

# **NCHRP**

## **REPORT 501**

**NATIONAL  
COOPERATIVE  
HIGHWAY  
RESEARCH  
PROGRAM**

### **Integrated Safety Management Process**

**TRANSPORTATION RESEARCH BOARD**  
*OF THE NATIONAL ACADEMIES*

## **TRANSPORTATION RESEARCH BOARD EXECUTIVE COMMITTEE 2003 (Membership as of March 2003)**

### **OFFICERS**

**Chair:** *Genevieve Giuliano, Director and Professor, School of Policy, Planning, and Development, University of Southern California, Los Angeles*

**Vice Chair:** *Michael S. Townes, Executive Director, Transportation District Commission of Hampton Roads, Hampton, VA*

**Executive Director:** *Robert E. Skinner, Jr., Transportation Research Board*

### **MEMBERS**

MICHAEL W. BEHRENS, *Executive Director, Texas DOT*

JOSEPH H. BOARDMAN, *Commissioner, New York State DOT*

SARAH C. CAMPBELL, *President, TransManagement, Inc., Washington, DC*

E. DEAN CARLSON, *Secretary of Transportation, Kansas DOT*

JOANNE F. CASEY, *President, Intermodal Association of North America*

JAMES C. CODELL III, *Secretary, Kentucky Transportation Cabinet*

JOHN L. CRAIG, *Director, Nebraska Department of Roads*

BERNARD S. GROSECLOSE, JR., *President and CEO, South Carolina State Ports Authority*

SUSAN HANSON, *Landry University Professor of Geography, Graduate School of Geography, Clark University*

LESTER A. HOEL, L. A. *Lacy Distinguished Professor, Department of Civil Engineering, University of Virginia*

HENRY L. HUNGERBEELER, *Director, Missouri DOT*

ADIB K. KANAFANI, *Cahill Professor and Chairman, Department of Civil and Environmental Engineering, University of California at Berkeley*

RONALD F. KIRBY, *Director of Transportation Planning, Metropolitan Washington Council of Governments*

HERBERT S. LEVINSON, *Principal, Herbert S. Levinson Transportation Consultant, New Haven, CT*

MICHAEL D. MEYER, *Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology*

JEFF P. MORALES, *Director of Transportation, California DOT*

KAM MOVASSAGHI, *Secretary of Transportation, Louisiana Department of Transportation and Development*

CAROL A. MURRAY, *Commissioner, New Hampshire DOT*

DAVID PLAVIN, *President, Airports Council International, Washington, DC*

JOHN REBENDSORF, *Vice President, Network and Service Planning, Union Pacific Railroad Co., Omaha, NE*

CATHERINE L. ROSS, *Executive Director, Georgia Regional Transportation Agency*

JOHN M. SAMUELS, *Senior Vice President-Operations Planning & Support, Norfolk Southern Corporation, Norfolk, VA*

PAUL P. SKOUTELAS, *CEO, Port Authority of Allegheny County, Pittsburgh, PA*

MARTIN WACHS, *Director, Institute of Transportation Studies, University of California at Berkeley*

MICHAEL W. WICKHAM, *Chairman and CEO, Roadway Express, Inc., Akron, OH*

MIKE ACOTT, *President, National Asphalt Pavement Association (ex officio)*

MARION C. BLAKEY, *Federal Aviation Administrator, U.S.DOT (ex officio)*

REBECCA M. BREWSTER, *President and CEO, American Transportation Research Institute, Atlanta, GA (ex officio)*

THOMAS H. COLLINS (Adm., U.S. Coast Guard), *Commandant, U.S. Coast Guard (ex officio)*

JENNIFER L. DORN, *Federal Transit Administrator, U.S.DOT (ex officio)*

ELLEN G. ENGLEMAN, *Research and Special Programs Administrator, U.S.DOT (ex officio)*

ROBERT B. FLOWERS (Lt. Gen., U.S. Army), *Chief of Engineers and Commander, U.S. Army Corps of Engineers (ex officio)*

HAROLD K. FORSEN, *Foreign Secretary, National Academy of Engineering (ex officio)*

EDWARD R. HAMBERGER, *President and CEO, Association of American Railroads (ex officio)*

JOHN C. HORSLEY, *Executive Director, American Association of State Highway and Transportation Officials (ex officio)*

MICHAEL P. JACKSON, *Deputy Secretary of Transportation, U.S.DOT (ex officio)*

ROGER L. KING, *Chief Applications Technologist, National Aeronautics and Space Administration (ex officio)*

ROBERT S. KIRK, *Director, Office of Advanced Automotive Technologies, U.S. Department of Energy (ex officio)*

RICK KOWALEWSKI, *Acting Director, Bureau of Transportation Statistics, U.S.DOT (ex officio)*

WILLIAM W. MILLAR, *President, American Public Transportation Association (ex officio)*

MARY E. PETERS, *Federal Highway Administrator, U.S.DOT (ex officio)*

SUZANNE RUDZINSKI, *Director, Office of Transportation and Air Quality, U.S. Environmental Protection Agency (ex officio)*

JEFFREY W. RUNGE, *National Highway Traffic Safety Administrator, U.S.DOT (ex officio)*

ALLAN RUTTER, *Federal Railroad Administrator, U.S.DOT (ex officio)*

ANNETTE M. SANDBERG, *Deputy Administrator, Federal Motor Carrier Safety Administration, U.S.DOT (ex officio)*

WILLIAM G. SCHUBERT, *Maritime Administrator, U.S.DOT (ex officio)*

## **NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM**

### *Transportation Research Board Executive Committee Subcommittee for NCHRP*

GENEVIEVE GIULIANO, *University of Southern California, Los Angeles (Chair)*

E. DEAN CARLSON, *Kansas DOT*

LESTER A. HOEL, *University of Virginia*

JOHN C. HORSLEY, *American Association of State Highway and Transportation Officials*

MARY E. PETERS, *Federal Highway Administration*

ROBERT E. SKINNER, JR., *Transportation Research Board*

MICHAEL S. TOWNES, *Transportation District Commission of Hampton Roads, Hampton, VA*

---

# **NCHRP REPORT 501**

---

## **Integrated Safety Management Process**

**GENI BAHAR**  
**MAURICE MASLIAH**  
**CALVIN MOLLETT**  
iTRANS Consulting  
Richmond Hill, Ontario, Canada

**AND**

**BHAGWANT PERSAUD**  
Ryerson University  
Toronto, Ontario, Canada

**SUBJECT AREAS**  
Safety and Human Performance

---

Research Sponsored by the American Association of State Highway and Transportation Officials  
in Cooperation with the Federal Highway Administration

---

**TRANSPORTATION RESEARCH BOARD**

WASHINGTON, D.C.  
2003  
[www.TRB.org](http://www.TRB.org)

## **NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM**

Systematic, well-designed research provides the most effective approach to the solution of many problems facing highway administrators and engineers. Often, highway problems are of local interest and can best be studied by highway departments individually or in cooperation with their state universities and others. However, the accelerating growth of highway transportation develops increasingly complex problems of wide interest to highway authorities. These problems are best studied through a coordinated program of cooperative research.

In recognition of these needs, the highway administrators of the American Association of State Highway and Transportation Officials initiated in 1962 an objective national highway research program employing modern scientific techniques. This program is supported on a continuing basis by funds from participating member states of the Association and it receives the full cooperation and support of the Federal Highway Administration, United States Department of Transportation.

The Transportation Research Board of the National Academies was requested by the Association to administer the research program because of the Board's recognized objectivity and understanding of modern research practices. The Board is uniquely suited for this purpose as it maintains an extensive committee structure from which authorities on any highway transportation subject may be drawn; it possesses avenues of communications and cooperation with federal, state and local governmental agencies, universities, and industry; its relationship to the National Research Council is an insurance of objectivity; it maintains a full-time research correlation staff of specialists in highway transportation matters to bring the findings of research directly to those who are in a position to use them.

The program is developed on the basis of research needs identified by chief administrators of the highway and transportation departments and by committees of AASHTO. Each year, specific areas of research needs to be included in the program are proposed to the National Research Council and the Board by the American Association of State Highway and Transportation Officials. Research projects to fulfill these needs are defined by the Board, and qualified research agencies are selected from those that have submitted proposals. Administration and surveillance of research contracts are the responsibilities of the National Research Council and the Transportation Research Board.

The needs for highway research are many, and the National Cooperative Highway Research Program can make significant contributions to the solution of highway transportation problems of mutual concern to many responsible groups. The program, however, is intended to complement rather than to substitute for or duplicate other highway research programs.

---

**Note:** The Transportation Research Board of the National Academies, the National Research Council, the Federal Highway Administration, the American Association of State Highway and Transportation Officials, and the individual states participating in the National Cooperative Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

## **NCHRP REPORT 501**

Project G17-18(5) FY'01

ISSN 0077-5614

ISBN 0-309-08770-8

Library of Congress Control Number 2003111401

© 2003 Transportation Research Board

**Price \$29.00**

## **NOTICE**

The project that is the subject of this report was a part of the National Cooperative Highway Research Program conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council. Such approval reflects the Governing Board's judgment that the program concerned is of national importance and appropriate with respect to both the purposes and resources of the National Research Council.

The members of the technical committee selected to monitor this project and to review this report were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the project. The opinions and conclusions expressed or implied are those of the research agency that performed the research, and, while they have been accepted as appropriate by the technical committee, they are not necessarily those of the Transportation Research Board, the National Research Council, the American Association of State Highway and Transportation Officials, or the Federal Highway Administration, U.S. Department of Transportation.

Each report is reviewed and accepted for publication by the technical committee according to procedures established and monitored by the Transportation Research Board Executive Committee and the Governing Board of the National Research Council.

Published reports of the

## **NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM**

are available from:

Transportation Research Board  
Business Office  
500 Fifth Street, NW  
Washington, DC 20001

and can be ordered through the Internet at:

<http://www.national-academies.org/trb/bookstore>

Printed in the United States of America

# THE NATIONAL ACADEMIES

## *Advisers to the Nation on Science, Engineering, and Medicine*

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. William A. Wulf are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is a division of the National Research Council, which serves the National Academy of Sciences and the National Academy of Engineering. The Board's mission is to promote innovation and progress in transportation through research. In an objective and interdisciplinary setting, the Board facilitates the sharing of information on transportation practice and policy by researchers and practitioners; stimulates research and offers research management services that promote technical excellence; provides expert advice on transportation policy and programs; and disseminates research results broadly and encourages their implementation. The Board's varied activities annually engage more than 4,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. [www.TRB.org](http://www.TRB.org)

[www.national-academies.org](http://www.national-academies.org)

## COOPERATIVE RESEARCH PROGRAMS STAFF FOR NCHRP REPORT 501

ROBERT J. REILLY, *Director, Cooperative Research Programs*  
CRAWFORD F. JENCKS, *NCHRP Manager*  
CHARLES W. NIESSNER, *Senior Program Officer*  
EILEEN P. DELANEY, *Managing Editor*  
BETH HATCH, *Assistant Editor*  
ANDREA BRIERE, *Associate Editor*

### NCHRP PROJECT G17-18(5) PANEL Field of Traffic—Area of Safety

THOMAS E. BRYER, *Camp Hill, PA (Chair)*  
LEANNA DEPUE, *Central Missouri State University*  
ADELE DERBY, *Alexandria, VA*  
BARBARA HARSHA, *Governors Highway Safety Association, Washington, DC*  
BRUCE IBARGUEN, *Maine DOT*  
MARGARET “MEG” MOORE, *Texas DOT*  
KIM F. NYSTROM, *Nystrom Consulting, Gold River, CA*  
PETER F. “PETE” RUSCH, *FHWA*  
RUDY UMBS, *FHWA*  
ANTHONY D. WYATT, *North Carolina DOT*  
JESSE BLATT, *NHTSA Liaison Representative*  
RAY KRAMMES, *FHWA Liaison Representative*  
KEN KOBETSKY, *AASHTO Liaison Representative*  
RICHARD PAIN, *TRB Liaison Representative*

### AUTHOR ACKNOWLEDGMENTS

The authors thank the following professionals, who contributed greatly to the development of the ISMSystem. Thanks to the efforts of each of the members of this multidisciplinary team, a remarkable product was produced. They are:

John J. Zogby  
Transportation Safety Management Systems  
Duncannon, Pennsylvania

David Brown  
University of Alabama  
Tuscaloosa, Alabama

Raymond C. Peck  
R.C. Peck and Associates  
Folsom, California

Daniel H. Mayhew  
Traffic Injury Research Foundation  
Ottawa, Ontario, Canada

Mary Chipman  
University of Toronto  
Toronto, Ontario, Canada

Gregg Gullickson  
PricewaterhouseCoopers  
Fairfax, Virginia

Alison Smiley  
Human Factors North  
Toronto, Ontario, Canada

Margaret Parkhill  
iTRANS Consulting, Inc.  
Richmond Hill, Ontario, Canada

# FOREWORD

*By Charles W. Niessner  
Staff Officer  
Transportation Research  
Board*

This report presents the findings of a research project to develop an integrated safety management process. The process is a tool to assist in integrating safety-related implementation actions by proposing a method for bringing together agencies within a jurisdiction that are responsible for highway safety. The report will be of particular interest to safety practitioners with responsibility for implementing programs to reduce injuries and fatalities on the highway system.

---

In 1998, AASHTO approved its Strategic Highway Safety Plan, which was developed by the AASHTO Standing Committee for Highway Traffic Safety with the assistance of the Federal Highway Administration, the National Highway Traffic Safety Administration, and the Transportation Research Board Committee on Transportation Safety Management. The plan includes strategies in 22 key emphasis areas that affect highway safety. The plan's goal is to reduce the annual number of highway deaths by about 5,000 to 7,000. One of the 22 emphasis areas is "Creating More Effective Processes and Safety Management Systems."

State organizations carry out a number of independent safety initiatives that individually help to reduce injuries and fatalities on highways. Although highway safety responsibilities are divided among multiple agencies (DOT, motor vehicle administration, state police, emergency service, etc.), most states do not have a comprehensive strategic approach. Many initiatives focus only on strategies that the particular agency is responsible for implementing and do not effectively address the entire safety problem. A coordinated, comprehensive management approach to integrating engineering, education, enforcement, and emergency service efforts is needed to more effectively address major crash problems and achieve a greater reduction of overall injuries and deaths.

Under NCHRP Project 17-18(5), "Integrated Management Process to Reduce Highway Injuries and Fatalities Statewide," iTRANS Consulting, Ltd., developed an integrated safety management process. The integrated management process comprises the necessary steps for advancing from crash data to integrated action plans. The process includes the following six steps: (1) review highway safety information; (2) establish emphasis area goals; (3) develop objectives, strategies, and preliminary action plans to address the emphasis areas; (4) determine the appropriate combination of strategies for identified emphasis areas; (5) develop detailed action plans; and (6) implement the action plans and evaluate performance. The process includes methodologies to aid the practitioner in problem identification, resource optimization, and performance measurements. For those states and jurisdictions just starting to develop an integrated management approach, the steps to establishing it are fully described. In addition, a detailed description of the roles and functions forming the organizational structure of such an integrated management approach is provided. The process is sufficiently flexible to allow individual states and jurisdictions to customize the process

according to their political and organizational culture, resource constraints, and safety needs.

While this report provides an overall framework for coordinating a safety program, the volumes of *NCHRP Report 500* provide a series of guides for dealing with particular crash scenarios. Each of the volumes includes a brief introduction, a general description of the problem, and the strategies and countermeasures to address the problem. Together, the management process and the guides provide a comprehensive set of tools for managing a coordinated highway safety program.



# CONTENTS

<b>1</b>	<b>SUMMARY</b>
<b>3</b>	<b>CHAPTER 1: Introduction</b>
<b>5</b>	<b>CHAPTER 2: The Integrated Safety Management System (ISMSystem)</b>
2.1	Overview of the ISMSystem within the Transportation System, 5
2.2	Mission and Vision Statements, 6
2.3	Components of the Organizational Structure, 7
2.4	ISMPProcess, 13
<b>23</b>	<b>CHAPTER 3: Establishing the ISMSystem</b>
3.1	Requirement 1: The SPL Is Formed, 23
3.2	Requirement 2: The SPL Determines the Coalition's Mission, 23
3.3	Requirement 3: The SPL Creates a Memorandum of Understanding, 23
3.4	Requirement 4: The SPL Develops a Communication Plan, 25
3.5	Requirement 5: The SPL Appraises the Existing Safety Management System, 25
3.6	Requirement 6: The SPL Establishes the Administrative Structure of the ISMSystem, 26
3.7	Requirement 7: The SPL Appoints an Operations Manager, 26
3.8	Requirement 8: The SPL and Operations Manager Assemble the RAE Group, 26
3.9	Requirement 9: The SPL Sets a Vision, 26
3.10	Requirement 10: The RAE Group Links the Agencies' Databases, 26
<b>27</b>	<b>CHAPTER 4: Summary</b>
<b>28</b>	<b>GLOSSARY</b>
<b>31</b>	<b>REFERENCES</b>
<b>33</b>	<b>ATTACHMENT 1: Description of Optimization Software and Spreadsheets in Attached Diskette</b>
<b>A-1</b>	<b>APPENDIX A: State Feedback</b>
<b>B-1</b>	<b>APPENDIX B: Some Guidelines on Establishing the ISMSystem</b>
	Appendix B1: Recommendations for Improving Coordination and Communication, B-2
	Appendix B2: Best Practice Suggestions for Databases, B-4
	Appendix B3: Appraisal Criteria for an ISMSystem, B-7
	Appendix B References, B-13
<b>C-1</b>	<b>APPENDIX C: Step-by-Step Walk-Through Example of the ISMPProcess</b>
<b>D-1</b>	<b>APPENDIX D: Tools and Examples</b>
	Appendix D1: Methodologies for Identifying Crash Concerns and Developing an Effective Combination of Strategies to Support Jurisdictional Goals, D-2
	Appendix D2: Tools to Estimate the Effectiveness of Promising or Innovative Strategies About Which There Is Insufficient Information, D-37
	Appendix D3: Performance Measurement and Evaluation Tools to Determine the Level of Implementation and Success in Meeting the Goals of the ISHSPlan, D-51
	Appendix D References, D-85

# INTEGRATED SAFETY MANAGEMENT PROCESS

## SUMMARY

An integrated management process is described, referred to as the Integrated Safety Management System (ISMSystem). The ISMSystem is an organizational structure that is supported by resources and defined in terms of

- Leadership;
- Mission and vision statements; and
- Processes equipped with tools for managing the attributes of road, driver, and vehicle.

The intended audience for this report is highway safety professionals integrating the planning, optimization, and implementation steps in highway safety-related activities in order to maximize safety. The ISMSystem is one part of the larger Surface Transportation System, and by that connection it is related to asset management programs (1, 2), construction and system preservation, and other transportation systems including air, water, and pipeline transportation systems. The ISMSystem is, for the most part, a governmental management process; however, it encourages participation by the private sector through government channels.

Though much of this report is on the Integrated Safety Management Process (ISM-Process), which identifies the steps necessary to maximize highway safety, the ISM-Process is only one of several components within the ISMSystem. The other critical components include leadership (which represents the safety champions who are proponents of highway safety causes regardless of discipline), mission and vision (which stress the importance of having a clear and united purpose among all integrated agencies), and tools (which include the methodologies, databases, and analytical techniques necessary to provide the basis for good decision making and the most cost-effective implementation programs).

Fundamental to the ISMSystem is an interdisciplinary organizational structure formed through a coalition of highway safety agencies that allocates different responsibilities to specific groups or people who must work together in order to maximize safety. Day-to-day management responsibility falls upon the Operations Manager, an appointee by the coalition, while interagency coordination and communication, integration of goals, and

priorities are the responsibility of the Safety Program Leadership (SPL) group, which is made up of the top management of the different agencies involved in highway safety. In setting priorities, the SPL selects emphasis areas, those areas that have been identified as safety concerns for which resources within the jurisdiction are allocated to develop and implement action plans forming the Integrated Strategic Highway Safety Plan (ISHS-Plan). Task Teams represent the professionals, drawn from agencies involved in specific emphasis areas, who will develop strategies and action plans. The Risk Analysis and Evaluation (RAE) group provides the quantitative analysis of highway data and gathers information to support the decision-making process of the SPL, Operations Manager, and Task Teams at each one of the major steps of the ISMProcess.

There is enough flexibility in the ISMSystem to allow individual states and jurisdictions to customize the process according to their political and organizational culture, resource constraints, and safety needs. The ISMSystem was developed to support safety plans, such as the AASHTO Strategic Highway Safety Plan, as adopted in 1999.

The ISMProcess is broken down into six major steps designed to describe those elements in the process that are required to maximize safety. These steps are

1. Review highway safety information;
2. Establish emphasis area goal;
3. Develop objectives, strategies, and preliminary action plans to address the emphasis areas;
4. Determine the appropriate combination of strategies for identified emphasis areas;
5. Develop detailed action plans; and
6. Implement ISHSPlan and evaluate performance.

The ISMProcess starts with the SPL setting overall highway safety mission and vision statements for the jurisdiction. Next, emphasis areas, corresponding to the jurisdiction's major safety concerns, are identified, using available data on vehicles, persons, and infrastructure elements associated with crash data and other relevant information. The selected emphasis areas are allocated to the Task Teams for development. Equipped with the AASHTO Strategic Highway Safety Plan Implementation Guides (AASHTO Implementation Guides, published as *NCHRP Report 500*) and aided by analytical tools proposed in this report and other applicable techniques, the Task Teams develop objectives and strategies and propose activities, performance indicators, and work plans or schedules. The RAE group and Operations Manager prioritize the strategies developed by the different Task Teams, using cost-benefit optimization techniques, to achieve an integrated implementation plan for the jurisdiction. This optimization is achieved by minimizing duplication of effort in two or more emphasis areas and by selecting the best combinations of strategies to the highest levels of return for a given budget. Using the strategies that are selected and the respective final detailed action plans, the Operations Manager and the Task Teams compile an ISHSPlan that works in a synergetic fashion whereby the net safety effect is greater than what the effect of the strategies would be independently.

---

## CHAPTER 1

# INTRODUCTION

Imagine several people from the same family trying to move a large heavy sofa from one room to another. If one person lifts, one person pushes, and another person holds a door open, this may perhaps at first sound like a reasonable approach. However, if those who are pushing and lifting are going in different directions, and the third person plans on opening the door tomorrow, obviously the sofa will not get moved very far.

Now imagine several different agencies from the same jurisdiction, each with the responsibility of improving highway safety. Should not there be some sort of management system in place so that the agencies' combined efforts at improving safety are greater than each agency working independently? Obviously, each agency is different with its own organization, members, and resources. Each agency representing engineering, enforcement, education, or emergency services is different in size, yet each has a critical role to play, just like holding the door open is critical to the success of the sofa-moving task. An agency with the best training, methodologies, and resources will still not maximize safety for its jurisdiction if it works independently and in isolation. The agencies responsible for highway safety are separate entities, but their management efforts and decision-making processes should be integrated to both improve safety and reduce costs.

The AASHTO Standing Committee for Highway Traffic Safety, with the assistance of the Federal Highway Administration (FHWA), the National Highway Traffic Safety Administration (NHTSA), and the Transportation Research Board Committee on Transportation Safety Management, has developed the AASHTO Strategic Highway Safety Plan (AASHTO SHSP). The AASHTO SHSP addresses highway safety in terms of the driver, the occupant, the vehicle, the road, and postcrash responsibilities in noninfrastructure areas. The AASHTO Strategic Highway Safety Plan Implementation Guides (AASHTO Implementation Guides, published as *NCHRP Report 500*) are in development for the 22 emphasis areas selected for this plan (3). However, in order for agencies to implement emphasis areas and their strategies as described in the implementation guides, a safety management system is necessary (4, 5).

The Integrated Safety Management System (ISMSys-tem) defines a system, organization, and process for managing the attributes of the road, driver, and vehicle to

achieve the highest level of safety by integrating the work of disciplines and agencies involved in highway safety within a jurisdiction. These disciplines include the planning, design, construction, operation, and maintenance of the roadway infrastructure; injury prevention and control (emergency medical services [EMS]); the design and manufacture of vehicles; and those disciplines involved in controlling and modifying road user behavior (education, enforcement, and department of motor vehicles [DMV]). In order to manage the ISMSys-tem and to achieve the level of integration required to meet the highest levels of safety, two key components are needed. The first is an organizational structure that will allow for the integration of the agencies involved in highway safety. The second is a formal management process that will direct activities of these agencies in a manner that will efficiently achieve the mission and vision of the ISMSys-tem.

This report should be seen as a detailed outline and not as a regulatory blueprint. That is, all the parts described within this system are necessary, but there is flexibility for states and other jurisdictions to customize the organizational structure and management process according to their political and organizational culture, budgetary constraints, levels of data integration and availability, and jurisdictional safety goals and needs. Often the term "jurisdiction" is used instead of "state" to recognize that the application of the ISMSys-tem is not restricted to the state level.

Chapter 2 describes the ISMSys-tem, while Chapter 3 provides guidance for establishing an ISMSys-tem. Information regarding the establishment of an ISMSys-tem can be found in Appendix B, which includes recommendations for improving coordination and communication, best practice suggestions for databases, and appraisal criteria for assessing the main aspects of an existing safety management system within a jurisdiction. For an established ISMSys-tem, Appendix C is a step-by-step walk-through example of the ISMProcess. Appendix D contains some tools necessary to implement the ISMSys-tem. These tools include

- Methodologies for identification of crash concerns for developing effective combinations of strategies for supporting jurisdictional goals,

- Methodologies for estimating the effectiveness of promising or innovative strategies that have insufficient information,
- Methodologies for evaluating performance measures to determine the level of implementation and success in meeting the goals of the ISHSPlan,

- Optimization approaches, and
- Applications and examples.

A disk that contains some useful spreadsheets and a simple optimization program for demonstration purposes is attached to this report.

---

## CHAPTER 2

# THE INTEGRATED SAFETY MANAGEMENT SYSTEM (ISMSYSTEM)

The components of the ISMSystem, depicted in Figure 2-1, are Organizational Structure, Leadership, Mission & Vision, Resources, Integrated Safety Management Process (ISM-Process), and Tools. Achieving improvements in highway safety requires aligning responsibilities for safety with vision, goals, and supporting strategies. Through a coalition, agencies within a jurisdiction form the integrated organization structure and provide the resources to manage and implement the ISMSystem.

The ISMSystem is a system for managing highway safety. Just as all other professional functions have specific roles tied to responsibilities—such as traffic operational analysis, which is the responsibility of traffic engineers, and the development of statutes, which is the responsibility of legislatures—there must be positions and roles that have the responsibility for safety. Although all staff should have a safety mindset (the explicit consideration of safety impact in all activities), it is still necessary that safety responsibility be explicitly defined and delegated to specific roles within the organization and taken up by professionals who have the expertise and passion to promote safety.

At the same time, it is necessary to have everyone's involvement in order to maximize safety because simply allocating a few people in leadership positions, even with the appropriate responsibility and authority, cannot do it alone. Even groundskeepers who are responsible for cutting back weeds and brush on roadsides, in addition to performing landscaping, are improving safety by increasing visibility for drivers and removing potential obstacles from the roadway. If those involved in the day-to-day tasks of maintaining a highway system are aware of the safety implication of their actions (i.e., have a safety mindset), this may influence how they perform their responsibilities, which will lead to increased highway safety. State and local government agencies, industry, and private citizens all have a role to play in reducing the severity and number of highway crashes.

In order to recruit everyone's involvement in maximizing highway safety, it is necessary to promote and foster a culture that recognizes and places the highest priority on safety. Supporting such a philosophy toward safety requires the combined effort of leadership and public policy. Those professionals with the explicit responsibility for safety must initiate appropriate policy to make a safety philosophy the official rule of their organization. Funding must be allocated so that the pro-

cedures are in place for planning, designing, organizing, implementing, maintaining, monitoring, and assessing safety status and processes. Every individual's following well-designed, effective procedures in everyday practice will result in a sustainable ISMSystem.

Given the level of integration required within the ISMSystem, responsibilities should be developed and apportioned to ensure both depth and breadth of safety awareness and alignment of actions across all participating agencies. Every individual within the systems relating to safety (e.g., enforcement, engineering, emergency medical services, and education) should feel accountable for safety and empowered to positively affect safety within his or her sphere of responsibility. Developing an organizationwide safety mindset (i.e., culture and associated behaviors) and implementing supporting processes and programs to achieve safety goals should be one of the highest priorities of safety system leaders.

## 2.1 OVERVIEW OF THE ISMSYSTEM WITHIN THE TRANSPORTATION SYSTEM

The ISMSystem is a subsystem of the surface transportation system (Figure 2-2) and, as such, is constrained by the limits of this larger system. The ISMSystem does not include air, water, or the pipeline transportation systems under its jurisdiction, except perhaps in a limited way (i.e., in their connection with the highway system). The ISMSystem is, for the most part, a governmental management process; however, it encourages participation by the private sector through government channels. The goals and objectives of the ISMSystem must be compatible with the goals and objectives of the surface transportation system and be integrated with the additional systems required by the surface transportation system to manage the highway transportation infrastructure. The ISMSystem, in this regard, is affected by (a) highway transportation budgets as defined in federal transportation law, (b) the respective state and local government budgets and available resources, and (c) the limitations mandated by federal and state law in the use of these resources.

Most transportation system resources are defined in federal and state law and limited in the budget documents at each level of government. Management considerations of existing human and information technology resources will be

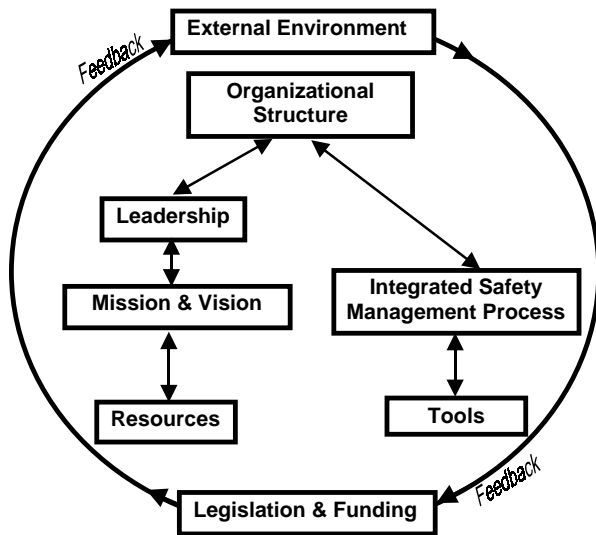


Figure 2-1. Components of the ISMSystem.

significant in determining the magnitude and scope of the ISMSystem developed and implemented. Requisite expertise and training are essential to the human resource component. Optimal use of information technology resources—especially in the area of integrated databases with relevant, complete, and accurate data; easily accessible report generation capabilities; and application of appropriate methodologies—can have a decisive impact.

Additional resources, while desirable, are not always available and in many cases are unnecessary. Existing resources can be enhanced through various changes such as reallocating key resources to new programs, adding training for personnel, and upgrading information technology (including hardware and software). For example, there are crash reduction and avoidance safety benefits of various infrastructure-based intelligent transportation systems (ITS) such as ramp metering, video enforcement, and weather-monitoring technologies (6).

The trend toward integration of disciplines responsible for transportation safety and the trend toward integration of planning for the movement of people and goods will demand a broader perspective for tomorrow's transportation safety manager. To meet the challenge, an ISMSystem is required, the expected results of which will optimize existing resources. There is also an anticipated benefit because of the synergy created as a result of the various highway safety disciplines working together to achieve a common vision.

There are two major sources for federal funds for traffic safety purposes: NHTSA and the FHWA. NHTSA provides funding for safety programs in nine different sections under TEA-21 legislation. Most of the funding categories require state- and local-level decision makers to determine the countermeasures to be implemented. Some funding programs allow the amount awarded to be distributed to other uses.

According to the NHTSA website, \$2.3 billion was authorized for highway safety grant programs over a 6-year period

(fiscal years 1998 to 2003). These grants are intended to provide support to identify highway safety problems, set goals and performance measures for improvements, start new programs, support existing programs, and fund analyses for determining the progress of safety improvement.

## 2.2 MISSION AND VISION STATEMENTS

The mission statement defines the purpose of the ISMSystem, what it does, and what it is all about. A mission statement usually does not change, and it sets the culture of the organization. On the other hand, the vision statement can and does occasionally change. The vision statement defines what the ISMSystem is striving for and what it hopes to achieve in a specified amount of time. The vision statement helps an organization set its priorities. While a vision may be long term (more than 20 years), a shorter 4- to 5-year vision statement is recommended. A shorter-duration vision statement is more likely to be achieved.

Within an organization, having a clear agreement on the safety system's mission and vision is key to having an effective and efficient ISMSystem. If the members of an organization involved in safety do not have the same goals or do not have a clear understanding of their roles, performance will suffer, and, in the case of highway safety, not as many lives will be saved. Division within ranks and between organizations as to what is the purpose of the ISMSystem is an institutional barrier. Without a clear and agreed-upon mission, different groups may find themselves working toward opposing objectives or with cumbersome, redundant, and/or ineffective processes. Misaligned vision can lead to political differences even within the same organization. Such a scenario may also lead to a public relations failure. If the members within a safety system are unclear as to their mission, there will be confusion for those agencies that must coordinate efforts. The general public will be very confused, having no clear idea what problems the organization is addressing or who to turn to given a specific problem.

Gaining clarity and agreement on the purpose of the ISMSystem's leadership is a critical first step in helping to ensure that expectations are established and supported by those who will be involved and who have the power to effect change. Therefore, consensus should be obtained on a mission statement that captures the purpose of the ISMSystem. The first responsibility of the ISMSystem leadership is to address this fundamental issue by creating a mission statement from which long- and short-term goals will emanate. The mission statement should answer two questions:

- What is the mission of the ISMSystem? (Why does the ISMSystem exist?)
- What are the underlying principles and beliefs of the organization and how are these translated into behaviors, culture, and organizational structure? (What type of relations, integrations, and coalitions underlie the organization?)

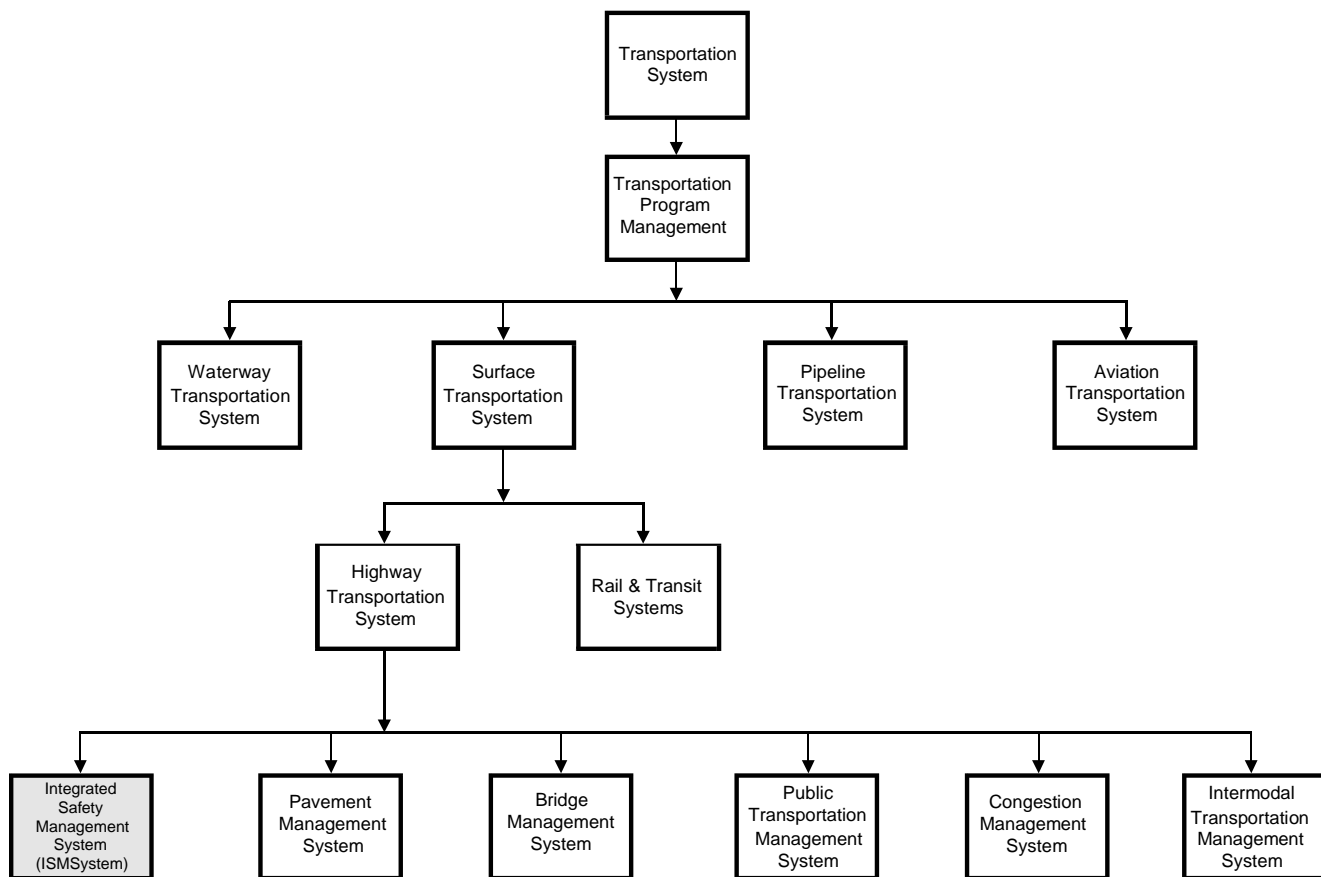


Figure 2-2. Overview of the transportation system.

In addition, an ISMSystem has a vision statement. While the mission statement describes the purpose and generally stays constant, the vision statement is focused on the short term and may change with changes in the driver population, crash data, or new administrative direction. A vision statement should specify a realistic goal within a realistic timeframe. The vision statement should answer two questions:

- What is the vision of success for the organization? (What will success look like?)
- What is the organization doing to fulfill its mission? (What is provided by way of products and services and to whom is it provided?)

A model mission statement, intended for customization, could be as follows:

This organization is an integrated organization formed by fully committed agencies, members of a coalition, aiming to maximize transportation safety for existing and future transportation networks, all road users, and supporting systems (enforcement, emergency services, and so forth). These agencies will prepare, implement, and evaluate an annual comprehensive integrated strategic highway safety plan that, as much as is realistically possible, minimizes the economic and human loss that results from traffic crashes.

An example of a vision statement is the goal of AASHTO SHSP: “To reduce the number of fatalities from traffic crashes by 5,000 to 7,000 lives annually.”

### 2.3 COMPONENTS OF THE ORGANIZATIONAL STRUCTURE

The ISMSystem organizational structure is not made up of roads, traffic laws, and countermeasures, but of people. These people may represent many different organizations with varying goals and purposes. Yet, in an ISMSystem, they must work together on many projects. Clearly, this can only happen if there are clear, open, and uninhibited lines of communication, coordination, and cooperation. Open lines of communication lead to cooperation and an understanding between agencies that allow people to go beyond identifying problems as merely “turf” issues. As a prerequisite for integrating agencies, Appendix B1 provides recommendations for improving communication and coordination. Funding, human resources, information, and technical expertise are four resources that should be shared across agencies in order for the ISMSystem to fulfill its mission (7).

In most cases, the mission and goals of the agencies composing the ISMSystem will be compatible. However, in



some instances safety objectives may appear to be in conflict with the other goals of an agency, such as focusing on efficiency savings as opposed to overall system effectiveness, of which safety is an important feature. A realistic ISMSystem must be sensitive to these other goals and avoid threatening the accomplishment of these goals (8). A threat can be viewed (and will be viewed) as anything that can negatively affect any of an agency's resources.

Figure 2-3 depicts an example of an organizational structure for a model ISMSystem. The Safety Program Leadership (SPL) represents the top management of the ISMSystem and gives direction and support to the organization formed by the coalition of the agencies represented in the SPL. However, in order for the organizational structure shown in Figure 2-3 to operate effectively, day-to-day management is required by the Operations Manager. The Operations Manager also acts as the focal point for coordination of ISMSystem. Quantitative and other safety analysis is the responsibility of the Risk Analysis and Evaluation (RAE) group. Unlike the SPL and Operations Manager, the Task Team members fulfill nonpermanent responsibilities. Task Teams are formed and disbanded as the state's safety needs and priorities change. Task Team members are recruited from the state's implementing agencies, as required for given emphasis areas. The agencies include those related to engineering, enforcement, injury and control, education, drivers, vehicles, and private safety associations.

Instead of a new bureaucracy, the organizational structure proposed in the ISMSystem is dependent upon existing agencies forming a coalition, usually as the result of a memorandum of understanding. The memorandum of understanding sets in writing how different implementing agencies will make commitments and share resources for the explicit purpose of integrating efforts aimed at reducing costs while improving

highway safety. More details on establishing the organizational structure may be found in Chapter 3.

### 2.3.1 SPL

The agencies responsible for highway safety, like any other organization in industry, need capable leadership in order to be able to adapt quickly to new situations, unite all levels of management, and promote realistic visions. Capable leadership provides a mission and visions that, among other attributes, (a) create greater cooperation and communication among agencies leading to integration of actions, (b) start new safety initiatives, and (c) promote safety consideration in all aspects of highway programs. The leadership of ISMSystem should comprise all (or a majority subset) of the disciplines involved in providing a safe highway system. These typically include what is commonly referred to as the 4Es of highway safety: engineering, enforcement, education, and emergency services. The leadership requires a partnership of these major disciplines, at different government agency levels, and the private sector.

Leadership is critical to building any system and a key determinant in the success of the ISMSystem. A questionnaire recently completed by a number of agencies has provided some quantitative evidence that, in the view of highway safety professionals, a person who is in a leadership position is "very important" to the future of the specific group's success or failure. While, by definition, the top-level management is thought of as the leadership of an organization, for the ISMSystem it means more than just being in charge of a group. It demands the qualities involved in being an advocate for highway safety or a safety champion. A safety champion cannot simply be replaced by someone with a similar resume because the way a person carries out a job and the associated priorities are not

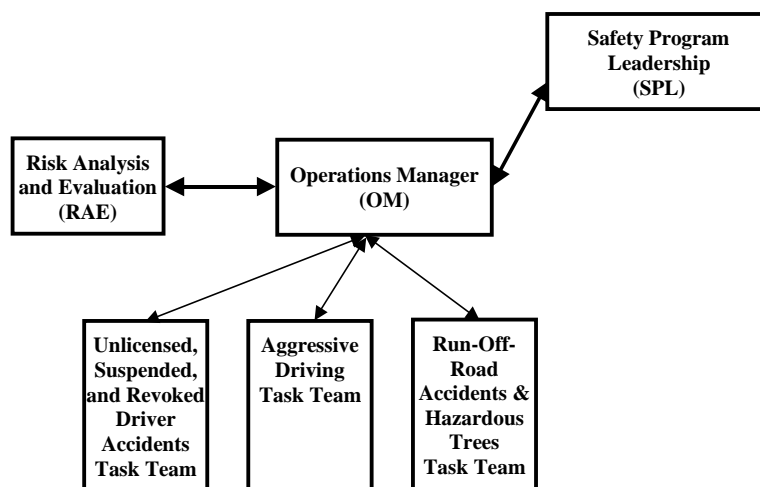


Figure 2-3. Example of an organizational chart for a model ISMSystem.

usually reflected in a technical skill set description. Safety champions may be found at all levels of the ISMSystem.

For the success of the ISMSystem, the SPL should be led by dedicated and responsible safety advocates (9) or safety champions. These safety champions are very important for acting as a focal point for safety advocates, improving inter-agency cooperation and communication, initializing new safety initiatives, increasing safety awareness, and reducing political turf issues.

The SPL ensures that team members and stakeholders at the jurisdictional level, which can be state or local, remain focused on their mission and vision. This team coordinates the development and implementation of goals (leading to emphasis areas) and supporting actions, facilitates the acquisition of needed resources, and provides whatever support (e.g., tools) is needed. The SPL facilitates the capturing of lessons learned and their use in improving the safety system performance. The SPL is responsible for the organization's strategic planning process and for how emphasis areas, objectives, strategies, and action plans are developed.

The SPL members should be representatives of agencies with highway safety responsibilities. The success of the ISMSystem is entirely dependent upon having top management from each of the stakeholders directly and personally involved in the SPL and having the authority to act on behalf of the stakeholder agencies. Selecting a chair for the SPL is one of several aspects of the ISMSystem that is determined by each state (or jurisdiction) individually depending upon its needs and structure. For example, in one state the chair may be fixed as the governor's highway safety representative, and in another the chair may rotate on a regular basis from agency to agency.

The selection of the SPL should include members at all levels of government (e.g., state, county, and metropolitan planning organizations) and should be tailored to each state's list of key stakeholders who are responsible for ensuring the safety of the traveling public. There should be frequent communication between SPL members and their respective agencies regarding key decisions, such as the setting of goals and the selection of emphasis areas. By definition, the representatives of the agencies who make up the SPL must be top-level management with the authority to make the necessary commitment.

Some examples of the agencies that may have representatives in the SPL are the following:

- **Transportation or highway agencies:** operation, planning, design, road construction, maintenance, and improvement.
- **Governor's highway safety agencies:** state highway safety planning, public information, education, and grant management.
- **Law enforcement agencies:** driver and vehicle safety surveillance.

- **Health agencies:** injury prevention, emergency medical care, alcohol and drug safety programs, and the Centers for Disease Control and Prevention (CDC).
- **Judiciary agencies:** adjudication of highway safety rules of the road.
- **Driver licensing agencies:** qualification and control of the driver licensed population.
- **Vehicle registration agencies:** qualification and control of the vehicle population.
- **State legislatures:** enactment of effective highway safety legislation.
- **Education agencies:** driver education and kindergarten through high school safety education.
- **Metropolitan planning organizations:** regional highway safety improvement programs.
- **State and local railroad regulatory agencies:** grade crossing safety and freight movement safety.
- **Federal and state commercial vehicle agencies:** commercial vehicle and driver safety inspections and safe hazardous materials movements.
- **Federal representatives from national, regional, or divisional offices of the U.S. Department of Transportation (USDOT):** The FHWA, the Federal Motor Carrier Safety Administration, and NHTSA.
- **Private safety organizations:** to include the American Automobile Association (AAA), Mothers Against Drunk Drivers (MADD), the National Safety Council, the American Trucking Association, and so forth.

There is some concern that as a top-down model, local jurisdictions may be overlooked in the planning and budgeting process. The straightforward solution is to ensure that representatives from various different agencies, specifically including rural and county-level commissioners, are active and involved in the planning process. Local agencies must be involved in the decision-making process; otherwise, statewide implementation will not be successful. Another possibility is the establishment of a two-tier ISMSystem, one at the state level and one at the metropolitan planning organization level.

Including a representative of remotely affected agencies on the SPL is not necessary. A membership of those agencies with major continuing safety programs (engineering, enforcement, education, and emergency services) could be the core, with representation from other agencies having membership as appropriate in the Task Teams, given the jurisdiction's particular needs and concerns. Examples of multidisciplinary collaboration can be found in the Safety Management System (SMS) Steering Committee in Florida, the SMS Coordinating Committee in Iowa, the Leadership Team and Safety Coordination Group in Ohio, the Traffic Alliance for a Safe California, the Traffic Improvement Association (TIA) of Oakland County in Michigan, and the Governor's Traffic Safety Committee of New York.

During the initial start-up of the ISMSystem, the SPL should meet frequently (e.g., monthly); however, progress

meetings would generally be on a quarter-annual basis, and budget meetings would be on an annual basis in line with existing budgetary cycles. The SPL is responsible for defining the overall highway safety goal, providing resources and other support, and ensuring learning and improvement. In addition, the SPL should lead in activities for the development of its own safety emphasis areas and strategic safety improvement plan, similar to the 22 major safety emphasis areas of the AASHTO SHSP.

### *SPL Responsibilities*

SPL responsibilities are as follows:

- Lead the establishment of the ISMSystem as outlined in this report.
- Set an overall highway safety goal for the jurisdiction, with the endorsement of the ISMSystem agencies.
- Lead in the preparation and justification for the budget to sustain the ISMSystem.
- Select the appropriate emphasis areas for the jurisdiction.
- Commit to and follow up on the promise to provide resources on behalf of those agencies assembled as part of the ISMSystem as outlined in the memorandum of understanding.
- Integrate the resources and activities of those agencies that have committed themselves to the ISMSystem.
- Ensure that the ISMSystem works efficiently and maximizes highway safety.

### *SPL Functions*

SPL functions are as follows:

- Acquire and manage the resources necessary to sustain the organizational process.
- Develop the memorandum of understanding, including formal funding and administrative structures.
- Formulate mission and vision statements for the ISMSystem; update vision as the system progresses.
- Identify, through an analysis of existing data sources and other knowledge sources, those major safety concerns (such as the AASHTO emphasis areas and/or other areas defined by the SPL) for further investigation and evaluation.
- Prioritize, allocate, and optimize all resources (people, infrastructure, funds, and information).
- Approve strategies proposed by the Task Teams and Operations Manager.
- Develop and fund the RAE group.
- Identify Task Team leaders with assistance from the Operations Manager.
- Evaluate the impact of the ISHSPlan.
- Appraise the performance of the ISMSystem.

### **2.3.2 Operations Manager**

While the SPL is a group that meets quarterly, or perhaps even monthly, there is a need to have a person who will manage the day-to-day tasks and the organization of the ISMSystem. The Operations Manager is the safety champion responsible for directing the daily activities, coordinating the efforts of various teams, acting as the focal point for the ISMSystem, and providing the SPL with support in planning and implementing highway safety system improvements. Among other responsibilities, the Operations Manager also ensures the professional development of immediate staff involved in the ISM-Process by providing experience and training opportunities.

The Operations Manager should also develop good public relations regarding all the positive impacts the ISMSystem has had within the jurisdiction (or should coordinate media information through existing public relations personnel in the various safety agencies). This public relations promotion should occur not just within the organizational structure, but should specifically include the public media both inside and outside the jurisdiction.

The Operations Manager serves as staff to the SPL. The Operations Manager is provided resources (funds, personnel, and materials) from the SPL coalition. It is the Operations Manager's responsibility to provide the SPL with information in a manner easily understood to allow interaction with the SPL members and staff for the determination of the mission and the development of the ISHSPlan. Many states have an office of highway safety or traffic safety bureau whose office could serve as the Operations Manager and supporting staff. The Operations Manager should have the total support and endorsement of top management and stakeholders.

