashes per year) in the treatment group.

Using the spreadsheet template in Table 16, enter the total "before" crashes in the treatment group, the desired level of significance (0.10 in this example), and the variance of the odds ratio (0.001 in this example). Now, vary the input for the expected reduction until the last yellow cell (upper bound of confidence interval) drops below 1.0. Start with a 5 percent reduction. In this case, the effect is too small because the upper bound of the confidence interval is greater than 1.0. Now, try 10 percent. Again, the 10 percent reduction is not detectable because the confidence interval includes 1.0. Now, try 20 percent. A 20 percent effect is detectable at the 0.1 significance level because the confidence interval does not include 1.0. At this point, work back from 20 in increments of 1 until the upper bound of the confidence interval dips below 1.0. For this example, the analyst could detect no less than a 17 percent effect with the available

sample at the 0.1 significance level (90 percent confidence). The analyst would need to consider if the detectable effect size is reasonable for the given countermeasure based on other studies or anecdotal evidence. The analyst would also need to collect a similar sample for a comparison group because the available data is for the treatment group only.

APPENDIX D: ADDITIONAL RESOURCES

Appendix D provides an overview of resources that provide additional information and assistance regarding data collection, management, analysis, and evaluation techniques. There is a brief overview, indication of the relevance to evaluation, and a link to the resource or additional information for each of the following resources:

1. **2010 HSIP Manual**[1]

   - **Overview**: This manual provides an overview of the HSIP and offers practitioners with a review of current standards, new and emerging technologies, and noteworthy practices for each step in the HSIP process. This is an updated version of the original 1981 manual and supplement.

   - **Relevance**: Overview of HSIP and Evaluation.

2. **A Guide to Developing Quality Crash Modification Factors**[20]

   - **Overview**: This guide provides the reader with an introduction to CMFs, helps the reader learn to assess the quality of existing CMFs, explains numerous methodologies to develop CMFs, and helps the reader to select an appropriate evaluation method based on the objectives and available data.

   - **Relevance**: Countermeasure Evaluation.

3. **Developing an Effective Evaluation Plan**[4]

   - **Overview**: This document provides guidance for developing a living "Evaluation Plan" for public health programs. In the case of highway safety, State transportation agencies may consider the knowledge and guidance in this document for management, evaluation, and improvement of their HSIP. The document helps program administrators identify answers to three questions about their program: "What?", "How?", and "Why it matters?". The six-step process presented in the document will help provide these answers as well as assist program managers with improving their plan.

   - **Relevance**: Program Evaluation.

4. **Developing an Effective Evaluation Report**[29]

   - **Overview**: This document builds upon the "Developing an Effective Evaluation Plan" document by providing guidance for producing a final evaluation report. This product summarizes how the program was monitored and evaluated as well as the review's findings and possible improvements to be implemented into the program.

   - **Relevance**: Program Evaluation.

5. **Manual of Transportation Engineering Studies**[21]

   o **Overview**: This manual is an updated and expanded version of the Manual of Traffic Engineering Studies, 4th Edition. The primary focus of this manual is on "how to conduct" transportation engineering studies in the field, particularly for conducting non-crash-based evaluations. While these evaluations are well-suited for certain situations (i.e., proactive projects and projects involving staged countermeasure implementation), they do not replace crash-based evaluations. Surrogate measures such as conflicts, speeds, and violations can serve as means to evaluate intermediate project effectiveness, but the ultimate measure of safety effectiveness should be the change in crash frequency and severity.

   o **Relevance**: Non-crash-based evaluations.

6. **Recommended Protocols for Developing Crash Modification Factors**[22]

   o **Overview**: This guide provides an overview of CMF development that focuses more on methodology than FHWA's A Guide to Developing Quality CMFs. Specifically, users will learn to understand the strengths and weaknesses of various methods and how to identify the appropriate method for CMF development given certain conditions. This document can serve as a companion to "A Guide to Developing Quality Crash Modification Factors," providing more details on potential sources of bias, opportunities to address sources of bias, and important information to document in evaluation reports.

   o **Relevance**: Countermeasure Evaluation.