### *Operations Manager Responsibilities*

Operations Manager responsibilities are as follows:

- Serve as staff to the SPL.
- Serve as the focal point of the ISMSystem.
- Provide day-to-day management of the ISMProcess.
- Provide the SPL with information in a manner easily understood to allow interaction with the SPL members and staff for the determination of jurisdictional priority emphasis areas.
- Ensure that the SPL mission and vision are clearly understood throughout the ISMSystem organization.
- Develop and administer the safety budget under the direction of the SPL.
- Integrate and prepare the ISHSPlan.

### *Operations Manager Functions*

As a staff assistant to the SPL, the Operations Manager will coordinate the activities of the periodic meetings and undertake many other functions, including the following:

- Develop the meeting agenda.
- Provide information as requested to the SPL.
- Ensure that the SPL mission and vision are clearly understood throughout the ISMSystem organization.
- Facilitate communication and coordination horizontally across the disciplines and vertically from the Task Team leaders up to the SPL.
- Manage the operation of the RAE group with respect to its work for the ISMSystem.
- Provide progress reports on the implementation of strategies.
- Manage and coordinate the Task Teams in the development of the various strategies.
- Select the best combination of strategies for implementation through optimization procedures performed in collaboration with the RAE group and Task Team leaders. Present these strategies to the SPL for approval.
- Select innovative safety strategies for pilot studies, funded research, or limited implementation in collaboration with Task Team leaders.
- Recruit, in cooperation with the Task Team leaders, members to form multidisciplinary Task Teams to further address selected emphasis areas.
- Monitor progress of implementation of statewide ISHS-Plan and make recommendations for improvement.
- Request and report the findings of process (administrative) and impact performance evaluations (outcomes) of the ISHSPlan.

### 2.3.3 RAE Group

The RAE group conducts the quantitative analysis and evaluation and assembles additional safety information as requested by the Operations Manager, SPL, or Task Teams. The RAE group is critical to the success of an ISMSystem because it provides the safety profile from which the SPL determines the emphasis areas of concern to be pursued. The RAE group responsibilities cover two aspects of data and information management: (a) the ability to collect, store, and retrieve data and relevant information and (b) the ability to analyze those data and information for local applications.

Data are usually stored in electronic databases. Simply put, without good data it is not possible to make good decisions. Given the amount, variety, and scope of data that must be collected and maintained, the linkages required among databases and the quality of data have a fundamental role in the success of the ISMSystem. Appendix B2 contains a list of best practice suggestions for creating or improving ISMSystem databases. Analysis refers to what is done with the collected data, which can range from basic number counts to sophisticated statistical manipulations. Very few jurisdictions would not benefit from improved analytical techniques and associated training. Even if a jurisdiction's knowledge is currently up to date, this may not be the case in the future as ongoing research develops new and advanced methodologies.

Members of the RAE group should be proficient in the use of analytical tools, relevant software, database manipulation techniques, and report writing. For most states, existing staff with these proficiencies could serve as the RAE group. Three options for staffing the RAE group exist. Two ideal staffing options for the RAE group would be the staff of the Operations Manager or an outside research institution (university, private, safety institutions, and so forth). Iowa State University's Center for Transportation Research and Education and the Iowa Traffic Records Advisory Committee are examples of groups that demonstrate the characteristics of a model RAE group. A third possible option for staffing the RAE group is having the RAE personnel remain in their existing organizations while having their job responsibilities modified to fit the required responsibilities and functions, as described below.

#### *RAE Group Responsibilities*

RAE group responsibilities are as follows:

- Provide highway safety information that is accurate, consistent, timely, and complete for the development of a statewide ISHSPlan.
- Provide highway safety information for the SPL and Operations Manager to identify major safety concerns, as needed for determining vision goals, budget, and emphasis areas.
- Bring to the attention to the Operations Manager and/or the SPL deficiencies and obstacles that may exist in the state's information system.

#### *RAE Group Functions*

RAE group functions are as follows:

- Obtain data from a number of sources of highway safety information, including crash, road, driver, police, emergency services, and hospital databases.
- Evaluate the adequacy of the highway safety information systems, including roadway, crash, vehicle, driver, EMS, and violation files.
- Assist in the electronic linkage of the safety databases of all coalition agencies.
- Process/analyze highway safety data and information to determine major safety concerns (such as the AASHTO emphasis areas) for consideration by SPL and Operations Manager.
- Perform cost-benefit and optimization analyses and prioritization process.
- Assist the Operations Manager and the Task Teams in the formulation of impact (outcomes) performance measures and guidelines for process and impact performance evaluations.

- Perform pre-implementation review of performance measurement strategies and baseline data.
- Perform impact (outcomes) evaluation.

### 2.3.4 Task Teams

Task Teams are composed of existing cadre in various agencies who are called on to address a specific safety problem or emphasis area under the direction of the SPL and Operations Manager. As emphasis areas are phased in or out by the SPL, Task Teams may be formed or disbanded. The selection of Task Team members depends upon the emphasis area. Different Task Teams will have different members from various agencies, depending upon the emphasis area. Private-sector advocates, where applicable, should be invited (by both the Operations Manager and the Task Team leader) to participate in a Task Team. Task Team membership requires commitment during the development of a given emphasis area, but members still belong to their original agencies and continue to work with their agency's resources throughout the development and implementation of the ISHSPlan. Being a member of the Task Team does not typically require a full-time commitment.

Each Task Team will require a leader who is closely associated with and experienced in the emphasis area. A leader will be selected by the Operations Manager, the SPL, and the agency where the leader now works. Once the Task Team leader has been selected, the remaining Task Team members are recruited by the Operations Manager and Task Team leader. It is important to engage all the disciplines and role-players that are encompassed within the scope of the selected emphasis area. This is achieved by establishing multidisciplinary, multiagency Task Teams.

Multidisciplinary, multiagency teams should be established for each of the emphasis areas, since differing levels of experience and expertise will be required for program development and implementation. Task Team members should be selected with experience in, and responsibility for, the emphasis area selected for the jurisdiction. The use of multidisciplinary teams not only provides the wide expertise needed for the various highway safety countermeasures, but also should create a synergistic impact on accomplishing the ISMSystem's vision.

At the time of writing this report, the 22 major safety emphasis areas of the AASHTO SHSP are being developed to form the AASHTO Implementation Guides. It is recommended that a Task Team assigned to 1 of these 22 emphasis areas use the corresponding AASHTO Strategic Highway Safety Plan Implementation Guides (published as *NCHRP Report 500*) as a primary source of information.

The membership of the Task Teams differs from the membership of the SPL (as given in Section 2.3.1); there is some overlap with respect to which agencies the members represent. The following are potential agencies and groups that may participate as members in a Task Team:

- **Transportation or highway agencies:** operation, planning, design, road construction, maintenance, and improvement.
- **Law enforcement agencies:** driver and vehicle safety surveillance.
- **Health agencies:** injury prevention, emergency medical care, alcohol and drug safety programs, and the CDC.
- **Driver licensing agencies:** qualification and control of the driver licensed population.
- **Education agencies:** driver education and kindergarten through high school safety education.
- **Metropolitan planning organizations:** regional highway safety improvement programs.
- **Federal and state commercial vehicle agencies:** commercial vehicle and driver safety inspections and safe hazardous materials movements.
- **Private safety organizations:** AAA, MADD, the National Safety Council, the American Trucking Association, and so forth.

While it may not be practical or efficient to include every discipline on each Task Team, the key discipline representative of a specific safety emphasis area should be in the lead role, referred to as the Task Team leader. The emphasis area to be investigated may appear on the surface to be solely the province of one discipline (e.g., improving the design and operation of highway intersections—engineering), but the mitigating solutions will certainly require other disciplines to bring success to their implementation (human factors, enforcement, and education in particular for new, untested, and/or innovative strategies). The Task Teams should develop the strategies' attributes with the use of information technology (shared databases, hardware, and software devices). This suggests that a highly integrated environment of management and technology is required for the benefits of the ISMProcess to be achieved.

For example, in Iowa, a number of task forces have been created to address areas of specific safety concerns, and in Florida, there exists a Safety Management System (SMS) subcommittee examining a variety of issues, including traffic records, drinking under the influence, legislation, and education. These groups demonstrate two key characteristics of successful Task Teams: multidisciplinary composition and prolific publication of findings.

### *Task Team Responsibilities*

Task Team responsibilities are as follows:

- Provide feedback on the appropriateness of emphasis area objectives, which are set by the SPL, to the Operations Manager.
- Develop strategies and action plans for the selected emphasis areas to achieve the ISMSystem's mission. Develop details of action plans (resources, interdepen-

dencies, responsibilities, and time schedules) for implementation.

- Coordinate with the Task Team's own particular agencies for (a) the allocation of resources for the implementation of relevant actions in the action plans and (b) the integration of other local plans with those of the ISHSPlan.
- Facilitate the implementation of the strategies' action plans.

### *Task Team Functions*

Task Team functions are as follows:

- Perform in-depth problem identification analysis in conjunction with the RAE group for the emphasis area assigned by the Operations Manager to the Task Team.
- Identify innovative strategies with associated preliminary action plans for submission as pilot projects to the Operations Manager.
- Review, if applicable, the corresponding AASHTO Implementation Guides for the emphasis area for additional strategies.
- Confirm that plan performance indicators are suitable for quantitative assessment of the success of different safety strategies.
- Assess the alternative strategies, and recommend a preferred alternative to the Operations Manager for implementation.
- Review and modify the action plans using feedback from the implementing agencies.

## **2.4 ISMPROCESS**

A crucial component of the ISMSystem is an organized process through which the ISMSystem's vision is achieved. This ISMProcess, as shown in Figure 2-4, identifies the steps necessary to achieve the overall highway safety goal identified by the SPL. Generally, there are three ways of institutionalizing safety improvements within a jurisdiction:

- Site-specific, safety-motivated projects in high-crash locations;
- Systematic improvements throughout a jurisdiction (for all project types); and
- Integration into existing highway projects (for rehabilitation, resurfacing, and restoration [3R] and rehabilitation, resurfacing, restoration, and reconstruction [4R] projects).

The process described in this section allows for the integration of safety improvements by identifying opportunities for different agencies to work together. Viewing safety improvement from a process perspective provides insight into the safety processes that cut horizontally across organizational

boundaries. Given the functional structure of most organizations, focusing on the process view of how safety work is done will provide insights into the interrelations between organizations and relevant functions that frequently are not readily apparent. Additionally, examining processes can help shed light on inefficiencies that, when understood and addressed, can lead to improved performance (e.g., non-value-added steps, lack of appreciation of internal process, customer needs and priorities, and bottlenecks). Focusing on processes also helps an organization's members see the bigger picture of which they are a part, leading to new ways of thinking that can be followed and new ways of acting (e.g., taking ownership of the process as a whole and viewing problem identification and rectification as a professionally enjoyable and rewarding learning process).

The process allows for an interaction among the personnel involved in management of highway safety databases, strategy development, analytical and evaluation processes, all phases of the infrastructure life cycle, education and enforcement of users, and external and/or internal policy direction. A comprehensive ISHSPlan for a jurisdiction can be organized with staff expertise and analytical methodology that identifies problems, designs remedies, and iteratively optimizes the design using realistic implementation indicators. Implementing the ISM-Process assumes that an organizational structure has already been established with sufficient resources for the operations. These resources include staff (the SPL, Operations Manager, RAE group, and supporting staff), facilities, and equipment required for the work. A detailed example of a step-by-step implementation of the ISMProcess is given in Appendix C.

The ISMProcess can be summarized as follows. First, an overall highway safety goal for the jurisdiction is set by the SPL. Emphasis areas and their objectives corresponding to the jurisdiction's major safety concerns are identified. For each objective, one or more strategies are developed, and, if required, pilot or research studies are proposed. Each strategy is developed with a corresponding preliminary action plan. From all strategies, the SPL and Operations Manager use cost-benefit analysis and optimization techniques to select action plans for further development into detailed action plans. The final ISHSPlan includes detailed action plans that work in a coordinated fashion. Selected products of the ISM-Process are shown in Figure 2-5. The six major steps of the ISMProcess are as follows:

1. Review highway safety information.
2. Establish emphasis areas and goals.
3. Develop objectives, strategies, and preliminary action plans to address the emphasis areas.
4. Determine the appropriate combination of strategies for identified emphasis areas.
5. Develop detailed action plans.
6. Implement ISHSPlan and evaluate performance.

These ISMProcess steps are described in more detail in the following sections.

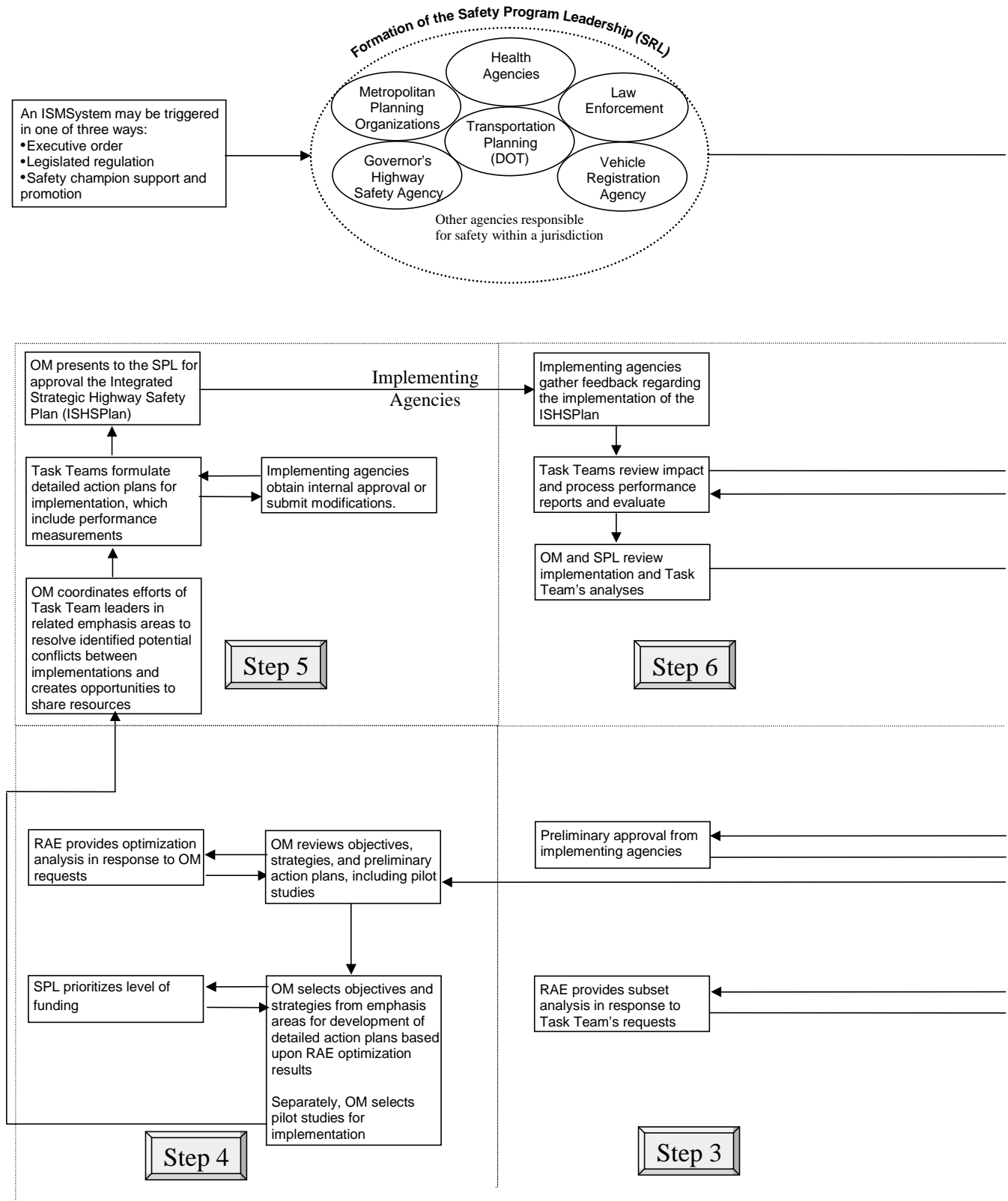


Figure 2-4. ISMProcess.

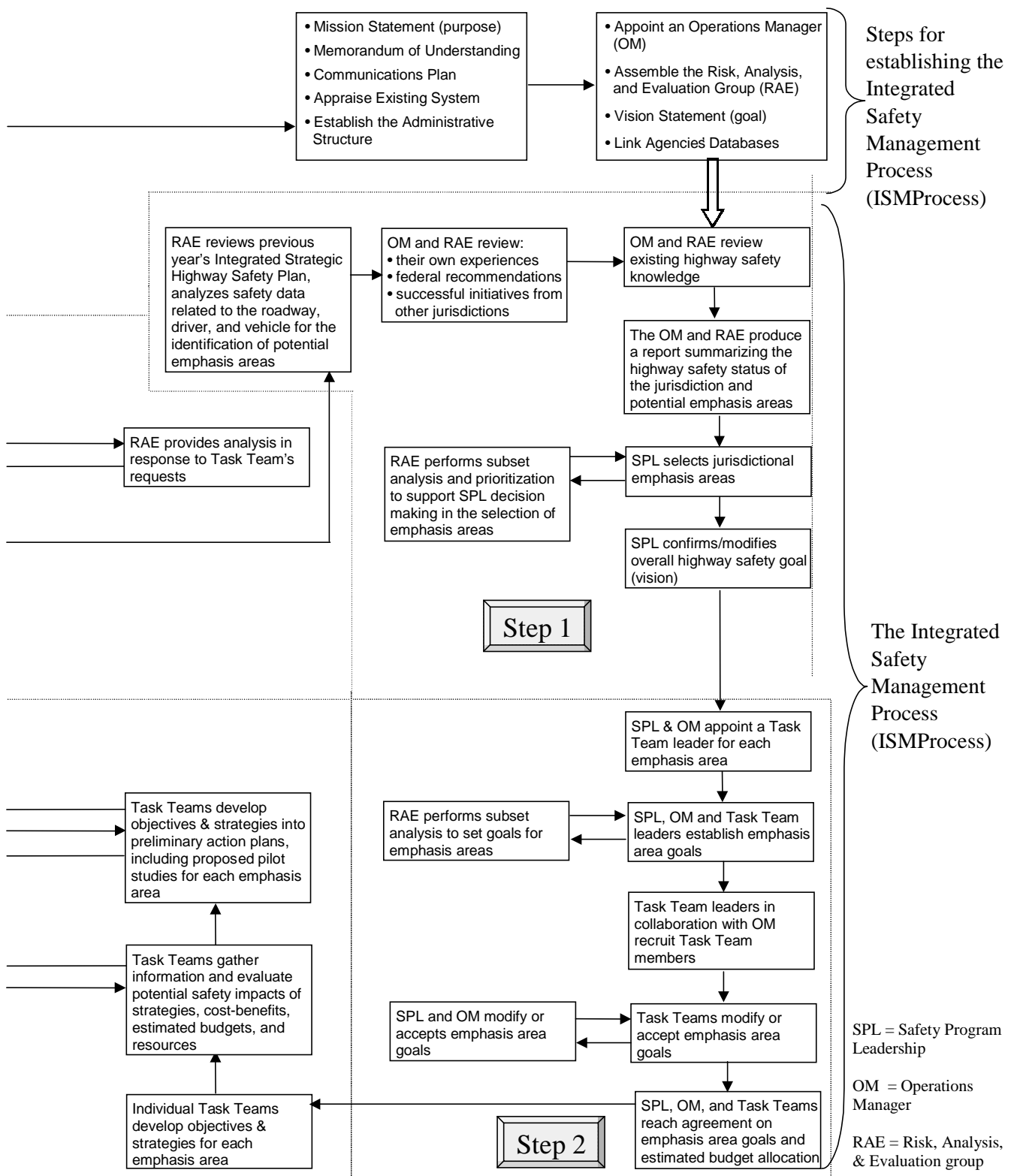


Figure 2-4. Continued.



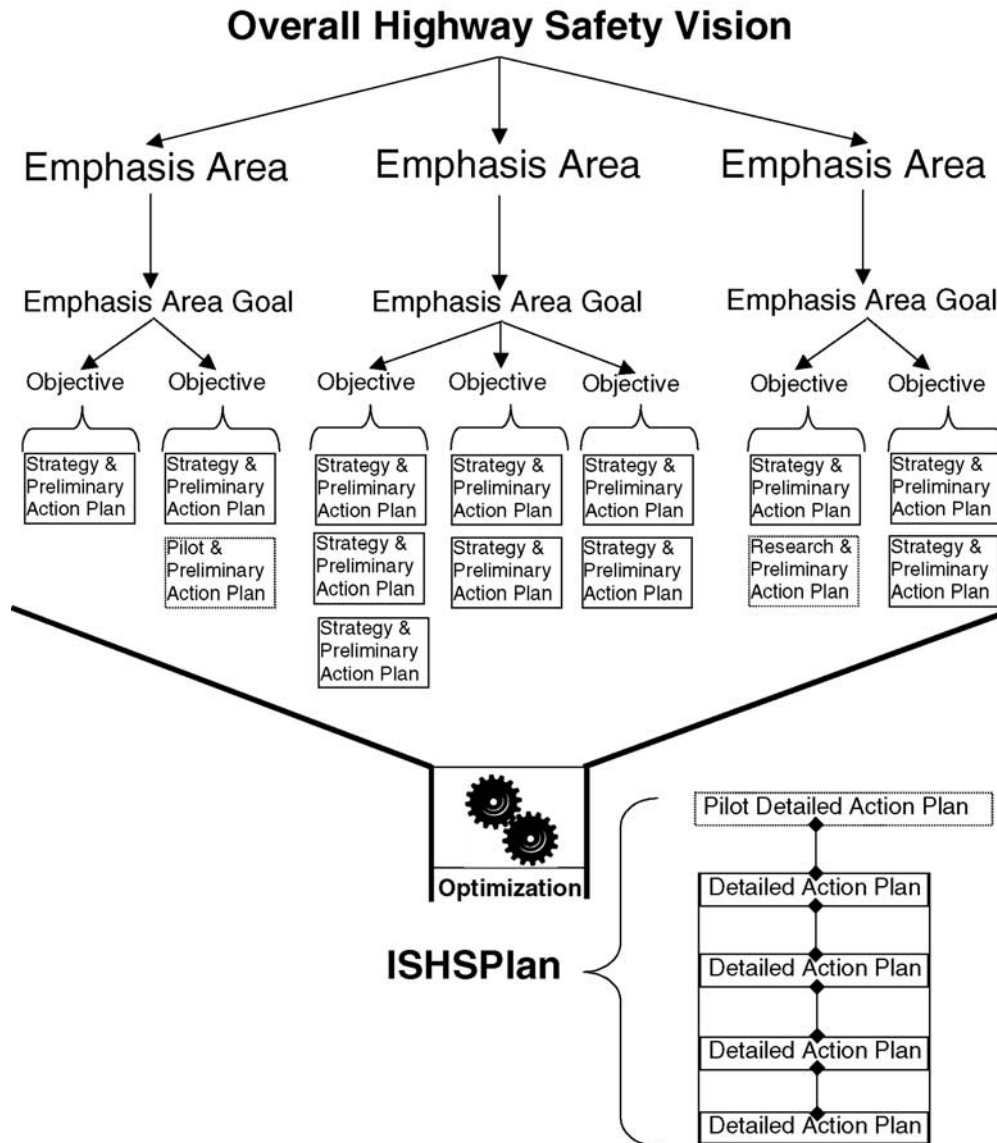


Figure 2-5. Selected products of the ISMProcess.

#### 2.4.1 Step 1: Review Highway Safety Information

To organize the problem identification tasks, highway crashes are organized into classifications (e.g., pedestrian, alcohol, and speed). Certain types of crashes are of more value than other types, either because they have a higher potential for general crash reduction or because they apply to special target groups (e.g., children). Other types of crashes are disregarded because of legal constraints, as is the case with categorical funding that restricts which types of activities can be funded. One might know intuitively that pedestrian, bicycle, and motorcycle crashes are of greater severity than most other “typical” crashes; however, in order to intelligently allocate limited funds, it is essential to determine just

how much worse different crashes are and just how often different kinds of crashes occur. There are ways of determining the frequency and severity of crash types, and information that does not measure quantitatively is usually dismissed as anecdotal evidence.

The aim of this step is to identify and define highway safety concerns, which, in turn, are used to identify the emphasis areas within the jurisdiction that will be the focus of attention of the ISMSystem. This step also involves identification of safety-related jurisdictional initiatives, including maintenance and construction, in order to provide explicit integration and consideration of safety implications. Significant data analyses are required to identify the safety concerns. While much of the analyses will involve reviewing data collected during the previous ISMProcess iteration or any other safety process pre-

viously undertaken by the involved agencies prior to the adoption of this ISMSsystem, this step is the opportunity for the SPL to view the big picture and introduce new emphasis areas for consideration. The following planning questions are addressed during this step:

- Where are the safety problems and opportunities?
- What is the scope of each problem and opportunity (e.g., how big is the problem and what parts of the safety system impact the problem)?
- If not all of these problems or opportunities can be addressed, which ones offer the greatest potential for improvement?
- How do the recent and ongoing safety initiatives in the jurisdiction—and those recommended by NHTSA, the FHWA, the Federal Motor Carrier Safety Administration (FMCSA), and other government agencies—address these safety problems and opportunities?
- What has recently been accomplished in other jurisdictions?
- What input has been received from highway safety practitioners, private groups, and the public?

The SPL and Operations Manager, supported by the RAE group, select the emphasis areas that will be the focus of the safety efforts for the jurisdiction. Emphasis areas are selected on the basis of cost and safety effectiveness, as well as social, economic, and political considerations. Additional consideration in establishing the emphasis area is given to highway safety-related programs and projects currently under implementation by the involved agencies.

### *Products*

Step 1 products are as follows:

- Overall highway safety goals and vision (i.e., the goals that apply to the entire jurisdiction, such as “reduce the total number of highway fatalities by 10%”);
- A report that describes the safety situation in the jurisdiction, identifies the major safety concerns, and produces a list of potential emphasis areas for further analyses; and
- A complete list of the jurisdictional emphasis areas that will be the focus of ISMSsystem.

### *Internal Process*

The Step 1 internal process is as follows:

1. The SPL annually asks the Operations Manager and the RAE group to conduct a review of highway safety information and ascertain the safety concerns of the jurisdiction.

2. The Operations Manager and RAE group analyze crash data and archival data relating to the roadway, driver, and vehicle to determine the type and magnitude of safety concerns, obtaining inputs from highway safety practitioners and the public.
3. The Operations Manager and RAE group survey existing highway safety research material applicable to the jurisdiction.
4. The Operations Manager and RAE group review safety initiatives recommended by the USDOT.
5. The Operations Manager reviews its own successful initiatives and others implemented by different jurisdictions.
6. The RAE group gathers information (both research literature and state crash data) and conducts the review of highway safety information using, among other methodologies, the proposed methodologies to identify major crash concerns. Appendix D1 provides a recommended methodology for the identification of problem areas leading to potential emphasis areas. This methodology has been applied by a number of states in recent years.
7. The Operations Manager and RAE group prepare a report that clearly identifies and describes the highway safety problem areas and concerns, thus leading to the definition or adoption of emphasis areas in a jurisdiction. This report should allow the SPL to easily understand the RAE group’s key findings, conclusions, and recommendations and should help the SPL decide which emphasis areas should be pursued further.
8. SPL team members meet with the Operations Manager, supported by the RAE group, and review and discuss the report findings to ensure that there is clear agreement on conclusions and recommendations. Where necessary, refinements are made and additional data are gathered.
9. The SPL reviews the final report with its recommendations and decides on the emphasis areas that should be studied further to meet the overall crash reduction goal of the jurisdiction. This review process should include a methodology of selection and priority setting of emphasis areas with promise of positive results. Where possible, this methodology should be analytical. If analytical methods are not possible, a brainstorming technique should be used to reduce the errors that may arise when aggregating individual judgments. A group decision takes full advantage of the multidisciplinary nature of the SPL. Such brainstorming techniques include the Nominal Group Technique and the Delphi Method. Appendix D2 provides an overview of subjective assessment techniques for estimating the effectiveness of new measures and describes when such knowledge is not deemed reliable for adoption. These subjective assessment techniques can also be applied to the

identification of problems when analytical methods are not feasible.

10. The SPL establishes an overall highway safety goal (i.e., the vision of the ISMSystem).

### *Requirements*

Step 1 requires the following:

- Sufficiently skilled and experienced staff to conduct a review of highway safety data and information. In terms of the proposed organizational structure, the RAE group could be in the office of the Operations Manager, a university, a research institute, or a private consulting group.
- The availability of, and easy access to, reliable highway safety data and information (i.e., research material; crash data; and data relating to the driver, vehicle, roadway, injury prevention and control, and enforcement) in computer databases that allow for multiple queries.
- Suitable methodologies to perform detailed analyses to identify crash concerns that take into account the type and quantity of available data in the jurisdiction.

### **2.4.2 Step 2: Establish Emphasis Area Goals**

The SPL and Operations Manager together recruit appropriate safety champions to lead and coordinate each Task Team. A Task Team leader should be a professional working in the implementing agency that is the largest stakeholder for the particular emphasis area. For each emphasis area, with the assistance of the corresponding Task Team leader, the SPL sets preliminary emphasis area goals or targets for each emphasis area to achieve. Once the preliminary emphasis area goals have been established, the Operations Manager and Task Team leaders recruit multidisciplinary, multiagency members to form a Task Team for each emphasis area. These preliminary emphasis area goals will form the basis from which the multidisciplinary Task Teams will identify objectives and formulate strategies in Step 3 of this process. The emphasis area goals serve to provide direction and purpose to the activities of the Task Teams.

The emphasis area goals that the Task Teams work with should be specific, be measurable, and align directly with the overall vision set forth by the SPL. Feasibility of achieving the emphasis area goals depends in part on funding available, the timeframe in which the goal is to be achieved, and the degree of complexity of the program. Each emphasis area goal should be an attainable target level with a basis in scientific research, previous experience, or both. The Task Teams will identify impact performance indicators that can be used to measure the extent to which the emphasis area goal has been achieved. Each emphasis area goal should be supported by at least one indicator to measure the resulting performance.

The AASHTO Implementation Guides are a primary resource and should be used by the SPL and the Task Teams to develop the emphasis area goals. In addition, Appendix D3 describes methodologies recommended for development of impact performance measures and provides some practical applications as examples.

The Task Teams may conclude that changes ought to be made to the preliminary emphasis area goals. These proposed emphasis area goals are submitted to the SPL for approval before proceeding to Step 3.

### *Products*

Step 2 products are as follows:

- Selection of Task Team leader,
- Emphasis area goal for each emphasis area,
- Recruitment of Task Team members, and
- Consensus of SPL and Task Teams on emphasis area goals.

### *Internal Process*

The Step 2 internal process is as follows:

1. The SPL and Operations Manager appoint a Task Team leader for each emphasis area.
2. The SPL, Operations Manager, and Task Team leaders establish emphasis area goals with support from the RAE group. The RAE group performs subset analysis to determine these goals, as described in Appendix D1.
3. Each Task Team leader, in collaboration with the Operations Manager, recruits Task Team members, forming a multidisciplinary, multiagency team.
4. The Task Teams may propose changes to the emphasis area goals through a collaborative process and reach consensus with the Operations Manager and the SPL.
5. The SPL, Operations Manager, and Task Teams reach an agreement upon the emphasis area goals and the estimated budget allocation for each emphasis area.

### *Requirements*

Step 2 requires the following:

- Identification by the SPL of the jurisdiction's emphasis areas,
- Recruitment of a Task Team leader for each emphasis area by the Operations Manager and SPL,
- Recruitment of Task Team members by the Task Team leader and the Operations Manager, and
- Estimated budget allocation from the SPL.

### 2.4.3 Step 3: Develop Objectives, Strategies, and Preliminary Action Plans to Address the Emphasis Areas

The Task Teams identify objectives that must be achieved in order to meet the emphasis area goals. For example, the emphasis area goal could be a 10% reduction in fatalities caused by crashes with trees. One of the possible objectives may be “eliminate the hazardous condition,” and one of the corresponding strategies for this objective may be “modify the roadside clear zone in the vicinity of trees.” In general, there may be multiple objectives for each emphasis area, and each objective may have multiple strategies.

Each Task Team will identify objectives and strategies to address its specific emphasis area. The Task Teams devise preliminary action plans for strategies for achieving the emphasis area goals. Preliminary action plans describe the required resources, activities, and timeframe for implementation of a strategy, along with estimates of the expected impact performance of the plan. Preliminary action plans may not contain many of the details necessary for actual implementation and are not coordinated with the action plans of other Task Teams at this point.

Proper strategy formulation draws upon the unique knowledge, skills, and experiences of the multidisciplinary Task Teams and requires a collaborative effort. If the Task Teams require assistance in performing quantitative analysis, a request for assistance from the RAE group should be made through the Operations Manager, depending on the jurisdiction’s definition of the RAE group (i.e., internal to an agency or an outside consultant). As part of the development of strategies, Task Teams should propose various deployment levels for each strategy, based on the systematic data analyses undertaken by the RAE group. The various deployment levels of personnel and financial resources should correspond to different forecasted safety benefits. There is a potential need to go beyond the traditional funding needs in order to have the necessary impact to meet the vision or goal. The option of having various deployment levels will be very useful to the Operations Manager and SPL in selecting the best overall set of strategies across all the emphasis areas in the next step (Step 4) of the ISMPProcess.

For those strategies where there is insufficient information regarding effectiveness, it may be necessary for a pilot study to be conducted before full-scale implementation. This may be especially true for new or innovative strategies. The AASHTO Implementation Guides are a primary resource to be used by the Task Teams to develop the objectives, strategies, and preliminary actions. In addition, Appendix D2 discusses evaluation tools to define the effectiveness of promising or innovative strategies that have insufficient information. Preliminary actions plans in the form of pilot or research studies may be proposed by the Task Teams and submitted to the Operations Manager for consideration.

### *Products*

Step 3 products are as follows:

- Objectives for each emphasis area,
- A set of potential strategies for each objective,
- Impact and process performance measures and data requirements for each objective,
- A preliminary action plan for each strategy, and
- Preliminary action plans for pilot or research studies designed to determine the effectiveness of promising or innovative strategies.

### *Internal Process*

The Step 3 process is as follows:

1. The Task Team develops strategies to achieve the emphasis area objectives. Strategies, for which there are insufficient data to assess their impact, are submitted as pilot or research studies. A pilot study is one where the implementation’s purpose is to evaluate the effectiveness of the strategy.
2. In response to requests from Task Teams, the RAE group provides subset analyses.
3. The Task Team develops a preliminary action plan for each strategy, including duration, interdependencies, resource requirements, and organizations involved.
4. The Task Team develops preliminary action plans for pilot and research studies.
5. The Task Team estimates the impact of each preliminary action plan in quantitative terms with respect to crash reduction, behavior modification, enforcement activity, process improvements, and so forth (depending on implementation levels).
6. The Task Team expresses the impact in terms of benefits to be realized.
7. The Task Team estimates costs to implement the preliminary action plan.

### *Requirements*

Step 3 requires the following:

- Multidisciplinary members of a Task Team led by a safety champion Task Team leader. Together, the team should have the skills, experience, and knowledge of the implementing agencies involved in the specific emphasis area.
- Access to the highway safety information and support of the RAE group for the analyses required for the development and selection of objectives, strategies, and performance measures.
- Preliminary approval from the implementing agencies.

#### 2.4.4 Step 4: Determine Appropriate Combination of Strategies for Identified Emphasis Areas

In Step 4, the Operations Manager will examine the strategies to ensure that redundancies are eliminated before commencing the process of optimizing resources required for implementation. The methodology for determining the appropriate combination of strategies involves an iterative process through subset comparative analysis, estimation of costs and benefits, and an optimization procedure. Subset comparative analysis is a type of analysis that compares different subsets of crash data broken down by the crash severity level in order to identify potential strategies within each emphasis area. An example of subset comparative analysis would be an examination of the number of alcohol crashes per population by area, age group, and time of day. An overrepresentation (e.g., by young male drivers in urban areas late at night) might suggest a specific strategy. Estimating benefits involves the application of accident modification factors (AMFs) to target crashes for different levels of implementation of a strategy, while associated costs for each implementation level are ideally based on experience. In some cases, it may be necessary to base costs, benefits, or both on experts' opinions.

Optimization is generally conducted using computer software and is performed several times during Step 4, which is described in more detail in Appendix D1. From the various strategies with multiple deployment levels, the Operations Manager must optimize the selection of those strategies at specific deployment levels across the emphasis areas that will reduce the highway fatalities and injuries to achieve the vision or goal. The purpose of optimization is to maximize safety across the jurisdiction while staying within the social, political, and practical constraints. Optimizations are conducted first to select objectives, and then to select strategies and activities, and a third time to optimize the level of implementation of the selected activities. Optimization may be conducted for one emphasis area, but the intent is to achieve integration by encompassing all emphasis areas simultaneously with a view to selecting not only the most effective combination of strategies but also the optimum implementation level for each selected strategy. The RAE group would generally conduct the analyses under the direction of the Operations Manager. Appendix D1 provides a description of the recommended methodologies involved in subset comparative analysis, estimating costs and benefits, and optimization.

The development of detailed action plans (in ISMProcess Step 5) with performance measures requires AMFs for cost-benefit analysis of various strategies. The relationship between the estimation and measurement tools in the ISMProcess is illustrated in Appendix D2. The figure shows that for the development of preliminary action plans, Task Teams with the support of the RAE group need the AMFs to determine the cost and benefits of various strategies. Where available, the AMFs come from the AASHTO Implementation Guides. If not available, such as when dealing with innovative strategies or others with insufficient information, the Task Teams must

seek recent sound evaluation studies or, at the absence of any such studies, must seek to estimate a temporary AMF. AMF information is added or updated as detailed action plan implementations are evaluated. Evaluation procedures are presented in Appendix D3.

Pilot studies and research projects for measuring the effectiveness of innovative strategies may be incorporated in the optimization process using some preliminary measures, or they may be allocated a dedicated budget for these special projects.

#### *Products*

Step 4 products are as follows:

- Objectives and strategies approved for implementation, along with the corresponding resources and funding that have been allocated toward each strategy, and
- Pilot or research studies approved for implementation, along with the corresponding resources and funding that have been allocated toward each study.

#### *Internal Process*

The Step 4 process is as follows:

1. The Operations Manager reviews the objectives, strategies, activities, and the additional material developed by the Task Teams for the preliminary action plans. Task Teams may be asked to modify or expand in certain areas of the preliminary action plans.
2. The Operations Manager reviews the proposed pilot and research studies.
3. The Operations Manager, with the support of the RAE group, applies a selected optimization technique to allocate the strategies on the basis of the available funds in the most cost-effective manner. The same optimization technique is applied separately on the pilot studies. The optimization techniques are described in Appendix D1.
4. The SPL confirms the level of funding available as part of the individual agencies' budget cycle.

#### *Requirements*

Step 4 requires the following:

- Strategies and corresponding preliminary action plans developed by each Task Team to address a specific emphasis area. The preliminary action plans should contain information on each strategy, including costs and benefits for various implementation levels, timeframe, resources required, and expected impact measure and outcome.

- The Operations Manager's assessment of all the pilot and research studies and their corresponding preliminary action plans proposed by the Task Teams.
- The RAE group's provision of results of the optimization procedure, including selected strategies, pilot studies, and respective required resources.
- The SPL's prioritization of levels of funding necessary for implementation. The individual agencies will be required to review and approve this funding as part of the agency's annual budget cycle.

#### 2.4.5 Step 5: Develop Detailed Action Plans

Using the results of the optimization procedures, the Operations Manager identifies preliminary action plans that will be implemented given the constraints of budgets and resources. For each strategy that is selected by the Operations Manager, the Task Teams expand upon the preliminary plans and develop detailed action plans. A detailed action plan may be thought of as a "cookbook list of instructions" for implementing agencies to carry out a strategy, whereas the preliminary action plan contains detail only sufficient for determining costs and benefits. The main purpose of detailed action plans is to decide who will do what, when, and how to achieve a specific outcome at the operating and implementation level. Each detailed action plan requires consideration of costs (materials, equipment, and so forth), personnel requirements, agency involvement, timeframes, legal and legislative implications, and other impacts. In addition, the detailed action plans specify how the performance measures or indicators (both process and impact) of the strategy will be assessed.

The detailed action plan also contains the performance measures, accompanied by a description of which data are required and how they will be collected for

- Process performance measures and
- Impact performance measures, including proxy (surrogate) measures.

Task Team leaders, through a coordinated effort by the Operations Manager, are made aware of related strategies and the need for integration. The Task Team leaders and Operations Manager will identify areas where resources may be shared and conflicts avoided. Through constant communication, the Task Teams develop detailed action plans that are interdependent and coordinated as necessary.

An action plan indicates activities for the various agencies involved, with milestones (times) of accomplishment and measures of performance. In most instances, an action plan developed by a Task Team will coordinate the efforts of various state and/or local safety agency personnel. Some details, as appropriate, may be modified or expanded by the implementing agencies. Action plans should contain information specifying the implementing agencies, funding (budget), personnel, training, scheduling and data requirements, equipment and material requirements, and legislation requirements

where appropriate. The performance measurement sections of the action plan should detail the goals, short- and long-term impact measures, and process measures. Appendix D3 provides useful guides for developing impact and process measures as part of action plans. Before the detailed action plans are submitted to the Operations Manager, the implementing agencies will individually consider the relevant components of the action plans and approve or propose modifications to the action plans. With the agencies' feedback, the Task Teams will formalize the detailed action plans.

The Operations Manager assembles all the detailed action plans for final approval and submission to the SPL. The Operations Manager does not have the authority or control of participating agencies' budgets; rather, the Operation Manager conducts the comparative analysis with recommended strategies and financial requirements for submission to the SPL for acceptance. In many cases, members of the SPL will have to take the recommendations and costs back to their individual organizations for review, approval, and funding. Ideally, this detailed action plan approval process should coincide with the organizations' budget cycles so that the funding levels may be considered during budget deliberations. Those action plans that are approved will form the ISHSPlan.

#### *Products*

Step 5 products are as follows:

- Detailed action plans for each strategy, pilot study, and/or research study, which include
  - Legislative needs,
  - Potential funding sources,
  - An impact and process performance measurement plan and schedule, and
  - A list of resources (funding, personnel, equipment, and material).
- The ISHSPlan.

#### *Internal Process*

The Step 5 process is as follows:

1. The Task Team members collaboratively formulate detailed action plans that are suitable for implementation at the operating and implementation level defined by the Operations Manager.
2. Prior to implementation, the Operations Manager will review the completed detailed action plans. At this stage, the Operations Manager will determine whether the action plans can be implemented based on the submitted documentation and/or whether any modifications are required. Should modifications be required, the Operations Manager will ask the Task Team to address them.
3. The final detailed action plans are combined to form the ISHSPlan. This comprehensive plan is submitted to the SPL for approval.

## Requirements

Step 5 requires the following:

- A list of strategies approved for implementation, along with the corresponding resources that have been allocated toward each strategy.
- A list of pilot or research studies approved for implementation, along with the corresponding resources that have been allocated toward each study.
- Integration of the detailed action plans forming the ISHS-Plan. Task Team leaders will coordinate their action plans and communicate to each other any modifications to prevent unexpected conflicts or ineffectiveness.
- Sign-off by the implementing agencies on the detailed action plans.

### 2.4.6 Step 6: Implement ISHSPlan and Evaluate Performance

Regardless of how thorough the detailed action plans are, issues will evolve at the implementation step that will require adjustments. In fact, as implementation timeframes get larger, the chances increase that some adjustments to the detailed action plans will be needed. Minor changes to the detailed action plans can be handled by the implementing agencies, while complex changes require coordination and communication with the Operations Manager, Task Teams, and other implementing agencies. Task Teams should meet periodically to monitor and appraise the progress of strategy implementation. The frequency of those meetings depends on the implementation timeframe of each strategy. Tentative meeting dates were entered into the detailed action plans. As feedback to future implementations, Task Teams should address the following question: “If the program were re-implemented today, what aspects would be kept the same and what aspects would be changed?” The Task Team leader and implementing agencies provide the feedback for progress reports to the Operations Manager and SPL, while the RAE group provides both process and impact evaluations. In addition, the SPL should continue having regularly scheduled meetings to monitor the overall progress and to review and approve any major changes recommended by the Task Teams.

The primary aim of the impact performance evaluation is to determine whether a program of actions has been successful in achieving its intended goals and objectives. This information is required to ensure that existing resources are used in the most effective and efficient manner. Performance measurement can (a) provide the basis for an agency to assess how well it is progressing toward its predetermined goals and objectives; (b) help the agency identify areas of strengths and weaknesses; and (c) help the agency decide on next steps, with the ultimate goal of improving performance. Performance mea-

surement can also provide the data necessary for showing how activities support broader goals, for supporting requests for additional resources, or for supporting new initiatives.

While individual jurisdictions have no authority over other jurisdictions, they would benefit from sharing with other jurisdictions their results, progress status, recommendations, observations, successes, and failures. Dissemination of knowledge can occur through the provision of web-based information, newsletters, conference presentations, in-house reports, and journal publications.

The results of the process and impact evaluations are the inputs for decision making during Step 1 at the next cycle of the ISMProcess.

## Products

Step 6 products are as follows:

- Periodic impact and process performance measurements collected by the implementing agencies and submitted to the Task Teams, Operations Manager, RAE group, and SPL;
- Review and revision of the action plans;
- Database management by the RAE group; and
- Preparation for the next ISHSPlan.

## Internal Process

The Step 6 process is as follows:

1. Feedback is gathered from the implementing agencies that enact the detailed action plans. The corresponding Task Team monitors and analyzes the feedback.
2. Impact evaluation of the strategies is conducted by Task Teams with support from the RAE group.
3. Process evaluation is conducted by the implementing agencies reporting their results to the Task Teams, Operations Manager, and SPL at the periodic meetings.
4. Highway safety databases are updated with the latest information.

## Requirements

Step 6 requires the following:

- The ISHSPlan,
- The implementing agencies’ continuous participation as planned in the detailed action plans,
- Periodic progress meetings, and
- Data collection for impact and process evaluation.

## CHAPTER 3

# ESTABLISHING THE ISMSYSTEM

In order to implement the ISMProcess described in the previous chapter, it is first necessary to establish the ISMSystem within a jurisdiction. The requirements for establishing an ISMSystem are depicted in Figure 3-1. The “trigger” for the establishment of an ISMSystem most likely would be one of three potential scenarios, described as follows:

- **Executive order**—An executive order issued by the jurisdiction’s chief executive (i.e., governor or mayor) stating the composition of the ISMSystem, stating the overall mission of the ISMSystem, and providing for the ISMSystem’s inclusion in the jurisdiction’s budgeting process.
- **Legislation**—The legislative branch of the jurisdiction (i.e., general assembly or city council) mandates an ISMSystem, giving broad authority that will, in many instances, require formal legal regulations for the elements of implementation. The mandate should also provide a funding mechanism for the development, implementation, and continuing maintenance of the ISMSystem.
- **Safety champion support and promotion**—A strong safety champion will have success with or without a formalized ISMSystem. However, to ensure sustainability of the ISMSystem in future times, a system must be established formally and safety must be institutionalised in the jurisdiction by one of the two scenarios described above.

In moderate to large states, there exists the potential to develop a two-tiered ISMSystem. The upper tier would be established at the state level, while the lower tier would consist of multiple ISMSystems at the metropolitan planning organization level. This lower level would focus primarily on regional safety goals aligned with the state ISMSystem vision and implementation.

Ten major requirements are identified here as necessary in order to successfully establish the ISMSystem. While the requirements depicted in Figure 3-1 are not all that is necessary to have a fully functional system, they form the foundation for the ISMSystem. Full implementation of the ISMSystem can be expected to take 2–5 years.

### 3.1 REQUIREMENT 1: THE SPL IS FORMED

Deciding which agencies compose the SPL is the first step in forming the ISMSystem. In an ISMSystem, the major agen-

cies in the highway safety community work in a coalition to develop, implement, and administer an ISHSPlan for the state or local jurisdiction. The number of partners in the coalition depends on the political, organizational, and management environment of the jurisdiction.

A coalition suggests that the many safety programs currently operating under different agencies in a jurisdiction can be coordinated to optimize resources and program results. The current independent planning and implementation of activities at the national, state, and local levels of government may result in redundant or conflicting efforts targeted to the same objective. These independent safety activities can and should be coordinated. A key requisite of the coalition is a strong commitment from each member agency to support the initiatives developed by the coalition and to provide ongoing funding to execute and administer the ISMSystem and develop the ISHSPlan. This commitment involves the formalization of commitment to allocate and share resources for the ISMSystem through a memorandum of understanding.

### 3.2 REQUIREMENT 2: THE SPL DETERMINES THE COALITION’S MISSION

The SPL sets the mission statement for the jurisdiction. This mission is developed and promoted as a long-term goal for that jurisdiction. The mission of an ISMSystem will be expressed as a consensus statement reflecting the aggregate of the safety missions existing in the coalition agencies.

### 3.3 REQUIREMENT 3: THE SPL CREATES A MEMORANDUM OF UNDERSTANDING

The major components of a memorandum of understanding are listed in Table 3-1. These components will form the foundation of the coalition. Each jurisdiction should tailor the memorandum of understanding to best suit the needs of the jurisdiction. The SPL should ensure that all aspects of the coalition are detailed in the memorandum of understanding. All roles and functions of the ISMSystem should also be clearly defined. Among those roles, the chair of the SPL will be identified in the documentation. For example, the chair could be the chief executive of the agency with the largest resource commitment. If the resource commitment of the SPL agencies is relatively equal, the chair can be decided by



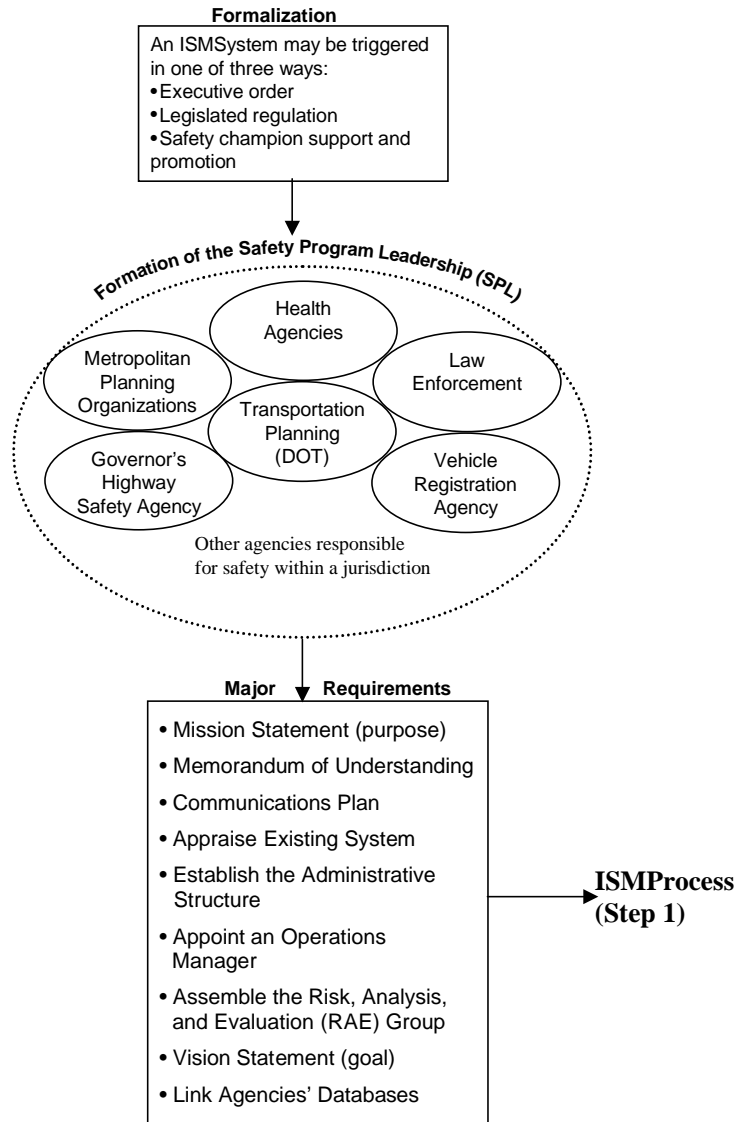


Figure 3-1. Establishing the ISMSystem.

TABLE 3-1 Proposed headings for memorandum of understanding

Heading	Description
Purpose	Reason for the memorandum of understanding: sustainability of the ISMSystem.
Membership	List of agencies involved in the ISMSystem and SPL and parties agreeing to memorandum of understanding.
Roles of SPL	Roles are outlined in this document, focusing on facilitating the implementation and sustainability of the ISMSystem.
Administration of Integrated Safety Management System	Clarification of all administrative issues (e.g., physical location of Operations Manager's office, procedures for resource allocation, and formalized position for the Operations Manager within the administrative structure).
Communication Channels and Protocols	Coordination of existing communication and marketing services to serve as public relations (external communications) for the ISMSystem in an integrated manner (a coordinated protocol); establishing internal procedures for communication and meetings among all agencies of the ISMSystem.
Funding	Formalization of the funding structure, ensuring sustainability.
Databases	Establishment of links among the relevant databases, and continual updating and maintenance of databases.

consensus of the members or by a rotation mechanism developed by the SPL members.

Policy development should be the province of the SPL and not be delegated to the day-to-day administrators of the ISMSystem. Therefore, this memorandum should also contain a policy statement outlining the budgeting process for the development and sustainability of the ISMSystem. The policy should also indicate the process for inclusion of the costs of the recommended strategies into the operating budgets of the implementing agencies. Risk management concerns will be appropriately included in policy statements and be in concert with existing risk management policies and tort liability statutes.

### **3.4 REQUIREMENT 4: THE SPL DEVELOPS A COMMUNICATION PLAN**

Public relations and marketing services should be employed at the earliest opportunity in order to stress the importance of the program and achieve public awareness and support. A marketing strategy is necessary not just to produce a slogan and a logo but also as part of a public education campaign. The public, in general, is unaware of the level of coordination or lack of coordination that exists between various government agencies. Part of a public campaign should be to bring this to the public's attention and to emphasize the leadership position that a jurisdiction is embarking upon. If politicians, senior management, and the public understand the purpose and benefits of an ISMSystem, they will strongly support it.

### **3.5 REQUIREMENT 5: THE SPL APPRAISES THE EXISTING SAFETY MANAGEMENT SYSTEM**

Establishing an ISMSystem is only the starting point and not the answer to improving highway safety. The answer to the question "What gives you the confidence that your state is doing everything possible to maximize safety on your highways?" is not "Well, we have established an ISMSystem" (10). The confidence will be expressed through improvements in crash statistics and safety-related cost savings. The ISMSystem presents an organizational structure and a process requiring effective management and staff activity for improving highway safety.

Appraising the status of the current safety systems is a requirement in evaluating what changes will be required to establish an effective ISMSystem. System changes should move in incremental steps in order to understand the impact of each change on the system and to use the insights gained to make refinements to the initial changes. Depending upon the scope of the changes made, one may expect an ISMSystem culture to develop over a period of several years before

an ISMSystem, as described here, is in place. However, integrated and systematic safety planning and programming can begin immediately.

The literature suggests a need for large systems to conduct regular appraisals in order to evaluate current practices and make necessary improvements. AASHTO has a history of conducting peer reviews, a type of assessment used by state DOTs, to identify internal strengths, weaknesses, external opportunities, and constraints that can be used to prioritize and align improvement initiatives. Predefined criteria for assessment have been applied in a variety of different systems, both in government and industry, such as the Baldrige National Quality Program "Criteria for Performance Excellence" (11). Osborne and Gaebler (12) present examples of federal, state, and local government organizations using performance assessment, among other techniques, to enhance their operations. Some states, such as Virginia, have a government review committee that conducts studies and appraisals that are used for assessing performance and making improvements (13, 14). Corporations around the world that subscribe to ISO 9000 standards use regular assessment procedures for certification (15, 16).

In order to have confidence in a system, it is necessary to conduct an assessment. For this purpose, appraisal criteria have been developed for all components forming the ISMSystem. These appraisal criteria are presented in Appendix B3 and are supplemented with expected appraisal results for a working system. These criteria are the basis for an assessment of whether or not an individual jurisdiction implementing the ISMSystem is meeting the requirements and relevant attributes stated in this report. Each jurisdiction should conduct an internal appraisal on a regular basis and should arrange for an external appraisal of the ISMSystem periodically, at least every 4 years. The internal appraisal is straightforward and inexpensive, since it can be conducted by those already familiar with the system. An external appraisal provides an independent view for comparison and, perhaps more importantly, introduces new ideas from other jurisdictions. Without answering the appraisal questions, it is extremely difficult to have the confidence that everything possible is being done to maximize highway safety.

For the internal appraisal, an internal group with appropriate authority and expertise should assess the various ISMSystem components. This group should include representatives of the different coalition agencies at different levels of the agencies, ensuring that there is at least one high-ranking staff member to provide the breadth of understanding and commitment to implementing improvements to the ISMSystem. The scope and success of the self-assessment will likely depend on the rank of the appraisal group's highest-ranking member within the ISMSystem leadership. Simply put, the higher the rank, the more likely the appraisal will lead to significant improvements in safety.

### **3.6 REQUIREMENT 6: THE SPL ESTABLISHES THE ADMINISTRATIVE STRUCTURE OF THE ISMSYSTEM**

Regardless of the process used for establishing the organizational structure, a critical part of the process is the allocation of funding to support and maintain the organizational structure. This means that funding must be allocated to (a) establish the positions for the Operations Manager and supporting staff and (b) ensure that funding will systematically continue to be budgeted for the ISMSystem.

While the functional structure for the Operations Manager has been specified, the administrative requirements, such as who will pay the Operations Manager's salary, must be worked out. All bureaucratic approvals must be obtained to create a full-time position for the Operations Manager of a sufficiently high management level to coordinate and manage the day-to-day operations of the ISMSystem.

### **3.7 REQUIREMENT 7: THE SPL APPOINTS AN OPERATIONS MANAGER**

Day-to-day operation of an ISMSystem will require a full-time Operations Manager to act as the focal point of the system. It is most desirable to have this person appointed immediately at the establishing stage. Until such an appointment takes place, staff from one of the coalition agencies may need to be appointed to this position, in the interim, as an Acting Operations Manager.

Most states have a person in similar position as the Safety Management Focal Point, a designation called for in the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA), or as the Governor's Highway Safety Representative (or coordinator) called for in the Highway Safety Act of 1966. The Operations Manager will serve as staff to the SPL and will direct the daily activities of the ISMSystem.

Human resources currently in place in the individual coalition agencies can serve as the personnel supporting the Operations Manager. These resources will not be organized in a separate agency, but rather will jointly develop, design, and implement an ISHSPlan for the jurisdiction. This ISHSPlan will aim to incorporate the existing safety programs operat-

ing in the coalition agencies to which these human resources are currently dedicated. The underlying philosophic foundation of an ISMSystem is that existing resources will work together on common highway safety initiatives (emphasis area's strategies), yielding effective results with less total resource expenditure.

### **3.8 REQUIREMENT 8: THE SPL AND OPERATIONS MANAGER ASSEMBLE THE RAE GROUP**

An effective ISMSystem requires analytical and evaluation expertise. The RAE group may be a research or analytical unit existing in one of the coalition agencies, a unit of a university research center, or contracted with a private research entity. Most likely, the RAE group will be made up of the current analysts and statisticians of the coalition agencies, potentially supplemented by outside professionals for specific expertise.

### **3.9 REQUIREMENT 9: THE SPL SETS A VISION**

A jurisdiction's vision should be ambitious and attempt to match the AASHTO SHSP. The SPL members must agree to this vision and obtain the endorsement of their respective agencies. The SPL should gather relevant information to develop such a vision. The vision should be set within a defined timeframe. Generally, visions will be in terms of reducing the total number of fatalities or the total number of fatalities and injuries. As such, the RAE group will conduct safety analyses, incorporating long-term budgets and population forecasts.

### **3.10 REQUIREMENT 10: THE RAE GROUP LINKS THE AGENCIES' DATABASES**

Each agency in the SPL possesses crash-related data that show a part of the status of highway safety. However, in order to see the big picture, it is necessary for the RAE group to analyze all relevant data, which means electronically linking the various databases. It is not necessary for one agency to act as a repository of the safety data; rather, it is sufficient merely to link the databases.

## CHAPTER 4

### SUMMARY

This report explains how to implement a management system designed to maximize highway safety. The ISMSystem describes how integrating the activities of agencies within a jurisdiction can result in a level of safety unobtainable by any agency working independently. In addition to highway safety benefits, implementing the ISMSystem should result in significant cost savings when the resources of different agencies are pooled and shared. The financial benefits of integration are quantified within the ISMProcess, which can be used as an additional argument to garner support for expanding the scope of an ISMSystem within a jurisdiction.

Currently, units within different agencies have various responsibilities and activities, some over-lapping and some mutually dependent, that have the potential to reduce the number of highway crashes, injuries, and deaths. Whether

agencies are reducing the number of crashes as much as possible highly depends on the level at which the different agencies are integrated in their problem identification, planning, optimizing, and implementation steps. The ISMSystem describes an organizational structure that includes a coalition of highway safety agencies that integrate their mission and vision at the highest level of management. Safety professionals with expertise in relevant emphasis areas are recruited into teams to address the specific goals set out by the coalition and, in doing so, produce action plans that integrate the activities of the implementing agencies. The ISMProcess describes the process for transitioning from crash data to integrated action plans (forming the ISHSPlan) using the best problem identification, optimization, and performance measurement tools available.

---

## GLOSSARY

*Accident Modification Factor (AMF):* An index of how much crash experience is expected to change following a modification in design or traffic control. AMF is the ratio between the number of crashes per unit of time expected after a modification or measure is implemented and the number of crashes per unit of time estimated if the change does not take place (17).

*Activity:* An action where a specific highway-related function or task is performed.

*Alpha:* The probability of making a Type I error (rejecting the null hypothesis when it is true).

*Annual Average Daily Traffic (AADT):* The counted (or estimated) total traffic volume in 1 year divided by 365 days.

*Appraisal Criteria:* A list of questions to be used in assessing the performance of an ISMSystem. While individual questions may apply to a specific component, the complete list of questions allows for a complete assessment of the ISMSystem (impact performance measures assess the performance of a strategy, whereas appraisal criteria assess the performance of the ISMSystem).

*ARIMA model:* Autoregressive, integrated, moving average model for the analysis of time-series data.

*Before-and-After Study:* A study in which the crash experience and other factors before a site or a group of sites is modified is compared with the crash experience after the modification in order to estimate the safety effect of the change.

*Cohort Study:* A study in which researchers compare two groups over a period of time. At the start of the study, one of the groups has a particular condition or receives a particular treatment, and the other does not. At the end of a certain amount of time, researchers compare the two groups to see how they did.

*Comparison Group:* A group of sites, used in before-and-after studies, that are untreated but are similar to the treated sites. The comparison group is used to control for changes in safety other than those due to a treatment.

*Control Group:* A comparison group of entities (e.g., drivers, roads, or vehicles) that have had a strategy applied to them and that can be used to compare with the entities that have not been affected by the strategy.

*Correlation:* The degree of linear association between two standardized variables; the slope divided by the perfect slope; ranges from minus one to plus one; measure of effect size.

*Covariate:* A measure that is correlated with the outcome but not affected by the treatment or the outcome.

*Crash Rate:* The number of crashes per unit of exposure. For an intersection, this is typically the number of crashes divided by the total entering AADT; for road sections, this is typically the number of crashes per million vehicle-miles traveled on a section.

*Cross-Sectional Study:* A study in which the crash experience of different sites is examined and differences in crash experience among sites are attributed to differences in specific site characteristics.

*Detailed Action Plan:* A specific plan that is under consideration to address a particular strategy, including a specification of budget and resource allocation (i.e., who, what, where, when, and how the program is going to be implemented). The action plan's activities are specified at the lowest level of detail for the ISHSPlan. An action plan may be made up of one or more ISHSPlan activities.

*Effect Size:* The magnitude of the standardized effect of a treatment variable on an outcome.

*Emphasis Area:* An area of safety concern identified through a comprehensive review and analysis of highway safety data and information (e.g., the AASHTO 22 emphasis areas or AASHTO Implementation Guides).

*Empirical Bayes (EB) Methodology:* A statistical procedure that is used to estimate the long-term annual number of crashes at a site using a weighted average of the site's short-term crash count and the average crash experience of similar sites.

*External Validity:* The generalizability of the results from a study; a threat is the interaction of treatment with another variable.

*Impact Performance Measures:* Metrics used to determine how well a strategy works to improve the safety (measured in number of crashes per unit of time) of a jurisdiction (compare with process performance measures, which evaluate the administrative aspects of implementation, such as expenditures and time schedules).

*Integrated Safety Management Process (ISMPProcess):* A process that, in an integrated fashion, transforms the inputs (resources, skills, information, and so forth) from the different disciplines and role-players who have an interest in safety by developing and implementing the safety strategies to address major safety concerns (i.e., emphasis areas). The integrated process consists of six steps:

1. Review highway safety information.
2. Establish emphasis area goals.
3. Develop objectives, strategies, and preliminary action plans to address the emphasis areas.
4. Determine the appropriate combination of strategies for identified emphasis areas.
5. Develop detailed action plans.
6. Implement ISHSPlan and evaluate performance.

*Integrated Safety Management System (ISMSystem):* A subsystem of the transportation system whose purpose is to minimize the occurrence and the consequences of automobile crashes. The individual components of the ISMSystem are leadership, mission and vision, organizational structure, resources, tools, legislation and funding, and the ISMPProcess.

*Integrated Strategic Highway Safety Plan (ISHSPlan):* A formal description of the strategies, related activities, performance measures, and goals of the ISMSystem. This document is a product of Step 5 of the ISMPProcess and will be implemented by the jurisdiction's engineering, enforcement, education, and EMS agencies, among others.

*Internal Validity:* Valid causal inference; estimating the effect due to a treatment; threatened by plausible rival hypotheses, such as regression-to-mean.

*Interrupted Time-Series Design:* A time series in which the initial observations serve as control and the observations made after a strategy is introduced are experimental.

*Measurement Error:* The random, unsystematic component in a measurement.

*Nonequivalent Control Group Design:* A design in which treatment and control groups are nonrandomly formed and subjects are pre- and posttested.

*Null Hypothesis:* The hypothesis that some population value (e.g., a mean difference, a correlation, or a regression coefficient) equals some particular value (usually zero).

*Objective:* A strategic component of achieving the vision or goal. An objective normally indicates anticipated levels of achievement and is time limited and quantifiable. An objective has direct relevance to an emphasis area.

*Operations Manager:* A person, appointed by the SPL, who is responsible for the day-to-day operations and management of the ISMPProcess.

*Power:* The probability of rejecting the null hypothesis when it is false; one minus the probability of making a Type II error.

*Pre-Post Design:* A design in which a group of people is measured before and after receiving a treatment.

*Problem Identification:* The process of transforming data from the highest levels into information that defines first emphasis areas and then objectives and strategies.

*Process Performance Measures:* Milestones used to assess the progress of the implementation of an action plan, usually administrative criteria.

*Program Areas:* The 6 areas from which 22 AASHTO emphasis areas emanated. These program areas relate to the driver, the vehicle, the roadway, highway safety management, injury prevention and control, and motor carrier safety.

*Random Assignment:* The assignment of people into treatment groups by a random rule; people have a fixed probability of being assigned to a treatment group.

*Random Selection:* The selection of people into the study randomly from some specified population.

*Randomized Experiments:* Studies in which units are randomly assigned to treatments.

*Regression Discontinuity Design:* A design in which people are assigned to treatment groups on the basis of a measured variable.

*Regression-to-Mean:* A phenomenon whereby sites with an unusually large crash count in one time period will, on average, experience a reduction in crashes in a subsequent period and vice versa without any intervention.

*Reliability:* The proportion of variance in a measure that is true, commonly estimated by an internal consistency measure.

*Resources:* People, materials, funds, and information databases.

*Safety Performance Function (SPF):* A mathematical equation that predicts the number of crashes, usually per year, at a site based on the site's traffic volumes and design and traffic characteristics.

*Safety Program Leadership (SPL):* A coalition of highway safety top management. The stakeholder agencies will assign

a top executive to provide leadership for the establishment, implementation, and administration of the ISMS system.

*Strategies:* Specific initiatives that emanate from the objectives, the implementation of which will meet the specified objective, thereby contributing to the achievement of the goal of the emphasis area.

*Structural Equation Modeling:* Modeling with a causal structure between latent variables.

*System:* A collection of interdependent components that interact with each other to achieve the purpose of the system. No one component can, on its own, achieve the purpose of the system; every component needs to interact with one or more of the other components of the system. Different systems have different components, and the nature and the extent of interaction required between the components to achieve the purpose of the system can vary.

*Task Team:* A multidisciplinary, multiagency group of safety practitioners with the primary responsibility of devel-

oping integrated and coordinated strategies and action plans to achieve the SPL's mission and vision. A Task Team comprises existing cadre in various organizations who are called on to address a specific emphasis area in which they have expertise. Members of a Task Team are recruited via the Task Team leader and the Operations Manager under the direction of the SPL. A Task Team is created to address a specific emphasis area and is disbanded either after the development of actions plans or after implementation and evaluation of the outcome.

*Tool:* Any "instrument" that assists in the execution of an activity (e.g., a management technique, software analytical package, documented procedure, documented guideline, or methodology).

*Type I Error:* Rejecting the null hypothesis when it is true; probability denoted as alpha.

*Type II Error:* Not rejecting the null hypothesis when it is false; its probability denoted as beta, and power equals one minus beta.

## REFERENCES

1. <http://www.fhwa.dot.gov/infrastructure/asstmgmt/amprimer.pdf>
  2. Markow, M. J., and Neumann, L. A., "Setting the Transportation Asset Management Research Agenda." (9-23-2001)
  3. [www.safety.transportation.org](http://www.safety.transportation.org)
  4. U.S. Department of Transportation, Federal Highway Administration (FHWA), and the National Highway Traffic Safety Administration (NHTSA), "Safety Management System: Implementation Workshop Proceedings." (3-1-1994)
  5. U.S. Department of Transportation, FHWA, "Safety Management System Workshop Proceedings: Managing Mobility Safely." (12-1-1991)
  6. McKeever, B. B., "Working Paper: Estimating the Potential Safety Benefits of Intelligent Transportation Systems." FHWA (1998).
  7. National Association of Governor's Highway Safety Representatives, National Association of State EMS Directors, NHTSA, and U.S. Department of Transportation, "Keys to Success: State Highway Safety and EMS Agencies Working Together to Improve Public Health." Vol. HS-809 111, Washington, D.C. (8-1-2000), pp. 1-50.
  8. Transportation Research Board of the National Academies, "Issues Surrounding Highway and Roadside Safety Management," *Transportation Research Circular 416* (10-1-1993).
  9. Zogby, J., "AASHTO Strategic Highway Safety Plan Case Studies DRAFT." NCHRP Project 17-18(6), Transportation Research Board of the National Academies, Washington, D.C. (8-1-2001).
  10. Johnston, S., "Five Reasons Why Safety Management Systems Don't Work." Queensland Mining Industry Health and Safety Conference Proceedings (1999), pp. 162-168.
  11. [http://www.quality.nist.gov/Business\\_Criteria.htm](http://www.quality.nist.gov/Business_Criteria.htm)
  12. Osborne, D., and Gaebler, T., "Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector." New York, New York, Penguin Books (1993).
  13. <http://jlarc.state.va.us>
  14. <http://jlarc.state.va.us/Summary/Rpt258/Highway.htm>
  15. <http://www.iso.ch/iso/en/iso9000-14000/iso9000/iso9000index.html>
  16. <http://www.ansi.org>
  17. Persaud, B. N., *NCHRP Synthesis of Highway Practice 295: Statistical Methods in Highway Safety Analysis*, Transportation Research Board of the National Academies, Washington, D.C. (2001).
-



## ATTACHMENT 1

### DESCRIPTION OF OPTIMIZATION SOFTWARE AND SPREADSHEETS IN ATTACHED DISKETTE

The diskette attached to this report contains an electronic attachment. The electronic attachment consists of the following files:

- **readme.txt:** a text file explaining the electronic attachment contents.
  - **Strategy Implementation Levels Worksheet.xls:** a Microsoft Excel spreadsheet used in preparation for optimization.
  - **Cost Benefit Analysis Worksheet.xls:** a Microsoft Excel spreadsheet for cost-benefit analysis.
  - **instructions.txt:** instructions for using the program costbenefit.exe.
  - **costbenefit.exe:** a Microsoft DOS program for cost-benefit analysis.
  - **in.txt:** a sample input text file.
  - **out.txt:** a sample output text file.
-

## **APPENDIX A**

### **STATE FEEDBACK**

Appendix A is not published herein, but can be obtained upon request from the NCHRP. Following is a description of the appendix.

State participation and feedback were part of the development of the ISMSystem. Results of an online questionnaire during the initial stages of development and a literature and website review provided the background information necessary for understanding the current status of the safety management systems in highway safety. These results are avail-

able as “Appendix A1: Survey Findings.” In addition, the questionnaire helped produce a contact list of front-line safety professionals and potential participants in safety workshops. After the development of a draft version of the ISMSystem, half-day workshops were conducted in three states to assess the feasibility of the proposed system. The results of the state feedback were incorporated into the final version of the ISMSystem. These results are available as “Appendix A2: State Workshop.”

---

## **APPENDIX B**

### **SOME GUIDELINES ON ESTABLISHING THE ISMSYSTEM**

This appendix, consisting of three subappendixes, includes guidelines that may be useful to jurisdictions desiring to establish an ISMSystem. These guidelines include recommendations for improving coordination and communication, best

practice recommendations for electronic databases, and appraisal criteria for assessing the status of an existing safety management system.

## Appendix B1: Recommendations for Improving Coordination and Communication

One of the most well-known catch phrases to describe institutional barriers is “turf issues.” This phrase is often used as a description and an explanation. However, simply identifying something as a turf issue provides neither understanding nor a solution. To break down an institutional barrier, the barrier must first be understood in terms of the underlying causes. For example, consider the situation of a state agency that is refusing, reluctant, or slow to share the data in its databases with another agency. Merely identifying this problem as a turf issue is not enough to sufficiently identify the root cause behind the refusal to share.

In the above case, the agency responsible may very well have guidelines and procedures stating that the data contained within the databases are confidential and cannot be released to unauthorized agencies or people. Guidelines and procedures within a government agency often carry the same weight as the rule of law, essentially acting as informal legislation. In the above example, the agency’s reluctance to share data happens to stem from the confidentiality policies in place. Once the true description of the barrier has been identified, does a solution become apparent? Following through with the example of confidentiality policies, the solution involves two steps:

1. Those responsible for maintaining and controlling access to the database must be properly informed regarding the legitimacy of the other agency’s request. This step will require communication between both agencies’ executives.
2. Assurance must be given to those providing the information to the requesting agency that sharing the data with unauthorized entities will not occur under any circumstance.

Neither of these steps could ever be implemented if the problem were merely categorized as a turf issue and no subsequent analysis of the causes were made. To reiterate, institutional barriers can be removed only when all parties involved understand them. Real institutional barriers are those that people have stopped trying to understand and have accepted as simply a byproduct of the system.

The following is a list of recommendations for improving coordination and communication. The list is based on a report intended to improve the relationship between state highway agencies and emergency medical services (EMS) (1). However, the recommendations hold true for many disciplines:

- Keep your communication lines open.
- Be forthright with the other agency about what is needed.
- Identify each need specifically.

- Understand the other agency’s needs. Keep talking until you do.
- Share credit. Don’t compete with one another.
- Look at what other states are doing and model your efforts after the best you find.
- Be prepared to both represent your own agency and respect the other agency’s needs.
- Be accessible as a resource. Offer your input.
- Work out details. Don’t leave things to chance.
- Make sure transfers and reimbursements of funds are prompt.
- Don’t overstate what you can do.
- Give awards or certificates of appreciation for those who go “above and beyond.”
- Stay in frequent contact.
- Stick to your timeline and your budget. If at all possible, coordinate state and federal budget cycles to reduce paperwork and administration.
- Don’t stray from your main objectives. Periodically review and reevaluate objectives to stay on track.
- Figure out how you will present the collaboration to the public. Shared credit can be confusing. Some state agencies are reluctant to be visible; others want to be visible.
- Make sure you’ve budgeted time and money for both benchmarking and evaluation so you can figure out what works and how. Measurable goals are critical.
- As new people join your agencies, keep them in the loop. It is especially important that grant administrators and contracting officers understand your collaborative effort and the value it brings to each agency.
- Reach outside your agency and take advantage of each other’s contacts. For instance, EMS may open the door to educating judges on injury prevention and highway safety, since EMS agencies are often involved in the in-service training of judges.
- A single point of contact (such as the Operations Manager, a Safety Program Leadership member, or a Task Team member) in an ISMSystem is critical to communications success.
- It is important to keep NHTSA and/or other grant agencies in the loop.
- If you have a formal Safety Program Leadership, Task Team, or other advisory-type committee meeting, keep minutes and be mindful of public record laws.
- Brainstorming sessions are very effective early in the process and then as needed to reinvigorate a program.
- Monthly reports are required under most grants and are a good way for everyone involved to stay on top of a project.
- Meet as often as is needed to be productive.

- Phone calls and face-to-face conversations are important, even when they are brief. They keep the “human touch” in a collaborative effort and are most useful in warding off impending problems.
- Lay everything out clearly from the beginning. If needed, the agency heads should have an informal, private meeting to work out any difficulties at the beginning of the process.
- Don’t try to hide anything (i.e., have no “hidden agenda”).
- Be mindful of state government politics and departmental reporting processes.
- If you report to an Operations Manager, Safety Program Leadership, Task Team, or other advisory committee, make certain that you maintain the original goals and that you are not sidetracked into other areas. Written goals and objectives will help.
- In most state capitals, everyone knows everyone else’s business. Be discreet and respectful of others; if you have a problem, address it directly and don’t let it linger.
- Keep “in the loop” anyone and everyone with an interest in the issue. Don’t forget to include all stakeholders, such as law enforcement and state legislators (especially those friendly to your issues). Minutes of the Safety Program Leadership meetings should be available online and circulated to all stakeholders.
- Be mindful of the natural tensions between civil (career) officials and political appointees who serve at the discretion of the governor.

## Appendix B2: Best Practice Suggestions for Databases

This subappendix contains best practice suggestions for maintaining databases listed in Table B-1. Descriptions of various safety databases are also provided in this subappendix.

The best practice suggestions are as follows:

- Ensure that the state databases are easily accessible by county and city jurisdictions. Ensure that the availability of the state databases to local jurisdictions is well known. Offer training as necessary to support local jurisdictions in their analysis needs.
- Improve, as necessary, crash data recording by emphasizing the following (1, 2):
  - One data entry for all crash databases,
  - Accurate crash location records using global positioning system (GPS) by enforcement agencies,
  - Reduction of under-reporting of crashes, and
  - Online accessibility of crash data (within privacy regulations and laws).
- Improve the sharing of crash and related highway safety data by making the data readily accessible in usable formats (1, 2).
- Coordinate safety management needs with ITS initiatives (1, 2). Obtain the involvement and support of all major agencies using the data record system, from the crash recording form to the analyst. This practice should result in buy-in to the entire process and should foster continuing and complete cooperation by all agencies in the state (3).
- Avoid innovative technology until it has been implemented and proven in the field. Failure to follow this suggestion could negatively affect database maintenance and lead to regrets such as the following:

In hindsight, it is clear that the decision to rely on [optical character recognition] OCR software to ‘read’ the crash report forms was one of the most detrimental to the success of the CRASH/CARS project. The OCR tech-

nology is simply unable to interpret accurately the variety of handwriting styles encountered on the crash report forms. The potential benefits of quicker data capture requiring fewer staff using OCR technology never materialized. Actually, the use of OCR software in this project resulted in just the opposite effect, as additional staff have been required by all three agencies to verify, repair and improve the often inaccurate and incomplete data generated by OCR (4).

- Run existing processes in parallel with new applications until proven stable (5).
- Do not justify new proposals on the basis of reduced staff levels due to technological innovations (5).
- Propose conservative implementation schedules with realistic deadlines (5).
- Develop crash record systems that are adaptable to future changes in data collection methods (5).
- Ensure that statewide crash databases contain complete historical data covering at a minimum the last 5 years.
- Promote more rapid availability of accurate data; delays are no longer acceptable (6).
- Improve data timeliness by releasing crash data quarterly, eventually in real time, for more responsive decision making. Evaluate frequently and correct more quickly (e.g., crashes due to icing on a road segment would initiate road crews dispatched to de-ice the segment) (6).
- Increase “cross-pollination” of data across disciplines. Allow the use of key data across various platforms, and address elements of roadway data, driver data, crash data, and so forth for a more complete analysis of crash factors and potential remedies (6).
- Coordinate public access to the data by developing and maintaining a data-contact list of key contact agencies or individuals for the various data sources. Post this list on the state website with active links where appropriate and available (6).

**TABLE B-1 Different safety databases and their descriptions**

Database	Description
Police Crash Archives	<p>Records on traffic crashes are maintained in state crash databases. The crash information is derived from the police report form that is usually completed by the investigating police officers at the crash sites. A typical crash report contains data on almost 100 different pieces of information that describe the crash, the people involved, and the vehicle.</p> <p>Although these archives contain a wealth of information on the driver, the vehicle, and the circumstances of the crash, a cautionary note is warranted. Crash data depend heavily on the subjective judgments of the police who attempt to reconstruct the crash after the fact. As such, the focus is typically on the proximate factors that can be easily identified at the crash scene and not on the more distal factors and events that may have propelled the behaviors causing the crash. As well, the apparent contributing or causal factors entered on police reports are often general and subjective, and it may be difficult to identify the underlying proximal behavior or action that precipitated the crash. For example, behaviors preceding the crash that appear to indicate aggressive driving (e.g., speeding or tailgating) may not have actually come from anger or frustration in driving (i.e., aggression). They could reflect other motives or causes, such as speeding so as not to be late for an appointment, or a lack of driving skills.</p> <p>Police-reported data may have considerable face validity (the appearance of accurately measuring an important factor) but are subject to error, especially subjective data that rely on interpretation and judgment (e.g., crash-casual variables). In this regard, Shinar et al. (7, 8), in an early U.S. investigation, used the data captured by a multidisciplinary investigation team on 124 crashes as the criterion to test the accuracy of the police-collected data. They found that "... the most valid police-reported data were those concerned with crash descriptors and the least reliable were driver/vehicle variables. The ability of the police to accurately attribute crash causes varied considerably across the different causes."</p> <p>Shinar and Treat (7) found that among crash-descriptive data, the police were most accurate on six variables: location, date, day of week, number of drivers, number of passengers, and number of vehicles. The least reliable police data concerned vertical and horizontal road character, crash severity, road surface composition, and speed limit.</p> <p>The major difficulty with driver and vehicle condition data was that the police either misclassified the information (e.g., on driver age and vehicle model year) or failed to provide any information on these driver or vehicle characteristics on the report.</p> <p>Shinar et al. (7, 8) also found that police sensitivity to crash causes was very low, in that investigating officers often failed to report factors that should have been reported. Police were more reliable in detecting human direct causes than vehicular, environmental, and human indirect causes. However, the authors also noted that in the area of human direct causes, the police performance was relatively good in identifying "failure to yield" and "failure to stop" but was relatively poor with respect to "speeding," "driving left of center," and "other improper turns."</p> <p>Data accuracy and data completeness also vary from state to state. Moreover, the National Highway Traffic Safety Administration reports, "... various sources suggest that about half of motor vehicle crashes in the country are not reported to police ..." (9). Unreported crashes typically involve only minor property damage and no significant personal injury.</p>
Department of Motor Vehicle Files	<p>Departments of motor vehicles maintain driver records of all licensed drivers in the state. Driver records are typically generated when a person enters the state licensing system to obtain a license or when unlicensed drivers have had a violation or crash in the state. The record contains basic identifiers (e.g., name, address, driver license number), demographic information on the driver (e.g., birth date, gender), and information relevant to license and driver improvement actions (e.g., license issuance and expiry/renewal dates, license class, violation dates, suspension periods). In some states, information on crash involvements (e.g., occurrence date, crash severity) is also available.</p> <p>Driver records are especially useful for examining issues related to driving history and rates of recidivism (e.g., re-offending for moving violations and traffic-related criminal convictions). However, many states purge the driver record of information on driving history after a certain period of time. Consequently, driver records are incomplete and drivers identified as first-time offenders may have had previous convictions for the same offense.</p>
DOT Highway Crash Database	<p>Departments of transportation are typically responsible for providing and maintaining the safe and efficient movement of people and resources on the state transportation network. Their highway crash database typically contains crash data, location codes, traffic volume information, and transportation network information, including highway and road inventory information. Some states also collect other types of roadway data, such as information on curves and intersections. Such information is used, for example, to locate hazardous sections and sites on roadways that require or are amenable to highway engineering and safety improvements.</p>
Vehicle Registration Files	<p>These files contain information on the vehicle identification number (VIN); plate number; and vehicle weight, model, make, and year. Vehicle registration data are typically used for transportation planning, safety strategies, environmental studies, and law enforcement.</p>
Ambulance Service and Pre-Hospital Care Data Sources	<p>State health departments or relevant other agencies are responsible for the transportation and pre-hospital assessment, diagnosis, and treatment of persons requiring emergency medical care. In some states, the ambulance crew must complete a report whenever a response number is generated and a patient is treated and/or transported by ambulance. Many persons requiring emergency medical care are the victims of motor vehicle crashes.</p> <p>Information collected and retained in the database includes the major complaint of the patient; a brief description of the mechanism of the injury; a recording of the vital signs; a diagnosis regarding the patient's most significant medical problem, a diagnosis about any associated problem or trauma score, and any additional comments (e.g., seat belts were worn, vehicle impact was head on, etc.).</p>
Hospital Care and Morbidity Data	<p>Health departments and/or hospitals maintain a health care morbidity database. The information contained in such databases typically consists of patient separation data, the level of care, the number of hospital days, type of diagnoses, and procedures performed on the patient. The international classification of disease coding (ICD9) can be employed to identify persons involved in motor vehicle crashes.</p>

(continued on the next page)

**TABLE B-1 (Continued)**

Database	Description
Trauma Centers	Information on seriously injured motor vehicle victims is typically retained at trauma centers. This may include details on the mechanism of the injury, blood alcohol concentration, length of hospital stay, and injury severity scale.
Federal Crash Archives	<p>The National Highway Traffic Safety Administration (NHTSA) maintains two crash databases: the Fatality Analysis Reporting System (FARS) and the General Estimates System (GES) (10).</p> <p>FARS contains annual data on a census of fatal traffic crashes within the 50 states, the District of Columbia, and Puerto Rico. According to NHTSA, "to be included in FARS, a crash must involve a motor vehicle traveling on a trafficway customarily open to the public, and must result in the death of an occupant of a vehicle or a nonmotorist within 30 days of the crash." FARS data are available annually back to 1975. FARS contains more than 100 data elements related to the driver, vehicle, involved persons, and the crash itself. FARS has proven to be a rich information source for research and program evaluation focusing on fatal crashes.</p> <p>GES data are obtained from a nationally representative probability sample selected from all police-reported crashes. According to NHTSA, "to be eligible for the GES sample, a police accident report (PAR) must be completed for the crash, and the crash must involve at least one motor vehicle traveling on a trafficway and result in property damage, injury, or death." GES began in 1988. Although useful for some research purposes, GES data are estimates of counts of crashes and injuries and are, therefore, subject to sampling errors. Moreover, statistical procedures are used to assign values to all the GES unknown data. Because of such limitations, NHTSA cautions that care should be taken in comparing GES estimates from one year to the next. This may preclude any trend analysis or the use of GES to examine fatal and injury crash involvement, before-and-after implementation of a program.</p>
Highway Safety Information System	Another useful crash database is the Highway Safety Information System (HSIS), a state safety database that is maintained by the Federal Highway Administration (FHWA). The HSIS uses data collected by eight states for the management of their highway systems and for the study of highway safety. The HSIS data include information of crashes, roadway design features, traffic, and other elements. All severity levels are included in the HSIS crash data, but often property-damage-only crashes are restricted to crashes over a specified amount of damage (e.g., \$500), which varies by state and year.



**Appendix B3: Appraisal Criteria for an ISMSystem**

This appendix contains appraisal criteria that may be used to assess the components of a safety management system. These criteria are appropriate for a jurisdiction looking to establish an ISMSystem that wishes to first assess its existing system. These appraisal criteria may also be used during periodic reviews of an existing ISMSystem in order to make improvements and to recommend changes. The appraisal cri-

teria are given in Table B-2 and are grouped by the ISMSys-tem components: leadership, mission and vision, organiza-tion, Operations Manager, Risk and Analysis (RAE) group, Task Teams, highway safety information systems, and ISMProcess. In Table B-2, the first column describes the issue to be assessed, and the second column describes the optimal scenario in a working ISMSystem.

**TABLE B-2 Appraisal criteria**

ISSUE	WHAT TO LOOK FOR
<b>LEADERSHIP</b>	
Is safety viewed from a systems perspective, one that aligns and integrates the actions of all functions and organizations (e.g., 4Es—engineering, enforcement, education, emergency medical services) involved in safety?	<p>Leaders should understand that each of the 4Es impact safety and know how to assess the overall ISMSsystem in terms of each, individually and collectively.</p> <p>The agencies that are responsible for highway safety engineering, traffic law enforcement, safety education, and emergency services or injury prevention programs are represented on the Safety Program Leadership (SPL) coalition.</p>
Are the representatives in the SPL recognized safety leaders or champions? Do they have the authority to make decisions committing resources on behalf of the agencies they represent? Is there an overall leader responsible for helping integrate and align the improvement actions that cut across organizational and programmatic boundaries (e.g., across each of the 4Es)?	<p>A statewide safety leader or champion is formally designated and has the competencies needed to understand and facilitate the improvement of the entire ISMSsystem, taking into account the needs of each of the agencies and integrating them to do what is best for the entire ISMSsystem (SPL chair and/or Operations Manager).</p> <p>The agencies that compose the SPL coalition each have identifiable committed resources to the ISMSsystem.</p>
<b>LEADERSHIP TRAINING</b>	
Do education and training support the expected accomplishment of key agency safety action plans? Do they address the agency's needs, including building safety knowledge, skills, and capabilities, and contribute to improved employee performance and development?	The ISHSPlan, strategies, and action plans describe the education and training needs of personnel within the ISMSsystem and the resources needed to accomplish the related strategies and actions.
Have training programs been developed that are designed to raise the safety consciousness of agency employees as to the safety implications of their decisions and actions?	The ISMSsystem is viewed from the perspective of the people who participate in the system. Safety competency requirements and safety training needs are identified and incorporated into the overall and individual organization strategic safety plans.
Are periodic training sessions on safety at all agency levels developed and conducted? Are employees permitted and encouraged to attend major conferences, seminars, and so forth for professional development? Do representatives from the state attend national conferences on safety issues and share knowledge and insights gained with key system stakeholders?	<p>Training plans are implemented.</p> <p>Conferences and seminars are used as learning resources, and processes are in place for ensuring adequate funding and for disseminating and using insights gained at such events.</p>
<b>MISSION AND VISION</b>	
Does the agency in charge of safety have mission and vision statements that encompass the entire ISMSsystem?	The SPL has a clear a mission statement and a vision statement that are formally adopted and shared with all organizations and stakeholders involved in the ISMSsystem.
Has the vision been defined in a specific timeframe for a specific reduction in fatalities or in fatalities and injuries?	The vision statement is explicitly defined in terms of timeframe and type of crash severity reduction.
Are the mission and vision statements clear and compelling, and are they understood and supported by all of the members of the SPL who participate in the ISMSsystem?	The mission and vision statements are easily understood and are known to be clear and compelling because they have been tested with key stakeholders. The mission and vision statements are clearly communicated and reinforced on a regular basis.
Have the mission and vision statements been reviewed and accepted by all the major stakeholders (especially the leadership) within the past 4 years?	The safety mission and vision statements are reviewed every 4 years and updated (or reaffirmed) on the basis of inputs received from key stakeholders.
Has highway safety been established as a goal of the ISMSsystem at the highest policy level?	Highway safety is incorporated in the formal priorities in the Governor's agenda, and each of the agencies has mutually reinforced goals within its individual strategic plan.
Have the mission and vision statements been translated into achievable goals for emphasis areas? Are responsibilities for accomplishing these goals shared and agreed upon by the implementing agencies that will have to take action?	<p>An ISHSPlan lists all of the goals and supporting strategies and actions.</p> <p>The ISHSPlan identifies the resources required to accomplish the plan, and the resources are in place to support the plan.</p>
<b>ISMSYSTEM ORGANIZATION</b>	
Does the jurisdiction have a formally designated SPL coalition?	An SPL is formally chartered by executive order, legislation, or memorandum of understanding among agencies. The formal charter of the SPL includes the necessary funding for running and maintaining the ISMSsystem.
Are all the major safety stakeholders represented on the SPL?	The SPL includes representatives from all key stakeholder groups that participate in and commit resources to the operation of the ISMSsystem.
Is the SPL fully knowledgeable of the day-to-day operations and results?	ISMSsystem senior leaders (the SPL) are regularly (e.g., quarterly) provided reports and information pertaining to the ISMSsystem operation (i.e., ISMSsystem performance results, using agreed upon performance measures, etc.).

TABLE B-2 (Continued)

ISSUE	WHAT TO LOOK FOR
Have the mission and vision statements, emphasis areas, objectives, strategies, roles, and responsibilities been documented, and are they comprehensive?	The SPL and Operations Manager produce a report that describes the safety situation in a jurisdiction, including identification of the major safety concerns, a list of potential emphasis areas, selected emphasis areas, and action plans.
Is there an organizational chart for the ISMSystem organization that clearly defines roles, responsibilities, and relationships? Is it easily available?	An organizational chart documents reporting relationships for everyone involved in managing and improving the ISMSystem. A system diagram shows the interrelationship among all the major components of the ISMSystem.
Is there an emphasis on safety, and is the emphasis reflected in the policies, standards, and work-related requirements?	Safety is formally designated as an organizational priority in all organizations that participate in the ISMSystem, and policies and standards are aligned to support safety goal attainment.
Are all agencies involved in the ISMSystem kept fully informed of approved action plans and progress, and are they systematically involved in coordinating improvements?	There is a formal process for ISMSystem stakeholders (represented by the SPL) to address the safety issues in their jurisdiction along with their information needs, and the communication and coordination necessary for decision making exist to determine whether communication and coordination are adequate and feedback received is used.
<b>OPERATIONS MANAGER</b>	
Has a person been designated and empowered to coordinate and operate the ISMProcess?	The Operations Manager's roles and responsibilities are designated in writing. SPL team members and the Operations Manager agree on performance expectations and how the expectations will be assessed from accountability perspectives. The Operations Manager has the organizational skills to understand how the current ISMSystem works, in terms of both a systems view (components of the system and the interrelationships between these components) and how well the ISMSystem is performing (e.g., knowledge of the systems performance, effects of improvement actions, accountabilities discharged, and barriers to improvement being addressed).
Does this person have adequate time and resources to devote to the ISMProcess?	The Operations Manager's primary job responsibility is safety management. Any other job responsibilities are assigned with the proviso that they are of secondary importance and cannot interfere with the effective conduct of managing the ISMProcess.
<b>RISK ANALYSIS AND EVALUATION (RAE) GROUP</b>	
Do analysis units exist on staff, or are they available through either contract or memorandum of understanding, to perform the analysis of highway safety information for (a) identification of problems, (b) development of emphasis areas and strategies, and (c) evaluation of programs or projects implemented through the ISMSystem?	Analysts with competence in highway safety program analyses and working knowledge of statistics, evaluation, and information systems are identified and formally engaged. The RAE group is provided with the technology needed to perform the group's roles and responsibilities.
<b>TASK TEAMS</b>	
Does the Task Team leader, who is selected by the SPL and/or Operations Manager, have experience in the emphasis area programs?	The Task Team leader should come from the agency that will most likely implement the strategy(ies) or have the greatest involvement.
Are Task Team members assigned to strategy action plans by associated discipline or agency?	The SPL and/or the Operations Manager should assign Task Team members for each defined emphasis area according to the members' expertise.
Do the Task Team members have highway safety expertise, especially in the areas to which they are assigned?	Task Team members should come from the SPL agencies with experience in the agencies' highway safety programs.
Is there a process in place for systematically soliciting feedback from Task Team members on organizational productivity?	A progress reporting system should be in place that responds to Task Team comments and suggestions.
Are the Task Teams sufficiently informed about the SPL's expectations with respect to emphasis area goals? Are the Task Teams receiving sufficient guidance with respect to policies, guidelines, and specifications?	Task Teams are formally informed with regards to what is expected from them and the support they will be provided in meeting expectations.
Is the representation of the Task Team sufficient? Will the inputs from the members result in integrated and coordinated strategies and action plans?	The SPL identifies all of the key stakeholders who will be affected by the Task Team's assignment and then ensures that the Task Team is made up of members who collectively can view their tasking from all required perspectives.
<b>HIGHWAY SAFETY INFORMATION SYSTEMS</b>	
Are accurate crash data files readily available and integrated into the agency's management information systems? Do the crash data files contain sufficient information about the crash, including, at a minimum, the severity of the crash, the type and exact location of the crash, contributing circumstances, time of day, and environmental conditions	Safety data, information, and knowledge needs are clearly articulated and incorporated into organized knowledge and information management processes and systems.

(continued on the next page)

**TABLE B-2 (Continued)**

ISSUE	WHAT TO LOOK FOR
(e.g., weather and roadway surface condition)? Do state databases contain, at a minimum, a set of crash data elements with standardized definitions that are relevant to injury control, highway and traffic safety (e.g., the Model Minimum Uniform Crash Criteria [MMUCC])?	
Are all the safety databases (traffic crash data, vehicle registration data, driver license and history data, roadway data, emergency medical services and trauma data, traffic citation and conviction data) linked? If not, is there a plan to link the databases being implemented, or can downloads from each database allow the creation of a safety database or research database?	The ISMSystem is viewed holistically with regards to data and information needs. These are then used to develop a comprehensive and integrated knowledge and information management process and system.
Are the SPL, Operations Manager, RAE group, and Task Team members able to get the data they need when they need it to perform their tasks? Are data and information used to predict process performance?	The key users of the information systems are involved in information needs identification, and their inputs are used to design the system. Data and information are used proactively to predict performance of the ISMSystem. Actual performance is compared against predictions to develop lessons learned and improvements.
Can the information contained in the safety databases be accessed online?	Technology to enable safety analysis is provided to people within the ISMSystem who are responsible for conducting safety analysis and developing findings, conclusions, and recommendations.
<b>ISMPROCESS STEP 1</b>	
Has the SPL or Operations Manager directed that a jurisdictionwide highway safety profile be prepared?	The RAE group within the ISMSystem organization or an outside agency, such as a university, management consultant, or a research institute, is periodically engaged to conduct an analysis of the jurisdiction highway safety information system to develop a jurisdiction highway safety profile.
Are the decisions made by the SPL based on a comprehensive and statistically defensible review of highway safety information?	The SPL is committed to using the scientific method to identify potential emphasis areas developed by the in-house or contract analysis unit.
Are the available data sufficient in quantity and quality for the recommended methodologies to be implemented?	Data requirements that align with safety methodologies are clearly articulated and used.
Are the analyses conducted according to recommended methodologies?	Agreed upon methodologies are used, and feedback on the methodologies is solicited annually or as required to improve methodologies and align data requirements.
Is there a process in place for reviewing the actions and outputs of Step 1? Are the review results incorporated into the next year's ISMProcess?	ISMProcess performance measures are established and used to assess process performance and make improvements. A report of the results or products on Step 1 (i.e., a Crash Analysis Report or a Highway Safety Profile Report) is published.
Are there policies, procedures, and processes in place for <ul style="list-style-type: none"> <li>Identifying sites with potential for improvement?</li> <li>Identifying target groups for potential countermeasures?</li> <li>Establishing profiles of high-risk groups?</li> <li>Identifying driver and/or vehicle failures?</li> <li>Analyzing crash information for different severity reduction strategies?</li> <li>Monitoring performance indicators?</li> </ul>	Formal policies, procedures, and processes are in place for identifying key factors that contribute to safety problems.
<b>ISMPROCESS STEP 2</b>	
Did the report resulting from Step 1 indicate highway safety emphasis areas of concern?	The SPL and Operations Manager develop highway safety emphasis areas based on the jurisdiction safety profile. These emphasis areas should be in areas that have promise for crash reduction based on the preliminary analysis conducted in Step 1.
Has the SPL, in conjunction with individual Task Teams, formulated goals for each emphasis area that directly align with the vision?	There is clear agreement among all SPL members and Task Teams on safety emphasis goals, and the emphasis goals are expressed in terms of reductions in fatalities or in fatalities and injuries within each emphasis area.
Do the AASHTO Implementation Guides and/or other resources exist for the emphasis area?	If the relevant AASHTO Implementation Guides exist, they are used as a primary resource by the Task Team.
Have the emphasis area goals been established and evaluated to form the basis for future decision making?	Emphasis area goals are designed to improve the ISMSystem as a whole and to align priorities for individual organizational safety improvements accordingly.
How does the SPL review the ISMSystem's performance and use the findings to identify opportunities for improvement and innovation?	Measures exist for (a) assessing ISMSystem performance, (b) using the measures to determine problems and opportunities for improvement, and (c) establishing improvement priorities.
Are the goals formulated by the SPL for an emphasis area suitable?	Criteria are developed for assessing goals and then used to ensure that the goals are the optimum ones needed to improve the ISMSystem.