7. **Systemic Safety Project Selection Tool**[25]

   o **Overview**: This document provides guidance for a crash potential-based approach to identifying and addressing systemic safety opportunities. It provides guidance for identifying and implementing systemic countermeasures and methods for assessing their effectiveness.

   o **Relevance**: Project Evaluation and Program Evaluation.

8. **The Art of Appropriate Evaluation: A Guide for Highway Safety Program Managers**[3]

   o **Overview**: This guide provides an overview of the traffic safety evaluation process, specifically for highway safety program managers. The guide helps the manager identify the appropriate evaluation method for a project as well as how to choose a well-qualified professional evaluator for a highway safety program.

   o **Relevance**: Project Evaluation and Program Evaluation.

The following are resources for assessing an HSIP program and identifying noteworthy practices:

1. **HSIP Assessment Toolbox**[10]

   o **Overview**: This toolbox provides support for agencies to conduct an assessment of their HSIP. Although not required, it is recommended that agencies review their program every five years to evaluate the success of the process. This toolbox provides strategies, methods, and best-practices for agencies to consider incorporating into their program.

   o **Relevance**: Program Evaluation.

2. **HSIP National Scan Tour Report**[11]

   o **Overview**: This report provides a summary of notable practices in the areas of HSIP administration, planning, implementation, and evaluation. Specific highlights include practices related to documenting HSIP processes, coordinating with internal and external partners, understanding the relationship between the SHSP and HSIP, making data-driven safety decisions, using advanced safety analysis methods and tools, addressing local road needs, and identifying opportunities to streamline project delivery.

   o **Relevance**: Program Evaluation.

3. **HSIP Noteworthy Practice Series**[12]

   o **Overview**: This series provides examples from around the United States of best-practices for various aspects of the HSIP process. It provides a series of case studies with noteworthy examples for other agencies to consider incorporating into their programs.

   o **Relevance**: Program Evaluation.

4. **HSIP Self-Assessment Tool**[13]

   o **Overview**: This tool provides a question-based method for managers to perform a self-evaluation of an agency's HSIP. Use of this tool will help agencies track progress, improve strategy development, identify opportunities for improvement with their current program, and ultimately improve their process.

   o **Relevance**: Program Evaluation.

An important element of the Guide is the data required to perform the evaluations. Chapters 4 through 6 describe the data requirements for project, countermeasure, and program level evaluations, respectively. Refer to the FHWA Safety Data and Analysis Toolbox to search for data-related resources. The toolbox allows users to perform customized searches and provides

an overview and summary of related capabilities to help the user determine if the resource is appropriate for their needs.

Training and technical assistance are other important elements related to HSIP evaluation. The following is a brief summary of training opportunities related to HSIP evaluation, including a summary, relevance to evaluation, and link to further information.

1. **NHI Course 380103: Highway Safety Improvement Program Manual**

   o **Overview**: This in-person instructor-led course introduces safety professionals to the HSIP process, including a basic introduction to evaluation at the project and program levels. It provides information on topics ranging from core safety concepts to detailed discussions of technical methods for data-driven safety planning which will result in successful HSIP efforts.

   o **Relevance**: Overview of HSIP Evaluation and Program Evaluation.

2. **NHI Course 380112: HSIP Project Evaluation**

   o **Overview**: This web-based course provides a description of safety effectiveness evaluation, an overview of fundamentals needed to perform safety effectiveness evaluation, and information about why safety effectiveness evaluation is important to a State's HSIP. Examples of project evaluation methodologies that account for regression-to-the-mean are discussed and participants are given an opportunity to calculate simple observational before-after studies, observational before-after studies with EB adjustment, and observational before after studies using comparison groups.

   o **Relevance**: Project Evaluation.

3. **NHI Course 380119: Developing Quality Crash Modification Factors**

   o **Overview**: This 30+ hour course features a blend of self-paced and instructor-led training, focusing on developing quality CMFs.

   o **Relevance**: Countermeasure Evaluation.

4. **NHI Course 380094: Science of Crash Modification Factors**

   o **Overview**: This web-based instructor-led course provides participants with the knowledge and skills needed to critically assess the quality of CMFs. The course covers concepts underlying the measurement of safety and the development of CMFs, and key statistical and methodological issues that affect the development of quality CMFs.

   o **Relevance**: Countermeasure Evaluation.