**TABLE B-2 (Continued)**

ISSUE	WHAT TO LOOK FOR
Has a Task Team leader been selected for each emphasis area?	A Task Team leader from the agency most likely to implement the major strategies is selected by the SPL and Operations Manager.
Have Task Team members been assigned to each emphasis area?	Task Team members are assigned to an emphasis area with experience in their agencies' programs related to the emphasis area.
<b>ISMPROCESS STEP 3</b>	
Are the strategies selected based on the consultation of the AASHTO Implementation Guides and/or highway safety data?	The strategies are based on the AASHTO Implementation Guides and available safety data.
Do the strategies reflect multidisciplinary input and actions?	The strategies are a result of the multidisciplinary Task Team and reflect the Task Team's judgment and analysis.
Are there monitoring criteria included for each action plan?	A Gantt chart is prepared that identifies all actions required to implement the strategies with identified milestones for completion.
Does the Task Team have adequate support from the Operations Manager and the RAE group for necessary analysis, technology, or other budget items required to develop strategies and preliminary action plans (i.e., travel, supplies, IT tools, etc.)?	Funding needs are developed as part of the overall planning process and are incorporated into budget processes. The Operations Manager establishes links between the Task Teams and the RAE group.
Is there a decision-making process developed for the Task Teams to use?	An accepted decision-making (i.e., problem identification) process is used by Task Teams in developing strategies and eventual action plans for each approved strategy.
Are all team members in clear agreement on their tasking and committed, involved, and satisfied?	At the beginning and end of each meeting, all actions will be reviewed and approved by consensus.
<b>ISMPROCESS STEP 4</b>	
Do the Operations Manager and his or her staff have sufficient guidance on how to assess the suitability of strategies?	Performance measures for strategies are developed by knowledgeable Task Team members and then used to assess performance, enhance learning, and make improvements.
Do the Operations Manager and his or her staff use the recommended optimization techniques to select strategies for further development and implementation? If not, how are strategies selected for inclusion in the jurisdictional strategic highway safety plan? Is this approach suitable?	Optimization techniques are fully understood and used to make recommendations for improving the ISMSystem.
Does a suitably skilled and qualified staff perform the optimization procedures?	The set of strategies for each goal is reviewed holistically to ensure alignment and efficiency in implementation. This process is repeated for the full suite of strategies tied to all goals.
Are emphasis area strategies and associated action plans clear and sufficient to accomplish the goals?	The SPL, with the support of the Operations Manager, views the preliminary action plans holistically and ensures that, in the goals designed to improve the ISMSystem, there are no gaps and that the goals selected mutually support each other in the attainment of the system's vision.
Is the allocation of funding across the agencies responsible for engineering, enforcement, education, and emergency services based upon a formal process that includes cost-benefit analysis?	Funds for safety improvement programs are clearly defined and incorporated into budgetary processes.
Is the ISMProcess understood? For example, is the linkage from safety goals to safety process clear, and are process goals supportive of broader organizational safety goals?	For every step in the implementation, the agency that is implementing the strategy understands the intended emphasis area goals and purpose of the actions. When there is disagreement or confusion as to actions or goals, the implementing agency has direct contact with the appropriate Task Teams in order to address any questions.
<b>ISMPROCESS STEP 5</b>	
Are performance measures established for determining the degree to which goals are achieved and strategies implemented?	Impact and process performance measures are developed for each goal.
Do action plans contain a suitable performance measurement plan?	Performance measures are developed and aligned from goals to strategies to actions.
Do action plans clearly identify actions required of transportation agency stakeholders, suppliers, and other transportation agency components critical to the agency meeting its goals?	Action plans explicitly define individual responsibilities for each task or action and are grouped by organization for ease of use and clarity of responsibility.
Are resources allocated to ensure that strategic goals and action plans can be accomplished?	Resource requirements for goals and associated strategies and actions are developed as part of the planning process. Then, they are integrated into the funding process. If there are funding shortfalls, plans (i.e., goals, strategies, and actions) are refined accordingly.
Were all ISMSystem stakeholders, especially the potential implementing agency, involved in the strategy development process, and do they fully support the action plan and the corresponding resource allocation?	All key stakeholders are involved in the planning process and there is an explicit process for testing for the level of understanding and support for safety plans for all of the key stakeholders.

(continued on the next page)

**TABLE B-2 (Continued)**

ISSUE	WHAT TO LOOK FOR
<b>ISMPROCESS STEP 6</b>	
Have sufficient resources been allocated to support the performance measurement effort?	The performance measurement process has adequate funding and staffing commitments to execute it successfully.
To what degree are these performance measures used to enhance accountability and to support organizational learning and improvement? Is there a process for improving the performance measures and the associated learning and improvement process?	The performance measures are used, at least quarterly, to assess progress, highlight accountabilities, and develop insights and improvements. The performance measurement and improvement process has its own measures that are used to assess progress, highlight accountabilities, and develop insights and improvements.
Are there sufficiently skilled and experienced staff to perform the processing and analysis of performance indicators?	The Operations Manager and his or her team members understand process improvement and how to use data to identify improvement opportunities and assess effectiveness.
Are the results of the performance measurement effort sufficiently documented and communicated to the SPL and all stakeholders?	Performance results are documented in formal reports and made available to all stakeholders. The reports addresses accountability and results, including what was done, by whom, when, where, and what were the outcomes.
Has an ISHSPlan been prepared and published for dissemination to all SPL members and other safety stakeholders?	An ISHSPlan has been published.

## Appendix B References

1. National Association of Governor's Highway Safety Representatives, National Association of State EMS Directors, NHTSA, and U.S. Department of Transportation, "Keys to Success: State Highway Safety and EMS Agencies Working Together to Improve Public Health." Vol. HS-809 111, Washington, D.C. (8-1-2000), pp. 1–50.
  2. Wilbur Smith Associates, "Vermont's Safety Management System Phase 1 Study." (5-1-2001).
  3. "Case Studies of Highway Safety Management Systems." FHWA-SA-97-061. (1997).
  4. <http://www.unm.edu/~dgrint/accf.html>
  5. Washington State Traffic Safety Meeting. <http://www.wsdot.wa.gov/mapsdata/tto/Concept.htm>, Washington State DOT (2000).
  6. Welch, T. M., and Stahlhut, M., "Safety Management in Iowa—After the Mandates." Mid-Continent Transportation Symposium 2000, Iowa State University, Ames (2000).
  7. Shinar, D., and Treat, J. R., "Tri-Level Study: Modification Task 3: Validity Assessment of Police-Reported Accident Data." National Highway Traffic Safety Administration, DOT-HS-805-087, Washington, D.C. (1979).
  8. Shinar, D., Treat, J. R., and McDonald, S. T., "The Validity of Police Reported Accident Data." *Accident Analysis and Prevention*, Vol. 15, No. 3 (1983).
  9. National Highway Traffic Safety Administration, "Traffic Safety Facts 2000." Washington, D.C. (2001).
  10. [www.nhtsa.dot.gov](http://www.nhtsa.dot.gov)
-

## **APPENDIX C**

### **STEP-BY-STEP WALK-THROUGH EXAMPLE OF THE ISMPROCESS**

Appendix C provides a hypothetical example of an application of the ISMProcess. This highly simplified example is intended not to be exhaustive but rather to provide some

additional information for decision makers applying the ISMSystem. The numbers used in this example are for illustration purposes only.



## STEP 1: REVIEW HIGHWAY SAFETY INFORMATION

During the first step of the ISMProcess, the Safety Program Leadership (SPL) group of a jurisdiction initiates an annual review of highway safety information and existing highway safety programs to identify potential emphasis areas for corrective action. The Operations Manager, working with the Risk Analysis and Evaluation (RAE) group, applies crash problem identification techniques (such as those described in Appendix D1) to the jurisdiction's databases.

For the purpose of this example, assume that four emphasis areas were identified after the high-level analysis of highway safety data by the RAE group and due consideration by the SPL of political, practical, and social issues. The four emphasis areas, in order of priority, are

- Run-off-road crashes,
- Aggressive driving,
- Crashes with trees, and
- Impaired driving.

Based on the high-level analysis and knowledge of the status of safety in the jurisdiction, the SPL also establishes an overall highway safety goal (vision) during this step. For this example, the vision statement is as follows:

*We envision a 15% overall reduction in injury crashes by 2008.*

For this hypothetical example, assume that the state experienced approximately 24,000 injury crashes in 2002 on all road types. A 15% reduction between 2003 and 2008 (5 years) implies a 3% reduction per year, or a reduction of 720 injury crashes per year.

## STEP 2: ESTABLISH EMPHASIS AREA GOALS

The CARE (Critical Analysis Reporting Environment) software package (1) is used to perform a subset comparative analysis of crash data. (Appendix D1 includes additional descriptions and definitions of CARE terminology). The analysis results indicate that the number of run-off-road injury crashes on two-lane rural highways is over-represented (i.e., *Over Rep* is significantly higher than 1). The maximum number of injury crashes that can be reduced by applying a combination of countermeasures is approximately 10% per year (*Max Gain* = 90) of the existing total number of run-off-road crashes on all two-lane rural highways. There were 905 run-off-road crashes on two-lane rural highways in the state in 2002, of which 180 were injury and fatality crashes.

When setting an emphasis area goal, it is essential to place the emphasis area in the context of the overall safety goal or vision of the ISMSys and continue identifying other emphasis areas or road types to achieve the overall vision or goal.

For this emphasis area, the SPL and Operations Manager invite the state DOT traffic and safety engineer to lead a Task

Team with the focus of formulating strategies and related action plans to reduce run-off-road crashes. The SPL, Operations Manager, and Task Team leader meet to assess and review existing information and past experience in the jurisdiction related to the run-off-road crash emphasis area in order to establish a goal for this emphasis area.

Based on the results of a comparative subset analysis, the SPL, Operations Manager, and Task Team leader decide on the following goal, which in their opinions is both measurable and attainable:

*Reduce run-off-road injury crashes on all two-lane rural highways by 10% per year.*

This goal implies a reduction of 18 injury crashes per year, or 2.5% of the overall vision of a reduction of 720 injury crashes per year.

The choice of agencies invited to join this Task Team is guided by a review of previous studies, analyses, and research. The run-off-road emphasis area can include a strategy aimed at making improvements to the roadway and roadside, thus requiring the involvement of highway design, traffic engineering, maintenance, and safety staff. Another strategy may be aimed at reducing vehicle-operating speeds, requiring the involvement of law enforcement. This emphasis area also deals with driver fatigue, driver perception and reaction, and other issues suggesting the involvement of human factor or driver safety professionals. Other strategies for run-off-road crashes may be found in the AASHTO Implementation Guides for this and related crash types.

The Operations Manager and Task Team leader, with the approval of the SPL, prepare a formal invitation letter addressed to the head of the unit from which each Task Team member is assigned. The letter outlines the emphasis area and the desired level of input required from the agency and the Task Team member. The following agencies are invited for the Task Team responsible for the run-off-road crashes:

- Federal, state, and local transportation officials (highway design, traffic operations, and maintenance divisions);
- State and local highway patrol or police;
- Emergency medical services;
- Governor's Office of Highway Safety;
- Driver licensing agency; and
- Human factor professionals from a university or consulting firm.

The newly formed Task Team reviews the existing information and available research results relating to run-off-road crashes. The Task Team members suggest adding the baseline from the last year of complete statistics (year 2002). The SPL and Operations Manager agree to this revision, so the goal for this emphasis area becomes the following:

*Reduce run-off-road injury crashes on all two-lane rural highways by 10% per year from a year 2002 baseline of 180 injury crashes.*

The SPL provides an estimated budget of \$50,000 per year for this emphasis area for 2003 and 2004. This budget does not include the salaries of the Task Team leader and members who are employed by any of the coalition agencies.

Similar activities take place in ISMProcess Step 2 for the other three emphasis areas identified in ISMProcess Step 1. The next steps exemplified in the following sections deal with the activities of Task Team for the run-off-road emphasis area.

### STEP 3: DEVELOP OBJECTIVES, STRATEGIES, AND PRELIMINARY ACTION PLANS TO ADDRESS THE EMPHASIS AREAS

The RAE group, as directed by the Task Team, examines the characteristics of run-off-road injury crashes in a greater level of detail. This detailed analysis is aimed at identifying objectives. The Task Team uses the AASHTO Implementation Guides. The run-off-road emphasis area is matched with the following three objectives:

- Keep vehicles from encroaching onto the roadside.
- Minimize the likelihood of crashing into an object or overturning if the vehicle travels beyond the edge of the shoulder.
- Reduce the severity of the crash.

The detailed analysis of run-off-road crash data, for this appendix's hypothetical state, reveals that 15% of vehicles that leave the roadway crash with an obstacle or overturn, and about 60% of those that crash result in injury to the occupants. (*These percentages are hypothetical and are only used for illustrative purposes.*) A review of literature reveals that these figures are very dependent on vehicle operating speeds, road geometry, adjacent road sideslope, and the presence of fixed objects close to the roadway. The detailed analysis also reveals that the roadside object that most errant vehicles collide with in the state is trees. The Task Team members decide that in order to achieve the emphasis area goal, the following objectives should be pursued:

- Decrease the likelihood of a vehicle leaving the roadway and crashing into a feature and/or overturning from 15% to 10%.
- Decrease the likelihood of a run-off-road crash causing injury from 60% to 40%.

The Task Team members note that these specific objectives are reflected in the three objectives from the AASHTO Implementation Guides. Therefore, the Task Team continues the process using the AASHTO objectives.

Once sufficient information becomes available, the Task Team members collaboratively identify and develop a number of strategies to address each objective.

At this stage, it is important to develop process and impact performance measures for each proposed strategy. Process and impact performance measures are discussed in detail in Appendix D3.

The Task Team identifies *engineering strategies* applying the AASHTO Implementation Guides for run-off-road injury crashes (Table C-1).

Other agencies can contribute to nonengineering strategies to realize the goal of this emphasis area. Vehicle speed is a major factor contributing to the severity of run-off-road crashes. The Task Team propose the following strategies to address the problem of excessive speeds on two-lane rural highways in the state:

- Increase law enforcement at locations that experience the highest incidence of run-off-road crashes.
- Launch a public information and education campaign targeted to the required driver behavior for these locations.

Many fatalities in rural crashes result from emergency medical services being unable to reach the crash location within a short timeframe. The following strategies are identified to improve emergency response times:

- Train first responders (possibly the general public) to accurately describe injuries and identify the crash location when reporting the crash.

**TABLE C-1 Engineering strategies to address run-off-road injury crashes**

Objective	Strategies
Keep vehicles from encroaching onto the roadside.	Install midlane rumble strips. Install shoulder rumble strips. Provide enhanced delineation of sharp curves. Provide improved highway geometry for horizontal curves.
Minimize the likelihood of crashing into an object or overturning if the vehicle travels beyond the edge of the shoulder.	Design safer slopes and ditches to prevent rollovers. Remove or relocate objects in hazardous locations.
Reduce the severity of the crash if the vehicle does crash into an object or overturn.	Improve the design of roadside hardware. Improve the design and application of barrier and attenuation systems.

- Position emergency response units geographically where feasible, based on prior crash and incident history.

Since other Task Teams are focusing on “Impaired Driving” and “Crashes with Trees,” the run-off-road Task Team decides not to formulate any strategies aimed at these emphasis areas. The three Task Teams coordinate their strategies throughout the ISMProcess.

The run-off-road Task Team gathers information about the potential safety impacts of the above strategies. This information is essential in order to obtain reliable estimates of (a) the expected benefits and costs of implementation and (b) the required resources for successful implementation.

For each strategy, the run-off-road Task Team investigates and determines the following information:

- Expected safety impact (accident modification factors);
- Resources required to implement;
- Cost of applying the resources to implement;
- Potential difficulties or barriers that could impact successful and effective implementation;
- Process performance measures;
- Impact performance measures;
- Data requirements for obtaining process and impact performance measures;
- Data collection processes;
- Resources required to collect, process, and analyze performance measurement data;
- Evaluation research design;
- Current organizational, institutional, and policy issues that need to be recognized before implementing;
- Issues that could affect the implementation time; and
- Any legislation that needs to be in place before implementation.

A preliminary action plan must be developed identifying the key activities to be undertaken to implement each strategy. Note that at this stage it is not necessary to schedule the activities with respect to time and location. This is done during the detailed action plan development in ISMProcess Step 5.

Three examples that encompass the development of strategies in the areas of engineering, enforcement and education, and emergency medical services, are provided below. These are applicable to the run-off-road Task Team.

#### **Engineering: Identify Segments for Improvement and Estimate the Safety Impact of the Improvements**

At the request of the Task Team, the RAE group develops Safety Performance Functions (SPFs) to estimate the expected number of run-off-road crashes per mile per year on two-lane rural highways in the state by severity. The SPFs, which were developed using 3 years of data assembled from the engineering database and police crash reports, are mathematical equa-

tions for estimating the expected number of crashes on a road section as a function of its traffic and geometric characteristics.

The SPFs were used in “network screening” to identify sections of highway with a high potential for improvement. The SPFs also aid the evaluation process by facilitating the use of the empirical Bayes approach to conduct longitudinal before-and-after studies. “Network screening” is an approach that uses SPFs to identify locations with the potential for safety improvement. A major ongoing FHWA project is providing states with practical state-of-the-art tools, including software, for conducting network screening as part of *SafetyAnalyst* (Software Tools for Safety Management of Specific Highway Sites). *SafetyAnalyst* is briefly discussed in Appendix D1.

In this example, imagine that 12 road segments, totaling 400 miles, are identified, through network screening performed by the RAE group, as locations with a potential to reduce run-off-road crashes. Also imagine that these segments experience a total of 805 run-off-road crashes over a 3-year period.

The Task Team conducts a detailed safety review of the 12 locations, finding the following:

- Eight of the sites included horizontal curves.
- Approximately one-third of the run-off-road crashes at all 12 locations involved drowsy drivers.
- Speed was a major contributing factor to run-off-road crashes.

Based on these findings, it is decided that improvements to curves, shoulder rumble strips, and targeted speed reduction countermeasures will have the most impact on run-off-road crashes in this jurisdiction. The Task Team also develops modest revisions to current guidelines with respect to the design of horizontal curves, sideslopes, and ditches for future improvements end construction.

Safety reviews are conducted at each of the 12 locations. These reviews recommended the following actions to meet the three objectives for run-off-road crashes:

- Install shoulder rumble strips (to a total of 400 miles of roadway).
- Improve delineation around curves (to a total of 4 miles) using chevrons and raised pavement markers.
- Construct geometric improvements at eight sharp curves.
- Remove 66 trees.
- Install guide rails (to a total of 6 miles of roadway).
- Install midlane rumble strips (a potential pilot project).

The installation of shoulder rumble strips is used here to illustrate the calculations to determine potential safety benefits. The data and the empirical Bayes estimate of the long-term mean crashes per year for the 12 locations is shown in Table C-2. Assuming an accident modification factor of 0.70 for target crashes, the application of shoulder rumble strips at the 12 identified locations will result in expected crash reductions per year shown in the last column of Table C-2 for a total annual saving of 70.1 crashes over all 400 miles of

**TABLE C-2 Data for treatment locations for run-off-road crashes**

Location	AADT (number of vehicles)	Length (miles)	Number of target crashes (n = 3 years)	Estimated long-term mean (m) per year	Expected crash reduction per year
1	9800	50	120	35.0	10.5
2	7200	40	75	21.6	6.4
3	8200	10	27	6.9	2.1
4	7200	50	88	26.5	8.0
5	6200	40	63	18.4	5.6
6	6100	20	38	10.0	3.0
7	10300	20	47	14.2	4.2
8	6000	25	38	11.3	3.5
9	10200	50	117	35.5	10.5
10	9500	15	40	11.1	3.3
11	6900	20	42	11.0	3.4
12	7000	60	110	31.8	9.6
<b>Totals</b>		400	805		70.1

AADT = annual average daily traffic.

roadway. Additional crash reductions may occur through other engineering strategies, enforcement, public information and education, and improvements to emergency services. All strategies tie in with the goal of a 10% reduction per year in the baseline total of 905 crashes. Crash severity has not been considered for the purposes of this example, but is normally factored in.

#### *Cost-Benefit analysis of shoulder rumble strips*

Table C-3 summarizes the information for the shoulder rumble strip preliminary action plan. Similar calculations and information are required for each of the proposed countermeasures (e.g., removal of trees and delineation measures).

The implementation of midlane rumble strips is considered, in this example, to be an experimental, innovative treatment that has not been evaluated to date. There is no reliable source of the expected effectiveness of midlane rumble strips. The Task Team compiles a justification report to the Opera-

tions Manager regarding a pilot study aiming to determine the effectiveness of midlane rumble strips.

#### **Enforcement and Education: Estimate the Safety Impact of Speed Enforcement and a Public Information Campaign**

As discussed earlier, speed is determined to be a major factor in run-off-road crashes. To assess the extent of excessive vehicle speeds on two-lane rural highways, a baseline survey is conducted. Ten representative points are selected on the road network, and a speed survey is conducted at each survey point. It is found that, on average, 30% of all vehicles exceeded the posted speed limit.

The Task Team decides on a strategy to increase the law enforcement on two-lane rural roads, with emphasis on locations with a high potential for safety improvement, which are identified during network screening. To support the law enforcement activities, it is decided to devise a public infor-

**TABLE C-3 Summary of information for the installation of shoulder rumble strips**

<b>Expected effectiveness</b>	AMF = 0.70 for target crashes (run-off-road crashes)
<b>Requirements</b>	Installation of 400 miles of shoulder rumble strips at 12 identified locations over 2 years
<b>Cost</b>	\$1,281,000
<b>Benefits</b>	\$40,690,000 (70 fewer crashes per year)
<b>Process performance measures</b>	Number of miles of road fitted with rumble strips
<b>Impact performance measures</b>	30% reduction in run-off-road crashes over treated road sections
<b>Potential difficulties</b>	State and county authorities do not have the milling equipment or the technical expertise to make rumble strips on existing roads
<b>Issues affecting time</b>	Purchasing equipment takes 2 months and training may require 2 weeks

AMF = accident modification factor.

mation campaign. The aim of the public information campaign is to promote lower speeds by informing drivers of the risks associated with high speeds and to increase the risk of apprehension by reporting on the results of law enforcement activities. The public information and education campaign will take the form of radio advertisements encouraging drivers to drive slower and newspaper advertisements regularly reporting on the results of speed law enforcement activities (e.g., number of citations and number of arrests).

After considering these strategies, the Task Team decides to consider three different levels of implementation for increased speed enforcement. A group of experts is convened to assess (a) the reduction in the speeding rate (where speeding rate is the percentage of drivers exceeding the speed limit) that can be attained for the three selected levels of implementation with a supporting public information strategy and (b) the effect that these reductions in speeding rates could have on crash frequency and severity (i.e., benefit). The information on which the experts reach consensus is listed in Table C-4.

#### **Emergency Medical Services: Estimate the Safety Impact of Decreasing Response Times**

It has been concluded that a large number of fatalities in rural crashes result from long emergency medical services response times. After analyzing the emergency services databases, it becomes apparent that the response time is longer than 40 minutes for more than 30% of the fatal and injury crashes on two-lane rural highways.

Following a procedure similar to the previous “Law Enforcement and Public Information and Education” example, levels of implementation are investigated for the two strategies (i.e., training of potential first respondents and positioning of emergency response units). A group of experts is convened to assess the impact of these two strategies on the severity outcome of the crashes. The Task Team, with the experts, concludes that this percentage can be reduced to 15%. Further investigation by the Task Team reveals that in part the long response time is the result of insufficient personnel and equipment to respond to all emergencies in a timely manner. A report is prepared to obtain feedback and buy-in from the implementing agencies prior to completing the preliminary

action plans. These reports document future requirements, such as the need for funding to increase the number of ambulances, trained paramedics personnel, and equipment.

#### **STEP 4: DETERMINE APPROPRIATE COMBINATION OF STRATEGIES FOR IDENTIFIED EMPHASIS AREAS**

After receiving all the cost and benefit information included in the preliminary action plans for all different strategies from all the different Task Teams, the Operations Manager computes an optimization procedure (as described in Appendixes D1 and D5). The result of this procedure reveals which combination of strategies, if implemented, will give the optimal economic returns on an investment based on a number of estimated resource implementation (deployment) levels. In addition, the Operations Manager receives a list of potential pilot projects for different strategies. The Operations Manager, in conjunction with the SPL, selects pilot projects for implementation.

For this example, the strategies selected during the optimization process are summarized in Table C-5. In addition, a pilot study for midlane rumble strips is approved through the prioritization procedure undertaken separately for all proposed pilot studies.

The expected benefits of these strategies at the selected levels of deployment should also be determined by the Task Team and explicitly stated (discussed further in Appendix D1). The sum of the benefits of all selected strategies should match the goal for the emphasis area. In this case, the sum of the benefits should equal a 10% reduction per year in run-off-road injury crashes on two-lane rural highways. The benefits of these strategies are not included here for brevity.

#### **STEP 5: DEVELOP DETAILED ACTION PLANS**

For each strategy (including pilot studies) approved by the Operations Manager for implementation, the Task Team develops a detailed implementation (i.e., action) plan. The detailed action plan outlines all the activities required to implement every strategy, with timeframes, responsibilities, and

**TABLE C-4 Annual level of implementation costs and benefits**

Level of implementation	Cost	Benefit	Speeding rate reduction
Level 1 (Additional 1000 hrs)	\$400,000	\$2,000,000	3%
Level 2 (Additional 1500 hrs)	\$600,000	\$2,600,000	8%
Level 3 (Additional 2000 hrs)	\$800,000	\$4,100,000	15%

NOTE: The figures are hypothetical and are for illustrative purposes only.

**TABLE C-5 Strategies and deployment levels selected through optimization for run-off-road crashes**

Objective	Strategies
Keep vehicles from encroaching into the roadside.	Install shoulder rumble strips: 400 miles. Provide improved delineation at locations identified during the safety review: 4 miles.
Minimize the likelihood of crashing into an object or overturning if the vehicle travels beyond the edge of the shoulder.	Remove roadside trees: total of 66.
Reduce the severity of the crash.	Install guardrails at locations selected during the safety review: 6 miles. Conduct additional speed law enforcement for the next 2 years on two-lane rural highways: 1,500 hours per year. Reduce emergency services response times through public education and resource management.

performance measurement criteria assigned to each activity. The detailed action plan includes the following:

- All preparation work needed to implement a strategy (e.g., resources for rumble strip installation),
- A performance measurement activities schedule (e.g., field observations for data collection), and
- Assignments of responsibilities for executing performance measures (e.g., processing and analysis of the data and preparation of reports).

During the scheduling of performance measurement activities, the desired sample sizes were determined in order to obtain reliable performance measures as specified in the evaluation designs that were compiled for each strategy. The required sample size is an important factor to consider during the determination of resource requirements in terms of staff hours and equipment to conduct performance measurement activities.

### Shoulder Rumble Strips

Figure C-1 is an example of a typical Gantt chart, which indicates when activities should be conducted and the relationship between activities.

The Task Team decides that 300 miles of shoulder rumble strips will be implemented by Agency A and Agency B combined during the first year. The remaining 100 miles are scheduled for implementation during the second year.

The detailed action plan contains monthly milestones (process performance measures) to ensure that 300 miles of rumble strips are implemented within the first 12 months. The plan requires each implementing agency to provide monthly feedback to the Operations Manager (or agency responsible for outlining the progress) with respect to the amount of rumble strips installed and at what cost. The detailed action plan schedules these performance reports and assigns responsibility for completion and submission in accordance with the implementing agencies.

Figure C-2 illustrates the planned monthly milestones for rumble strip installation per participating agency.

### Speed Enforcement

A detailed action plan was developed for the 2-year law enforcement program. The detailed action plan contained information on location, date, time, and the enforcement agency with responsibility for the jurisdiction in accordance with the implementing agency. The number of enforcement hours per month and number of citations issued are used to measure the process and impact performance of the law enforcement program.

Each law enforcement agency agrees that a feedback report on each activity performed will be completed on a weekly basis. The feedback reports are provided to the Operations Manager, who uses the reports to periodically compile an integrated progress report. This integrated report is part of the process measure evaluation. For example, from assessing the number of hours worked per month, the Operations Manager can determine whether the target of an additional 1,500 hours of law enforcement per year will be attained.

### Public Information and Education

The detailed action plan for the public information and education component is designed to support and enhance the law enforcement campaign to facilitate synergy between these two components. Radio advertisements are scheduled for airing at strategic times in advance of major law enforcement efforts. Newspaper ads are scheduled after high-profile law enforcement activities to report on the number of citations issued to motorists.

To measure the performance of the public information and education campaign, a baseline driver attitudinal survey is conducted prior to the onset of the campaign to determine drivers' attitudes toward speeding and their knowledge of the risks associated with excessive speeding. Similar surveys at

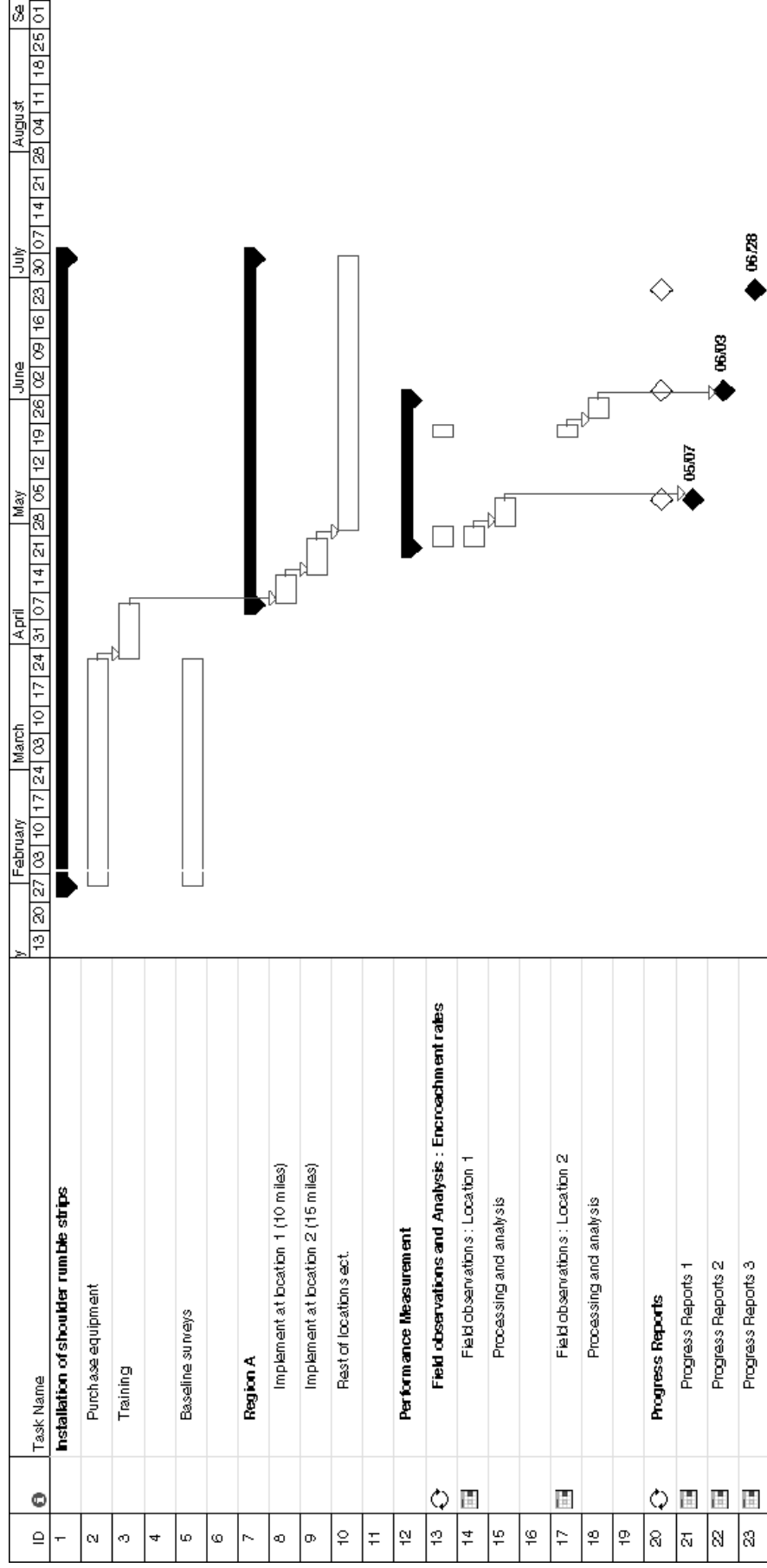


Figure C-1. Example Gantt chart.

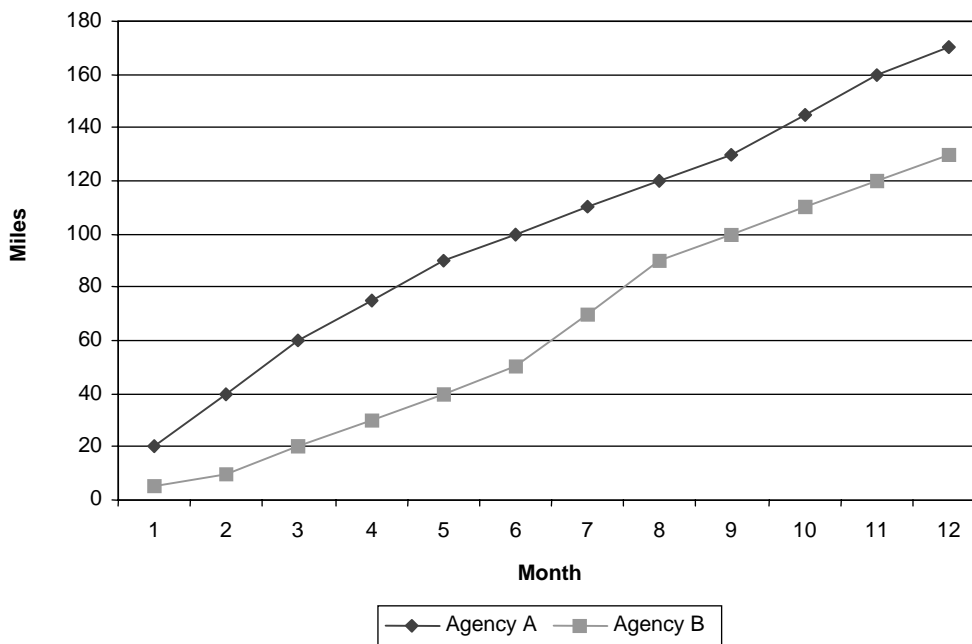


Figure C-2. Monthly agency milestones for rumble strip installation.

4-month intervals over a period of 2 years are scheduled. Responsibilities are assigned for the planning and implementation of these surveys.

### Emergency Services Response Times

The 2-year detailed action plan to reduce emergency response times outlines new methods for positioning emergency medical services resources and possible timelines for budget increases and acquiring additional resources. In addition, a public education and information campaign is developed, including radio and television ads. The intent of the campaign is to inform the public (potential first respondents to a crash scene) how to describe the crash location and the injuries when reporting the occurrence via the telephone.

### General

At regular intervals during the development of the detailed action plans, the Operations Manager meets with the Task Team leaders and implementing agency representatives. This continuous coordination and communication among the Operations Manager, Task Team leaders, and implementing agencies allows for early detection and elimination of conflicts. For example, some of the activities and strategies planned for the “crashes-with-trees” emphasis area require coordination between the activities and strategies planned for the “run-off-road crashes” emphasis area. There are ways to combine activities to reduce costs.

The completed detailed action plans of all strategies for all emphasis areas are presented to the SPL by the Operations Manager and incorporated into the ISHSPlan.

### STEP 6: IMPLEMENT ISHSPLAN AND EVALUATE PERFORMANCE

During implementation, the Operations Manager compiles a performance report on a monthly basis that provides details about the implementation of each strategy, the data collection efforts, and the interim results of process and impact performance measurement activities. Examples of such performance measurements are shown below.

### Speed Enforcement—Impact Evaluation

Speed surveys are conducted at 10 representative locations, on a monthly basis, to assess the impact performance of the law enforcement strategy (Table C-6).

At the end of the first year of implementation, it is concluded that the speed enforcement and public information and education programs were successful in reducing by 17% the percentage of drivers speeding. These programs did not reach the target reduction of 27% (i.e., a reduction from 30% to 22%), as expected when implementing Level 2 (additional 1,500 hours). This expected reduction from 30% to 22% is shown in Table C-4. After consultation with law enforcement implementing agencies, the second year’s level of implementation and expected target rate are reviewed to



TABLE C-6 Performance assessment from speed surveys

Location	% drivers speeding		Change	Statistically Significant?
	Before	After		
Site 1	28	23	18%	Yes
Site 2	32	30	6%	No
Site. . .				
TOTAL	30	25	17%	Yes

reflect the realistic results reached in the first year of implementation.

Shoulder Rumble Strips—Process Evaluation

The performance report identifies areas of, and reasons for, performance that is lower than planned and includes recommendations on how the lower performance can be corrected. Figure C-3 and Table C-7 show the results of the performance measurements of shoulder rumble strip installations. This report is reviewed and assessed by the Task Team, the Operations Manager, and the SPL, and corrective actions are initiated.

For the purpose of this example, the Operations Manager determines that the implementation of rumble strips is falling behind the original schedule and that the costs are greater than

anticipated for Agency A. The Operations Manager communicates with the implementing agency to assess the condition, and together they seek potential assistance (or modification to the planned schedule).

Run-off-Road Crash Treatment Sites—Impact Evaluation (Before-After Safety Evaluation)

One year after the commencement of the program, the Operations Manager and the RAE group conduct a formal evaluation to determine if the engineering strategies aimed at reducing run-off-road crashes are effective at the 12 treatment locations (Table 3-1) where no other nonengineering strategies were deployed. The evaluation used an empirical Bayes before-after method, which is briefly described in Appendix D2.

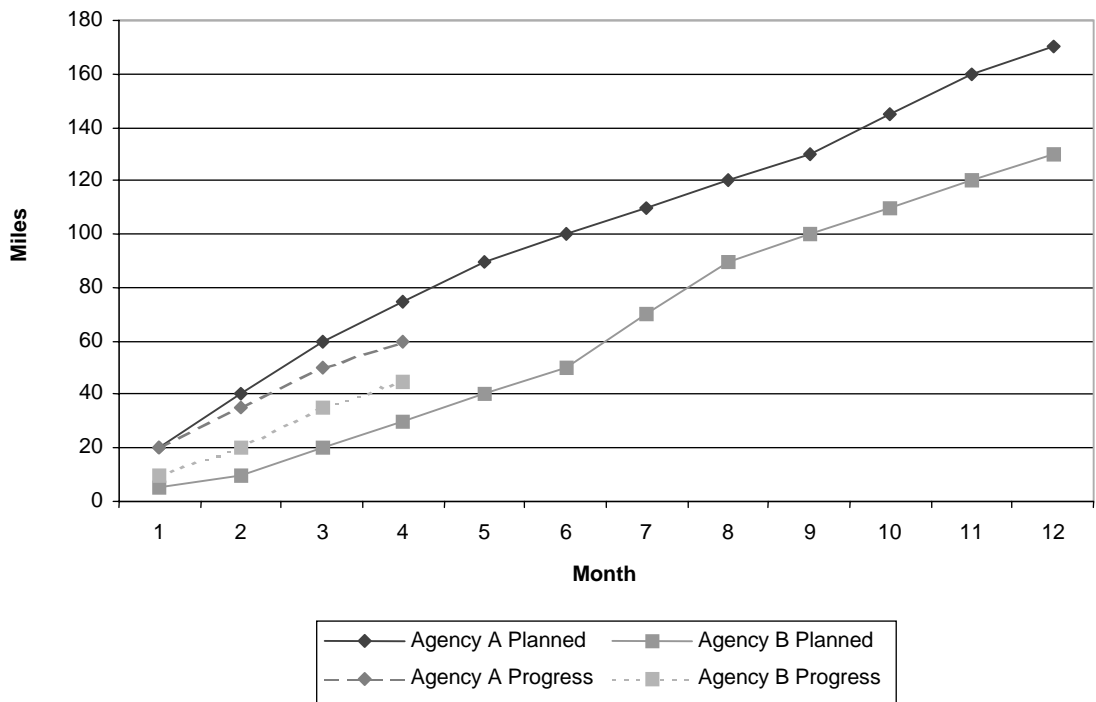


Figure C-3. Sample of example progress report (Month 1–4) for shoulder rumble strips.

**TABLE C-7 Sample of example monthly progress report**

Agency A	Planned	Actual
Miles of rumble strips implemented during month	15	10
Total miles of rumble strips implemented	75	60
Expenses for month	\$48,000	\$32,000
Total project expenses	\$240,000	\$192,000
% project completed	44%	35%
% of funds expended	44%	37%

Generally, the evaluation method depends on the emphasis area and the type and amount of available data. For example, the empirical Bayes approach may not be a suitable evaluation method to evaluate the effectiveness of strategies aimed at reducing impaired driving. For the evaluation of run-off-road countermeasures, the empirical Bayes approach is suitable because the evaluation is aimed at specific locations for which there are SPFs. Furthermore, 12 treatment locations were selected because of poor safety performance during the before period, and the empirical Bayes approach is required to eliminate the potential regression-to-mean bias.

For each treated section of highway, the number of crashes for the 3-year before period is known. The SPFs, developed by the RAE group, are used to estimate the long-term mean,  $m$ , for each of the 12 treatment locations. Crash totals for the year after implementation are available for each of the 12 treatment locations. There are five steps in the empirical Bayes evaluation. The first four steps pertain to calculations for a single location and are illustrated for Location 1. This is followed by a fifth step, in which a composite treatment effect over all locations is then estimated:

- **Step 1: Estimate the comparison ratio ( $R$ ) and its variance  $\text{VAR}(R)$ .** The comparison ratio is applied to the long-term mean based on before-period data to estimate the expected number of crashes in the after-period in the absence of treatment. This comparison ratio accounts for changes in traffic volumes in the before and after periods, as well as for time trends in factors such as weather and crash reporting practices.

A comparison ratio is determined from SPFs following the methodology given in Hauer (2). The Operations Manager is confident that the changes in traffic volumes and other potential confounding factors between the before and after periods can be accounted for by using such a method. The SPFs are calibrated based on all run-off-road crashes on all two-lane highway locations where there are no road improvements and where drivers are not exposed to the public information and education or law enforcement programs.

Thus for Location 1, the comparison ratio ( $R$ ) was calculated as equal to 1.07 with a variance of 0.00456 (Table C-8).

- **Step 2: Determine the expected number of crashes at the treatment locations during the before period ( $m_b$ ) and its variance ( $\text{VAR}[m_b]$ ).** The expected number of crashes during the before period at the treatment locations has already been determined prior to the treatment and is shown in Table C-2. For Location 1, this value is 35.0 crashes per year. The estimated variance for Location 1 is 25.6.
- **Step 3: Determine the expected number of crashes that would have occurred, in the after period, had no interventions taken place ( $B$ ) and its variance ( $\text{VAR}[B]$ ).** The expected number of crashes that would have occurred during the after period, had no interventions taken place, was determined by multiplying the expected number of crashes ( $m_b$ ) in the before period by the comparison ratio ( $R$ ). For Location 1,

$$B = m_b * R = 35.0 * 1.07$$

$$= 37.45 \text{ crashes per year}$$

The variance of  $B$  was determined as follows:

$$\text{VAR}(B) = B^2 * [\text{VAR}(m_b)/m_b^2 + \text{VAR}(R)/R^2]$$

$$= (37.45)^2 * [25.6/(35.0)^2 + 0.00456/(1.07)^2]$$

$$= 34.90 \text{ crashes per year}$$

- **Step 4: Determine the actual number of crashes during the after period ( $A$ ).** For Location 1, the actual number of crashes, in the after period, was 24 crashes per year.
- **Step 5: Determine the index of effectiveness of the interventions ( $\theta$ ) and its variance ( $\text{VAR}(\theta)$ ).** The values from Steps 1 to 4 for all locations are shown in Table C-8. The index of effectiveness ( $\theta$ ) of all the interventions applied to the study locations was estimated as follows:

– Obtain the sums, over all locations, of the numbers of crashes *expected without treatment* ( $\Sigma B$ ) and the number of crashes *recorded* in the after period ( $\Sigma A$ ). For the data in Table C-8, the number of crashes *expected in after period without treatment* =  $\pi = \Sigma B = 241.90$  with summed variance of 225.65.

TABLE C-8 Calculations for estimating treatment effects

Location	AADT (number of vehicles)	Length (miles)	# Target crashes  ( $n = 3$ )	Estimated long-term mean per year  $m_b (\text{VAR}(m_b))$	Comparison ratio  $R(\text{VAR}(R))$	Expected crashes per year after  $B(\text{VAR}(B))$	Recorded crashes per year after  $A$
1	8800	50	120	35.0 (25.6)	1.07 (0.00456)	37.45 (34.90)	24
2	7200	40	75	21.6 (18.4)	1.04 (0.00402)	22.46 (21.78)	14
3	8200	10	27	6.9 (5.5)	1.10 (0.00509)	7.59 (6.90)	5
4	7200	50	88	26.5 (20.9)	0.94 (0.00356)	24.91 (20.97)	12
5	6200	40	63	18.4 (15.2)	1.05 (0.00420)	19.32 (18.18)	14
6	6100	20	38	10.0 (8.7)	1.05 (0.00420)	10.50 (10.01)	6
7	10300	20	47	14.2 (12.1)	1.08 (0.00475)	15.34 (15.07)	12
8	6000	25	38	11.3 (9.4)	1.07 (0.00456)	12.09 (11.34)	8
9	10200	50	117	35.5 (27.2)	1.06 (0.00460)	37.63 (36.36)	28
10	9500	15	40	11.0 (9.4)	0.97 (0.00380)	10.67 (9.30)	9
11	6900	20	42	11.2 (9.5)	0.97 (0.00380)	10.86 (9.42)	7
12	7000	60	110	31.8 (25.3)	1.04 (0.00402)	33.07 (31.43)	21
<b>Totals</b>		400	805			$\pi = 241.90$ (225.65)	$\lambda = 160$

AADT = annual average daily traffic.

- Actual number of crashes *recorded* in the after period =  $\lambda = \Sigma A = 160$  with a variance of 160 (equal to the count, assuming a Poisson distribution), and estimated index of effectiveness =  $\theta = (\lambda/\pi)[1 + \text{VAR}(\pi)/\pi^2] = 0.664$ .
- The estimated treatment effect =  $100(1 - \theta) = 33.6\%$  fewer crashes in the after period than if no interventions had taken place.
- The variance of the estimate of the index of effectiveness is  $\text{VAR}(\theta) = \theta^2[\text{VAR}(\lambda)/\lambda^2 + \text{VAR}(\pi)/\pi^2]/[1 + \text{VAR}(\pi)/\pi^2]^2 = 0.00442$ .
- The estimated index of effectiveness has a standard deviation =  $\sqrt{0.00442} = 0.0665$ .

The results reveal that the combination of strategies has been successful in reducing run-off-road crashes at the locations where shoulder rumble strips and other road safety engineering countermeasures were undertaken. All the engineering strategies were implemented simultaneously; therefore, it is not possible to determine the effectiveness of each strategy individually. In this regard, the use of process performance measures can provide good information as to whether there is

a realistic expectation that a particular strategy contributed to the overall reduction in crashes.

Crash data can also be analyzed to determine the effectiveness of the law enforcement and public information and education strategies in reducing run-off-road crashes and severities on all other two-lane highways in the jurisdiction.

The treatment effect determined from the safety evaluation can be used to determine the cost-effectiveness (e.g., cost-benefit) of the combination of strategies that have been implemented. This information reveals whether it is worthwhile to proceed with the current combination of strategies or whether attention should be given to terminating or modifying the current approach to address the problem of run-off-road crashes.

## APPENDIX C REFERENCES

1. <http://care.cs.ua.edu>
2. Hauer, E., "Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety." Oxford, United Kingdom, Pergamon Press, Elsevier Science Ltd. (1997).

## **APPENDIX D**

### **TOOLS AND EXAMPLES**

Appendix D consists of three subappendixes covering the different analytical tools that may be used to identify crash concerns, estimate the effectiveness of strategies, select appropriate combinations of potential strategies, and optimize the

allocation of resources. In addition, examples of different methods for identifying crash concerns and of process and impact measures for evaluating strategies are presented.

**Appendix D1: Methodologies for Identifying Crash Concerns and Developing an Effective Combination of Strategies to Support Jurisdictional Goals**

Appendix D1 comprises methodologies for use during the ISMProcess Steps 1 to 4. These methodologies will address emphasis area identification, establishment of realistic jurisdictional goals, optimization of strategies, and level of implementation based on existing resources (financial and personnel) within a given timeframe. The methodologies presented here would normally be used in conjunction with safety strategies, such as the ones described in the AASHTO Implementation Guides.

Problem (emphasis area) identification, in simplest terms, involves categorizing all available data into different subsets—for example, grouping all fatalities by age of the driver. The categorized data are then graphed and visually examined for patterns or “high poles” (values that are much larger than the neighboring values, Figure D-1). Categorized data, as appropriate, are then subcategorized—for example, by type of roadway—until the emphasis area has been identified. Once emphasis areas have been identified, optimization is necessary in order to decide which resources to use and which strategies and countermeasures to apply. Optimization is a series of iterative calculations where different combinations of countermeasures and budget options are tried and tested under the criteria to maximize the potential safety benefit for a given budget. After optimization is complete, jurisdictional safety goals can be set.

The four methodologies outlined in this appendix will be primarily used by the Risk Analysis and Evaluation (RAE) group and can be summarized as in Table D-1.

Before presenting the methodologies, it is necessary to put some perspective on exactly what the methodologies can and cannot accomplish. Optimization techniques and other mathematical models are tools, and like any other tool, they can be used or abused. A hammer will not work to tighten a bolt, nor is a wrench most effective for pounding nails. It is important to put the optimization techniques and other mathematical models into proper perspective and to recognize that, while they have limitations, they also enable one to do things that otherwise would be impossible. In this case, it is not practical or possible for a person to evaluate and compare each of a large number of possible combinations of alternative activities without the help of some tool analogous to the ones that are presented in this document.

However, it is important that mathematical models be considered in light of the alternative to them. If estimates are not made at lower levels, they will ultimately have to be made over entire programs. Experience has shown that in the absence of an analytical technique, decision makers have no rational basis to move forward and thus can be expected to base their decisions on political considerations. In the presence of factual evidence, producing the greatest benefit becomes politically expedient. Sensitivity analyses can be performed easily to determine just how sensitive the results are to minor changes in the data. In this way, the ramifications of social and humanitarian considerations can be established based on the effect that they have on other potential programs. In other words, there is a basis, albeit imperfect, for evaluating decisions. The alternative is essentially to be flying blind.

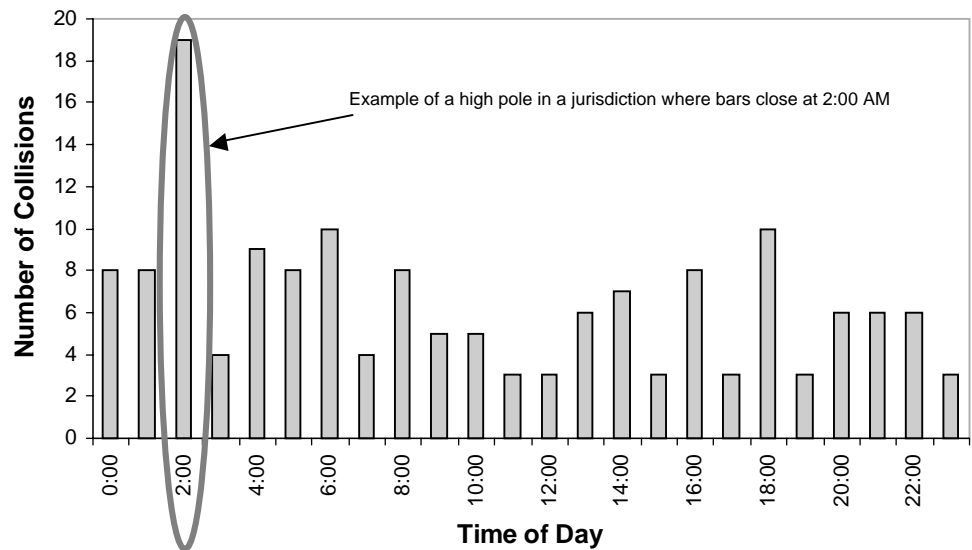


Figure D-1. Example of a high pole.

**TABLE D-1 Four methodologies for identifying crash concerns and developing an effective combination of strategies**

<b>Development of Jurisdictional Goals (ISMProcess Step 1)</b>	Methodology to support the development of jurisdictional goals in terms of reducing injuries and fatalities and relating the goals to the cost and level of implementation of sets of selected strategies.
<b>Problem (Emphasis Area) Identification (ISMProcess Step 2)</b>	Methodology to guide the selection of emphasis areas and the definition of objectives by performing detailed analyses that identify major crash concerns in the state and local agencies using available data on the vehicle, person, and roadway infrastructure.
<b>Strategy Identification (ISMProcess Step 3)</b>	Methodology to guide the selection of strategies within emphasis areas, such as those presented in AASHTO's Strategic Highway Safety Plan (SHSP) or others.
<b>Optimization of Strategies (ISMProcess Step 4)</b>	Methodology to define the most effective combinations of strategies considering available funds, costs, effectiveness, and other limitations. The methodology should also be capable of prioritizing emphasis areas, i.e., if all of the problems/opportunities cannot be addressed.

All analytical techniques can be criticized for being imperfect. Those who criticize the use of any given analytical method are duty-bound to come up with one that is better. Those who use analytical methods are duty-bound to input only the most accurate data available to them and to do everything in their power to see that these data improve in quality in the future.

#### **D1.1 OVERVIEW OF THE TOOLS THAT SUPPORT THE FIRST FOUR STEPS OF THE ISMPROCESS**

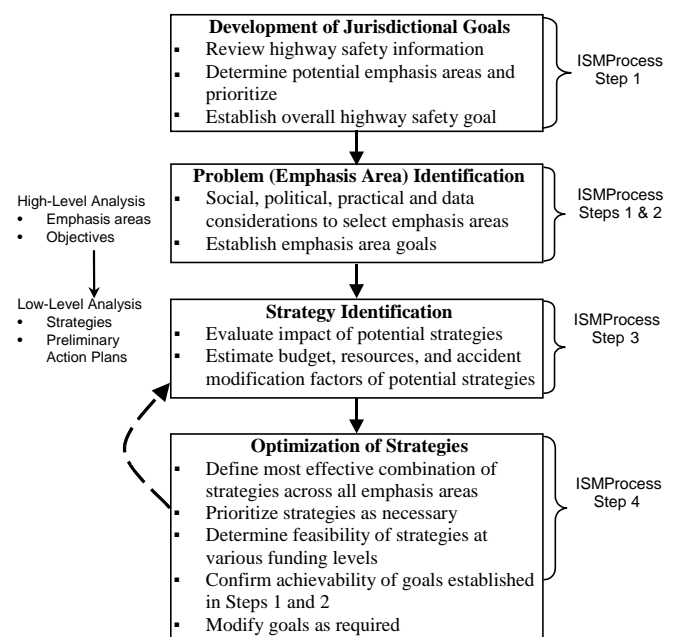
The ISMProcess identifies the steps necessary to achieve a jurisdiction's overall highway safety vision (or goal). This appendix covers the methodology involved in problem identification, the tools for subset comparative analysis, and optimization. The first four steps of the ISMProcess are shown in Figure D-2 showing more details about the process involved in effectively creating strategies to support jurisdictional safety goals.

The ISMProcess begins by applying a methodology to support the development of jurisdictional goals. This includes a review of existing highway safety knowledge in the jurisdiction and a report summarizing the status of safety in the state. Potential emphasis areas are also identified through this review.

The Safety Program Leadership (SPL) requires an initial subset analysis to identify jurisdictional emphasis areas and set an overall highway safety goal (or vision). Potential emphasis areas are identified by the review of existing highway safety knowledge, and the RAE group performs a subset comparative analysis on each emphasis area to the extent possible (i.e., to which data are available). Emphasis area goals will begin to surface, but the purpose of this first iteration of subset comparative analysis is to compare various emphasis areas. Based on this initial subset analysis, the SPL can establish jurisdictional emphasis areas and an overall highway safety goal (ISMProcess Step 1).

Next, Task Teams are formed (ISMProcess Step 2) to further explore each emphasis area, with the help of further subset analysis by the RAE group. A goal must be established for each emphasis area based on social, political, and practical considerations.

Once the emphasis area goals have been firmly established, Task Teams begin to identify potential safety strategies (ISMProcess Step 3). This includes estimates of costs and benefits, budgets, and resources, either quantitatively or by expert opinion. AASHTO Implementation Guides provide useful information in estimating the characteristics of proposed strategies. The costs and benefits of implementing proposed strategies may be estimated at multiple implementation levels and



*Figure D-2. Flowchart depicting the procedure for determining an effective combination of strategies to achieve a jurisdiction's safety vision.*

account for cost savings if combinations of strategies are implemented.

Optimization (ISMProcess Step 4) is conducted to select the most effective combination of strategies for implementation. The optimization process is repeated for a range of possible funding levels. For the selected funding level confirmed by the SPL, strategies are identified for development of detailed action plans (ISMProcess Step 5).

Optimization may identify the need to reduce the scope of the emphasis areas or objectives based on available resources. Task Teams may be required to develop additional preliminary action plans for a number of alternative levels of funding (ISMProcess Step 3).

## **D1.2 METHODOLOGY TO SUPPORT THE DEVELOPMENT OF JURISDICTIONAL GOALS (ISMPROCESS STEP 1)**

A methodology is required to support the development of jurisdictional goals in terms of reducing injuries and fatalities and for relating those goals to the cost and level of implementation of selected sets of strategies.

When the ISMSystem is first implemented, the development of jurisdictional goals must be based on a limited amount of information because the decision makers will only have a limited amount of information available to them. The initial jurisdictional emphasis areas and goals are identified through an analysis of existing data sources (such as the AASHTO emphasis areas and/or other areas defined by the SPL).

After the ISMProcess has progressed through all six steps, there will be sufficient information to provide a better basis for the identification of emphasis areas. The methodology outlined here is intended to occur after the completion of ISM-Process Step 6, when the ISMProcess is about to begin another iteration.

Four fundamental, interrelated questions provide the guidance to support the development of jurisdictional goals:

- **Question 1:** For a given strategy, how is the economic benefit of injury and fatality reduction related to the cost and level of implementation?
- **Question 2:** How is the potential economic benefit of injury and fatality reduction in a jurisdiction related to the overall level of investment in a set of selected strategies?
- **Question 3:** What is an optimal level of investment to obtain a jurisdictional goal?
- **Question 4:** What are the impacts of varying investment programming options?

Question 1 requires data on safety benefits of a strategy for various levels of implementation. This information serves as input to the optimization methodology (ISMProcess Step 4). Question 2 applies a similar methodology to the highest ISMSystem level to assess the potential economic benefit of injury and fatality reduction in a jurisdiction for various levels of investment. Part of ISMProcess Step 4 will include a selected set of strategies for each investment level and the optimum implementation level for each strategy. All of this information is then provided to decision makers (the SPL) who will address Question 3. Considerations for Question 3 will include available budgets, political and pragmatic considerations for specific strategy sets, and an assessment of marginal return for increasing investment levels. Then, the effects of various investment-programming options can be considered in addressing Question 4 if resource constraints require that strategies not be implemented all at once. Each investment programming option requires an adjustment of the jurisdictional goal.

### **D1.2.1 Question 1: For a given strategy, how is the economic benefit of injury and fatality reduction related to the cost and level of implementation?**

Table D-2 illustrates a possible output after addressing Question 1. In this example, there are three emphasis areas.

**TABLE D-2 The desired output for Question 1**

Emphasis Area	Strategy	Implementation Level 1		Implementation Level 2		Implementation Level 3	
		Cost	Benefit	Cost	Benefit	Cost	Benefit
E1	S1.1	C1.1.1	B1.1.1	C1.1.2	B1.1.2		
	S1.2	C1.2.1	B1.2.1	C1.2.2	B1.2.2	C1.2.3	B1.2.3
	S1.3	C1.3.1	B1.3.1	C1.3.2	B1.3.2	C1.3.3	B1.3.3
E2	S2.1	C2.1.1	B2.1.1	C2.1.2	B2.1.2	C2.1.3	B2.1.3
	S2.2	C2.2.1	B2.2.1	C2.2.2	B2.2.2		
	S2.3	C2.3.1	B2.3.1	C2.3.2	B2.3.2	C2.3.3	B2.3.3
	S2.4	C2.4.1	B2.4.1	C2.4.2	B2.4.2	C2.4.3	B2.4.3
E3	S3.1	C3.1.1	B3.1.1				
	S3.2	C3.2.1	B3.2.1	C3.2.2	B3.2.2	C3.2.3	B3.2.3

Emphasis Area 1 (E1) has three strategies, E2 has four strategies, and E3 has two strategies. Strategies 1.1 and 2.2 have two possible levels of implementation; Strategy 3.1 has one implementation level, while all other strategies have three implementation levels.

Adding or deleting rows or columns to Table D-2 allows for more or fewer strategies or levels of implementation. For each implementation level in Table D-2, there is a corresponding estimate of cost and benefit (reduction of crashes for each severity level) over the effective life of the strategy.

Throughout the process, the safety benefit should be measured in terms of the quantifiable goal for the emphasis area. The costs in Table D-2 are calculated from implementation, maintenance, and all other costs associated with the implementation level of each strategy. The benefits are determined from the value of crashes assigned to the target crash for the severity level expected to be reduced. These monetary values should be converted to equal time values. Methodologies for converting the time value of money can be found in most economic textbooks.

How does one get the input for such a table as Table D-2? This table is the result of estimates of the following:

- Number of crashes related to a given strategy;
- Expected crash reductions (safety benefits) that are (a) applied to the number of crashes relating to a given strategy and (b) resulting from various levels of implementation of the strategy;
- Number of years that crash reductions will remain in effect;
- Implementation, maintenance, and all other costs associated with the strategy at various implementation levels;
- Unit cost assigned to a crash; and
- Expected life of a strategy.

Safety benefits are accrued annually, while costs can be incurred initially, in the future, or annually. Thus, costs and benefits need to be normalized to a common base, either present values or annual values, by applying a discount rate and factors based on equations found in standard economic textbooks.

To determine the number of crashes related to a given strategy, manipulation of existing databases is required as outlined in the “Problem (Emphasis Area) Identification Methodology: Identify” subsets of data (ISMProcess Step 2).

Estimates of crash reductions associated with various levels of implementation of a strategy will ideally be made on the basis of evaluations conducted in ISMProcess Step 3 (Develop Objectives, Strategies, and Preliminary Action Plans to Address the Emphasis Areas). At the present time, there is little or no solid information that enables estimates of safety benefits (crash reductions over time) to be derived for various levels of implementation. An exception, which illustrates the complexity of safety benefit estimation, is an evaluation

of road safety advertising in Australia, in which safety benefits are related to the amount of advertising dollars expended each month (*I*).

In the absence of safety benefit information from formal evaluations, the safety cost and benefit estimates need to be determined using “Tools to Estimate the Effectiveness of Promising or Innovative Strategies that Have Insufficient Information” (Appendix D2).

The cost to implement the strategy at various implementation levels must be determined during ISMProcess Step 3, as the Task Teams develop preliminary action plans for implementation of all strategies.

#### **D1.2.2 Question 2: How is the potential economic benefit of injury and fatality reduction in a jurisdiction related to the overall level of investment in a set of selected strategies?**

Table D-3 illustrates the output after addressing Question 2 for Investment Level A. This output is the same as Table D-2 except that now the *shaded* cells identify the optimal strategy set for this investment level. In this case, for the optimum allocation of the investment at Level A, one strategy is selected from Emphasis Area 1 (S1.1 at Implementation Level 2), two strategies from Emphasis Area 2 (S2.1 and S2.3 at Implementation Levels 1 and 3, respectively), and no strategies from Emphasis Area 3.

The total costs and benefits of the set of selected strategies are estimated, as outlined in Question 1. Similar information could be obtained for a variety of possible jurisdictional investment levels.

How does one attain the optimal allocation of strategies for a specified level of investment and given the outputs from Question 1? A number of optimization approaches and software packages are available (reviewed in the “Optimization of Strategies Methodology” section at the end of this appendix) and some jurisdictions may already have access to such software. Optimization is further discussed in the “Optimization of Strategies Methodology” section below (ISMProcess Step 4).

#### **D1.2.3 Question 3: What is an optimal level of investment to obtain a jurisdictional goal?**

The output from Question 2 will be a set of strategies identified as optimal for a variety of investment levels. Since software is used to produce this output, the number of investment levels analyzed is not limited by resource constraints.

Additional guidance is necessary in order to select an optimal level of investment to obtain a jurisdictional goal. Once the investment level is selected, then the jurisdictional goal can be defined on the basis of the crash reductions indicated



**TABLE D-3 The optimal strategy set for Investment Level A**

Emphasis Area	Strategy	Implementation Level 1		Implementation Level 2		Implementation Level 3	
		Cost	Benefit	Cost	Benefit	Cost	Benefit
E1	S1.1	C1.1.1	B1.1.1	C1.1.2	B1.1.2		
	S1.2	C1.2.1	B1.2.1	C1.2.2	B1.2.2	C1.2.3	B1.2.3
	S1.3	C1.3.1	B1.3.1	C1.3.2	B1.3.2	C1.3.3	B1.3.3
E2	S2.1	C2.1.1	B2.1.1	C2.1.2	B2.1.2	C2.1.3	B2.1.3
	S2.2	C2.2.1	B2.2.1	C2.2.2	B2.2.2		
	S2.3	C2.3.1	B2.3.1	C2.3.2	B2.3.2	C2.3.3	B2.3.3
	S2.4	C2.4.1	B2.4.1	C2.4.2	B2.4.2	C2.4.3	B2.4.3
E3	S3.1	C3.1.1	B3.1.1				
	S3.2	C3.2.1	B3.2.1	C3.2.2	B3.2.2	C3.2.3	B3.2.3

in the table for that investment level. Special consideration is given to three factors:

- Available budgets,
- Political and practical considerations for specific strategy sets, and
- Assessment of marginal return for increasing investment levels.

The first two considerations require judgment based on factors in individual local situations. Therefore, guidance is not provided here, as it would be difficult to apply to all jurisdictions.

For the third consideration, assessing the marginal return for increasing investment levels will identify the point of diminishing marginal returns. If this criterion is desired for determining the optimum investment level, then one identifies the investment level with the largest net present value (where net present value is the present value of benefits minus the present value of costs). With the identification of the point of diminishing marginal returns, the cost of going to the next highest investment level is larger than the benefits derived.

Table D-4 depicts the information available so far. Note that at Investment Level A, the optimal set of strategies has been defined by answering Question 2. Different investment levels

will have different optimal sets of strategies, but the sets of strategies in Table D-4 are based on those listed in Table D-2.

The crash reductions per year for the selected strategy set would be estimated by adding the crash reductions estimated for the implementation levels of each strategy that is part of the optimal set. These values have been estimated in response to Question 1. Similarly, the costs and benefits of the optimal strategy set would be determined by combining the costs and benefits of the individual strategy implementation levels.

These marginal benefit-to-cost ratios are indicated in Table D-4. The last column in Table D-4 is the net present value of each strategy set, calculated from the present value of benefits minus the present value of costs.

Alternatively, one can recognize uncertainty in all of the inputs and can build in a factor of safety by looking at the marginal benefits and costs for each investment level and by basing a decision on the criterion that the marginal benefit-to-cost ratio must exceed, for example, a value of 2.0. For this selection criterion, one would simply go down the list and select the investment level that is indicated by a value of 2.0 (or the one larger than 2.0 that is closest to that value).

One can also combine these considerations. For example, the top six strategy sets could be identified on the basis of the net present value, and then the one with the most independent set of strategies could be selected on the basis of a “prag-

**TABLE D-4 Example optimal strategy sets for various investment levels**

Investment level \$ (in ascending order)	Optimal set of strategies (Emphasis area. Strategy. Implementation level)	Fatal and injury crash reductions per year for strategy set					Cost and benefit for strategy set (in present value)			
		Year				Total	Cost	Benefit	Marginal B/C Ratio	Net Present Value
		1	2	3	4					
A	1.1.2; 2.1.1; 2.3.3						C <sub>A</sub>	B <sub>A</sub>		NPV <sub>A</sub>
B	1.1.2; 2.1.2; 2.4.1						C <sub>B</sub>	B <sub>B</sub>	M <sub>AB</sub>	NPV <sub>B</sub>
C	1.2.2; 2.1.2; 3.1.1						C <sub>C</sub>	B <sub>C</sub>	M <sub>BC</sub>	NPV <sub>C</sub>

B/C = benefit/cost.

NPV = net present value.

matic” consideration. This would recognize the possibility that benefits might be lower than anticipated because a strategy set is composed of many strategies that tend to target the same crash type. If the selected set nevertheless contains such strategies, as is inevitable, one could do a sensitivity analysis on the selected set by doing a worst case benefit scenario and comparing this to the worst case scenarios of other highly ranked strategy sets. On this basis, it may be necessary to revise the jurisdictional goal that was initially obtained on the basis of the assumption that strategies in the selected set affect crashes independently.

Another pragmatic consideration might be a subjective assessment of the synergy among strategies. In this, two or more strategies in a set may produce a combined benefit that is larger than the sum of their individual effects. If the selected set contains such strategies, then the jurisdictional goal is subjectively revised to reflect the larger potential benefits resulting from this synergy.

#### D1.2.4 Question 4: What are the impacts of different investment programming options?

The methodology to this point assumes that all the strategies in the optimal set will be implemented at time zero. In reality, resource constraints dictate that strategy implementation be phased over time. Thus, the effects of various investment programming options need to be considered, since each option requires an adjustment of the jurisdictional goal. Specifically, the benefits to be derived over the next, say, 10 years will be reduced if some strategies are implemented several years into the future instead of at time zero.

The revision of the jurisdictional goal for a specific programming option is fairly straightforward, in that the reductions in injury and fatal crashes in the original goal are adjusted (Table D-4) to reflect the fact that the crash reductions for strategies implemented later in the program accrue over fewer years than originally expected.

#### D1.3 PROBLEM (EMPHASIS AREA) IDENTIFICATION METHODOLOGY (ISMPROCESS STEP 2)

The purpose of this section is to present a systematic methodology for performing problem identification using available crash and demographic data to identify emphasis areas in a state or local agency. This methodology supports the formulation of objectives, strategies, and activities for addressing the emphasis areas.

There are a great number of combinations of reports that can be generated from crash and demographic databases, so a “roadmap” is required to navigate through the data to arrive at good information to support the decision-making process. The objectives are to

- Foster the use of crash and demographic records at all levels (from choosing emphasis areas to developing strategies) and
- Promote the generation and use of this valuable information by all those who develop strategies and action plans.

This section provides a more detailed documentation of the subset comparison tabular outputs illustrated in Figures D-3 through D-9. Generally speaking, the objective of these tables is to redirect attention toward more important factors by arranging the output of the codes within each variable in a “worst-first” ordering. The following definitions give the meaning of each column of the tabular output, from left to right. The current subset is that subset of the data defined as being of interest (e.g., a particular county). The data structures used to specify subsets of the total database are called filters. Some definitions assume that the current subset is compared with its complement (e.g., a particular county against the rest of the state). These definitions are as follows:

- **Code (unlabeled):** An integer indicating the internal code of the item (e.g., if the variable is day-of-the-week, item Sunday would have code 1). This is for reference purposes.
- **Description (unlabeled):** The description for the item; usually what was actually marked on the crash report.
- **Subset Frequency (SubFreq):** The frequency with which the corresponding item occurs in the crash subset.
- **Subset Percentage (Sub%):** The percentage that the *Subset Frequency* is of the total records in the crash subset.
- **Other Frequency (OtrFreq):** The frequency of cases that fall into the item classification, but that are *not in the crash subset*. *Other* here might be the complement of the subset or any other subset that is chosen as appropriate for comparison, such as a control group.
- **Other Percentage (Otr%):** Analogous to the *Subset Percentage*, but it is the percentage for the *Other Frequency* rather than for the *Subset Frequency*. The comparison for over-representation must be made between the two *Subset* and *Other* percentage values, since a direct comparison of the raw frequencies has very little meaning.
- **Over Representation (OverRep):** A measure of the over-representation of each item. It is calculated by dividing the *Subset Percentage* by the *Other Percentage*. The higher this resulting number is, the greater the potential percentage reduction due to countermeasures applied to this specific item, all other things being equal. An asterisk appearing after this number indicates that the statistical test applied suggests that the difference in the two subsets of data being compared are statistically significant for the corresponding variable and value, as discussed in Table D-5.
- **Maximum Gain (MaxGain):** The maximum expected number of cases that could be reduced by applying countermeasures to this item. It can be viewed as the  
(continues on page D-11)

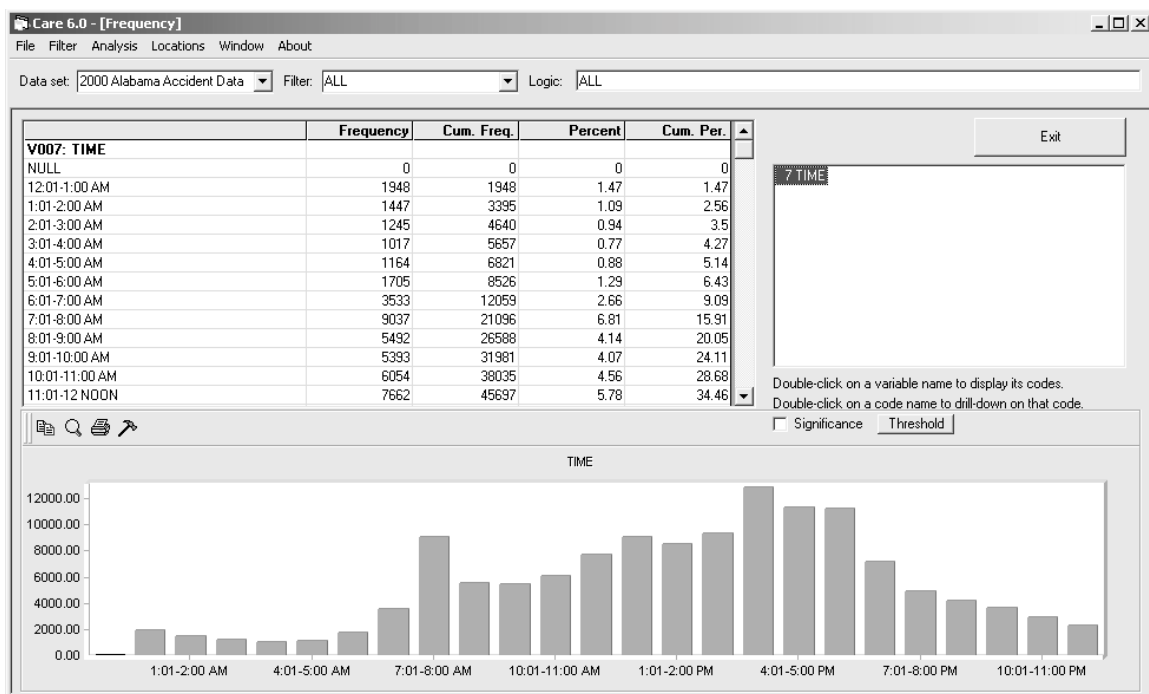


Figure D-3. Example frequency distribution—time of day for all crashes.

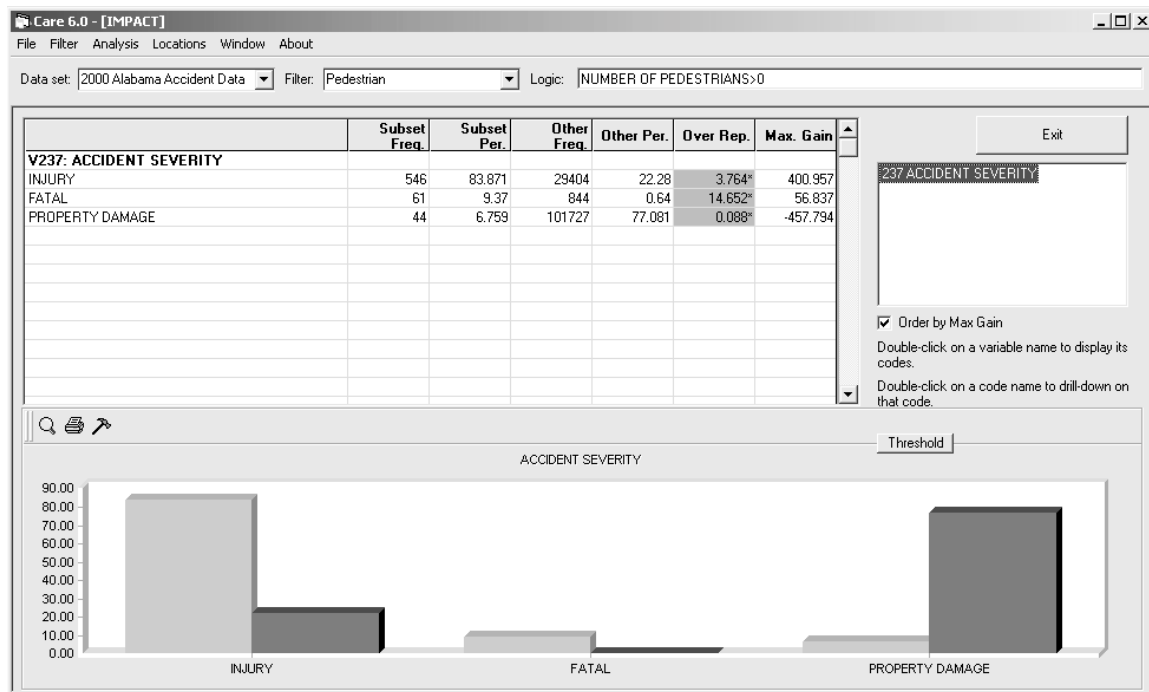


Figure D-4. Example severity comparison (Step C5)—pedestrian (light) versus nonpedestrian crashes (dark).

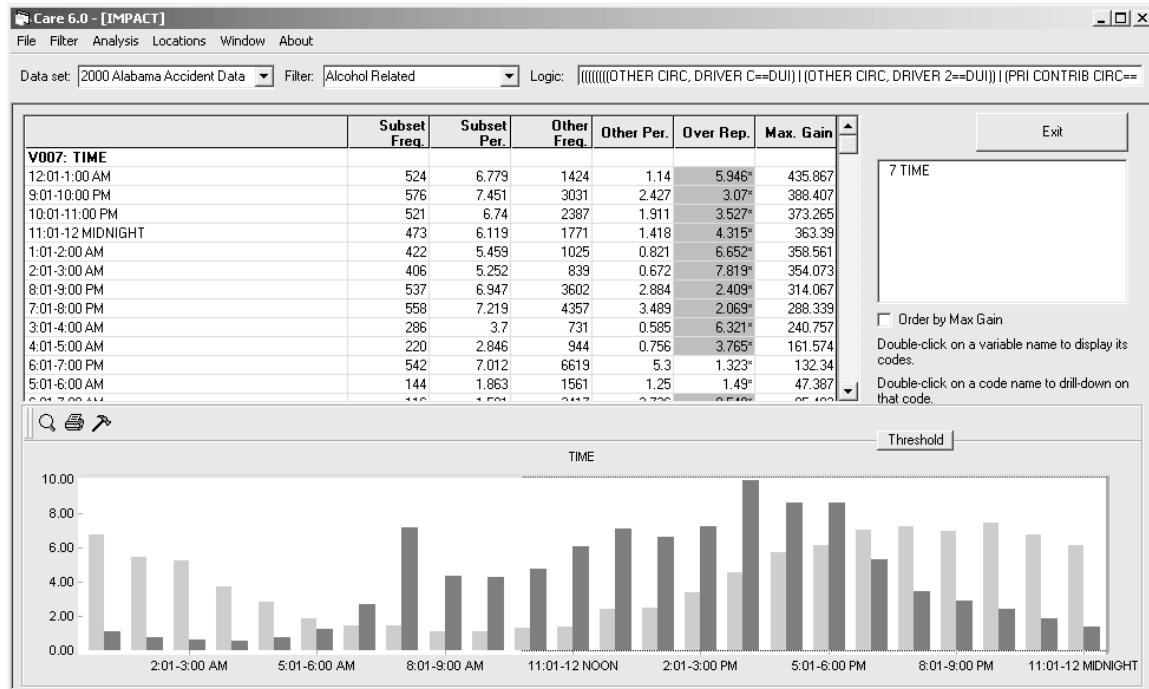


Figure D-5. Example time-of-day comparison—alcohol-related (light) versus non-alcohol-related crashes (dark).

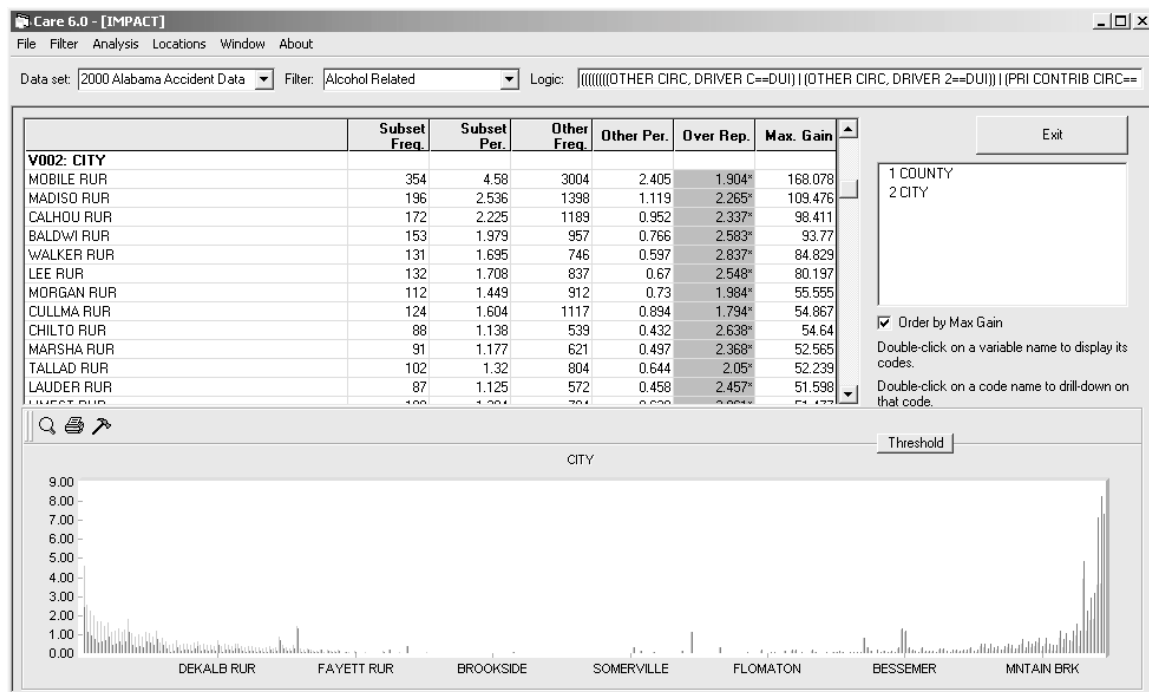


Figure D-6. Example area comparison—alcohol-related versus non-alcohol-related crashes (the Montgomery row of data is off-screen above the scrolling panel).

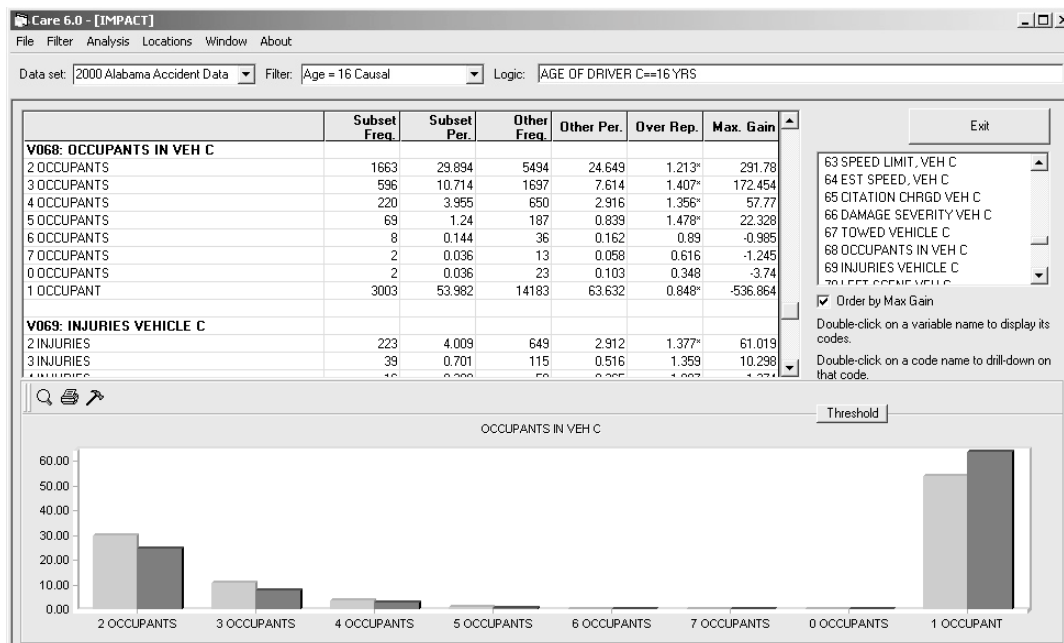


Figure D-7. Example comparison of two noncomplementary subsets—16-year-old causal drivers (light) versus 17- to 20-year-old causal drivers (dark)—occupants in causal driver vehicle.

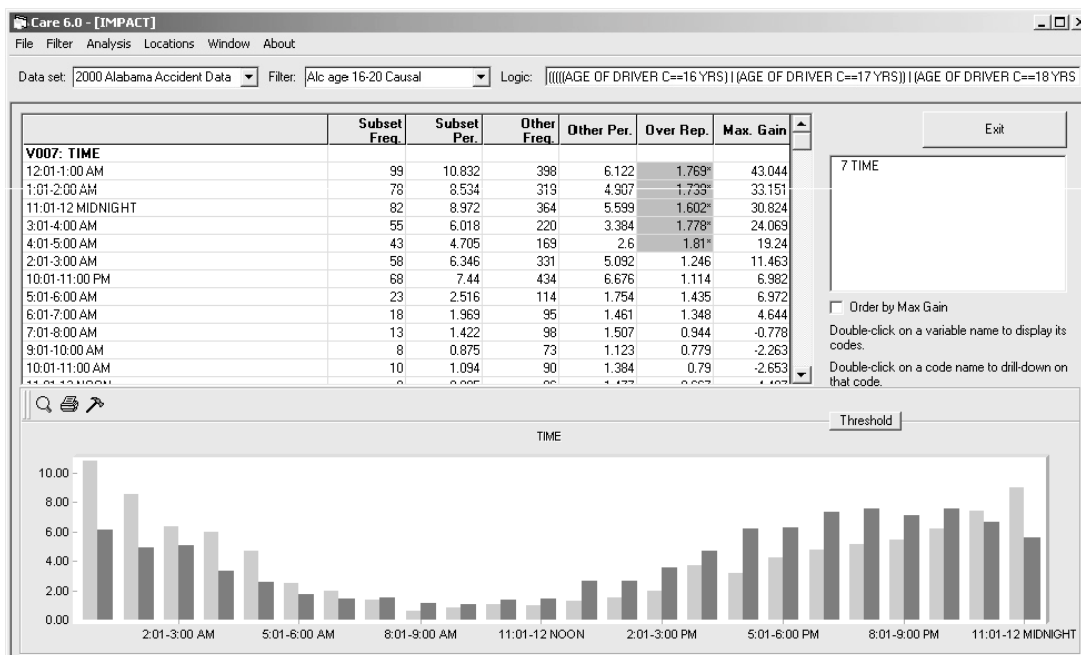


Figure D-8. Second example comparison of two noncomplementary subsets—16- to 20-year-old causal DUI drivers (light) versus 21-and-older causal DUI drivers (dark)—time of crash.

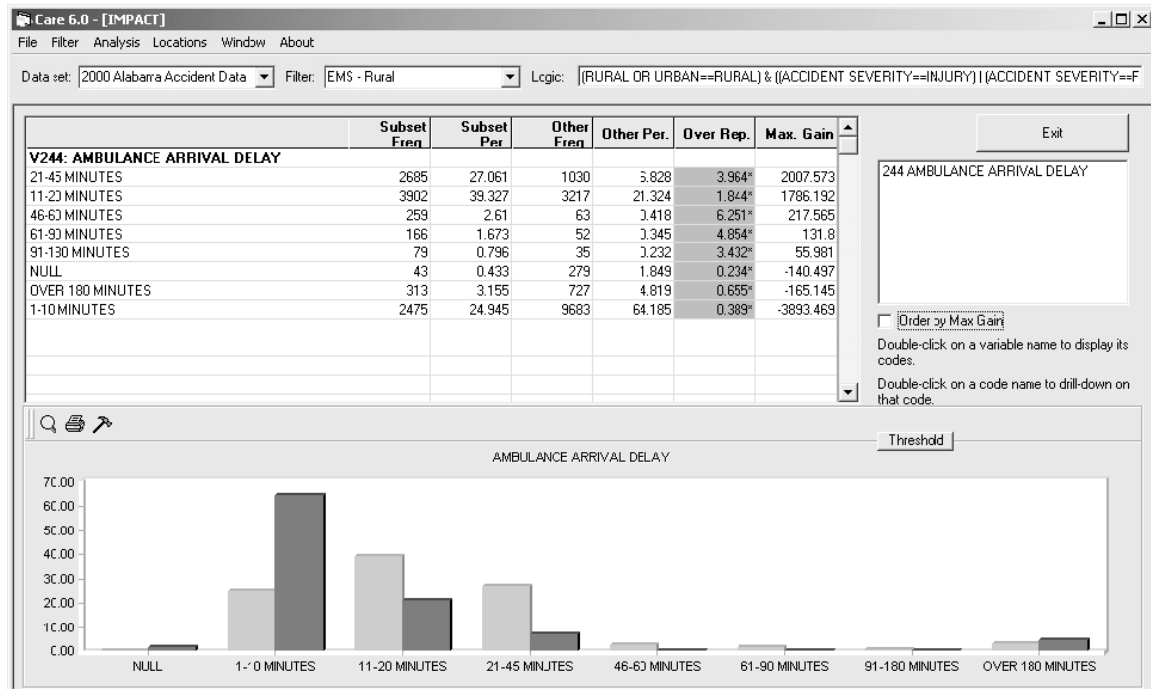


Figure D-9. Severity reduction example—rural (light) versus urban (dark)—ambulance delay time.

number of cases that would be eliminated if the *Subset Percentage* could be reduced down to its *Other Percentage* value (i.e., its expected value). The raw frequency of the subset cannot be reduced to zero; instead, it is assumed that the subset can be reduced only to the expected value as given by the comparison set of crashes (i.e., the control

group). The output for each variable is sorted according to *Maximum Gain*, as opposed to *Over Representation*, although the two are highly correlated. This is because sometimes items with very high over-representations have very few cases, and thus they do not have the highest potential for countermeasure implementation.

TABLE D-5 Statistical test employed in Critical Analysis Reporting Environment (CARE) subset comparison analysis

STATISTICAL TEST EMPLOYED	
<p>If the <i>Over Rep</i> factor has an asterisk assigned to it, this indicates that a statistically significant difference exists in that variable value between the test and the control subset. The test subset is the subset of concern, and the control subset is either its complement or a second subset chosen by the user to surface particular information. The statistical test is based on a weighted estimate of standard error for unequal sample sizes. A normal approximation of the binomial assumption was applied in order to perform the statistical test. The following calculations are necessary to apply this test:</p> $z = \frac{\left( \frac{x_1}{n_1} - \frac{x_2}{n_2} \right)}{\sqrt{p * (1 - p) * \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$ <p style="text-align: right;">where <math>p = \frac{x_1 + x_2}{n_1 + n_2}</math></p>	
Frequency Count Variable	Description
$x_1$	test subset for the particular value of the variable under consideration
$n_1$	test subset for all possible values of the variable under consideration
$x_2$	control subset for the particular value of the variable under consideration
$n_2$	control subset for all possible values of the variable under consideration
<p>A proportional difference is considered to be significant when the difference between the proportions being compared is greater than or equal to three standard errors. In this case, the <math>z</math> statistic calculated determines the number of standard errors of difference that exists between the two proportions. Thus, if a <math>z</math> value is greater than +3 or less than -3, an asterisk is assigned and placed after the <i>Over Rep</i> factor in the CARE program that corresponds to the comparison under consideration. A large value for significance is preferred, since the large sample sizes make a relatively small difference in proportions significant and the objective in problem identification is to surface the more dramatic of these differences.</p>	

The problem identification methodology has three major stages:

- **Stage A: Identify subsets of data** that will provide information on potential emphasis areas that have a high potential for crash reduction benefits.
- **Stage B: Formally specify emphasis areas** after assessing the subsets and considering social, political, and practical issues.
- **Stage C: Evaluate emphasis areas and formulate objectives** by comparing the characteristics of potential emphasis areas against those of their complementary subsets on a percentage or rate basis (e.g., time-of-day of alcohol crashes would be compared against time-of-day crashes that were not alcohol-involved). This step will also resolve emphasis areas and initiate the analysis of emphasis areas into objectives and ultimately into strategies and action plans.

The problem identification methodology is illustrated here using CARE (Critical Analysis Reporting Environment) (2), a software system for providing decision makers within the traffic and aviation safety communities with direct access to crash and incident information. For implementing the methodology in the AASHTO Strategic Highway Safety Plan, either CARE or a similarly designed tool is recommended, since the user does not need to have significant computer experience. The various options of CARE are incorporated into windows that provide a thorough guide to all desired output. Examples using the CARE tools are provided throughout this appendix.

#### D1.3.1 Stage A: Identify Subsets of Data

Identifying subsets means figuring out which subsets of the entire available dataset should be studied in depth. Choosing the right subset to analyze is important because all future decisions will be made based on the results from the selected data. Identifying subsets consists of the following steps:

- **Step A1:** Run frequency distributions over all crashes for every variable in the crash records database; review these variables to get an overall feel for the data and any reporting deficiencies. Look for the “high poles,” those significantly higher-frequency cells, on the bar chart outputs of the frequency distributions. Take note of all that might be of concern in the further analyses (e.g., variables and values where notable high-frequency characteristics occur), since this will be useful in creating the filters needed to define subsets in Step A3. As examples, screen shots from CARE are shown in Figure D-3 through Figure D-9. These screen shots depict Alabama crash data from the year 2000 as an example of a global frequency distribution output for time of day. Notice that the distribution of crashes in Figure D-3 follows the

typical rush hour volume pattern. Examples of other typical frequency distributions that might be generated during the identification procedure are day of the week, month, weather conditions, or any other variable that is captured on the crash police report form, including all aspects of the causal driver.

- **Step A2:** Rerun the frequency distributions, but restrict the analysis to just injury and fatal crashes. This introduces crash severity into the analysis. Fatalities are usually too rare to establish patterns, but injury (including fatal) crashes are an excellent proxy for fatal crashes. Once again, “high poles,” especially those that are different from those found in Step A1, are identified. All findings should be documented.
- **Step A3:** Synthesize the results of Steps A1 and A2 and produce a comprehensive list of the types of crashes that will be subjected to further analyses (e.g., a definition of the subsets of the crash records to be considered). This should include the definitions of the variables within the database that define these subsets. For example, if alcohol-related crashes could be disaggregated by several different variables within the database (e.g., time of day, location, age, and group), each one of the different variables and their respective values must be noted to produce the proper filter to create the alcohol-crash subset.
- **Step A4:** Recognize that the subsets selected are from a pool of crash subsets that will be aggregated to form the agency’s emphasis areas. If additional potential emphasis areas should be considered for any reason, they should be added to the list and their crash subsets defined for further analysis as per Steps A1 and A2.

At this point, the information value attained is little more than raw data. Some “denominator” is essential to transform these raw frequencies into usable information. This is done first to resolve the set of emphasis areas and then to define strategies and objectives to be considered. Ultimately, both cost and effectiveness must be considered before developing a preliminary action plan. Effectiveness must take into consideration both crash severity and the potential of the proposed activities to reduce frequency, severity, or both.

#### D1.3.2 Stage B: Formally Specify Emphasis Areas

At this point, a number of subsets of the crash data should be identified from Stage A. Emphasis areas will emerge by having the SPL suggest as many potential emphasis areas as possible based on an examination of the crash data subsets and other considerations, as discussed below. Thereafter, emphasis areas are evaluated relative to each other and prioritized.

While emphasis areas may be best identified based on economic cost and human-suffering reductions, the following considerations are also legitimate tradeoffs in developing a set of emphasis areas:

- **Social considerations:** Zero traffic fatalities can only be attained at the cost of a major restriction of transportation. It is clear that society has made a choice to accept certain risks in order to enjoy the freedom to travel, so the need for increased traffic flow and reduced congestion are legitimate considerations in establishing emphasis areas. Society makes certain subgroups, such as children, of greater concern, thus justifying emphasis areas such as proper use of child restraints even though the emphasis areas might not be justified from a purely economic (maximum return on investment) point of view.
- **Political considerations:** These considerations are a reflection of social considerations in a free society, since most politicians wish to please their particular constituencies in order to get re-elected. An example of a political consideration is promoting a graduated driver's licensing bill so that some positive legislation in this area will get passed. Few traffic safety advocates are totally satisfied with the legislation that gets passed, but it is recognized that compromises are necessary.
- **Practical considerations:** Decision makers who are using the approach presented in this document might not have the latitude to use their resources for the broadest range of safety programs. They would therefore have to restrict consideration to those emphasis areas that are within their purview. In some jurisdictions, it might be known from the personalities involved that certain types of strategies just will not work at this point in time, and common sense would dictate that the strategies be deferred until a new set of players arrives on the scene.

The previous considerations highlight the fact that the transition from problem identification to the development of emphasis areas is not a rigid scientific process. Considerable human judgment is required, but to the extent possible, it should be based on the overall increase in safety for roadway users. The quantitative techniques employed in the problem identification procedure can only serve to advance this general goal.

### D1.3.3 Stage C: Evaluate Emphasis Areas and Formulate Objectives

One method for evaluating a potential emphasis area and identifying its potential objectives is to compare its crash records subset against the complement of that subset over all of the variables. This comparison will identify those characteristics of the potential emphasis area that vary significantly from the norm. In this case, the complement of the subset acts as a proxy for the typical crash and thus for traffic volume in general. Taking alcohol-related crashes as an example, one should expect the same proportion of alcohol-related crashes to the state total in a given county as the proportion of non-alcohol-related crashes to the state total. The non-alcohol-

related crashes act as a proxy for the traffic volume in that county, since traffic volume is the primary factor in determining crash frequency in any area. When the proportion of alcohol-related crashes is significantly higher than the proportion of non-alcohol-related crashes, it is intuitive to suspect that there is a problem with alcohol-related crashes in that county.

While this analytical approach is not definitive, it cannot be ignored as an easily applied first indicator. The fact that comparing a subset to the complement of the subset can be applied easily to all variables and all subsets is also compelling because most demographics apply only to very few specific variables. For example, populations apply only to cities or counties, and further subdivisions are necessary if they are to apply to gender and age. Average daily traffic (ADT) counts apply to vehicles in general and might, at times, be subdivided by vehicle types. Therefore, these demographics, while quite useful for producing rich information from the few variables to which they apply, do not add any insight with regard to the vast majority of the variables in the database.