The following is a brief description of available technical assistance related to HSIP evaluation:

1. **FHWA Office of Safety Technical Assistance Program**
   - o **Overview**: This includes free technical assistance on policy, program, and technical issues to State and local roadway agencies. Agencies can request tailored technical assistance to help assess, develop, implement, or evaluate effective safety strategies and programs. Delivery mechanisms include remote assistance (telephone and email), one-on-one onsite assistance, training workshops, and facilitated peer exchanges.

GLOSSARY

After period: time after construction or implementation of the countermeasure.

Analysis: process of using qualitative and quantitative data to support evaluation. Qualitative analysis focuses on the strengths, limitations, and opportunities to improve processes and procedures. Quantitative analysis focuses on the cost, effectiveness, and resilience (how long it is effective) for each project, countermeasure, or program.

Before period: time before implementation of countermeasure, and typically includes the time before and during project development.

Before-after study: the evaluation of safety projects, comparing the frequency, severity, or rate of crashes before and after implementation. There are several different types of before-after studies.

Comparison group: a group of untreated sites used in before-and-after studies to control for changes over time other than the countermeasure of interest.

Components: the three general phases of the HSIP: planning, implementation, and evaluation.

Countermeasure: a roadway based strategy intended to reduce the crash frequency, crash severity, or both at a site (synonymous with treatment).

Crash Modification Factor: a multiplicative factor that indicates the expected change in crashes associated with a countermeasure.

Evaluation: a component of the HSIP, including the analysis of crash- and non-crash-based measures of effectiveness to develop quantitative estimates of the effect a project, countermeasure, or program.

Five-year rolling average: the average of five individual, consecutive annual points of data (e.g., the five-year rolling average of the annual fatality rate).

Implementation: a component of the HSIP process, including the construction of a project.

Interim period: time between site selection and project implementation.

KABCO: the coding convention system for injury classification established by the National Safety Council. Fatal injury (K), Suspected Serious Injury (A), Suspected Minor Injury (B), Possible Injury (C), and No Apparent Injury (O).

Method: the specific technique uses to perform an evaluation.

Non-motorist: consistent with 23 U.S.C. 217(j), non-motorists are those transportation system users who are not in or on traditional motor vehicles on public roadways. This includes persons traveling by foot, children in strollers, skateboarders (including motorized), roller skaters, persons on scooters, persons in wagons, persons in wheelchairs (both non-motorized and motorized), persons riding bicycles or pedalcycles (including those with a low-powered electric motor weighing under 100 pounds, with a top motor-powered speed not in excess of 20 miles per hour), persons in motorized toy cars, and persons on two-wheeled, self-balancing types of devices.

Number of fatalities: total number of persons suffering fatal injuries in a motor vehicle traffic crash within 30 days as a result of a crash.

Number of non-motorized fatalities: FHWA defines the fatally injured non-motorist person, i.e. the "person type," defined in FARS, to include the person level attribute codes for (5) Pedestrians, (6) Bicyclists, (7) Other Cyclists, and (8) Persons on Personal Conveyances.

Number of non-motorized serious injuries: FHWA defines the seriously injured person type as the codes and definitions for a (2.2.36) pedestrian or (2.2.39) pedalcyclist in the American National Standard (ANSI) D16.1-2007 Manual on Classification of Motor Vehicle Traffic Accidents.

Number of serious injuries: total number of persons suffering suspected serious injuries in a motor vehicle traffic crash.

Planning: a component of the HSIP process, including planning and programming, environmental review, and preliminary and final design.

Procedures: the possible ways in which an agency can perform each of the processes. For example, the procedures for conducting project evaluations include the simple before-after method and shift of proportions method.

Processes: the sequential elements within each component of the HSIP.

Rate of fatalities: ratio of total number of fatalities to the number of vehicle miles traveled (expressed in 100 million VMT).

Rate of serious injuries: ratio of total number of suspected serious injuries to the number of vehicle miles traveled (expressed in 100 million VMT).

Regression-to-the-mean: the situation when periods with relatively high crash frequencies are followed by periods with relatively low crash frequencies (and vice versa) simply due to the variability of crashes, and not the project in question.

Serious injuries: from April 14, 2016 to April 15, 2019, injuries classified as "A" on the KABCO scale through use of the conversion tables developed by NHTSA; and after April 15, 2019, "suspected serious injury (A)" as defined in the Model Minimum Uniform Crash Criteria.

Surrogate: an indirect measure of safety performance when crash-based measures are not available. For example, crash-based measures may be insufficient for projects completed within the last two years.

Treatment: a roadway based strategy intended to reduce the crash frequency, crash severity, or both at a site (synonymous with countermeasure.

ACKNOWLEDGEMENTS

The project team would like to thank the following individuals who provided information and examples related to HSIP evaluation.

- Adam Weiser, Delaware Department of Transportation
- Adriane McRae, Louisiana Department of Transportation and Development
- Brandon Darks, Tennessee Department of Transportation
- Charles Meyer, Colorado Department of Transportation
- Dahir Egal, Colorado FHWA Division Office
- Dan Hinton, South Carolina FHWA Division Office
- Darren McDaniel, Texas Department of Transportation
- David Morena, Michigan FHWA Division Office
- David Swenka, Colorado Department of Transportation
- Filiberto Sotelo, Illinois Department of Transportation
- Ida Van Schalkwyk, Washington State Department of Transportation
- Jarrod Stanley, Kentucky Transportation Cabinet
- Jeff Murphy, Tennessee Department of Transportation
- Jessica Rich, Tennessee FHWA Division Office
- Joe Santos, Florida Department of Transportation
- John Milton, Washington State Department of Transportation
- Karen King, Virginia FHWA Division Office
- Ken Kochevar, California FHWA Division Office
- Lance Johnson, Idaho FHWA Division Office
- Marc Thornsberry, Missouri FHWA Division Office
- Marcee Allen, Montana FHWA Division Office
- Matt Walker, Alaska Department of Transportation and Public Facilities
- Michael McNeill, Ohio Department of Transportation
- Mike Castellano, Pennsylvania FHWA Division Office
- Regina Doyle, New York State Department of Transportation
- Rob Limoges, New York State Department of Transportation
- Robert Peterson, California Department of Transportation
- Roland Stanger, Utah FHWA Division Office
- Ryan Tenges, Kentucky FHWA Division Office
- Scott Jones, Utah Department of Transportation
- Sean Raymond, Rhode Island Department of Transportation
- Shawn Troy, North Carolina Department of Transportation
- Stephen Read, Virginia Department of Transportation
- Thomas Schriber, California Department of Transportation
- Tracy Lovell, Kentucky Transportation Cabinet

REFERENCES

1. Herbel, S., L. Laing, and C. McGovern. *Highway Safety Improvement Program (HSIP) Manual*, Report FHWA-SA-09-029, Federal Highway Administration, Washington, DC, 2010.

2. US Department of Transportation. *Highway Safety Evaluation: Procedural Guide*, Report FHWA-TS-81-219, Federal Highway Administration, Washington, DC, 1981.

3. US Department of Transportation. *The Art of Appropriate Evaluation: A Guide for Highway Safety Program Managers*, Report DOT HS 811 061, National Highway Traffic Safety Administration, Washington, DC, 2008.

4. *Developing an Effective Evaluation Plan: Setting the course for effective program evaluation*. Atlanta, Georgia: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; Division of Nutrition, Physical Activity, and Obesity, 2011.

5. American Association of State Highway Transportation Officials (AASHTO). *Highway Safety Manual*, 1st Edition, Washington, DC, 2010.

6. *Local Road Safety: A Manual for California's Local Road Owners*. California Department of Transportation (Caltrans), Version 1.3, April 2016.

7. *District Highway Safety Guidance Manual*. Pennsylvania Department of Transportation (PennDOT), PUB 638 (12-14), December 2014.

8. *Highway Safety Improvement Program Manual*. Utah Department of Transportation (UDOT), Division of Traffic and Safety, March 2014.