The following is the recommended procedure for performing initial comparisons of potential emphasis area subsets:

- **Step C1:** Run frequency distributions for all of the variables in the crash database for one of the emphasis area subsets under consideration (e.g., alcohol-related crashes).
- **Step C2:** Run the same set of frequency distributions for the complement subset (e.g., crashes that have no evidence of alcohol involvement).
- **Step C3:** Compare the proportions of these two sets of distributions for all of their corresponding variables. This can be done graphically to provide visualization. It can also be quantified by dividing the percentage for each of the variable-value frequencies from the subset by its counterpart from the complement subset to form an over-representation factor. As an example of this step, comparing severity of pedestrian crashes shows that the fatality proportion is more than 14 times that of fatal crashes in the general population of crashes. This comparison is presented in Figure D-4. In this figure, an asterisk next to the value in the *Over Rep* column indicates that a statistically significant difference (at the 1% level of significance) exists in that variable value between the test and the complement (or control) subset (Step C5).
- **Step C4:** Perform a statistical test to determine if the differences between the two subset variable values are significant. The details of this statistical test are given in Table D-5.
- **Step C5:** Calculate the potential gain that would be obtained from crashes if the percentage over-representation (column *Over Rep*) of the emphasis area subset were eliminated (i.e., its percentage value was reduced to the percentage value of the complement subset). This will form a fair and conservative measure of maximum potential effectiveness (maximum expected gain), assuming



that potential strategies are not expected to reduce the proportion of the subset below the proportion of the comparison control group (referred to here as the complement subset).

- **Step C6:** Sort the results within each variable by this potential gain (column *Max Gain*) so that the output is ordered by the highest potential gain. At this point, objectives for each emphasis area could be formulated and one could begin to identify potential strategies for this emphasis area based on these results with a view to getting a better handle on costs and benefits of addressing the emphasis area. For example, if a particular subset of drivers and times appears to be predominant, the objective would be to decrease the predominance of this subset, and a strategy (possibly one of many) could be to conduct selective enforcement. This would identify only the selective enforcement strategy, not the specific activities or detailed action plan. While it is not necessary to identify strategies at this time, obvious potential strategies should begin to emerge and be documented by the Task Team.
- **Step C7:** Repeat the previous six steps for all emphasis area subsets under consideration and identify subsets (e.g., trucks, motorcycles, bicycles, pedestrian, alcohol, geographic areas, young drivers, older drivers, and emergency medical services). Note: The first seven steps are easily automated directly from the crash records.
- **Step C8:** Compare other demographics for proxy comparison groups as they are available. For example, areas might be compared on a crash-per-population basis, roadway locations might be compared on the basis of crash frequency for a given traffic volume, and age groups and vehicle types might be compared on a mileage-driven basis.
- **Step C9:** Provide the information acquired in Steps C1 to C8 to decision makers for the purpose of resolving the subset of emphasis areas that will be considered further should it not be possible to consider all promising emphasis areas (or all objectives in all emphasis areas).

Several states currently have the CARE program (e.g., Michigan, Tennessee, North Carolina, and Iowa).

Figure D-4 presents a comparison of two simple frequency distributions. The first two numeric columns are the frequency counts and percentages by severity of pedestrian crashes. The second two numeric columns represent the same counts and percentages of nonpedestrian crashes. The graph presents a visual comparison of the two subsets (pedestrian and nonpedestrian crashes) by severity as a percentage. A numerical comparison for each of the pair of bars in the bar chart is given in the table under the column labeled *Over Rep* (a measure of over-representation—Step C5). *Over Rep* is calculated by dividing the pedestrian crash subset percentage by the nonpedestrian crash subset percentage for each of the severity levels. In this example, for injury crashes, *Over Rep* is calculated from  $83.871 / 22.28 = 3.764$ . In essence, the *Over Rep* value is

a measure of how much the particular characteristic is either over- or under-represented. The greater the value of *Over Rep* is over 1, the larger the over-representation; the lower the value is under 1, the larger the under-representation. An *Over Rep* of 1 would be neither over- nor under-represented. The results in Figure D-4 also illustrate how objectives can be formulated. In this case, an objective could be to reduce the severity of pedestrian crashes based on the over-representation of the percentage of injury and fatal pedestrian crashes. However, the analysis so far, in this particular example, does not give any intelligence into potential strategies for achieving this objective.

The Problem (Emphasis Area) Identification Methodology (ISMProcess Step 2) calls for the consideration of all variables for all subsets. The previous example is just one variable (severity) for one subset comparison (pedestrian versus nonpedestrian crashes). When actually applying the methodology presented here, one should not limit the analysis to just the variables presented in the examples; rather, all possible variables should be carefully examined. For example, for pedestrian crashes, other variables may include age of pedestrian, location of crash, and time of day of crash.

Figure D-5 illustrates a second example that compares the time-of-day distribution for alcohol-related crashes against non-alcohol-related crashes. In Figure D-5, the non-alcohol-related crashes follow the rush hour pattern observed in Figure D-3, and alcohol-related crashes follow the pattern for those times when drinking and driving might take place (i.e., from 5 p.m. to 3 a.m.). Note that the column labeled *Max Gain* orders the tabular output. *Max Gain* is the number of crashes that would be reduced if some strategy could be applied to reduce all of the over-representation (in this example, just for the 1 hour ranked highest by *Max Gain*). By arranging the tabular output in this way, the most critical hours for a strategy such as selective enforcement rise to the top of the list (Step C6). So between 12:01 a.m. and 1:00 a.m., 435.867 alcohol crashes are over-represented, which is calculated by

$$\frac{\text{Subset Freq} * (\text{Subset Per} - \text{Other Per})}{\text{Subset Per}} = 524 * \frac{(6.779 - 1.14018)}{6.779} = 435.867$$

While the results from Figure D-4 that “pedestrian crashes are more severe” and from Figure D-5 that “alcohol-related crashes occur at night” are not startling revelations to the traffic safety community, these examples have been intentionally fashioned so that they illustrate the information generation capabilities of the proportional comparisons of variables. When all of the variables are reviewed in this manner, inevitably some counterintuitive information is generated, and a definitive quantification of this information is forthcoming from the tabular outputs. This review of variables increases the ability of decision makers to formulate emphasis areas, objectives, and strategies and to begin to estimate

the potential benefits that might be derived from the emphasis areas, objectives, and strategies.

As an illustration of the use of proxy comparison groups, as suggested in Step C8, consider Table D-6. It shows the use of readily available population figures in providing an alternative “denominator” in the comparisons. Populations are available for cities and for rural areas (all of which are listed under the *AREA* column). The *RATE* column of Table D-6 is the fraction of crashes (in this case alcohol-related crashes) per population. Notice that the output is sorted by this rate so that it comes out in a “worst-first” arrangement (all other things being equal, which they rarely are). Areas are grouped by population range since it is not appropriate to compare the rate of a densely populated city with that of a very small town (or rural area). Table D-6 is a small part of the total report for a typical state.

The selected demographic should relate to the variable under consideration. For example, population would apply to better understanding the number of crashes in a city or county area. However, population would be useless in the time-of-day or severity comparisons in Figure D-4 and Figure D-5. An analogy for the time of day might be the number of vehicle-miles driven at that particular time of day—a demographic that is generally quite difficult to obtain.

The value in these types of demographic analyses, however, is enormous in validating the easier-generated subset comparisons exemplified in Figure D-3, Figure D-4, and Figure D-5. The analogous subset comparison is given in Figure D-6, a comparison of alcohol-related versus non-alcohol-related crashes for the same areas (cities and rural counties) using the methods given in Steps C1–C7. If non-alcohol-related crashes were a perfect proxy for population, then the two summaries (Table D-3 and Figure D-6) will produce identical results. Given the deficiencies in using pure populations to create a rate, a good case can be made that non-alcohol-related crashes form a superior denominator in that they reflect miles driven in general far more accurately in an area than does its population. However, the fact that the same counties are identified to have the highest rates with each method shown in Table D-3 and Figure D-6 illustrates that non-alcohol-related crashes are a fair proxy for population, and both factors are fairly good proxies for traffic volume.

An additional argument in favor of subset comparisons is that the graphical outputs enable patterns to be seen in a way that raw numbers cannot easily identify. In the area comparisons for alcohol, for example, the graph indicates that the problem is in the moderate-sized areas as opposed to within the big cities for this particular state (other jurisdictions may have different results, so this is not a generalization). Further identification of the particular locations indicates that the problems are in the rural areas adjacent to the largest cities. Again, this information is useful for providing a basis for selecting the emphasis areas to be adopted by the agency. This information can also suggest objectives and preliminary potential strategies.

It is not the goal here to come up with the perfect metrics, since there is no such thing. Instead, the use of rates is encour-

aged in order to shed additional light on the validity of subset comparisons, which form the starting point of information generation due to their ease of generation from readily available crash data.

In summary, the problem (emphasis area) methodology focuses on the problem areas uncovered from raw frequencies (high poles), frequency comparisons between subsets using proportions, and the use of rates that can be created from available demographic data. Emphasis areas have been identified, as well as objectives, for addressing the safety concerns in those areas. There may also be some preliminary information on potential strategies for meeting those objectives through the completion of this methodology.

#### **D1.4 STRATEGY IDENTIFICATION METHODOLOGY (ISMPROCESS STEP 3)**

The purpose of the strategy identification methodology is to formulate potential strategies and corresponding activities by performing further analyses of those characteristics that are found to be significantly or practically over-represented on a percentage or rate basis.

The methodology is not rigid, but the following guidelines should be applied to each of the potential strategies and activities that are being formulated:

1. Ask the questions, “Is this information sufficient for action item development? If not, what further information is needed to act on this finding?”
2. Consider cross tabulations of two variables within the subset of data that pertains to the activities under consideration if one or more of the following types of conditions hold:
  - If the activities are time critical (e.g., all selective enforcement strategies), perform a time-of-day by day-of-the-week analysis. As an example, alcohol-related crashes will likely be over-represented in the early morning and on weekend days. A logical approach is to perform a cross tabulation of time-of-day by day-of-the-week to determine the best times and days for driving under the influence (DUI) selective enforcement. The goal of the procedure at this point is to determine additional details (who, what, where, when, and how) for those crash types identified by the analyses performed to this point.
  - If the over-represented variable is not constant over all crash severities, cross tabulate the variable by severity (e.g., nighttime, rural, and older-driver crashes tend to be more severe).
  - If the activities can be targeted to geographic location, age group, gender, race, or any other demographic factor within the crash records, consider these variables for cross tabulation with other over-represented variables.

**TABLE D-6 Summary of alcohol crashes per population by area (counties that are also visible in Figure D-6 are highlighted in gray)**

<b>AREAS OF POPULATION FROM 20001 TO 50000</b>				
<b>COUNTY NAME</b>	<b>AREA</b>	<b>POPULATION</b>	<b>CRASHES</b>	<b>RATE</b>
MONTGOMERY	MONTGOMERY RURL	21979	95	0.004322
CHILTON	CHILTON RURL	20701	88	0.004251
LEE	LEE RURL	33830	132	0.003902
LEE	AUBURN	30935	113	0.003653
MORGAN	MORGAN RURL	34765	112	0.003222
WALKER	WALKER RURL	41019	131	0.003194
DALLAS	DALLAS RURL	24141	77	0.003190
RUSSELL	RUSSELL RURL	20841	64	0.003071
SAINT CLAIR	ST. CLAIR RURL	28297	84	0.002969
TALLADEGA	TALLADEGA RURL	34901	102	0.002923
ETOWAH	ETOWAH RURL	24721	71	0.002872
BLOUNT	BLOUNT RURL	29634	85	0.002868
JEFFERSON	BESSEMER	31509	89	0.002825
LIMESTONE	LIMESTONE RURL	35612	100	0.002808
ELMORE	ELMORE RURL	31379	88	0.002804
JACKSON	JACKSON RURL	25539	69	0.002702
MARSHALL	MARSHALL RURL	34807	91	0.002614
CULLMAN	CULLMAN RURL	48128	124	0.002576
CALHOUN	ANNISTON	26623	64	0.002404
LAUDERDALE	LAUDERDALE RURL	39265	87	0.002216
ETOWAH	GADSDEN	42523	90	0.002117
LEE	OPELIKA	22122	45	0.002034
MORGAN	DECATUR	48761	99	0.002030
MOBILE	PRICHARD	34311	67	0.001953
DEKALB	DEKALB RURL	28838	55	0.001907
RUSSELL	PHENIX CITY	25312	48	0.001896
JEFFERSON	HOOVER	39788	73	0.001835
DALLAS	SELMA	23755	35	0.001473
DALE	DALE RURL	24622	33	0.001340
LAUDERDALE	FLORENCE	36426	48	0.001318
JEFFERSON	HOMEWOOD	22922	29	0.001265
COFFEE	ENTERPRISE	20123	23	0.001143
<b>AREAS OF POPULATION OVER 50000</b>				
MADISON	MADISON RURL	59770	196	0.003279
MOBILE	MOBILE RURL	115041	354	0.003077
CALHOUN	CALHOUN RURL	59796	172	0.002876
BALDWIN	BALDWIN RURL	58004	153	0.002638
TUSCALOOSA	TUSCALOOSA RURL	54491	141	0.002588
MADISON	HUNTSVILLE	159789	302	0.001890
TUSCALOOSA	TUSCALOOSA	77759	142	0.001826
HOUSTON	DOTHAN	53589	95	0.001773
MOBILE	MOBILE	196278	281	0.001432
SHELBY	SHELBY RURL	57274	63	0.001100
MONTGOMERY	MONTGOMERY	187106	200	0.001069
JEFFERSON	BIRMINGHAM	265968	283	0.001064
JEFFERSON	JEFFERSON RURL	338230	136	0.000402

3. Consider creating subsets of the data for additional comparisons where activities are to be targeted to a particular subgroup of the population. For example, insight into a graduated driver's license strategy can be obtained by comparing 16-year-old causal driver crashes against 17- to 20-year-old causal driver crashes. As another example, insight into youth alcohol enforcement activities can be attained by comparing alcohol-related crashes of 16- to 20-year-old causal drivers against alcohol-related crashes of their 21-year-old and older counterparts. Each of these types of comparisons can show differences between the respective subpopulations.
4. Use the results of each analysis to determine what further information is needed before the best decision can be made, and repeat the analysis with the additional information.
5. Persist and maintain a thread of evidence until the information available has been exhausted. If the information generated indicates a significant factor, create further subsets of the data (e.g., youth-pedestrian crashes), and repeat the entire analysis.
6. Reject any strategies and activities at this point that the data clearly show to be counterproductive (i.e., activities that will consume resources that could be better applied elsewhere). Maintain a list of all potential strategies and corresponding activities that will be subjected to further analysis in the optimization procedure.

Following these guidelines will result in a list of strategies and corresponding activities that will be subjected to further analyses. This list of strategies is not final, however, and there is no reason that it cannot be modified as more information continues to be developed.

Table D-7 presents an example cross tabulation of severity by the rural-urban indicator for alcohol-related crashes (subset). This cross tabulation was performed because both of these variables (severity and rural-urban) had at least one of their values highly over-represented in the alcohol-related versus

non-alcohol-related subset comparison. Rural crashes and the more severe classifications were both over-represented. The cross tabulation of these two variables strictly within the alcohol-related subset shows that rural alcohol-related crashes are over-represented in fatal and injury crashes.

Figure D-7 presents an example of a comparison between two noncomplementary subsets for the purpose of learning more about the differences between 16-year-old drivers and their 17- to 20-year-old counterparts to better design graduated driver's licensing laws. Figure D-7 demonstrates that causal 16-year-old drivers tend to have multiple occupants in their vehicles in greater proportion than their older counterparts do. This is just a single example; all of the variables in the database should be analyzed to obtain as much information as possible. For example, the number of occupants injured also supports the benefit of limiting the number of occupants for young drivers.

Figure D-8 presents a second example of comparisons between two noncomplementary subsets motivated by the development of a specific action item. In this case, the subject under consideration is youth-alcohol enforcement, and the time of day is of concern. A comparison of youth-alcohol-related crash times against alcohol-related crashes of older drivers shows that young drivers are involved in alcohol-related crashes later at night and later in the morning than are older causal drivers.

Some safety strategies may target crash severity reduction as opposed to crash frequency reduction. As an example, Figure D-9 presents the ambulance delay times compared for urban and rural areas.

At this point, it is assumed that each of the emphasis area objectives has had a comparison analysis conducted of the subset defined for it against its complement subset over all of the variables in the database. An example output for a comparison analysis is presented in Table D-8 for AASHTO Emphasis Areas 15 and 16 ("Keeping vehicles on the roadway" and "Minimizing the consequences of leaving the road"), which were combined for this analysis to reflect run-off-road crashes. Table D-8 represents only two of the possible variables (day of the week, rural or urban)—a similar output should be generated for all variables in the crash database. The CARE output was defined previously in the Problem (Emphasis Area) Identification Methodology section in this appendix.

The output of these subset comparisons is used to generate a set of strategies for each of the emphasis areas, using the guidelines discussed at the beginning of this methodology. Task Team members develop strategies in areas for which they have particular expertise. Input is sought from research literature, state traffic safety websites, and other sources of ideas for new and innovative safety programs. Analysis results of the impact of various strategies are distributed to the Task Team members, and the widest range of strategies is considered. Only those strategies that are clearly impractical are eliminated from consideration at this point.

**TABLE D-7 Example cross tabulation—alcohol-related crash severity by rural-urban indicator (CARE by default highlights the cells with the larger values)**

Severity	Rural (% of Rural Crashes)	Urban (% of Urban Crashes)	Total (% of Total Crashes)
Property Damage	2092 49.37%	2196 62.87%	4288 55.47%
Injury	1959 46.24%	1246 35.67%	3205 41.46%
Fatal	186 4.39%	51 1.46%	237 3.07%
TOTAL	4237 100%	3493 100%	7730 100%

**TABLE D-8 Impact analysis of two subsets for AASHTO Emphasis Areas 15 and 16**

<b>CARE System IMPACT Report</b> <b>2000 Alabama Accident Data: run-off-road vs. not run-off-road</b> <b>Generated by CARE on 02/20/2002</b> <b>5.976 11391 10.358 0.577* -993.02</b>							
<b>V008:</b>	<b>Day of Week</b>	<b>SubFreq</b>	<b>Sub%</b>	<b>OtrFreq</b>	<b>Otr%</b>	<b>Over Rep</b>	<b>MaxGain</b>
1	Sunday	3916	17.283	9093	8.269	2.090*	2042.46
7	Saturday	4230	18.669	13619	12.385	1.507*	1423.92
5	Thursday	3216	14.194	17610	16.014	0.886*	-412.4
2	Monday	2841	12.539	16686	15.174	0.826*	-597.01
4	Wednesday	2590	11.431	16195	14.727	0.776*	-746.85
6	Friday	3423	15.107	20720	18.842	0.802*	-846.18
3	Tuesday	2442	10.778	16045	14.591	0.739*	-863.94
<b>V009:</b>	<b>Rural or Urban</b>	<b>SubFreq</b>	<b>Sub%</b>	<b>OtrFreq</b>	<b>Otr%</b>	<b>Over Rep</b>	<b>MaxGain</b>
1	Rural	14835	65.474	24186	21.994	2.977*	9851.67
2	Urban	7823	34.526	85782	78.006	0.443*	-9851.67

\*Significantly over-represented.

In summary, the goal of the strategy identification methodology is to create and develop a set of preliminary action plans. The goal here is not to optimize resource allocation. Each preliminary action plan relates to a specific strategy for a specific emphasis area objective. Once a set of preliminary action plans is formulated for all emphasis areas, the estimation of costs and benefits must be completed. These estimates are essential for optimizing the strategies.

#### **D1.5 ADDITIONAL EXAMPLES OF METHODS FOR IDENTIFYING CRASH CONCERNS AND DEVELOPING STRATEGIES (ISMPROCESS STEPS 1, 2, AND 3)**

The above sections propose a set of methodologies for identifying crash concerns and formulating emphasis areas, objectives, and strategies as part of an integrated strategic highway safety plan. A number of applications that are similar in nature to these methodologies are reviewed in this section in varying levels of detail. These applications cover all crash concern areas, further emphasizing the potential of the methodologies for facilitating an integrated management process.

In the safety literature, a methodology presented for identifying crash concerns is sometimes referred to as an “induced exposure method.” This method is typically used for estimating the relative or absolute risk of driver groups or other elements, such as traffic control and roadway factors. The method uses relative involvement of each group or element in observed crashes compared with the relative proportion of that group or element in the entire population. The induced exposure method recognizes that comparisons between groups of entities (e.g., drivers and location types) require some measure

of the groups’ relative exposure to risk to provide meaningful results. However, in many cases, exposure data are simply not available.

The CARE IMPACT module uses an induced exposure method to make comparisons between subsets. As an example, suppose that a county were comparing itself to the rest of the state. There is no reason to expect a priori that the citizens of a county are any different in their driving habits from the rest of the state or that reporting is any different (although both of these factors might be exposed by this analysis). Thus, it is assumed that the total crash history of the county provides a proxy measure of the amount of driving that goes on in that county. If this assumption holds, then the expected proportions of the various crash characteristics in this county should be the same as it is statewide.

For those characteristics where the county is significantly different from the rest of the state, the conclusion can be drawn that the county is either over- or under-represented in that characteristic. For example, say alcohol-related crashes constitute 15% of all crashes in the rest of the state, and alcohol-related crashes constitute 20% of all crashes in the county. If this difference is statistically significant, then it can be reasonably concluded that this county has a significant over-representation of alcohol-related crashes and one should further investigate the cause. Assuming that crashes involving alcohol are reported uniformly across the state, reporting would not have to be perfect to expose this finding. For example, if only 50% of alcohol-related crashes in the entire state were being reported as such, the finding of over-representation within this example county would still be valid. Of course, if there is some reason to believe that reporting is not consistent across the state, then this should be taken into consideration when drawing conclusions.

Several applications of the induced exposure methodology are reviewed in the following sections. The last one—an Alabama application—is reviewed in considerably more detail than the others, since it is the most similar in principles and application to the methodology presented in the body of this report.

#### **D1.5.1 Application 1: Analysis of Serious Crashes and Potential Countermeasures on North Carolina Highways**

Huang et al. (3) investigated the effects of driver, roadway, environment, and crash factors on fatal and severe injury crashes on all road types in North Carolina. The analysis included examining distributions of the variations, descriptive statistics, and bivariate statistics. Factors associated with a proportion of these crashes that was higher than the average were flagged as higher risk. For example, urban two-lane roads experienced a higher severe crash rate (severe crashes per million vehicle-miles) than did any other road type. Further, the characteristics of severe crashes were explored to see if differences exist between the North Carolina data and a second dataset.

The binomial proportions test was used to determine which factors were associated with a higher-than-average proportion of crashes. First, the proportion of severe crashes for each factor was calculated and the standard error determined as follows:

$$SE = \sqrt{\frac{p * (1 - p)}{n}}$$

where

SE = an estimate of the standard error of the proportion of severe crashes for a given factor,

$p$  = the proportion of severe crashes for a given factor, and

$n$  = the total number of crashes associated with that factor.

A factor was considered a significant contributor to severe crashes when the percentage of severe crashes associated with that factor was greater than or equal to two standard errors more than the percentage of all severe crashes.

#### **D1.5.2 Application 2: Statewide Traffic Crash Analysis Using GIS in Michigan**

Datta et al. (4) present a process for evaluating and prioritizing traffic safety projects and programs as proposed by local agencies. High-priority counties in Michigan were identified for implementing safety initiatives.

The crash frequencies and crash rates using various exposure measures were calculated for all cities, townships, and counties in Michigan for 12 crash categories. Ranking was performed by a frequency-rate method to avoid the selection of areas with low crash frequencies but high crash rates. In

the frequency-rate method, the crash frequency is plotted on the horizontal axis and the crash rate on the vertical axis. Areas (counties, etc.) are then plotted into cells, which represent a particular range of crash frequency and crash rate. The upper and right-most cell then contains the most hazardous areas. Proceeding to the left and down, decreasing levels of risk are identified. In the analysis, the frequency-rate method was conducted for each crash category and separately for each exposure type.

#### **D1.5.3 Application 3: Using Induced Exposure to Investigate the Effect of Driver Factors**

Vitetta and Abdek-Aty (5) used quasi-induced exposure methods to identify high-risk driver groups through a relative crash involvement ratio. Quasi-induced exposure is based solely on multivehicle crashes where fault has been assigned to only one driver involved in a crash. The use of quasi-induced exposure assumes that the distribution of nonresponsible drivers in the population of two vehicle crashes closely matches the distribution of the entire population of drivers. The use of quasi-induced exposure also assumes that the driver type of nonresponsible driver is independent of the driver type of the responsible driver. Crash propensity was examined for different age groups (e.g., 15–24, 25–34, and 35–44), vehicle types, road types, divided and undivided roadways, straight versus curved roadways, and various other conditions.

In general, when the event being studied is quite rare (e.g., a crash) in comparison to the amount of exposure (e.g., number of miles driven), then the odds ratios may be used to estimate relative risk, or crash propensity. The odds ratio is as follows (6):

$$\text{Odds Ratio} = (F1/E1)/(F2/E2)$$

where

E1 and E2 = the exposure counts for Groups 1 and 2 and

F1 and F2 = the event counts for Groups 1 and 2.

Group 2 is used as the baseline, and the odds ratio compares the odds of an event in Group 1 to Group 2.

#### **D1.5.4 Application 4: Estimating the Exposure and Fatal Crash Rates of Suspended, Revoked, and Unlicensed Drivers in California**

DeYoung et al. (7) used quasi-induced exposure to determine if unlicensed drivers or drivers with suspended or revoked licenses were over-represented in crashes. Data from NHTSA's Fatality Analysis Reporting System (FARS) were used to calculate crash and exposure rates for drivers in California. Only multivehicle crashes in which one driver was assigned fault were used to comply with the methodology.

Involvement ratios were calculated by dividing the percentage of unlicensed, suspended, or revoked drivers in the “at-fault driver” subset by the percentage of unlicensed, suspended, or revoked drivers in the “not-at-fault driver” subset.

$$\text{Involvement Ratio} = \frac{\% \text{ of drivers in the at-fault group}}{\% \text{ of drivers in the not-at-fault group}}$$

An involvement ratio greater than 1 indicates over-involvement, and an involvement ratio less than 1 indicates under-involvement. This study indicated that unlicensed, suspended, and revoked drivers are over-represented in fatal crashes 2 to 5 times more than fully licensed drivers. Potential sources of bias in the methodology cited by the authors included the following:

- The proportion of unlicensed drivers was 33% higher in the entire dataset than in the subset of crashes where only one driver was deemed at fault.
- Drivers who die in a crash are less likely to be assigned fault, and if any of the groups is over-or under-represented in fatalities, the rates will be affected.
- If not-at-fault drivers have characteristics making them more likely to be involved in a crash, the exposure of these groups will be over-estimated.

#### **D1.5.5 Application 5: Use of the Induced Exposure Method to Study the Highway Crash Involvement of Driver Groups Under Different Light Conditions**

Dissanayake and Lu Jian (8) used Chi-square statistics and contingency table analysis to explore the relationship among driver age (ages 15–25, 25–65, and 65+), light condition (daylight, dusk, dawn, darkness with street lights, and darkness without street lights), and crashes. For two random classification variables  $X$  and  $Y$  (in this case age and light condition), where  $X$  has  $r$  levels and  $Y$  has  $c$  levels, the contingency table has  $(r \times c)$  cells. Each cell represents the observations of  $X$  and  $Y$ .

In contingency table analysis, the observed number of crashes under each category is compared with the expected number of crashes obtained by assuming a null hypothesis. The Chi-square goodness-of-fit test statistic determines if the observed and expected values vary significantly. If this is the case, the null hypothesis used in obtaining the expected values is determined to be false. Based on the critical Chi-square value, a decision is made regarding the acceptance or rejection of the null hypothesis.

Three types of analyses were undertaken. The first tested if driver age and light condition were independent considering crash involvement. The second tested if driver age and light condition were independent considering driving exposure. The third tested if crash frequencies were proportional to driving exposure.

To test if random variables  $X$  and  $Y$  are dependent on each other, the null hypothesis is that they are independent. Under this hypothesis, the expected frequencies for a cell are found by multiplying the total number of observations by the probability that  $X = x$  and  $Y = y$  (where  $x$  and  $y$  are the marginal distributions of the row and column variables, respectively). To consider driving exposure, an induced exposure measure was used. Only the not-at-fault drivers were considered in order to represent the amount of travel by that group. Considering at-fault drivers as well would introduce biases if some driver groups were over-represented in the crash data.

The analysis determined that driver age and light condition were factors in crash involvement. It was also found that driving exposure depends on driver age and light condition. Finally, analysis showed that crash involvement of drivers did not depend on driver exposure or light condition and that certain driver groups are over-represented in crash involvement for certain light conditions. Identification of over-represented driver age groups was measured as the ratio of the percentage of at-fault drivers to the percentage of not-at-fault drivers. Where this value is greater than 1, there is an over-representation of that age group in crashes.

#### **D1.5.6 Application 6: Problem Identification Methodology—Network Screening Applications**

“Network Screening” is a formal analytical procedure to identify “sites with promise” for more detailed analysis. “Sites with promise” are sites that will benefit the most in terms of safety by the expenditure of resources. A vast body of literature exists on these procedures. *NCHRP Synthesis of Highway Practice 91* (9) and *NCHRP Synthesis of Highway Practice 128* (10) provide a survey of common procedures used in state DOTs for identifying hazardous locations. *NCHRP Synthesis of Highway Practice 295*, recently completed by Persaud (11), has provided some additional insights in this regard.

More sophisticated procedures have been proposed in recent years. A major ongoing FHWA project will provide states with practical state-of-the-art tools, including software, for conducting network screening as part of *Safety-Analyst* (Software Tools for Safety Management of Specific Highway Sites). An interim product is expected to be available by 2004.

Below is a summary of more recently developed techniques and ongoing initiatives that will influence the *Safety-Analyst* tools.

##### **D1.5.6.1 Advanced Network Screening Procedures**

Flowers and Griffin (12) have identified some difficulties with conventional quality control methods. They suggest that the upper control limit should be stated in terms of crash fre-

quency rather than crash rate, which is used in conventional applications. This is in recognition of the nonlinear relationship between crashes and traffic volume. The authors further point out that, due to the randomness in the count of fatal crashes and the large weight attached to this count, gross inaccuracies would result if the *observed* crash severity for a road section were used for prioritization. In short, the authors suggest that EPDO (equivalent property damage only) crashes should *not* be used for ranking. To mitigate this problem, the empirical Bayes procedure is proposed (13, 14, 15).

There are several variations of the empirical Bayes method. The essence of these variations is that the hazard of a site should be based on an estimate of this site's long-term mean ( $m$ ) rather than on its short-term count ( $x$ ) or short-term crash rate. An estimate of the long-term mean is obtained from an empirical Bayes procedure that combines the crash count ( $x$ ) of a specific site with an estimate of the long-term mean ( $P$ ) based on the crash history of similar sites. This estimate of  $m$ , the long-term mean, is a weighted average of  $x$  and  $P$ .

In research by Hauer (14) and Pendleton (15), the safety performance (SP), or the predicted number of crashes, is estimated from a safety performance function. A safety performance function is a regression model with traffic and geometric factors as independent variables. The weights of  $x$  and SP are estimated from the mean and variance of the regression estimate. Safety performance functions for estimating SP and the associated weights for a variety of location types are available in recent literature. It is expected that a major effort in the development of *SafetyAnalyst* will be devoted to completing and upgrading the suite of safety performance functions that might be required for the purpose of network screening.

Persaud et al. (16) have evaluated the simpler empirical Bayes-based procedures using data for 1-km, two-lane highway sections in Ontario. For these procedures, the hazard of a site is indicated by one of the following:

- The value of the long-term mean ( $m$ ),
- The *potential for safety improvement* (i.e., the difference between  $m$  and what is expected for similar sites [ $m - SP$ ]), or
- The long-term mean crash rate (i.e.,  $m$  divided by traffic volume).

It was found that the empirical Bayes-based procedures outperform the conventional techniques of using the short-term crash count or crash rate to flag potentially hazardous sites. Of particular note were the poor results of the crash rate procedure, which showed a trend to identify sites with lower annual ADTs. The *potential for safety improvement* method has since been refined and extensively tested by Persaud et al. (17) and has recently been adopted for use in a few Ontario jurisdictions (18, 19).

There are other, more complex variations of procedures based on the empirical Bayes technique. Higle et al. (20), for

example, propose that hazardous sites be identified on the basis of the probability of the empirical Bayes crash *rate* of a site exceeding some threshold value.

#### D1.5.6.2 Current Network Screening Initiatives

Somewhat more sophisticated, but similar in principle, is a relatively recent application of empirical Bayes-based procedures in an FHWA-supported initiative in New York State (21). Following this work is another FHWA-supported initiative in Colorado that is being spearheaded by Hauer. This research effort recognizes that the overriding aim of the highway safety improvement process is to spend money where it achieves the greatest effect in terms of crash frequency and severity reduction. The implication is that money will tend to go to sites where there are many severe crashes or where the potential for crash reduction is large, not to sites where crashes are few but the crash rate is high because of low traffic volumes. Given these considerations, current initiatives are exploring the practicality of ranking locations for investigation using prospective cost-effectiveness of potential safety treatments. The most important product of the Colorado effort that will feed the development of the *SafetyAnalyst* will be a software package that implements advanced statistical methods for identifying "sites with promise."

#### D1.5.7 Application 7: Methodology for Developing Countermeasures for Hazardous Locations (*SafetyAnalyst*)

Traditional approaches to developing crash countermeasures have tended to address either the engineering or human factors in isolation. Such studies ignore the fact that human errors are more likely in some road and traffic situations than in others and that the likelihood and the consequences of those errors can be greatly mitigated through road design. Of late, several initiatives have offered an integrated approach to developing countermeasures. These are being culminated in FHWA's *SafetyAnalyst* project, which is developing state-of-the-art tools, including software, for all aspects of the road network safety management process.

The purpose of the diagnosis/countermeasure selection model under development as part of *SafetyAnalyst* is to guide the user in the diagnosis of safety problems and the selection of a possible array of countermeasures at a specific site. This site may be selected by a network screening tool or by the user on some other basis. The decision of which countermeasures actually get implemented is made with the use of economic analysis and priority-ranking tools. Capabilities planned for a network screening tool will include site identification, crash pattern identification, and the diagnosis of areas of concern by means of detailed office and field investigations. The diagnosis of areas of concern will be followed by the identification of countermeasures.



For each issue of concern, a list of potential countermeasures will be selected. Based on practical constraints at the site (e.g., ability to expropriate property to widen the road) and financial feasibility (e.g., ability to fund building an interchange), a reduced list of potential countermeasures will then be presented, together with target crashes and contra-indications (e.g., no rumble strips in high-density residential areas). Once the contra-indications have been used to eliminate any countermeasures, a final countermeasure list will be considered for implementation.

#### **D1.5.8 Application 8: Alabama Problem Identification Examples Based on an Integrated Roadway-Crash File—A Detailed Documentation**

This application pertains to the problem identification methodology discussed in previous sections of this report. This application includes the identification of roadway factors that may be crash concerns. Rather than providing references to assorted documents, detail is provided here to allow an interested reader to apply this methodology in his or her jurisdiction.

The objective of this section is to demonstrate the capability of roadway-crash file integration. This example only scratches the surface of the potential use of these data, and the displays below are strictly for exemplary purposes and not for drawing general conclusions. In fact, conclusions in one jurisdiction might not transfer to another, and so it is important that all jurisdictions move to establishing this integration.

In order to identify problems using a jurisdiction's roadway characteristic data, it is necessary to create two integrated databases. The first is a roadway data file to which crash information is added, the *roadway-crash* file. The second is a crash information file, to which roadway data are added (which is most commonly done), the *crash-roadway* file. Since distinguishing the two files could get confusing, here are formal definitions:

- **Roadway-crash file**—the file of roadway segment characteristics that has been augmented to include the frequency of crashes by severity for each of these segments. In this file, the number of crashes by severity can be determined for any segment and summarized for any subset of segments.
- **Crash-roadway file**—the file of all crashes to which geometric roadway data have been written so that the roadway characteristics (e.g., lane width, shoulder surface type, shoulder surface condition) can be determined for any crash and summarized for any subset of crashes.

To remember the distinction between the two files, recognize that the X-Y file is the X file to which the Y data have been added. Thus, the roadway-crash file is the roadway file to which crashes have been added.

While the distinction between these two files might seem subtle, it is quite significant. Without the roadway-crash file,

it is impossible to evaluate the effect of roadway characteristics since those geometric and other features where crashes did not occur cannot be determined. In effect, this roadway-crash file enables the comparison of “crash locations” with “noncrash locations,” which is impossible to do with the crash-roadway file.

To illustrate, the frequency distributions for only two of the variables in the roadway-crash file for Alabama are given in Figure D-10 and Figure D-11. These distributions look reasonable, and the data are complete for the mile-posted (i.e., state, federal, and interstate) roadways of Alabama. Note that in addition to the road characteristics and the crashes, ADT counts have been added to this file. Preliminary analyses determined that *it was impossible to draw any meaningful conclusions from these data without comparing roadways of comparable ADT*. This is because ADT is the primary and overriding factor in determining the rate of crashes at any location.

Another new variable that was not given in the crash-roadway file, at least as a categorical variable, is the road identifier. Its presence in the roadway-crash file enables identification of over-represented roadway types using the CARE IMPACT system. In essence, this gives the appearance that the worst locations are those with superior design characteristics, since these characteristics are usually correlated with the highest ADT categories.

##### *D1.5.8.1 Comparison of Crash-Roadway and Roadway-Crash Results*

The graph of crash frequency by average grade from the crash-roadway file is shown in Figure D-10. The vertical axis is crash frequency; the horizontal axis is roadway grade. Note that the top of the list in the table is the far left column of the graph (i.e., severe downgrade is the first column in the graph).

It might be reasoned that for every upgrade, there must be a downgrade, since the roadways are measured in both directions, and the opposing directions are usually close to each other. This reasoning would hold on two-lane roadways. It is expected that this graph would be nearly symmetrical. The conclusion drawn from the graph and tabularized data in Figure D-10 was that more crashes are occurring on downgrades than on upgrades.

Compare Figure D-10 with Figure D-11 for all segments in the roadway-crash database. In Figure D-11, as in Figure D-10, the vertical axis is crash frequency and the horizontal axis is roadway grade. Note that the top of the list in the table is the far left column of the graph (i.e., null is the first column in the graph).

Figure D-11 provides a control for the crash distribution given in Figure D-10. In other words, Figure D-11 includes slopes for all segments, while Figure D-10 was for slopes for only crash segments. The basic assumption (for every upgrade there is a comparable downgrade) holds very closely for all of the slope categories in Figure D-11 except the two “slight-

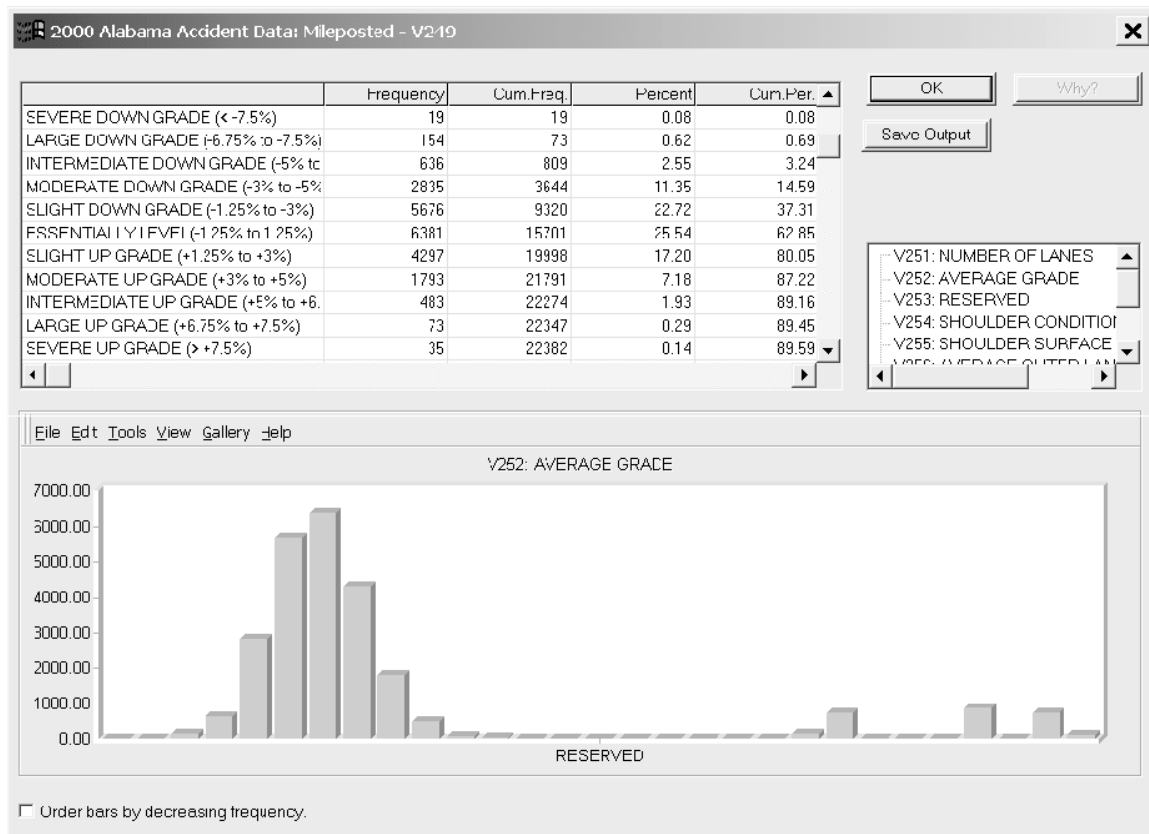


Figure D-10. Alabama crash data: Crash frequency by road grade, from the crash-roadway database.

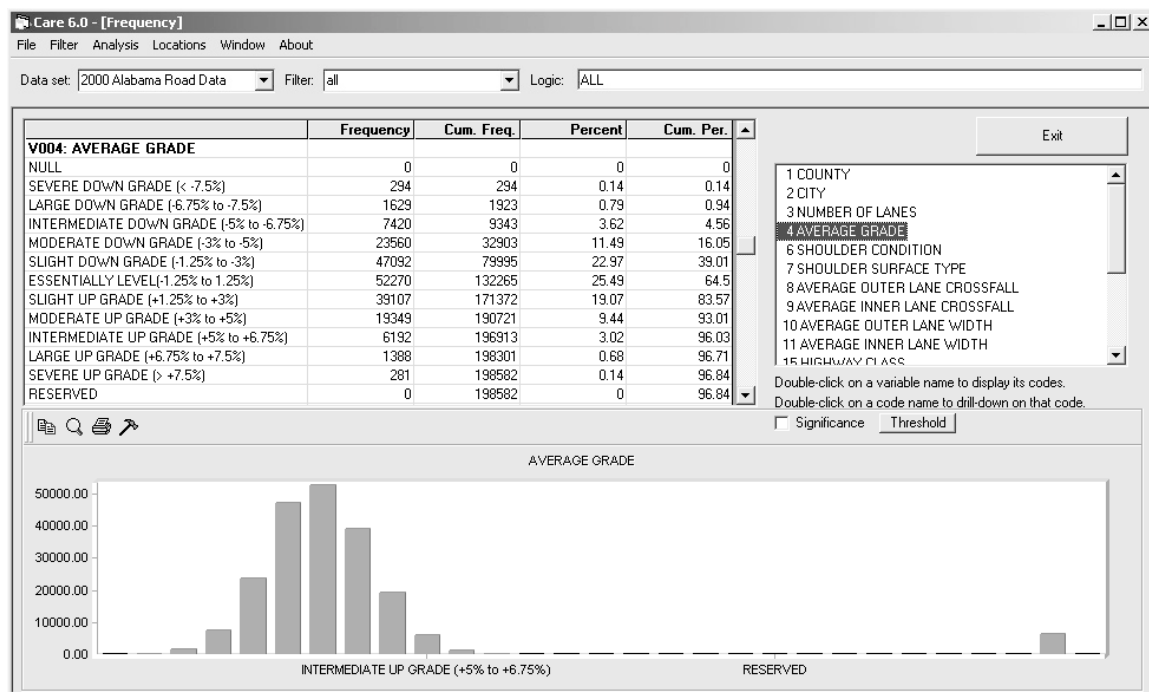


Figure D-11. Alabama crash data: Crash frequency by road grade, from the roadway-crash database.

grade” categories. This discrepancy could be due to incomplete data (e.g., some road segments were under construction) or to the accumulation of differences on some four-lane roadways that are separated enough to have differences in slopes. Regardless, the main purpose of this comparison is to demonstrate the value of having the type of data that is in the *roadway-crash* file.

#### D1.5.8.2 Example IMPACT Analysis within ADT Classification

Prior to the adding of ADT to the roadway-crash file, many CARE IMPACT analyses were run in an attempt to determine in general those roadway characteristics that could be causing crashes. All of these studies identified roadway characteristics at the upper ADT range to be characteristics with the most crashes. That is, those roadway characteristics identified as “high-crash characteristics” were inevitably some of the *best designed roadways in the state*. The reason is that the frequency of crashes increases with ADT, more so than with any other factor. Therefore, roadways with the highest ADTs, such as interstate highways, were identified. However, these high-volume roadways usually have the “best design” parameters. *In the absence of the ADT consideration, one might conclude that wider lanes, paved shoulders, and gradual slopes cause crashes.*

The addition of ADT to the roadway-crash file enables roadways to be compared *within ADT classifications*, thus removing ADT as the overriding causal factor (i.e., the pres-

ence of more vehicles is a causal factor). By comparing only roadway segments within a common range of ADTs, the effects of the roadway characteristics can be isolated. The next section presents an example for one ADT classification.

#### D1.5.8.3 Low-Volume Example: 2,501–5,000 ADT for All Crashes

All ADT classifications should be analyzed with all variables in the database, comparing those segments that have crashes (of various severities) with those that do not. This analysis requires a large number of computations. However, it is quite feasible with current technology. Some tools, such as the Alabama CARE system, can facilitate this analysis. The feasibility of the analysis will be illustrated here with just one ADT classification and one roadway characteristic.

For a relatively low ADT classification (2,501–5,000 vehicles per day), the roadways were almost exclusively (98.5%) two-lane roads. Several of the variables, including average grade and shoulder condition, showed no significant differences between crash and noncrash segments. Shoulder surface type analysis, however, showed that more than 11% more crashes than expected occurred on surfaces with grass shoulders, as shown Figure D-12. (*OverRep* value for grass shoulder surface is 1.111.)

In Figure D-12, the bar on the left is the proportion of the segments that had crashes with the particular shoulder surface type, while the bar on the right gives the proportion of segments that did not have crashes with the particular shoulder

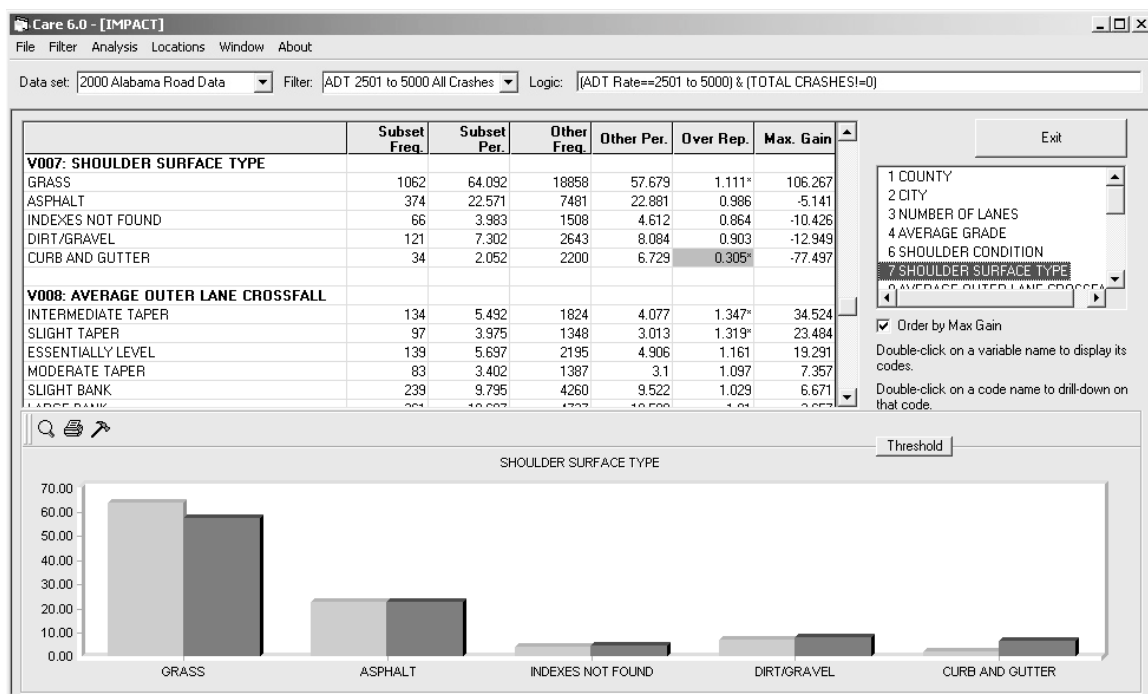


Figure D-12. Alabama crash data: All crashes on roadways with ADT between 2,501 and 5,000 vehicles per day.

surface type. If the particular characteristic (in this case shoulder surface type) has no effect, then the same proportion is expected in both bars. This is the case with most of the other shoulder type categories in Figure D-12, since the two bars are not significantly different. The “curb and gutter” category is under-represented in crash segments, perhaps because of lower speeds in urbanized areas resulting in fewer crashes. This analysis could be repeated with “curb and gutter” excluded for a sensitivity analysis.

This is just one example of the analysis of different variables in the database, and the objective of this example is not to conclude that grass shoulders are inferior. In other ADT classifications, grass shoulders were not found to be significantly different, and in other states this result might not be replicated. This is an example of the simple problem identification methodology previously given (ISMProcess Step 2) applied to determine the significance of a roadway characteristic in crash causation. A thorough analysis of all such characteristics over all ADT classifications is required in order to perform an effective and comprehensive problem identification of roadway characteristics in a jurisdiction.

Another useful comparison is within the roadways themselves. All of the roadways that were more than 50% higher than the expected crash values for this ADT classification (2,501–5,000 vehicles per day) are listed in Table D-9. (Since this information is nondiscoverable, the roadway identifiers have been altered for this example.) One would expect that

the crash percentage and the noncrash percentage for each roadway would be about the same:

- “Crash percentage” is the percentage of the crash segments on the roadway of all crash segments in this ADT classification.
- “Noncrash percentage” is the percentage of the noncrash segments on the roadway of all noncrash segments in this ADT classification.