9. *Highway Safety Improvement Program Implementation Guidelines*. Virginia Department of Transportation (VDOT), Traffic Engineering Division, 2015.

10. Tang, E. *Highway Safety Improvement Program Assessment Toolbox*, 2nd Edition, Report FHWA-SA-15-015, Federal Highway Administration, Washington, DC, 2015.

11. Kuznicki, S., B. Ibarguen, J. Symoun, and P. Pavao. *Highway Safety Improvement Program National Scan Tour*, Report FHWA-SA-16-024, Federal Highway Administration, Washington, DC, 2016.

12. US Department of Transportation. *HSIP Noteworthy Practice Series*, Report FHWA-SA-11-02, Federal Highway Administration, Washington, DC, 2011.

13. Tang, E. *Highway Safety Improvement Program Self-Assessment Tool*, 2nd Edition, Report FHWA-SA-15-014, Federal Highway Administration, Washington, DC, 2015.

14. *Alaska Highway Safety Improvement Program Handbook: Methodology for Identifying, Prioritizing and Evaluating Highway Safety Improvement Program Projects*. 17th Edition, Alaska Department of Transportation and Public Facilities, January 1, 2017.

15. *Highway Safety Improvement Program Reporting Guidance*. Federal Highway Administration, December 29, 2016.

16. Srinivasan, R., F. Gross, and G. Bahar. *Reliability of Safety Management Methods: Safety Effectiveness Evaluation*. Report FHWA-SA-16-040, Federal Highway Administration, Washington, DC, 2016.

17. Tillman, R., G. Khan, and X. Qin. *Wisconsin Highway Safety Improvement Program (HSIP) Evaluation: Year-Four Support*, Final Report, Traffic Operations and Safety (TOPS) Laboratory, August 2009.

18. *Before/After Safety Analyses II*. FHU Reference No. 114271-13, Colorado Department of Transportation (CDOT), Traffic and Safety Engineering Branch, December 2016.

19. Robinson, B. *Spot Safety Project Evaluation of the Sight Distance Improvements and Installation of a Four-Way Stop at the Intersection of SR 1126 (Forlines Rd) and SR 1127 (Frog Level Rd) Pitt County*. Spot Safety Project # 02-02-243, North Carolina Department of Transportation, Safety Evaluation Group, August 13, 2010.

20. Gross, F., B. Persaud, and C. Lyon. *A Guide to Developing Quality Crash Modification Factors*, Report FHWA-SA-10-032, Federal Highway Administration, Washington, DC, 2010.

21. Schroeder, B.J., C.M. Cunningham, D.J. Findley, J.E. Hummer, and R.S. Foyle. *Manual of Transportation Engineering Studies*, 2nd Edition, Institute of Transportation Engineers, 2010.

22. Carter, D., R. Srinivasan, F. Gross, and F. Council. *Recommended Protocols for Developing Crash Modification Factors*, National Cooperative Highway Research Program, Transportation Research Board, February 2012.

23. Hauer, E. (1997), *Observational Before-After Studies in Road Safety*, Pergamon, Elsevier Science Ltd.

24. Pendleton, O. *Evaluation of Accident Analysis Methodology*. Report No. FHWA-RD-96-039, Federal Highway Administration, Washington, DC, 1996.

25. Preston, H., R. Storm, J.D. Bennett, and B. Wemple. *Systemic Safety Project Selection Tool*, Report FHWA-SA-13-019, Federal Highway Administration, Washington, DC, 2013.

26. *Solutions for Saving Lives on Texas Roads*. Texas Department of Transportation, June 30, 2016.

27. *Estimating the Costs of Unintentional Injuries*. National Safety Council, 2014.

28. Council, F., Zaloshnja, E., Miller, T., and Persaud, B. *Crash Cost Estimates by Maximum Police-Reported Injury Severity within Selected Crash Geometries*. Report FHWA-HRT-05-051, Federal Highway Administration, McLean, VA, 2005.

29. *Developing an Effective Evaluation Report: Setting the course for effective program evaluation*. Atlanta, Georgia: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Division of Nutrition, Physical Activity and Obesity, 2013.