One would expect that any given roadway would have the same percentage of crash and noncrash segments. When these percentages vary significantly, then the roadway should be subjected to further analysis to determine specifically where the problems might be. Note that because the data were gathered directionally, each roadway direction has a separate analysis (e.g., S-9 West is separate from S-9 East).

#### D1.6 OPTIMIZATION OF STRATEGIES METHODOLOGY (ISMPROCESS STEP 4)

The optimization tools given in this section can be applied at a variety of levels of the Problem (Emphasis Area) Identification Methodology. At the highest level, it might be applied to select emphasis areas in situations where all potentially promising areas cannot be addressed. This section assumes that the previous steps have established emphasis areas, objectives,

**TABLE D-9 Sample of roadways with more than 50% over-representation in crashes for the 2,501–5,000 ADT classification**

Road Identifier (Not Actual)	Crash Frequency	Crash Percentage	Noncrash Frequency	Noncrash Percentage	Over Rep.	Max. Gain
S-9 West	37	1.516	209	0.467	3.246*	25.602
S-6 South	57	2.336	665	1.486	1.572*	20.733
S-13 East	62	2.541	762	1.703	1.492*	20.443
S-11 East	49	2.008	580	1.296	1.549*	17.368
S-41 East	28	1.148	210	0.469	2.445*	16.547
S-74 North	31	1.27	267	0.597	2.129*	16.439
S-9 East	28	1.148	218	0.487	2.355*	16.111
S-74 South	29	1.189	269	0.601	1.977*	14.329
S-29 East	33	1.352	363	0.811	1.667*	13.203
S-13 West	55	2.254	769	1.719	1.311	13.061
S-26 North	40	1.639	499	1.115	1.47*	12.786
S-41 West	23	0.943	215	0.481	1.962*	11.274
S-27 West	45	1.844	626	1.399	1.318	10.860
S-26 South	38	1.557	501	1.12	1.391	10.677
S-6 North	47	1.926	675	1.509	1.277	10.187
S-29 West	30	1.23	366	0.818	1.503	10.039
S-58 South	32	1.311	433	0.968	1.355	8.385
S-170 North	19	0.779	195	0.436	1.787	8.365
S-116 North	10	0.41	32	0.072	5.73	8.255
S-221 East	14	0.574	109	0.244	2.355	8.055
S-179 East	16	0.656	146	0.326	2.009	8.038

\*Significantly over-represented.

potential strategies, and preliminary action plans. The objective is to take these inputs and use crash and demographic data to establish an optimal set of strategies. Possible applications at higher levels will be explained once the methodology is presented.

Once the emphasis areas have been established and a potential set of strategies is formulated, optimization can be subdivided into three major steps, analogous to the general problem identification methodology. The difference here is that the goal is to identify target crashes and potential crash reductions for emphasis areas, objectives, and strategies. Optimization consists of the following major stages:

- **Stage A: Identify subsets of data** that will provide information on the emphasis areas and the objectives identified. For each of these subsets, determine crash frequency by severity for the total subset of crash records that could be affected by the proposed strategies.
- **Stage B: Analyze the crash characteristics of emphasis area subsets** to generate the potential benefits for a set of strategies for each objective of each emphasis area.
- **Stage C: Determine the most effective combination of strategies** by using an optimization procedure that yields a final implementation plan considering available funds, costs, effectiveness, and other limitations.

Once optimization is complete, the Task Team should assess the achievability of the goals set out in ISMProcess Steps 1 and 2. It is desirable to confirm that the goals will be achievable through with the available resources. If the goals are unreachable with the current level of resources, the goals must be revised with the approval of the SPL during ISM-Process Step 4.

#### D1.6.1 Stage A: Identify Subsets of Data

In contrast with the subsets defined in the Problem (Emphasis Area) Identification Methodology, the emphasis areas, objectives, strategies, and preliminary action plans that have already been defined provide a definitive pathway to move forward.

The defined emphasis areas enable one to use crash records to estimate the potential crash reduction of strategies within each objective and emphasis area.

Subsets of the crash records can be defined in a variety of ways depending on the emphasis area or objective. All subsets share one thing in common: they are defined by variables within the crash records. Logical combinations of variable attributes within the database define the subsets, sometimes combined with Boolean “AND”s and “OR”s. Table D-10 lists selected AASHTO emphasis areas and related subset definitions. Note that some emphasis areas have multiple subsets and that some subsets may apply to more than one emphasis area.

The following rules provide overall guidance in specifying subsets:

- If the subset can be defined by a single variable, use a logical OR operator to include all of the values in that variable that apply. As an example, the subset of crashes affected by graduated driver’s licensing would be a logical combination of the causal driver age within the range applicable for the state (e.g., ages 16–19), as in Table D-10, Emphasis Area 1.
- If more than one variable is required to define the subset, use a logical OR operator to include all of the values of each variable that applies. For example, alcohol-related crashes might be identified by several different variables. To form a comprehensive subset of alcohol-related crashes, it is necessary to identify all of the ways that alcohol could be indicated and use a logical OR operator to include them. Also, a decision must be made in this case as to whether “alcohol” also includes drug-related crashes. If so, these crashes will be included as well with logical OR (Table D-10, Emphasis Area 5).
- A logical AND combination is required in those cases where two variables must be present simultaneously in order for the crash to qualify. For example, all restrictions that involve both a geographic area and a specific variable will require a logical AND operator with the geographical specification and the unqualified specification of the variable. Another example is where the subset is only referring to a limited severity classification, as is typically the case with restraint emphasis areas (Table D-10, Emphasis Area 8).
- It is important to distinguish between “crashes caused by” and “crashes involving.” As an example, the emphasis area of “Improving motorcycle safety and increased motorcycle awareness” requires that all crashes involving motorcycles be included as opposed to just those caused by motorcycles (Table D-10, Emphasis Area 11).
- Some emphasis areas or objectives may have corresponding categories that are not clear-cut; these emphasis areas or objectives may require additional effort in order to define the categories. For example, an “emergency medical services” emphasis area might consider only those crashes for which the response time was greater than some minimum expected time (Table D-10, Emphasis Area 20). Coming up with a reasonable minimum response time may require a combination of expertise and trial and error.
- Some combinations require creativity and knowledge of the way that the crash data are structured. For example, just the portion of the vehicle damaged cannot determine “head-on crashes.” While some crash databases might have a code for head-on/cross-over crashes, it is not critical. Head-on crashes can also be selected by making sure that the two vehicles involved in the crash were traveling in opposite directions (Table D-10, Emphasis Area 18).

**TABLE D-10 Example subset definitions for selected emphasis areas**

EMPHASIS AREA	SUBSET DEFINITION
<b>I Drivers</b>	
1. Instituting Graduated Licensing for Young Drivers	Causal driver aged 16–19
3. Sustaining Proficiency in Older Drivers	Causal driver aged >64
4. Curbing Aggressive Driving	Primary contributing circumstance is improper passing, improper lane change/use, following too closely, failure to heed sign/signal, failure to yield right-of-way, or driving on wrong side of road
5. Reducing Impaired Driving	Any indication of alcohol or drugs on the part of any driver
6. Keeping Drivers Alert	Causal driver fatigued (includes apparently sleepy)
8. Increasing Seat Belt Usage and Improving Airbag Awareness	a) Any child occupant aged 5–9 in an injury (fatal + nonfatal injury) crash b) Causal driver not restrained
<b>II Special Users</b>	
9. Making Walking and Street Crossing Safer	a) Involving pedestrians b) Pedestrians or bicycle crash at intersection or interchange
10. Ensuring Safer Bicycle Travel	Involving bicycles
<b>III Vehicles</b>	
11. Improving Motorcycle Safety and Awareness	Involving motorcycles
12. Making Truck Travel Safer	Causal vehicle is a commercial vehicle
<b>IV Highways</b>	
14. Reducing Vehicle-Train Crashes	Involving trains
15. Keeping Vehicles on the Roadway	Single vehicle crash off-roadway
16. Minimizing the Consequences of Leaving the Road	
17. Improving the Design and Operation of Highway Intersections	Crash occurred at intersection
18. Reducing Head-On and Across-Median Crashes	a) Head-on crash on two-lane road b) Head-on crash on divided road
<b>V Emergency Medical Services (EMS)</b>	
20. Enhancing Emergency Medical Capabilities to Increase Survivability	EMS arrival time was greater than 20 minutes

- Some emphasis areas or objectives cannot be identified with any subset. For example, better management and improvements in the traffic records system cannot be isolated to any particular type of crash. In these cases, recognize that potential strategies will have an impact, though small, on all crashes. Thus, the entire set of crash records should be used as a base, recognizing that the percentage crash reduction might be estimated as quite small.

The above rules are not comprehensive, but they are sufficient to get most of the subsets defined. This must be done in conjunction with personnel (RAE group and Task Teams) who are familiar with the crash database.

#### **D1.6.2 Stage B: Analyze the Crash Characteristics of Emphasis Area Subsets**

The purpose of analyzing the crash characteristics of emphasis area subsets is to determine the potential benefits for a set of strategies for each objective of each emphasis area. The analyses may be facilitated by an examination of the preliminary

action plans related to a strategy. The crash records that define these subsets are analyzed in two ways:

- According to the severity of the crash subset for the entire emphasis area or objective and
- Using further analysis of the potential strategies by severity for various implementation levels.

The following steps are recommended to determine the potential benefits for a set of strategies:

- Step B1:** For each emphasis area or objective, generate a crash-frequency-by-severity table. Table D-11 gives example frequencies by severity for two emphasis areas.
- Step B2:** For each strategy, generate a description of how the emphasis area subset might be further modified so that it will better reflect the strategy (as in the second column of Table D-12).
- Step B3:** For each of the strategy subsets defined in the previous step, determine the logical specification to encompass the crash data (e.g., AND, OR). (The strategy subsets formulated and the considerations given in Stage A still apply.) Use the subset definitions to produce

**TABLE D-11 Example emphasis area crash frequencies by severity**

Emphasis Area	Fatal		Injury		PDO		Total
	Number	Percent	Number	Percent	Number	Percent	Number
1. Instituting Graduated Licensing for Young Drivers	124	0.54	5659	24.72	17113	74.74	22896
3. Sustaining Proficiency in Older Drivers	101	0.86	2694	22.90	8971	76.25	11766

PDO = property damage only.

a frequency distribution for each of the strategies by severity, as illustrated in Table D-12.

- **Step B4:** Reiterate Steps B2 and B3 for all of the emphasis areas so that the result will be a frequency distribution by severity for each strategy.
- **Step B5:** Determine or estimate accident modification factors (AMFs) (e.g., obtain AMFs from AASHTO Implementation Guides) for each strategy, costs, and benefits at different levels of implementation.

The strategies have now been quantified according to crash frequency by severity. However, before an optimal preliminary action plan can be attained, the strategies have to be transformed into mutually exclusive projects, and some estimate of effectiveness must be made.

#### **D1.6.3 Stage C: Determine the Most Effective Combination of Strategies**

While a few areas involved subjective judgment, most of the strategies and approaches were guided by the creation

and comparison of the subsets of the crash records (Stage B of this optimization methodology).

Before continuing, it would help to view an example output of determining the optimal preliminary safety action plan (Table D-13). This table was generated by the “Strategy Implementation Level Worksheet” (contained in the Electronic Attachment—Optimization Software and Spreadsheets) used to assist in the final steps that involves optimization—the selection of activities to implement within the total resource constraints to bring about maximum safety returns.

For example, at Implementation Level 1a of “Provide education, outreach, and training,” the expected AMFs are 0.98, 0.96, and 0.94 for fatal, injury, and property damage only (PDO) crashes, respectively. The related crashes determined from subset analysis were 124 fatal, 5,659 injury, and 17,133 PDO crashes. The present value of the cost of this implementation level is estimated to be \$20,000. The annual benefits are estimated to be \$32,643,000, calculated through the expected reduction in crashes (2.5 fatal, 226 injury, 1,027 PDO). The costs per crash type used for this example are \$3,000,000, \$100,000, and \$2,500 for fatal, injury, and PDO

**TABLE D-12 Example crash frequencies by severity for selected run-off-road strategies**

Strategy (from AASHTO Implementation Guides)	Description of crash data subset	Fatal		Injury		PDO		Total
		Number	Percent	Number	Percent	Number	Percent	Number
a. Shoulder rumble strips	All run-off-road crashes	397	1.75	8270	36.50	13991	61.75	22658
b. Inner shoulder rumble strips	Run-off-road in median	10	1.84	166	30.51	368	67.65	544
c. Enhanced delineation of sharp curves	Crash occurred on horizontal curve	155	1.95	3201	40.23	4600	57.82	7956
d. Improved highway geometry	All run-off-road crashes	397	1.75	8270	36.50	13991	61.75	22658
e. Improved pavement markings	All run-off-road crashes	397	1.75	8270	36.50	13991	61.75	22658
f. Elimination of shoulder drop-offs	Low shoulder roadway defects present	1	1.39	38	52.78	33	45.83	72
g. Skid-resistant pavement surfaces	Wet pavement conditions	55	0.93	1722	28.97	4167	70.10	5944

PDO = property damage only.

**TABLE D-13 Example cost-benefit calculations for selected strategies**

Strategy	Expected Number of Crashes/AMFs			Costs and Benefits Values in Present Values (PV)				
	Fatal	Injury	PDO	Cost in 1,000s	Benefit/Year in 1,000s	Life (Yrs)	Sum Benefit in 1,000s	Cost/Benefit *1,000
<b>1 Provide education, outreach, and training</b>	<b>124</b>	<b>5659</b>	<b>17113</b>					
1a Implementation Level 1	0.98	0.96	0.94	\$20	\$32,643	10	\$252,060	0.079
1b Implementation Level 2	0.96	0.94	0.92	\$50	\$52,257	10	\$403,510	0.124
1c Implementation Level 3	0.93	0.91	0.90	\$80	\$81,249	10	\$627,385	0.128
<b>2 Ensure drivers are licensed and competent</b>	<b>728</b>	<b>24058</b>	<b>75452</b>					
2a Implementation Level 1	0.99	0.965	0.965	\$100	\$112,645	2	\$209,455	0.477
2b Implementation Level 2	0.98	0.955	0.955	\$200	\$160,429	2	\$298,305	0.670
2c Implementation Level 3	0.075	0.94	0.94	\$500	\$210,265	2	\$390,970	1.279

AMFs = accident modification factors.

PDO = property damage only.

crashes, respectively. These values are estimated from the 1994 FHWA crash costs. Estimating crash costs is further discussed later in this appendix (Table D-17).

The effectiveness of this strategy is expected to extend over 10 years. The present value of the benefits (\$252,060,000) is calculated using an interest rate of 5%. The cost-to-benefit ratio is then calculated ( $20,000/252,060,000 = 0.000079$ ). This procedure is then repeated for each implementation level of each strategy.

The application of the above procedure at a higher level involves applying the AMFs to crashes by severity at that higher level. For example, at the highest level, this procedure could be applied to resolving an emphasis area by determining crashes by severity targeted in the entire emphasis area. Then expert estimates must be obtained as to the crash reduction that could be expected from that emphasis area in general and at what cost. Of course, this would be a very broad estimate and would have to assume some typical strategies.

In terms of solving the optimization problem, a number of approaches and software packages are available and some jurisdictions may already have access to such software. The approaches and software are reviewed at the end of this section. A sample software package that optimizes for two levels of implementation accompanies this document (Electronic Attachment—Optimization Software and Spreadsheets). The adaptation of that application to nonroadway improvement strategies or activities involves viewing each of the strategies or activities as a “location.” The alternatives being considered are implementation levels within these strategies or activities. These would correspond to the alternative countermeasures or activities being considered at a location.

To formulate this in a more general sense, recognize that optimization (for implementation levels or specific countermeasures) can be performed if the following information is available:

- **A subdivision of the total problem into recognizable stages** (these stages might be the emphasis areas, the strategies, or the action plans, depending on the level at which the optimization is being applied).
- **A number of alternatives for each stage of the total problem.** Of these alternatives, only one will ultimately be selected. These alternatives might be implementation levels at higher optimization levels or specific activities or countermeasures at the lower optimization levels. The “do nothing” alternative is implicit with each stage and need not be specified. The optimization procedure will specify it when no other alternative at that stage is superior to those given at other stages.
- **A cost and a benefit estimate for each alternative** (within each of the stages).

The model being employed to estimate crash reductions (or crash severity modifications) assumes independence between stages (emphasis areas, objectives, strategies, activities, etc.). That is, crash reduction at one stage is assumed not to affect decisions made in another stage. However, since crash causes are generally multifaceted and since it is impossible to come up with totally mutually exclusive crash categories (where a stage includes all such categories), this assumption does not always hold. It would tend to hold more in location-specific applications, especially if the locations were all far removed from each other, than in the implementation of other types of strategies, such as seat belt programs.

Adjustments need to be made if interdependence between stages within the optimization procedure is significant. Interdependence will generally always exist, but most often will not be significant enough to alter policy in any way. However, when it does, there are some easy ways to account for this. Consider one or more of the following:

- **Combine stages:** if two stages address the same problem, perhaps the categorization used can be improved.



For example, some strategies address both run-off-road and obstacle-off-roadway crashes.

- **Change the structure of the alternatives:** consider combining two strategies to produce three alternatives: A, B, and A plus B (simultaneously). This structural change would generally be necessary when stages are combined, but might also be considered when a strategy affects more than one stage (e.g., a public information and education program might be targeted at both speed- and alcohol-related crashes). The combined countermeasure would take into account the effect of implementing both simultaneously, either adding benefit for the synergistic effect or removing cost due to the overlap.
- **Rerun the optimization under various structures:** a strategy might be redundant in more than one stage. Once the strategy appears in the final policy, the source data spreadsheet could be altered to remove the strategy from the stage that is not funded at the lower optimization level. Then, rerun the optimization starting at a higher minimum cost. Such a hybrid policy would take into account that earlier funding already covered the strategy.

Again, the spreadsheet model and optimization technique are merely tools aimed at *minimizing* subjectivity in decision making, and not at *replacing* it entirely. Once decision makers understand how results are generated, these results can be used in a flexible manner to take into account interactions among the various stages in order to maximize safety for a

given level of funding (i.e., reduce fatalities and injuries in the most effective way).

#### D1.6.4 Illustrative Application of Optimization Methodology for Site-Specific Improvements on Roadways

This section is intended for use in conjunction with the electronic version of this report (Electronic Attachment—Optimization Software and Spreadsheets), which includes spreadsheets and data files illustrating software that may be used to optimize the allocation of resources for road safety improvements. A spreadsheet format has been used to input the cost and benefit information described above into Table D-14. For each implementation level for each strategy, enter the following estimates/information in the “Strategy Implementations Levels Worksheet.xls” spreadsheet:

- A cost for the level of implementation of the activities.
- A crash reduction percentage that will be expected if the action item is implemented at the cost level specified. This reduction may not necessarily be the same for all severity levels.
- A life in years that the action item is expected to bring about these benefits.

The spreadsheet automatically calculates a benefit per year and a present worth of the total benefits over the life of the

**TABLE D-14 Summary of cost-benefit information for three potential projects, taken from the summary sheet in the “CostBenefit Analysis Worksheet.xls” file**

Reference Number: 1249						
		Cost	Benefit	Maintenance Cost	C/B Ratio	Total Cost
Alt 1	Install traffic signal	\$140,000.00	\$1,096,444.07	\$6,000.00	0.1277	\$146,000.00
Alt 2						
Alt 3						
Reference Number: 1250						
		Cost	Benefit	Maintenance Cost	C/B Ratio	Total Cost
Alt 1	Protected dual left-turn lanes from EB to NB	\$125,000.00	\$10,350,977.44	\$100,000.00	0.0121	\$225,000.00
Alt 2						
Alt 3						
Reference Number: 1251						
		Cost	Benefit	Maintenance Cost	C/B Ratio	Total Cost
Alt 1	Improve radius for dual left-turns	\$40,000.00	\$153,353.86	\$4,000.00	0.2608	\$44,000.00
Alt 2						
Alt 3						

C/B = cost/benefit.  
EB = eastbound.  
NB = northbound.

action item. Once the costs and benefits are available, optimization procedures can be employed to select the set of strategies that will bring about maximum total benefit within the budget constraint.

The adaptation of this spreadsheet to nonroadway improvement-type strategies or activities involves viewing each of the strategies or activities as a “location.” The alternatives being considered are implementation levels within these strategies or activities. These would correspond to the alternative countermeasures (or activities) being considered at a location.

As an illustration, alternative hazard roadway improvements are used in the file “CostBenefit Analysis Worksheet.xls” to illustrate the general application of optimization to traffic safety resource allocation. When conducting an optimization, one should recognize that it is impossible to bring all locations to the highest standards attainable. Given this limitation, there are two primary considerations:

- There are literally thousands of locations on public roadways within most jurisdictions that could be improved by some countermeasure.
- In any given year, jurisdictions have a limited amount of safety funds available for these improvements.

In other words, funding for safety improvements is limited, and analytical procedures are necessary if benefits from these improvements are to be maximized. Traffic safety funds should be allocated to those locations or strategies that will produce the maximum return in terms of safety (both crash frequency and severity reduction). It is not a matter of trading money for lives, but rather saving as many lives as possible with the resources that are available. It is not possible to determine if any location or strategy will qualify for funding by considering only that location in isolation of all other candidate locations. Optimization requires that all potential locations be considered simultaneously. This can be accomplished by the following general steps:

1. Identify the locations within the jurisdiction that have the highest apparent potential for crash reduction; these will be subjected to further investigation.
2. Perform investigations at these candidate locations to propose countermeasures, and estimate the countermeasures’ costs and benefits.
3. Process the data to generate an optimal budget allocation.

*Optimal budget allocation* is, by definition, that subset of all of the proposed countermeasures or activities that will maximize the total benefit produced. A problem identification approach should be used for determining potentially hazardous locations. The following steps are recommended to determine those locations that will be investigated in order to generate cost and benefit information:

1. Include locations that crash records indicate have above a given threshold criterion as a function of crash frequency and severity.
2. Determine locations that have excessive crash rates (with severity considerations).
3. Include locations that have been reported by the public, police, or other governmental officials to be potentially hazardous.

The number of locations considered for improvement should be about three times greater than that which can be funded at this point. No questionable location should be excluded. Each of these locations will be investigated to generate the data that will drive the remainder of the process.

Once a set of locations for potential improvement has been determined, reports are generated to summarize the crash records for each of these locations. These reports should include all of the variables that will be useful for investigation, including such variable summaries as time of day, day of the week, crash type, primary harmful event, weather conditions, and all other roadway and environmental factors. In addition to this summary, the original hard copies for all of the crashes occurring at the location should be made available to the investigation team. An interdisciplinary local investigation team is best, including state and local officials and crash investigation specialists. Local participation is essential, and those who have first-hand knowledge of the crashes (first responders) can be invaluable. However, the team should also include those who have experience in countermeasure evaluation so that reasonable estimates of benefit can be made. Several teams might be required for statewide or large area optimizations.

The investigation team should determine the following for each location:

- A categorization of each crash that occurred at the location by cause,
- A summary of the crash frequency and severity by cause, and
- A statement of one or more potential roadway improvement countermeasures for each crash cause.

In turn, the expert investigation team will call upon the literature and personal experience to generate the following for each of the potential countermeasures that the team proposes for the location:

- The countermeasure cost (out of the budget that is to be allocated),
- The number of the crashes that have occurred at the location that will be reduced if the countermeasure is implemented (this can be indicated by a percentage reduction within that given cause), and
- The expected life for which the countermeasure will bring about benefits (before replacement or other major costs are required).

With these data, the cost-benefit information required for the optimization process is generated. Table D-14 provides an example of this information for three locations that are used later to demonstrate the inputs and outputs of the optimization software.

The process at this point is one of getting the data, exemplified by Table D-14, into a format that can be entered into the optimization routine. A cost and benefit estimate is required for each proposed alternative. The benefit estimate is obtained by taking the present value of the crashes saved (by severity) according to the agency's method of costing crashes. A pooling of injury and fatal crashes is recommended in order not to skew the results toward those locations that might have had a fatality. (In most cases, an injury crash is as good a proxy for future fatality crashes as a fatality crash.) The benefit calculation routine can be programmed into the spreadsheet (Cost-Benefit Analysis Worksheet.xls), as was done to produce Table D-14, so that the benefit estimate can be read directly from it.

At this point, all of the data are available to drive the optimization process. The optimization process is a stagewise decision process where the stages here are defined by the locations (i.e., each location is a stage). The problem resolves itself to allocating a portion of the budget (including zero if no countermeasure is implemented) to each location in such a way that the combination of these allocations produces the maximum total benefit. Several optimization techniques can be applied to accomplish this, and these are reviewed at the end of this section (ISMProcess Step 4). The only constraint on using the computer program included with this report is that there cannot be more than two alternatives at any location. (In practice, this has not been found to be limiting.)

Table D-15 presents a portion of the cost-benefit data formatted for input to the computer program. The major part of the data (after the first three lines) consists of three numbers per line: a location number, a present value cost, and a present value benefit. A location in this version of the software can have up to two countermeasures in this computer program. (At most, one countermeasure will be selected for a location; however, a countermeasure could be a combination of improvements, or a strategy could be achieved by a combination of activities.)

**TABLE D-15 Data formatted for the optimization process**

1			Output flag parameter
2000000	3500000	250000	Budget parameters: lower; upper; increment
68			Number of locations
1217	1250001	734949	Location number; PV of cost; PV of benefit
1220	112000	361081	
1220	350000	2166486	
1221	550000	476858	
1223	116000	4134549	

PV = present value.

Notice that the second line of Table D-15 tells the program to begin outputting results for budgets that range from \$2 million up to \$3.5 million in increments of \$250,000. These parameters can be altered to produce allocations over any range of budgets. Table D-16 demonstrates the outputs from just three budgets generated for this example. This illustrates that as the budgets are increased incrementally, additional countermeasures are added to the optimal solution. A zero cost and benefit indicates that no alternative was selected for the corresponding location.

The cost and the benefit over the range of budgets can be plotted as in Figure D-13. This enables comparisons between different optimizations over time or even allocations to various types of countermeasures. It also enables those with flexibility in postponing or accelerating expenditures to determine if additional funds should be allocated to (or removed from) the current budget.

#### *D1.6.4.1 Adjusting Safety-Related Costs and Benefits to Current Dollar Amounts*

Usually, the most recent official cost-benefit publications need to be adjusted to current prices before they may be used for highway safety calculations. For example, the costs of PDO and injury crashes for the value of life (as defined in "Revision of Departmental Guidance on Treatment of the Value of Life and Injuries" [22]), were provided in FHWA's 1994 Technical Advisory (23).

To convert a historical cost to a current cost using the methodology given in the Office of the Assistant Secretary for Transportation guidance document (22), take the ratio of the most recently available gross domestic product deflator over its corresponding historical value.

Thus, to convert the 1994 FHWA cost of \$2,000 for a PDO crash to 2002 dollars, the ratio of the 2002 (third quarter) and 1993 (fourth quarter) gross domestic product deflators (110.73/94.79) is multiplied by \$2,000, resulting in a value of about \$2,300. The gross domestic product deflators can be found on the Department of Commerce's Bureau of Economic Analysis website, [www.bea.gov](http://www.bea.gov) (Select News Releases > Select Gross Domestic Product > Scroll down to "Table 5. Quantity and Price Indexes for Gross Domestic Product"). Following this prescribed methodology, the crash costs by severity for 2002 are shown in Table D-17.

#### **D1.6.5 Review of Other Optimization Approaches**

In reviewing mathematical optimization approaches of relevance, it is recognized that these approaches are mainly geared toward the prioritization of specific projects (e.g., allocating a budget for site improvements). Generally, the projects are independent in that specific crashes are targeted by only one project. Therefore, available procedures may need to be adapted for the problem at hand—allocating a general safety budget among emphasis areas, strategies, and activities (each a different stage of optimization), which may overlap in terms

**TABLE D-16 Partial output of optimization process**

Location #	Cost (\$)	Benefit (\$)	C/B Ratio
1306	10000	6007070	0.001665
1305	0	0	
1304	0	0	
1303	0	0	
1302	0	0	
1301	3000	454501	0.006601
1300	0	0	

C/B = cost/benefit.  
 Budget = \$2200000  
 Total benefit = \$114164856

of crashes targeted. There are two alternatives to optimize projects that may overlap:

- Conduct additional research, and very likely change data structures and the way crash data are collected.
- Apply conventional procedures, recognizing the potential for and the implications of project overlap, and make appropriate adjustments.

This review of optimization approaches is based on the second alternative.

It should be recognized that Task Teams will not be able to optimize their own strategies and action plan activities because they possess the knowledge of only a particular subset of the budget. The Operations Manager, on the other hand, has the complete picture. With the support of the RAE group, the Operations Manager has tools designed to help optimize the subset of the total safety resources.

Several methods, covering a wide range of complexity, are available for priority ranking of sites/improvements. A number of relevant sources have been reviewed in order to determine the current state of knowledge. While the summary on

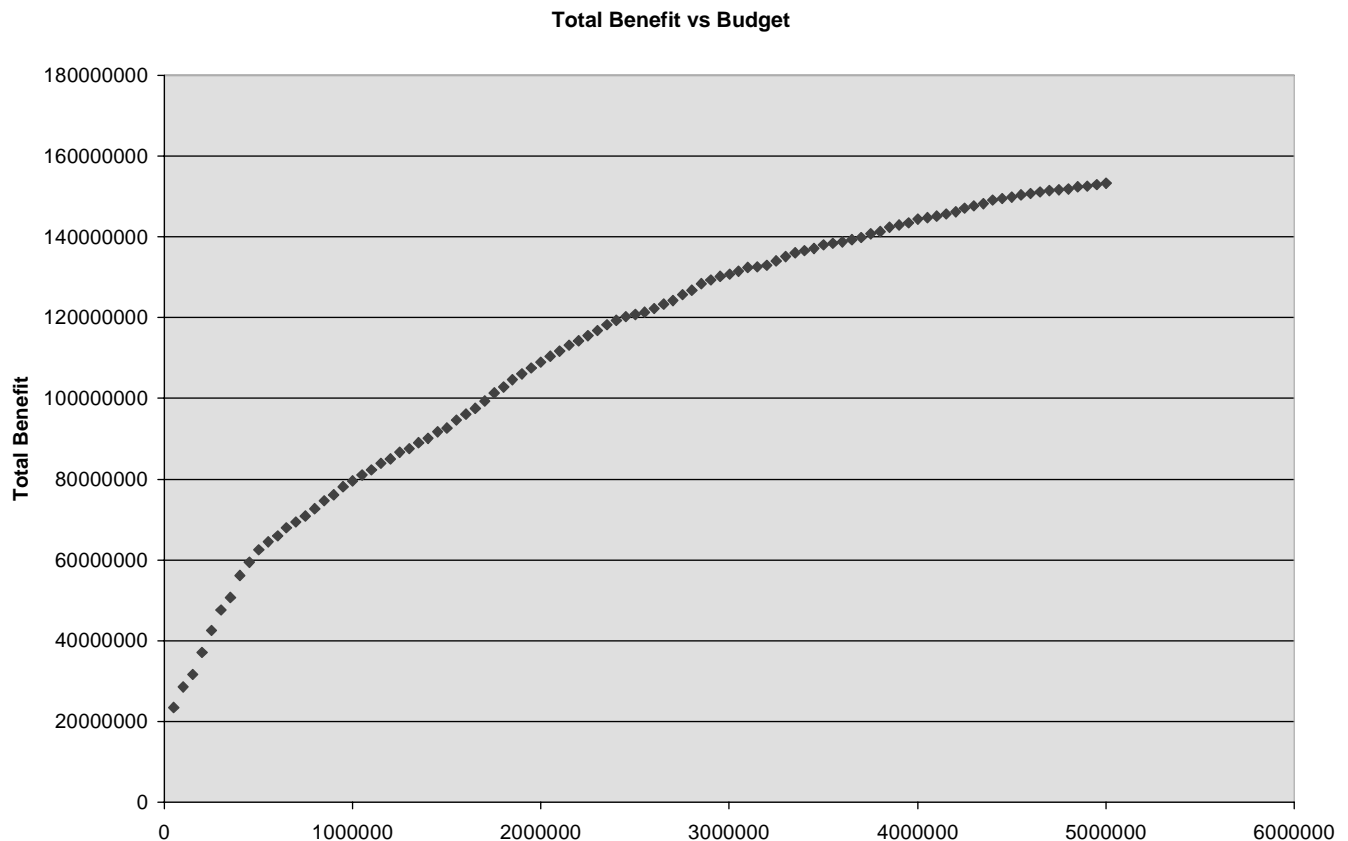


Figure D-13. Example diminishing marginal return curve.

**TABLE D-17 FHWA crash costs converted into 2002 dollars**

Crash Severity	1994 Value	Conversion	2002 Value
Fatality	\$2,600,000	$\frac{110.73}{94.79} = 1.168$	\$3,000,000
Incapacitating	\$180,000		\$210,000
Evident	\$36,000		\$42,000
Possible	\$19,000		\$22,000
Property Damage Only	\$2000		\$2,300

optimization methods in the following paragraphs cannot be regarded as exhaustive, it can be seen as a reasonable representation of current practice, as well as some new ideas.

#### *D1.6.5.1 Conventional Mathematical Approaches*

A useful summary of safety resource allocation methods is provided in two reports (24, 25), the first of which reviews a number of promising procedures for selecting safety improvements to result in the maximum safety benefits per dollar spent. These procedures are grouped into two categories:

- Weighting methods, including the cost-benefit and cost-effectiveness methods, and
- Mathematical programming methods, including dynamic and integer programming.

**Weighting methods.** There are two main types of weighting methods:

- **Project development ranking** ranks projects according to net benefit, cost-effectiveness, rate of return, or other economic measures and then selects from the list until available funds are depleted. The procedure is simple and popular but is not regarded as effective where there are a number of alternatives at each location or where the priority listing is constantly revised.
- **Incremental benefit-to-cost ratio** allows a project to be selected if the extra expenditure is less than the extra expected benefits when compared with the next lower-priced alternative. This widely used procedure reduces the impact of very-low-cost projects (compared with the cost-benefit ranking method) while enhancing consideration of additional improvements based on expected additional benefits.

A recent report by FHWA examined current and potential roles of cost-benefit analysis in the transportation infrastructure field (26). The way in which recent research advances can be incorporated into the theoretical base and practical techniques of cost-benefit analysis were explored. Insights were provided into the use of cost-benefit analysis in program- and project-level analyses and into the appropriate cost-benefit

decision criteria for particular applications. The report synthesized the following information from various other sources:

- Estimation of parameters, such as the appropriate discount rate, value of life, value of time, and externalities, when applying cost-benefit analysis to the transportation field (27).
- Estimation of relevant costs and benefits, including life-cycle concepts and nonmarket costs (28).
- Recent and ongoing research into the significant external benefits (such as productivity improvements) that result from transportation network improvements (29).
- Techniques for assessing multimodal and multiple-objective alternatives (1).
- Future directions, knowledge gaps, opportunities for improvement, and future research necessary for cost-benefit analysis (30).

**Mathematical programming methods.** Five methods were considered in the FHWA Highway Safety Improvement Program (24): goal programming, network analysis techniques, linear programming, integer programming, and dynamic programming. The last three methods were recommended for consideration in the allocation of highway safety funds and are summarized below:

- **Linear programming (LP)** defines a class of problems in which the decision variables are nonnegative, the criterion (or objective function) for selecting the best values of the decision variables is a linear function of these variables, and the constraints (e.g., resources) can be expressed as linear equations or inequalities. The LP method is said to be the most widely used method of mathematical programming, although it was not common or easy to employ in the selection of safety measures for specific locations. However, the FHWA report (24) suggested that the LP formulation could possibly be used to allocate safety funds among safety programs.
- **Integer programming** is a linear programming problem in which some or all of the decision variables are restricted to integer values. The FHWA report about the Highway Safety Improvement Program (24) states that limited progress has been made in the solution of large-scale highway safety problems using this method.

However, recent optimization software, such as LINDO (31), can handle very large problems involving up to 4,500 variables and 800 constraints.

- **Dynamic programming** is an optimization technique that transforms a multistage decision problem into a series of one-stage decision problems. At the single-stage level, a single project with several alternatives is evaluated. At the multistage level, selection is made among several projects, each with several alternatives. At the time of the FHWA report about the Highway Safety Improvement Program (24), two states (Kentucky and Alabama) were actually using this technique and more states were considering using it.

In Alabama (28), the dynamic programming procedure begins with a computer search of crash records over the last several years to provide a list of candidate locations for safety improvement. The data are then summarized and sent to divisional investigation teams, where engineers familiar with the location generate possible safety improvement alternatives. The divisional investigation teams are encouraged to add locations to the list that may not have had enough crashes to be included, but are considered to be potentially hazardous. Investigations at each site are then conducted, and standardized forms are sent to the central office for accuracy and consistency checks. The forms are processed by an algorithm, which generates cost and benefit data for each alternative at each candidate location. The benefits are based on the specific crashes that the investigators believe will be addressed by each proposed countermeasure. These costs and benefits (of perhaps multiple proposed countermeasures per location) provide the input into the dynamic programming optimization routine, which produces the maximum benefit possible over a wide range of potential budgets.

Brown et al. (29) presented an update to the dynamic programming procedure used in Alabama (28). They developed a so-called “branch-and-bound” technique to handle larger sets of project data more efficiently and to guarantee optimality.

#### *D1.6.5.2 Comparison of Conventional Mathematical Approaches*

With regard to the choice of safety projects that maximizes total benefits for a fixed budget, the FHWA report about the Highway Safety Improvement Program (24) ranked the possible procedures in the following order:

1. Integer Programming,
2. Dynamic Programming,
3. Incremental Benefit-to-Cost Ratio, and
4. Benefit-to-Cost Ratio.

Despite its number one ranking, the use of the integer programming approach is rare, perhaps because of its complex-

ity. Fortunately, as the aforementioned FHWA report suggests (24), incremental benefit-to-cost ratio (ranked third) gives approximately the same choice of projects as the two programming methods. The main difference is in the choice of marginal projects within the budget. This is not seen to be an important shortcoming, since safety budgets are often not precisely specified, and human judgment should be applied on those projects competing for the last dollars to be allocated.

#### *D1.6.5.3 Alternative Approaches*

Alternative approaches have been developed to overcome difficulties experienced by highway agencies in applying conventional approaches. Three fairly recent initiatives provide a flavor for these alternative optimization approaches.

McGeehan and Samuel (32) evaluated the road improvement prioritizing system being used at the time by the Virginia DOT and sought to develop an alternative method. This alternative method is a sufficiency rating system that evaluates proposed projects on the basis of points assigned for a number of variables representing cost, safety, traffic intensity, and road classification. Values are assigned to a variable based on an estimate of its significance in the prioritizing decision. Points for each variable are awarded to a project according to its level of deficiency in that variable, with a maximum number of points assigned for the most serious condition. The sum of points awarded to a project yields a number for comparison with other projects in prioritization and optimization. In this new method, each variable is divided into three data ranges representing values that are high (above a threshold), medium (marginal), and low (adequate). All proposed projects are initially sorted into three groups of high, medium, and low as defined by the ranges of a primary variable. The grouped projects are then evaluated using ranges of a second variable. Projects in the high range of the second variable may be elevated to the next highest group. The projects are ranked within the high, medium, and low classification by the ranges of the individual variables as well. In the final prioritized list, the first project in each group would be the most deficient as assessed by all variables.

The above approach is particularly attractive when good benefit estimates are not available (e.g., for new and innovative programs or at the level of an emphasis area or strategy) and where many traffic safety experts are available to make the ratings.

Chowdhury et al. (33) recognized that the decision of which improvements should be programmed should be based on credible methodology that accounts for multiple objectives. After identifying the appropriate countermeasures on each road segment and the functional relationships between the crash rate and the cost of implementing the identified countermeasures, Chowdhury et al.'s multiple-objective methodology selects the best countermeasures and allocates the available resources optimally.

Finally, and of especial relevance, is a Delphi-style decision-making procedure used by Wisconsin to prioritize emphasis areas. In this procedure, two ratings for each emphasis area in the AASHTO Strategic Highway Safety Plan (in addition to two emphasis areas pertinent to Wisconsin) were developed. Rating 1 is based on the importance of the emphasis area in terms of improving safety—essentially reducing the target crashes. Rating 2 is based on the expected ability of safety agencies to influence the problem. These two factors were rated by 160 attendees at a conference and were multiplied together to get an overall rating. The conference attendees included representatives from the American Automobile Association, the Department of Public Instruction, the University of Wisconsin, NHTSA, FHWA, the American Association of Retired Persons, the courts, the media, law enforcement, and the legislature.

Through this conference, called the “Traffic Safety Strategic Change Event,” the original Wisconsin list of 24 emphasis areas was pared to 7 emphasis areas that participants felt were not only important, but also could be influenced by actions taken by Wisconsin DOT. Groups were formed to develop strategies that included specific projects and policy recommendations for each of the seven emphasis areas.

#### *D1.6.5.4 Optimization Software*

Commercially available software packages can be classified as either generic optimization software or software specific to safety resource allocation:

- **Generic optimization software.** Numerous packages are available, many with free downloadable demo versions. A very useful summary with Internet links can be found at <http://www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide>.
- **Software for safety resource allocation: The Safety Resource Allocation Program (SRAP) package.** FHWA initiatives form the basis of an FHWA-developed software package, SRAP (34). A companion

report to the user manual for SRAP describes implementation experience in Iowa (35).

SRAP’s input data requirements include the following:

- Number of locations to be considered;
- Overall budget available, in dollars; and
- Number of countermeasure alternatives to be considered at each location.

Each alternative at each location requires the following:

- Initial construction costs, in dollars and
- Present worth of annual net benefits, in dollars.

The inputs in the form of costs and benefits need to be converted to present values using standard economic analysis equations. Analysis requires assumptions on discount rate and project life. The SRAP user manual recommends applying an annual increase for maintenance and operating costs.

The SRAP procedure can be used to prioritize all projects within a given budget, not just safety-oriented projects, as long as it is possible to estimate the present value of costs and benefits of each project. The output of the program is a list of projects, prioritized to provide the maximum possible benefit from the expenditure of a specified budget. Three different prioritization methods can be used and selected by the software user: incremental cost-benefit analysis, integer programming, or dynamic programming.

The main difficulty in applying the SRAP package is that it can only handle one level for each alternative, so it is ill suited for solving a problem in which different levels of implementation of an alternative need to be considered. The only way to apply the SRAP for different implementation levels is by artificially adjusting the inputs iteratively if the selected optimal strategy set contains one or more strategies at more than one implementation level. However, this adjustment is cumbersome, and it may be far easier to adopt a generic software package that specifically addresses this difficulty.

## Appendix D2: Tools to Estimate the Effectiveness of Promising or Innovative Strategies About Which There Is Insufficient Information

### D2.1 INTRODUCTION

Before making a decision on a safety countermeasure or strategy, decision makers must have information about the expected safety benefits. Unfortunately, this important information is often not available, is highly variable, or is even contradictory among different safety professionals. Results of studies completed using data from one region may not be applicable to another.

When assessing an innovative strategy (or a strategy with no reliable or known AMF), it is also important to consider the strategy within the context of the entire Highway Safety Improvement Program. The Operations Manager and SPL should ask the following questions:

- What effect could implementing this strategy have on other safety initiatives?
- Could other initiatives suffer in any way because of the implementation of the new strategy?
- Could resources be drawn away from other initiatives?

The relationship between estimation and measurement tools in the ISMProcess is shown in Figure D-14. This appendix focuses on methods to study the effect of new, innovative, or existing strategies for which there is otherwise insufficient information to estimate their effect. The tools discussed here include the following:

- Formal evaluation studies.
  - Experimental studies.
  - Quasi-experimental designs and observational studies (cohort studies, case control studies, before-after studies, time-series analysis).
- Alternatives for formal evaluation studies.
  - Bayesian analysis studies.
  - Systematic reviews.
  - Subjective assessments.

The purpose of these tools is to quantify the effectiveness of different strategies. Using these tools often requires specific data to be collected before, during, and after the implementation of a countermeasure. A formal statistical experiment is appealing to the researcher because the logic of a simple experiment is often easiest to understand, with the results and conclusions being the most compelling.

An alternative design to the formal experiment is the quasi-experimental observational study, which can take advantage of the surveillance practices used in many jurisdictions associated with the licensing of drivers, the registration of vehicles, and the documentation of traffic crashes.

Case studies of the different evaluation tools are provided toward the end of this appendix.

#### D2.1.1 Assessing Whether Information Is Insufficient

Assessing whether information on the safety performance of a measure is insufficient can be done only after a systematic review of available knowledge. Therefore, a systematic review is an evaluation tool to define effectiveness of promising, new, or innovative strategies. The conduct of systematic reviews is covered later in this appendix.

Information can be insufficient for two primary reasons:

- There have been no safety evaluation studies.
- The studies undertaken do not provide reliable information.

How does one decide if the available studies are reliable? There are two alternatives:

- Case 1: The studies are inconclusive because the data available are insufficient to reach a reliable conclusion.
- Case 2: The evaluation studies may be flawed.

The second case presents a challenge in that it requires a certain amount of statistical expertise and experience to undertake a critical assessment of an evaluation study. Such studies, or the knowledge obtained from studies, may be flawed for a variety of reasons, including the following:

- Not accounting for selection bias (regression-to-mean), or doing so improperly;
- Failure to separate the safety effects of other changes (e.g., traffic volumes, other countermeasures, crash reporting);
- Use of comparison groups that are unsuitable due to differences in population, location, or weather conditions;
- Incorrect interpretation of accuracy of estimates or presentation of results without statements of accuracy;
- Incorrect interpretation of the results of cross-sectional studies where differences between two groups may be due to factors other than the measure of interest;
- Publication bias (the tendency to publish only favorable results); and
- Selective citing of results (the tendency to ignore negative aspects of results such as declining effects over time or increases in nontarget crashes).

Specific guidance on the conduct of a critical assessment of evaluation studies is beyond the scope of this document, since



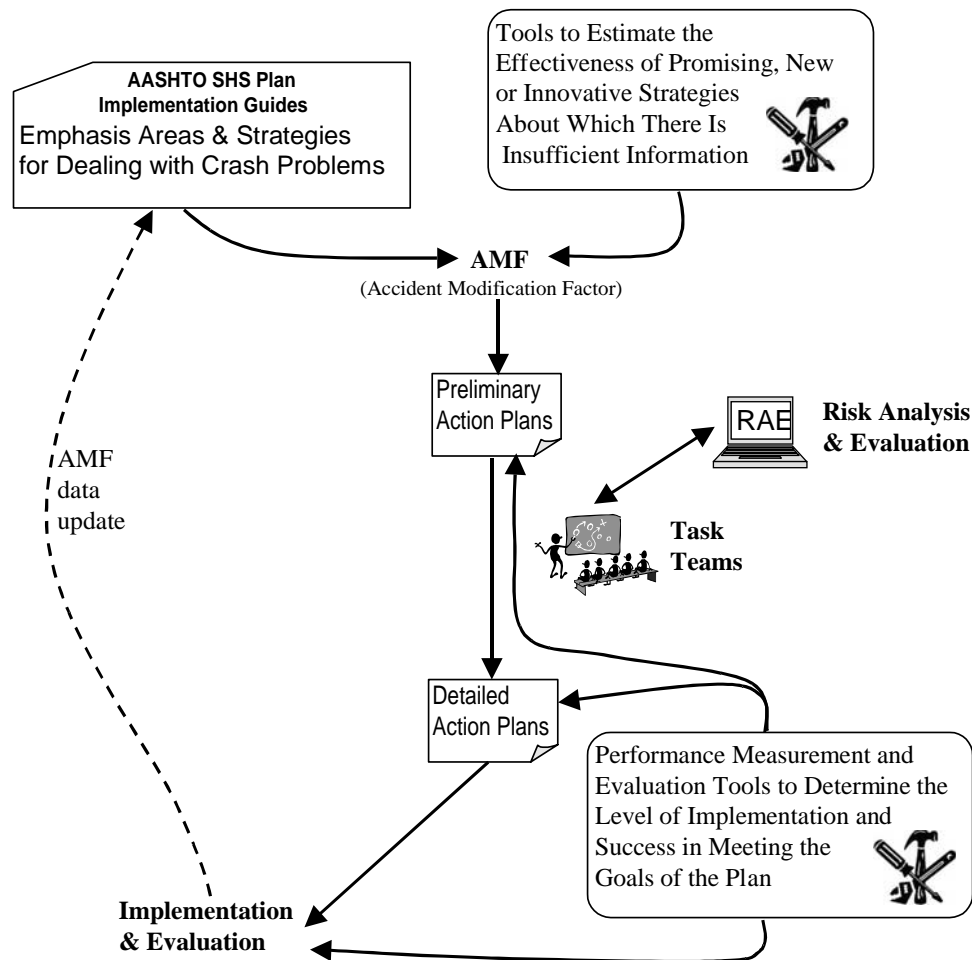


Figure D-14. Relationship between the estimation tools and measurement tools in the ISMProcess.

no amount of documentation can substitute for the required training and experience. Good examples of critical assessments can be found at [www.roadsafetyresearch.com](http://www.roadsafetyresearch.com). This website currently has systematic reviews for several measures conducted to provide information on AMFs for the FHWA's Interactive Highway Safety Design Model.

#### D2.1.2 Deciding on the Size and Scope of an Evaluation: Decision Analytic Approach

Just as strategies must be prioritized, potential research projects to estimate the safety effects of strategies can be prioritized, and in fact justified, based on an analysis of expected costs and benefits of the evaluation study. Hauer et al. (13) described a decision analytic approach to this problem and applied it to designing a pilot study to estimate the safety effects of daytime running lights. In Hauer's proposed study, a randomly selected fleet of vehicles would be selected, and for each pair of vehicles, one would be equipped with daytime running lights. Crash data of the treatment and control vehicles would then be compared after 1 year of driving. The

decision analytic approach was used to determine how many treatment vehicles should be considered.

The conceptual background for the approach is as follows. A strategy can be erroneously administered in two ways:

- By implementing measures when the costs exceed the benefits or
- By not implementing measures that would have positive benefits in relation to costs.

Larger studies cost more, and, at some point, diminishing returns are achieved from larger study budgets. That is, the increase in the value of information per unit of increase in cost reaches a limiting value where investment in other studies or programs would provide a higher return.

The complex equations developed will not be shown here, but the general concept is described to demonstrate that the method is well founded in statistical theory. The starting point is to determine what is presently known about the strategy of interest—for example, Hauer's proposed study of daytime running lights. Sixteen estimates of the AMF for daytime running

lights were obtained from an international team of experts, with a mean value of 0.951 and a standard deviation of 0.03. These parameters were used to fit a gamma cumulative probability function to the data. If a study is conducted, the results of the study add to the knowledge about the AMF and modify the probability distribution for the AMF, which will now describe the probability of the AMF given the study results.

The value of a study will be positive if this new knowledge reverses an erroneous implementation decision and negative if the decision stays the same. For all possible outcomes of study results, the value of the study can be calculated as the difference between the costs of implementing the measure and the expected crash savings, calculated using the total number of target crashes and the expected value of the AMF given the study results. The expected value of the strategy can then be calculated by the summation of all study values given the study results multiplied by the probability of observing the study results.

The probability function for the observed study results is a function of the probability of a target crash to occur in the control group integrated over all possible values of the AMF. In a study by Hauer et al. (13), the results indicated that although daytime running lights were expected to reduce crashes, the costs of implementation were greater than the expected benefits, and the expected value of information gained from a study was less than the study costs.

### D2.1.3 Procedure for Estimating AMFs

In order to determine which strategy to select for implementation, either on a jurisdictionwide scale or on a smaller pilot study scale, the strategy's AMF or function must be used in a cost-benefit analysis to compare the strategy in question with alternatives. The AMF is the ratio of crash frequency expected after a strategy is implemented to the expected crash frequency had the strategy not been implemented. However, for many strategies, even those that may have been implemented in the jurisdiction for many years, the necessary data for calculating the AMF may be unreliable, missing, unavailable, or simply never collected in the first place. Innovative strategies are even more likely to lack the necessary data given that the strategies may never have been implemented.

The procedure for estimating an AMF, depicted in Figure D-15, can involve a variety of different methodologies, including quantitative analysis, systematic reviews, meta-analysis, Bayesian analysis, and subjective assessment. The procedure is iterative, with AMF estimates for new or unevaluated strategies first being developed through a review of external data sources such as journals, publications, and other jurisdictions or developed by subjective assessment using expert focus groups. Estimates of the AMF should be updated as new data are collected from either pilot or full-scale implementations of the strategy.

Determining an AMF involves decision making by the RAE group starting with deciding whether there are sufficient data available for calculating the AMF (Figure D-15). Subsequently, the RAE group determines if the available data are reliable. Unfortunately, there is no guarantee that the data from a previous evaluation will necessarily be useful due to problems with reliability, validity, or the incorrect type of data. The less that is known about a safety strategy, the more important the strategy's evaluation will be both within the ISMS system and to the implementing agencies.

## D2.2 FORMAL EVALUATION STUDIES

Formal evaluation studies may be experimental or quasi-experimental. In experimental studies, the investigator has ultimate control over the assignment of entities (drivers, vehicles, road locations, etc.) to strategies or control conditions. In quasi-experimental studies, the investigator does *not* have control over assignment for many or all factors of interest, and instead the assignment must be used "as-is," in which case the investigator assumes a more passive role, merely observing and attempting to understand the variety of different factors at play in the study. This difference in control means that sometimes the investigator cannot separate two factors of interest that produce a single confounded measure where the individual contribution of each factor cannot be determined. For example, a study in which students are selected to enroll in a driver education course would not be effective because it would probably confound the effect of the driver course with the effect of the students' ages, since most people who enroll in driver education courses are young prospective drivers. This sub-appendix includes a discussion of experimental studies and quasi-experimental studies. Included in the quasi-experimental studies are cohort studies, case control studies, before-after comparisons, and time-series analysis.

Different study designs each have their strengths and weaknesses, which should be recognized by those undertaking studies of risk and safety evaluations. The intent here is not to take issue with designs used by different jurisdictions but rather to provide some documentation of "best practices" in this regard. Methods for the development of "best practices" in clinical medicine and public health have been ongoing for a number of years. The best known and best established is the Cochrane Collaboration (36). The Cochrane Group has established a formal hierarchy of designs for the evaluation of the efficacy of strategies in public health and preventive medicine. The Cochrane Collaboration has a group dedicated to the review of studies related to injury. These studies include reviews of research on the effectiveness of the following:

- Helmets for preventing head and facial injuries in bicyclists,
- Traffic calming for preventing traffic-related deaths and injuries, and

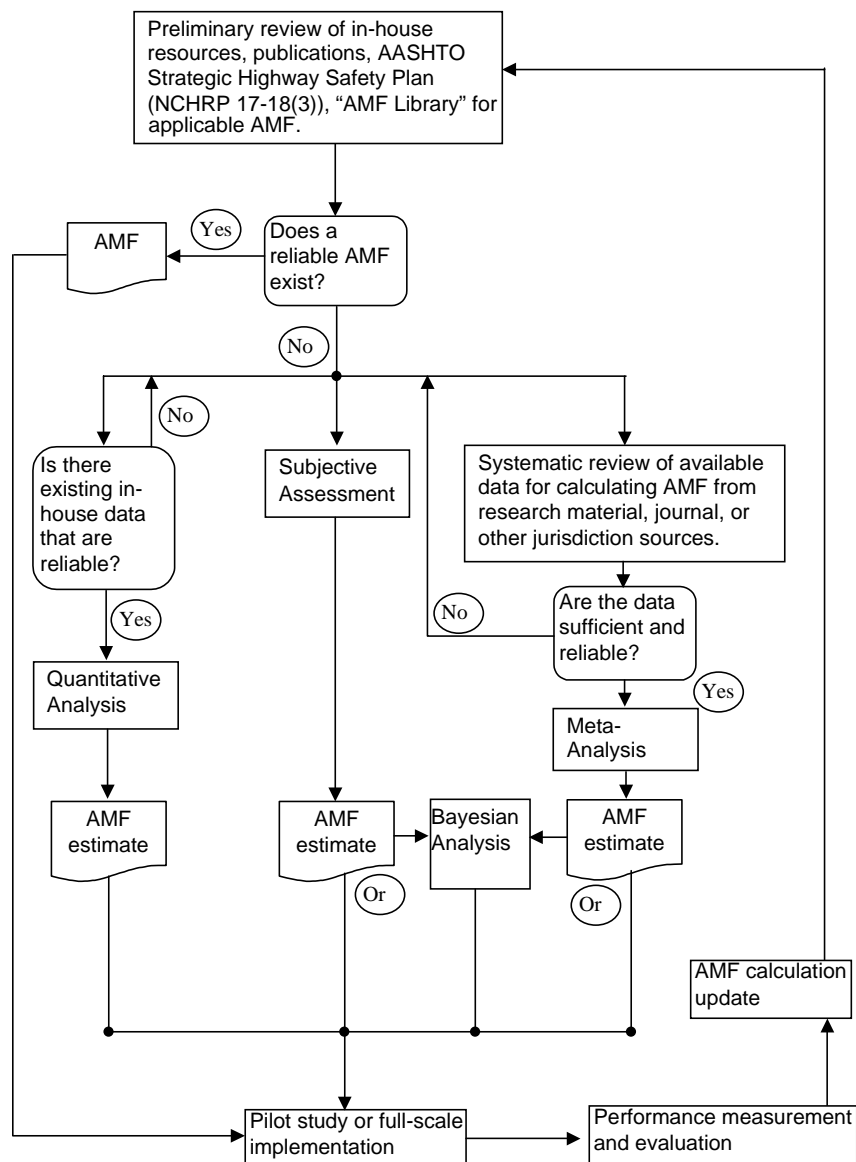


Figure D-15. Procedure for determining the AMF.

- Graduated licensing for reducing motor vehicle crashes among young drivers and others.

Some of these reviews have been completed and are available. Others are still in progress, but subscribers to the Cochrane Library can examine the review protocol. The Injuries Group has its own website at <http://www.Cochrane-injuries.Ishtm.ac.uk>.

### D2.2.1 Experimental Studies

In experimental studies, entities (drivers, vehicles, or road locations) are first selected according to a clear set of inclusion and exclusion criteria. Eligible entities are randomly assigned to one of several treatment or exposure groups (which may, in

some cases, include a nontreatment group). Examples include (1) drivers receiving either formal driver training or no formal driver training and (2) intersections with or without red light cameras. This assignment to groups ensures that chance alone determines which treatment (typically one new measure and one standard or control measure) each entity receives. Entities are treated and followed to see if there are differential safety effects that indicate that the countermeasure is better (or worse) than the control condition. Because of the random allocation of entities to groups, it is assumed that statistically significant differences in safety performance between the groups can be attributed to differences in treatment.

Experiments in the laboratory, such as testing the materials and effectiveness of seat belts, motorcycle helmets, or vehicle door frames, are relatively easy to conduct from a statisti-

cal perspective and have been very important in the setting of standards for equipment on which traffic safety programs rely. On the other hand, the quantification of the change in safety for improvements at road locations can be very difficult to determine, since sufficient data are usually not available to detect typically small changes. This is obviously a problem when an entirely new safety improvement, or one for which there is insufficient information, is being considered. Some researchers have attempted to circumvent the need for crash data by conducting field experiments and observing noncrash measures. As a proxy, what is measured is the behavior of drivers (e.g., conflict situations, speeding, red light running) that are believed to be related to crash risk. The difficulty with such studies is that they usually fall short of establishing the necessary statistical association between unsafe behavior and crash frequency.

Another difficulty with experimental studies is that in experimental traffic safety research involving humans, ethical issues may arise that are often analogous to those that arise in clinical medicine. Participants must consent voluntarily after being given a clear description of the risks and benefits of their participation. The process of obtaining informed consent will vary by circumstance and may be uncertain or impossible in some circumstances.

It is possible to conduct simple experiments on specific aspects of a safety program, such as features of the driver education component of graduated licensing, or the issuing of warning announcements for the location of red light cameras. It is usually more difficult to conduct an experiment that tests the overall efficacy of a program, since all individuals in a jurisdiction are exposed to the program and a control group in the same jurisdiction is not available.

It is very difficult to conduct experiments to assess the difference in risk of road crashes. This is because the risk is often not susceptible to randomization, such as the risk due to the age or gender of the driver. Even if random assignment is technically feasible, it may be ethically difficult to justify arbitrarily assigning people to the group that one expects to be at increased risk of a harmful outcome. The next section discusses quasi-experimental designs for studies, where the investigator is unable to assign entities to a group.

## **D2.2.2 Quasi-Experimental Designs and Observational Studies**

This section discusses four types of quasi-experimental designs (also referred to as observational studies [Hauer, 37]): cohort studies, case control studies, before-after comparisons, and time-series analysis.

### *D2.2.2.1 Cohort Studies*

In the evaluation of strategies, it is sometimes possible to identify two otherwise comparable groups, or cohorts, that dif-

fer by some variable of interest (e.g., exposure to a particular program or engineering feature) and follow them over time to assess differences in consequences (e.g., their crash record). Although the study entities have not been formally randomized as they would be in an experiment, they are often assumed to be alike enough to be compared after controlling for measured confounding variables. One can use administrative databases of traffic convictions or crash records (for example) to obtain direct estimates of risk of conviction or crash.

Case Studies 1 and 2 in Section D2.2.3 are examples of cohort studies.

The lack of random assignment means that preexisting differences between the cohort groups may be at least partly responsible for the results observed. However, cohort study designs are stronger than case control studies (described in the next section) because it is easier to ensure that the groups are defined and selected independently of the outcome that is of interest.

The cohort study design is among the most common designs used in the evaluation of strategies targeted at drivers. Even studies conducted many years ago have often employed sophisticated statistical analyses. Though not an example of an evaluation study, a good example of a cohort study design is provided by Brezina (38). Brezina examined a sample of approximately 49,000 Ontario drivers over a 39-month period in the late 1960s. He examined patterns of crash and conviction occurrence associated with gender, age, degree of experience, and size of community, using life table analyses for first crash, multiple regression to predict crashes, and measures of association between convictions and crashes. Brezina's study is a good example of an early attempt to identify the characteristics of high-risk drivers in a jurisdiction.

The cohort study design is also commonly used in evaluations of safety measures targeted at road locations. In this context, the evaluations are also referred to as "cross-sectional" or "with and without" studies. They are thought to be particularly useful for situations where information is insufficient to conduct the preferred observational before-after study. As is the case for studies involving drivers, it is usually likely that differences in crash experience between two cohort groups could be partly due to factors other than the measure being evaluated—factors that could not be controlled for. For example, Sebastian (39) examined the crash rates of signalized intersections in Wisconsin with various types of left-turn treatments and concluded that fully protected left turns are the safest and that protected/permissive phasing is less safe than permissive only. The last result is counterintuitive and is likely due to differences in characteristics other than left-turn treatments between the groups of intersections. This finding shows the difficulty in interpreting results of cohort studies.

Differences in traffic volume and other characteristics can be controlled for through the use of a regression model in which crashes are related to a variety of characteristics, including exposure. The safety effect of making a change in one or more variables can then be estimated using the regression

model to calculate the resulting change in crashes. For example, Council and Stewart (40) evaluated the safety effects of converting rural two-lane roadways to four-lane roadways based on regression equations relating crashes to the annual average daily traffic (AADT) for roads with these two types of cross sections. The estimate of safety effects so obtained can be tricky to interpret, because it may not be possible to control statistically for all possible confounders, such as speeds, sight distance, and geometry. This difficulty is the reason why these studies may lead to counterintuitive conclusions (e.g., left-turn lanes or illumination are bad for safety). Council and Stewart found that, contrary to conventional wisdom, conversion from two-lane to four-lane undivided highway could increase crashes.

The studies to examine the effectiveness of graduated licensing programs (GLPs) include a number of cohort studies. For example, a cohort of newly licensed drivers before a GLP was introduced may be compared with a cohort of newly licensed drivers after implementation of the GLP. Cohort study methodology has been used to evaluate GLPs in Ontario (41), Florida (42) and Nova Scotia (43). Many, if not all, of these studies have relied exclusively on administrative databases to identify suitable new drivers (i.e., drivers eligible for the GLP in their jurisdiction, whether subject to it or not) and examined crash frequencies as recorded in the driver records for a defined period of time.

#### *D2.2.2.2 Case Control Studies*

Cohort studies are not always feasible for evaluation. Often groups exposed to different risks of crash or injury are not well defined, or the outcome of interest is so rare that large numbers must be followed for extended periods to accumulate a reasonable number of outcome events. In these circumstances, the case control design may be advocated as an effective alternative. In general, case control studies require much smaller sample sizes than other designs; this difference is important when data must be collected for each individual involved in the study, such as the amount and type of driving done by each individual. The difference is also a useful feature when designing pilot studies for measures for which there was insufficient information for full-scale implementation.

The case control study design identifies subjects who have the outcome of interest, such as crashes (cases), and compares them with a group of people who have not experienced this outcome (controls) with respect to factors that might change the risk of the outcome. For example, one could look at intersections with crashes involving elderly pedestrians and compare the characteristics of these intersections with characteristics for intersections that have not experienced such crashes. Analyses might find a number of differences between cases and controls, so a simple cause-and-effect relationship that implicates one independent variable is not clear. Statistical analyses that are able to examine the association with one factor while controlling for other factors, such

as multiple logistic regression, are commonly used to clarify these relationships.

Haddon et al. (44) provide a good example of a case control study. Haddon et al. carried out a case control study in New York City to examine factors that might be associated with fatal injuries to adult pedestrians. The cases involved adults hit by motor vehicles at a known site and time in Manhattan who died of their injuries and were autopsied. For each case, a team of investigators returned to the site of the crash at the same time of day and day of week to recruit four pedestrians of the same gender as the case. Sites varied from outside Grand Central Station in the evening rush hour to a street in Harlem at 2:00 a.m. Each control pedestrian was asked to provide information that had been obtained for the case in the investigation of the fatality (e.g., age, race, place of residence, travelling alone or in company, origin and destination of the trip). At the end of the interview, a breath sample was collected to assess blood alcohol content (BAC) to compare with BAC obtained at the autopsy from each case. For this study, matching on gender and site was not sufficient. The controls were much younger than the cases, and age was likely to confound other factors. The statistical analyses available to Haddon et al. when these data were collected are very different from what is possible today; nevertheless, a strong association with alcohol, after stratifying for age in the study's comparisons, was demonstrated.

Another case control example is by Jones and Stein (45), who studied the relationship between fatigue and tractor-trailer crashes. For each crash, three control tractor-trailer drivers passing the crash site at the same time of day and day of week were selected. Based on data from crash-involved and non-crash-involved drivers, those who had been on the road for more than 8 hours were found to have double the crash risk of those who had been driving for less than 8 hours.

Neither of these studies of long-distance truck drivers or Manhattan pedestrians would have been possible as cohort studies. Nor would they be ethical to study as an experiment. The exposure (driving longer than 8 hours or pedestrian alcohol impairment) could not easily be used to identify individuals for a follow-up study, and the outcomes (truck crashes and pedestrian fatalities) are rare, so large numbers of exposed and control groups would have to be followed to generate a reasonable number of events. Only after the events have occurred is it possible to identify individuals to take part in such an investigation.

Despite their utility when the outcome of interest, such as a fatal crash, is rare, case control studies are not as heavily used in traffic safety research as might be expected. Unlike some cohort studies, case control studies rarely rely solely on the data already available in administrative databases. Such databases may provide part of the data needed, but case control studies usually involve additional data collection from the cases and controls. For example, it has been suggested that this study design could be very useful in some aspects of the evaluation of GLPs, especially when factors that are not part

of what is routinely available on all drivers, such as a driver's employment status or annual distance driven, are examined.

#### *D2.2.2.3 Before-After Comparisons*

There are two distinct methodologies for conducting before-after comparisons: conventional before-after comparisons and the empirical Bayes procedure. The state of research in before-after evaluation methodology is well covered in a recent book by Hauer (37). The book identifies the special problems created by the peculiarities of crash and related data and presents the latest methods for accommodating these problems in the proper conduct of observational before-after studies. Fundamental to the concepts presented is a recognition that some or all of the observed changes in safety following a treatment can be due to factors other than the treatment and need to be separated from the treatment effect. These factors include traffic volume changes, seasonal trends in crash occurrence, and random fluctuation in crash counts. Three examples of before-after studies are presented in Case Studies 3, 4, and 5 in Section D2.2.3.

The conventional before-after study, which examines the same entities before and after some treatment and observes the change in crash frequency, is often used for the evaluation of safety measures targeted at road locations. For example, Griffith (46) used two approaches to evaluate shoulder rumble strips installed on freeways. These approaches were a before and after evaluation with yoked comparisons and a before and after evaluation with a comparison group. Griffith considered the empirical Bayes approach to evaluate shoulder rumble strips, but assumed that there was "no selection bias of treatment sites based on crash history" (i.e., regression-to-mean). The basis for this assumption was a statistical examination of the similarity between the crash experience of the comparison and treatment groups before the rumble strips were installed. The accuracy of the conventional before-after study relies on the validity of the assumption of similarity between the treatment and comparison sites. The validity of this assumption is often questionable, especially in the case of observational studies to evaluate measures applied to road locations.

The empirical Bayes approach overcomes the difficulties in conventional before-after studies by controlling for regression-to-mean and facilitating the proper accommodation of traffic volume or exposure changes and time trends in crash experience in a jurisdiction. By not relying on the use of a comparison group to accomplish these manipulations, the empirical Bayes method overcomes a major difficulty with conventional observational before-after studies. The objective, as it is in the conventional before and after comparison, is to estimate the number of crashes that would have been expected in the after period had there been no treatment. The treatment effect is the difference between this estimate and the number of crashes actually recorded after treatment.

The number of crashes that would have been expected in the after period had there been no treatment is a weighted average of information from two sources:

- The number of crashes in the before period and
- The number of crashes expected for entities with similar exposure (e.g., traffic) and physical characteristics.

The latter estimate is based on a reference population of similar entities. Where sufficient data are available, the crash experience of the reference group is used directly. However, traits that make entities similar in safety tend to be many and continuous in nature. Therefore, identifying a reference population of similar entities is usually problematic. It is becoming commonplace to instead use a multivariate model, or safety performance function, that relates crash experience to traffic and to physical characteristics of the population of entities from which the treatment entity is drawn. There are two drawbacks of the multivariate model: suitable reference population data for calibrating the models are rare in practice, and the task of calibrating a multivariate model can be daunting, even for those with substantial statistical knowledge. To overcome these drawbacks, some analysts seek to adapt models developed by others for reference populations similar to those of interest. To this end, considerable research is underway to develop a comprehensive suite of models for a variety of reference populations.

Persaud et al. (47) recently applied the empirical Bayes method to evaluate the conversion of conventional intersections to roundabouts.

#### *D2.2.2.4 Time-Series Analysis*

Time-series analysis is often used routinely on data of traffic crashes over time to assess the effects of changes such as seat belt legislation, an increase in enforcement for speeding violations, or a DUI checkpoint program. Changes observed in the frequency of traffic crashes may be due to legislation or the enforcement of a program, but they may also be due to other changes occurring concurrently over which the investigator has no control. Sometimes these other changes can be measured and controlled in the analysis, but often they cannot.

Time-series analysis typically relies exclusively on administrative data when used in traffic safety research. Many investigators will conduct a time-series analysis to seek support for a hypothesis about a new and promising measure before embarking on the more expensive and difficult efforts to collect data to confirm the hypothesis. Case Studies 6 and 7 in Section D2.2.3 are examples of time-series analysis.

According to McBurney (48), an interrupted time-series design is a "research design that allows the same group to be compared over time by considering the trend of the data before-and-after experimental manipulation." The strategy (or manipulation) interrupts sequential measurements of the per-

formance measure (dependent variable). Differences in the data series, before and after the strategy, are examined to determine whether the strategy caused the dependent variable to change.

The interrupted time-series approach requires a relatively long series of equally spaced observations both before and after the implementation of a strategy. Monthly numbers of traffic crashes are often used for this purpose. The analysis involves examining the data statistically for the presence of long-term trends and cycles in the series and determining whether these trends and cycles changed when the strategy was introduced.

For example, Voas et al. (49) used an interrupted time-series study design to determine the effectiveness of 0.08 BAC law enacted in 1997 in Illinois. The states surrounding Illinois that still had a 0.10 BAC law were used as a comparison group. Scopatz (50) and Tarko et al. (51) reported that this type of comparison between states is not always appropriate and that ideally the comparison group should be drawn from the same jurisdiction as the treatment group. Case Study 7 in Section D2.2.3.7 contains more details of the Voas BAC study.

### D2.2.3 Formal Evaluation Case Study Examples

#### D2.2.3.1 Case Study 1: Cohort Study of a Driver Program

In a study by Marsh (52), over 17,000 drivers who had reached a certain number of demerit points were randomly assigned to one of six treatment programs or to the control group. Four treatments involved variations of a group educational meeting. Two treatments involved nonclassroom strategies, one of which was an incentive in return for homework and maintenance of a clean driving record. One of the educational meeting treatments proved more effective than the others in that it was associated with a significantly lower rate of crashes in the first 6 months compared with the control group. For various practical reasons, planned experiments of this type, in which chance alone dictates which subjects or entities get which treatment, have been rare in the evaluation of the strategies.

#### D2.2.3.2 Case Study 2: Cohort Study of Another Driver Program

Ulmer et al. (42) conducted a study of Florida's GLP for 15–17 year olds. The program placed restrictions on unsupervised nighttime driving that varied by age, penalties for traffic violations (resulting in an extension of the time until full licensure), and a zero tolerance policy toward driving with a BAC of 0.02%. The study compared per-capita crash rates of 15, 16, 17 and 18 year olds in 1997, the first full year of graduated licensing, with crash rates in 1995, the last full year without graduated licensing. Florida crash data were

also compared with data from Alabama, where full licensure is available at age 16, for the same years.

Only fatal and injury crashes were analyzed due to the variance in crash reporting between states and over time. Crash rates based on the number of fatal crash or injury crash drivers per 10,000 people were calculated for each age group. Crash rate ratios were then calculated by dividing the crash rate of each group by the crash rate of the reference group, which was defined as 25- to 54-year-old drivers for the same period. This standardization of crash rates allowed for comparisons across states. Differences in crash ratios before and after the GLP were tested using the standard test for normally distributed variables:

$$z = (R_a - R_b) / (V_a + V_b)^{0.5}$$

where

$R_a$  and  $R_b$  = crash rate ratios in the after and before periods and

$V_a$  and  $V_b$  = variances of the after and before crash ratios.

Crash rate ratios for 15, 16, and 17 year olds experienced statistically significant declines, while the 18-year-old group experienced no significant change. No significant changes were found for any age group in Alabama. Overall, the crash reduction for 15–17 year olds was 9% (19% reduction for 15 year olds, 11% reduction for 16 year olds and 7% reduction for 17 year olds). At the same time, learner license applications rose by 29% and 14% for 15 and 16 year olds, respectively, indicating that reductions in crash rates were not due to fewer people holding licenses.

#### D2.2.3.3 Case Study 3: Before-After Study with Random Assignment—Conventional Evaluation of Crash Surrogates

In field evaluations by Miller (53) and Retting et al. (54), two methods for restricting right-turn-on-red (RTOR) were evaluated:

- Traffic signs restricting RTOR at specified times and
- Traffic signs restricting RTOR when pedestrians are present.

Fifteen signalized intersections were randomly assigned to three groups: control; no RTOR from 7 a.m. to 7 p.m., Monday to Friday; and no RTOR when pedestrians are present. At each restricted location, a sign was placed near the stop line on the right side of the street and a second mounted overhead next to the traffic signal. Each site was observed for 10 hours before and 10 hours after sign installation during the same time of day and day of week, and no observations were made during poor weather conditions. Data for the after period were collected several weeks after the installation of the signs.

Each vehicle that arrived at the stop line unimpeded by other vehicles was observed and recorded if

- The vehicle stopped at the marked stop line,
- The vehicle turned right on red,
- The vehicles that turned right on red did so without coming to a full stop, or
- The pedestrians yielded to the RTOR vehicle.

Significance testing was applied to the percentages of observations for the four measures at the control and test sites for the before and after periods. The testing was applied separately when pedestrians were present, when pedestrians were not present, and combined. It was concluded that signs prohibiting RTOR during specific hours were very effective at increasing driver compliance with stop lines and reducing the number of pedestrians yielding to RTOR vehicles. Signs prohibiting RTOR when pedestrians are present were not deemed very effective.

#### *D2.2.3.4 Case Study 4: Before-After Study with Random Assignment—Conventional Crash Evaluation*

For an ongoing study to evaluate the safety effects of pavement markings at hazardous horizontal curves (54), many rural curve sites with prior crash histories were identified in Pennsylvania. Data were assembled for each PennDOT district participating in the study. Crash history indicated that 200 control and 200 treatment sites were needed for the study. Based on this sample size requirement and following engineering field inspections, 400 eligible sites were randomized into treatment and control groups, by district. Pavement markings were installed at treatment sites in the fall of 2001. Crash data will be monitored at treatment and control sites over time.

#### *D2.2.3.5 Case Study 5: Before-After Study with Random Assignment—Empirical Bayes Procedure*

U.S. experience with modern roundabouts is rather limited, but in recent years there has been growing interest in their potential benefits and a relatively large increase in roundabout construction (55). This interest created a need for the evaluation of the safety of U.S. roundabouts as a promising new measure. Although there were several evaluations in other countries, the available information was deemed to be insufficient, since it was unclear whether the experience in other countries was transferable to the United States.

A paper by Persaud et al. (55) evaluated changes in motor vehicle crashes following conversion of 23 intersections from stop sign and traffic signal control to modern roundabouts. The settings, located in seven states, were a mix of urban, suburban, and rural environments. The urban sample

consisted of both single-lane and multilane designs, and the rural sample consisted of only single-lane designs. A before-after study was conducted using the empirical Bayes procedure, which accounts for regression-to-mean and the traffic volume changes that usually accompany conversion of intersections to roundabouts. Overall, for the 23 intersections combined, this procedure estimated highly significant reductions of 40% for all crash severities combined and 80% for all injury crashes (AMFs of 0.60 and 0.20, respectively). Reductions in the numbers of fatal and incapacitating injury crashes were estimated to be about 90% (AMD of 0.10). The results were said to be consistent with numerous international studies and to suggest that roundabout installation should be strongly promoted as an effective safety treatment for intersections. Since the empirical Bayes approach is relatively new in safety analysis, a secondary objective of the paper was to demonstrate the potential of this methodology in the evaluation of safety measures.

#### *D2.2.3.6 Case Study 6: Time-Series Analysis of a Photo Radar Program*

Chen et al. (30) evaluated the effect of mobile photo radar on vehicle speeds at implementation sites and nonimplementation sites and on crashes and injuries. Photo radar units were deployed in high-crash locations and locations requested by local communities where there was a perceived speeding problem. Speed data at implementation sites were collected from the photo radar units. Speed data at a number of nonimplementation sites were collected by induction loop detectors embedded in the pavement. Sites for induction loops were selected to ensure free-flow traffic conditions. The change in vehicle speed at the implementation sites was examined using a before-and-after comparison of the percentage of vehicles exceeding the speed limit.

A 50% reduction in the percentage of speeding vehicles was found within 7 months of introducing photo radar. A reduction of 75% was found for vehicles traveling at least 16 km/h (10 mph) over the speed limit in the same period. The speed data at the nonimplementation sites were then examined for general changes in speeds associated with the introduction of photo radar. A similar reduction in the percentage of vehicles that were exceeding the speed limit was found but on a smaller magnitude.

A pooled cross-sectional time-series analysis was conducted on the mean speeds of the nonimplementation sites. A model of one-way cross-sectional effect was estimated, with the different nonimplementation sites serving as the cross sections and the strategy of the photo radar program as a dummy variable. Simply put, the model predicts the mean speed dependent on the characteristic of interest: the site, and the before- or after-photo-radar period. The effect of photo radar on mean speeds at these locations was estimated from the model to be 2.4 km/h (1.5 mph).



Interrupted time-series analysis was used to study the effect of the photo radar program on the number of monthly traffic crash victims carried by ambulance- and police-reported fatalities. Only crashes involving an unsafe speed of travel as reported by the investigating police officer were included. Nighttime crashes were also not included because an overlapping program aimed at drinking and driving occurred that would affect crashes in that period.

This type of model is referred to as an ARIMA model (AutoRegressive Integrated Moving Average model) and introduces strategies, in this case photo radar implementation, as step functions in the models. This model can control for crash trends over time and seasonality effects. Gasoline sales were also included in the model as a surrogate for fluctuations in the amount of driving over time. Daytime unsafe speeds and injuries were found to have decreased following the introduction of the photo radar program.

#### *D2.2.3.7 Case Study 7: Time-Series Analysis of BAC Intervention*

Voas et al. (49) examined the effectiveness of the 0.08% BAC law enacted in 1997 in Illinois regarding the number of drinking drivers in crashes. The previous legal limit was 0.10%. Data from the Fatality Analysis Reporting System (FARS) for Illinois were used and also compared with data of surrounding states. The ratio of drinking drivers to non-drinking drivers was analyzed using a time-series analysis. Nondrinking drivers were included in the analysis to account for general changes in miles driven, the relative safety of the roadway and vehicle, and other factors. ARIMA intervention models were developed with the data to compare the monthly number of alcohol-positive (BAC > 0.01%) drivers involved in fatal crashes for the years 1988 to 1998. A total of 114 pre-0.08 BAC law months were compared with 18 post-0.08 BAC law months. This interim report concluded that the number of drivers in fatal crashes with a positive BAC decreased by 13.7%, while surrounding states with a 0.10 law experienced an increase of 2.5%.

A similar evaluation of a new BAC law points to a problem with these studies: the difficulty of identifying a suitable comparison group that is similar to the treatment group in all possible factors that could influence safety. A recent paper by Scopatz (50) points to the difficulties of fulfilling this need by examining the results from Hingson et al. (56). Hingson et al. found that lowering legal BAC limits to 0.08% resulted in a 16% reduction in the probability that a fatally injured driver would have a BAC above 0.08%. The treatment group consisted of states that passed a lower legal BAC law, while the comparison states retained a 0.10% BAC legal limit. Scopatz (50) points out that there are numerous differences other than legal BAC limits between 0.08-BAC law and 0.10-BAC comparison states. Therefore, it is impossible to conclude that the passage of a law was the sole source of

the result and not some other uncontrolled-for factor. To support this point, Scopatz (50) showed that if logically valid but different comparison states are chosen, the results change dramatically and in most cases are consistent with a conclusion of “no effect.” Tarko et al. (51) found that a similar situation would arise even in analysis confined to a single state, in which the treatment group is in one county and the comparison group is in another.

Ideally, the comparison group should be drawn from the same jurisdiction as the treatment group. The difficulty is that the pool available for the comparison group could be too small if most or all elements are treated or affected by the treatment. In the above BAC cases, the law applied to all drivers in a state. As discussed earlier, measures such as red light cameras are believed to have significant spill-over effects to untreated sites.

### **D2.3 ALTERNATIVES TO FORMAL EVALUATION STUDIES**

Formal evaluation studies often require considerable resources—data, time, money, and skilled analysts. One or more of these resources are often so constrained that a formal evaluation study cannot be undertaken. In the case of new and promising initiatives for which there is no readily available knowledge representing some sort of consensus about its effects, the choices are as follows:

- Abandon consideration of the initiative.
- Consider alternatives to formal evaluations.

Assuming that the first choice is undesirable, it is of interest to consider alternatives to formal evaluations. The alternatives discussed in this section range from purely subjective (expert opinion) to more quantitative methods (Bayesian analysis, meta-analysis).

#### **D2.3.1 Bayesian Analysis Studies**

Bayesian analysis attempts to combine information from previous studies with information on local conditions to arrive at an estimate of safety effectiveness specific to the area of interest.

A Bayesian analysis study requires a prior estimate of an AMF that is then transformed into a more reliable posterior estimate of an AMF by combining the prior estimate with information on local conditions and experiences. Spiegelhalter et al. (57) describe different approaches for obtaining prior estimates for use in Bayesian analysis studies. Two common approaches are the solicitation of expert opinion and the meta-analysis of a systematic review of available research material. These approaches are discussed later in this appendix. Spiegelhalter et al. note that obtaining prior estimates is often based on judgment and that a certain degree of subjectivity is unavoidable. They recommend that a sensitivity analysis be performed using a range of prior estimates.

Melcher et al. (58) describe the development of a Bayesian analysis framework for refining AMFs from previous studies using local engineering evaluations to develop likelihood functions of AMFs, denoted by  $\theta$ . In this application, five persons knowledgeable in road safety engineering were used to review data for a number of local crashes and rate 12 countermeasures as to their effectiveness for each crash. This rating could include “no effect” if it was believed that a countermeasure would have had no effect on reducing the risk of the crash. By evaluating countermeasures over random crash data, the result was a statistical estimate of  $\theta$  that reflected the local roadway and driver environment. These estimates of  $\theta$  served as the observed local data. By combining this local estimate with knowledge gained through previous studies, a stronger estimate of local effectiveness was obtained. The Markov chain Monte Carlo method, a type of statistical simulation method, was used to obtain an estimate of the most likely value of the AMF. In applying this method for the subjective estimates, the application must be repeatable and developed logically.

### D2.3.2 Systematic Reviews

It is rare for a single study to influence policy development on its own. What is much more common is for a body of work, conducted in a variety of settings, to be brought together to inform policy makers, program developers, legislators, and others involved in the development of traffic safety management. The process of assimilating knowledge from the documented information is referred to as a systematic review, also known as a study of studies. A systematic review is a type of scientific study that tries to answer a clear question by finding and describing all published and, if possible, unpublished work on a topic.

In conducting a systematic review, the strategy for searching for available information should be as thorough as possible within arbitrary limits for language of publication, dates of publication, source of material, and so forth. For example, some reviewers look only at papers published in the peer-reviewed literature while others seek to access technical reports, unpublished material, and so forth in the so-called grey literature (i.e., literature that does not follow a formal publication pattern, such as reports, technical notes and specifications, preprints, translations, and trade literature). Often, several levels of review are used based on established criteria for papers or studies that address the strategy in question. Starting with titles, abstracts, or executive summaries, one uses these criteria to exclude a number of clearly unsuitable papers before obtaining the remainder for more detailed review. This review is based on a formal abstraction document and ideally involves two reviewers, with additional reviewers to resolve disagreements. Studies judged to be of deficient quality (using criteria previously established and agreed to) are excluded from estimates of the effectiveness of the strategy.

Key questions to be asked in assessing study quality have been documented by Elvik (59):

- How were study units sampled? Does the study describe the sampling technique used?
- Do the data collected in the study refer to primary study units or to aggregates of primary study units?
- Was crash or injury severity specified?
- Were study results tested for statistical significance or their statistical uncertainty otherwise estimated?
- Did the study use appropriate techniques for statistical analysis?
- Did the study find a statistical association between the road safety measure and the outcome variables of interest?
- Can the causal direction between treatment and effect be determined?
- How well did the study control for confounding factors?
- Did the study uncover the causal mechanism through which the treatment influenced the outcome variables of interest?
- Did the study have a clearly defined target group, and were effects found in the target group only?
- Are study results explicable in terms of well-established theory?

Averaging the estimates from different studies assumes that the estimates each reflect a common value, the “true” effect. Sometimes the estimates vary so widely that one cannot assume there is a single “true” value for the effect of the strategy. There is much debate about how to handle this situation. The most conservative approach is simply to acknowledge the variability in effect sizes and interpretation.

Often, the results are presented simply as a table, listing all usable studies and their salient characteristics and results. In addition, one may present the size of the effect of the strategy in each study graphically, distributed around some average of the effect size derived from these studies. The simplest form of average is the median, but many systematic reviews that have a quantitative estimate of effect size use a weighted average. The weight used may reflect the precision of the estimates from each study (i.e., an estimate from a study with a large sample size and limited random error would be given a greater weight than one from a smaller study with more “noise” in the data).

#### D2.3.2.1 Systematic Review Examples

**National programs.** Recently, the Centers for Disease Control and Prevention (CDC), NHTSA, the National Association of Governor’s Highway Safety Representatives, and many community organizations concerned with public health and traffic safety conducted a series of systematic reviews of evidence for strategies to reduce deaths and injuries to motor vehicle occupants. This series of reviews is part of a larger

program to develop guidelines for community preventive services.

The methods used have been described in detail for the general program (28). Miller (53) considered the economic efficiency as well as effectiveness for the evaluation of traffic safety strategies. These methods have been described more recently (60). Briefly, the process comprises the following tasks:

1. Develop a conceptual framework that identifies factors that contribute to injuries, strategies that could modify these factors, strategy selection criteria, and outcome performance measures for which evidence is sought. Selected strategies are those thought to be modifiable and expected to have the greatest effect on reducing the burden of injury from traffic crashes. The choice of suitable outcomes requires consideration of the nature of events (crashes, injuries, or usage), how these will be measured (self-reported or observed), who will be affected (children or drivers), and where the strategy will take place.
2. Develop an analytical framework for each strategy that identifies the specific issues that require further investigation, how evidence will be sought, and the criteria for studies to be included in the review.
3. Conduct, for each strategy, a detailed search for studies by consulting with experts and by searching through computerized databases. Each study is assessed for its eligibility based on the criteria formulated during the development of the analytical framework.
4. Evaluate and assess each study for its quality in design and execution. Only studies that are considered good or fair should be considered for evidence of effectiveness.
5. Summarize the results and findings of the studies. Quantitative summaries are often called meta-analyses. In a paper by Zaza et al. (60), the median of all effects was estimated and the degree of variability between effects was presented as a range or an interquartile range, depending on the number of effects considered.
6. Summarize research gaps that become apparent during the review and assessment of research studies.

**Meta-analysis.** Meta-analysis is a statistical technique that combines the independent estimates of safety effectiveness from separate studies into one estimate by weighting each individual estimate according to its variance.

In a study on the safety effects of median barriers, guardrails, and crash cushions, Elvik (61) applied the meta-analysis technique to 32 individual studies. In this application, the “log odds” meta-analysis method was used.

The effect on crash rate of the three safety devices,  $E_i$ , was estimated by the odds ratio, defined as

$$E_i = \frac{ACC(G)/ACC(W)}{ACC(W)/VKT(W)}$$

where

ACC = total number of crashes;

VKT = vehicle kilometers of travel;

G = presence of a median barrier, guardrail, or crash cushion; and

W = lack of a median barrier, guardrail or crash cushion.

The statistical weight for each study result is defined as  $w_i = 1/[1/ACC(G) + 1/ACC(W)]$ .

The estimated mean effect on crashes using all studies is calculated using the log odds method:

$$E = \exp \frac{\sum_{i=1}^n \ln(E_i)(w_i)}{\sum_{i=1}^n w_i}$$

where

$E_i$  = estimated effect of study  $i$  and

$w_i$  = statistical weight assigned to study  $i$ .

Elvik also analyzed the effects on crash severity, defined as the change in the probability of a fatal or injury crash, given that an injury occurred. There are different odds ratios and statistical weight formulas for these effects. The reader is referred to Elvik's study (61) for these formulas.

Elvik's study also tested for publication bias in the studies used. Publication bias occurs when research results are not published, often due to the results being counterintuitive (e.g., an increase or no effect on crashes when a decrease was expected). Publication bias was investigated using a graphical method called the “funnel graph” method. Each study result is plotted on a graph in which the horizontal axis shows each result and the vertical axis shows the statistical weight assigned. If there is no publication bias, a scatter plot of results should resemble an upside down funnel. As sample size increases, the dispersion of estimates should converge, since larger sample sizes should give more accurate results. If the tails of the funnel are not symmetrical, then publication bias may exist. Publication bias was not found to be an issue in this study.

Hagenzieker et al. (62) is an example of a meta-analysis of studies of incentive programs to encourage seat belt use. The effectiveness of incentive programs on short- and long-term seat belt use was sought. Studies were excluded from the analysis if

- Incentive was given not to individuals but to organizations or policymakers,
- Incentives were nonmaterial (e.g., praise),
- Enforcement was also a part of the program,
- No behavioral measures were reported, or
- The article was a review.

A total of 34 articles were included in the analysis. Twenty-three variables related to the study background (e.g., year,

location, and presence or absence of seat belt law), research characteristics (e.g., study design and observation periods), and incentive program characteristics (e.g., duration of strategy period and type of reward) were used to describe the studies. Two authors independently coded each of the studies for the variables and attempted to deduce information that was not explicitly given. Both unweighted and weighted means of effect size were calculated. The study weights were assigned based on the number of observations, as follows:

- <250 observations, assigned weight 1;
- 250–1,000 observations, assigned weight 2;
- 1,000–10,000 observations, assigned weight 3; and
- >10,000 observations, assigned weight 4.

The estimates for short-term effects were 20.6% and 12% increases in seat belt usage for unweighted and weighted estimates, respectively. The estimates for long-term effects were 13.7% and 9.6% increases in seat belt usage for unweighted and weighted estimates, respectively.

The authors also found that the effectiveness of incentive programs varied by the type of population involved, the immediacy of receiving the reward, and the initial rate of seat belt use. Campaigns aimed at elementary schools showed larger effect sizes, as did programs offering immediate rewards and programs implemented in areas with a low rate of seat belt usage.

**Deriving AMFs for use in FHWA’s Interactive Highway Safety Design Model.** Expert panels were assembled at a 2-day workshop to critically evaluate the findings of published and unpublished works related to the safety effects of various geometric design and traffic control elements on two-lane highways. Each panel selected reliable studies and, on this basis, derived AMFs, where possible, for application in FHWA’s Interactive Highway Safety Design Model (IHSDM).

A related effort is a series of reviews conducted by Hauer to also provide information on AMFs for use in FHWA’s IHSDM. At the time of the report’s writing, reviews have been provided for lane width and number of lanes, shoulder width and paving, access, road grade, horizontal curves, and highway medians (available on Hauer’s website, [www.roadsafetyresearch.com](http://www.roadsafetyresearch.com)).

### D2.3.3 Subjective Assessments

It may not be possible to know accurately the effect of a strategy, but it may be possible to estimate what is not the effect. For example, if a new strategy is estimated to result in a 10–20% improvement in safety, then one may take a conservative chance and assume that the strategy will *not* have a 30% effect. Once certain best-case assumptions are made, then cost-benefit analysis may be conducted. It may be known

that a 30% improvement in safety can be achieved by expending the same resources in another area.

Formal evaluations and systematic reviews sometimes cannot provide any clues to the likely safety impact of new and promising measures. One can, with care, resort to more subjective assessments, recognizing the limitations of this type of information. Three potential sources of knowledge can be tapped: expert opinion, “front line” opinion, and public opinion. These three sources can be used individually or in conjunction with each other.

#### D2.3.3.1 Expert Opinion

One can consult a group of experts with knowledge of similar safety initiatives. This group may include high-profile safety champions, group leaders, recognized consultants, and university professors. These experts will have opinions on a variety of subjects, but they may also have access to a library of information that may be relevant to the strategy in question.

#### D2.3.3.2 Front Line Opinion

One can solicit the opinions of safety professionals who are not necessarily experts but who are on the “front line.” These individuals include local traffic engineers, police, and health care professionals involved in safety. They are the ones who are implementing strategies on a daily basis and/or may have knowledge on driver behavior and crash causation and what impact a potential safety initiative may have. The information may not be in quantitative terms, but it may provide a qualitative indication (e.g., is there likely to be reduction, increase, or no effect on crashes).

#### D2.3.3.3 Public Opinion

The public’s attitude toward safety and risk, reckless driving, and observing the rules of the road are related to safety. A strategy that may not have any affect on the road conditions but does in fact influence public attitudes is one worth investigating. Ways of assessing public opinion include mail-, web-, and telephone-based surveys. Statistical aspects of opinion and attitude polls must be taken into account if the data are to be trusted, including procedures for unbiased estimation of whatever characteristics are measured in these polls. In addition, it is important to ask questions of different sources of opinion.

#### D2.3.3.4 Subjective Assessment Example

An example of a subject assessment methodology on a large scale is Wisconsin’s method of ranking emphasis areas in the AASHTO Strategic Highway Safety Plan. Rating 1 is based on the importance of the emphasis area in terms of

improving safety—essentially, the target crashes. Rating 2 is based on the ability to influence the problem—essentially, the subjective perception of the safety impact of potential strategies. These two factors, which were rated by all 160 attendees at a recent conference, were multiplied together to get an overall rating. The attendees included representatives from the American Automobile Association, the Department of Public Instruction, the University of Wisconsin, NHTSA, FHWA, the American Association of Retired Persons, the courts, the media, law enforcement, and the legislature.

## D2.4 SUMMARY

This appendix briefly describes a number of types of studies useful in traffic safety research. The focus is on methods to study the effect of new, innovative, or existing strategies for which there is insufficient information. Some methods require more intensive effort, and while they may use established databases in each jurisdiction, they may also require the collection of data specific to the study. Other methods may be conducted with administrative databases exclusively.

Experiments are the most appealing for many reasons. The logic of a simple experiment is often easiest to understand, so the results and conclusions from such a study are more compelling. On the other hand, experiments have logistical difficulties when used to evaluate population-based safety programs that, by definition, cannot be selectively administered. A more popular design is the cohort study, which can take advantage of the surveillance practices used in many jurisdictions associated with the licensing of drivers, the registration of vehicles, and the documentation of traffic crashes.

Before-after studies are appealing but difficult to interpret because of the lack of random assignment and confounding factors. The empirical Bayes procedure overcomes many of these concerns, but should only be conducted by those with proper knowledge of statistical techniques as applied to such studies. Also popular—and popular for many of the same reasons—is the time-series study, tracking changes in the frequency of events during a series of time intervals and

associating these changes to strategies implemented in the jurisdiction.

Alternatives to formal evaluations such as the systematic review (or meta-analysis) are growing in popularity. These alternatives demand few of the normal resources usually required for traffic safety research, but do require access to a good library and qualified experts and researchers to conduct detailed reviews.

Which designs are feasible? In jurisdictions with the resources, all of these designs will prove useful in different situations. Where conducting a study may not be feasible because of data, time, staff, or budget constraints, then consulting experts, other professionals involved in safety, or the public at large may yield valuable insights into the potential effectiveness of a safety initiative.

Regardless of which methodology is used to determine an AMF, it is important for the Operations Manager and the SPL to realize that the procedure for determining an AMF is iterative. As new data are collected from either a pilot study or a full-scale implementation, additional performance measures should be taken and evaluated. AMFs change over time with different populations and even vary from jurisdiction to jurisdiction. The AMFs must be updated in order to be valid representations of the safety contribution of different strategies.

The process for determining an AMF emphasizes the extreme importance of publishing the results of integrated strategic highway safety plans. Data that have been collected but have not been analyzed, presented, reported, or shared have no value. Whether or not a new strategy has been successful, it is very important to report the results so that other jurisdictions considering the same strategy have additional information to consider. Reporting that a specific countermeasure has not been successful, is just as important as reporting the success of a strategy.

Finally, consider that it may not be possible to know accurately what the effect of a strategy is, but it may be possible to estimate what effect the strategy is *not*. Systematic reviews and the use of subjective assessments can be used to estimate the safety impact of strategies. The limitations of these studies must be recognized when applying the information gathered.

## Appendix D3: Performance Measurement and Evaluation Tools to Determine the Level of Implementation and Success in Meeting the Goals of the ISHSPlan

This appendix presents performance measurement and evaluation tools to determine the level of implementation and success in meeting the goals of a jurisdiction's ISHSPlan. The two types of performance measurement, process and impact, are reviewed in regards to their data collection, evaluation, and reporting requirements. Only after data collection and assessment is it possible to evaluate the performance of strategies and determine any necessary changes to improve the strategies' implementation.

The mechanisms of data collection for performance measures should be designed during the development of the preliminary and detailed action plans (ISMPProcess Steps 3 and 5, respectively). The required data and analytical methods should be determined before any strategies are implemented. The Task Teams, with support as needed from the RAE group, have the responsibility for identifying what, when, and by whom different performance measures will be collected. Task Teams also ensure that the selected performance measures will in fact meet the needs of the ISMSystem.

The Task Teams and RAE groups require more than just this appendix alone to have the complete knowledge necessary for developing evaluation designs, observational studies, and analysis methodology for strategies. This appendix provides sufficient knowledge for *management* to understand the major issues involved in performance measurement. For management (and for Task Team members) who desire additional understanding, references have been provided for further reading throughout the following text.

### D3.1 OVERVIEW OF PERFORMANCE MEASUREMENT

Conducting *performance measurement* means examining, appraising, and judging the worth of a strategy or a combination of strategies by answering the following questions:

- Have the right decisions been made?
- Is the strategy or combination of strategies working?
- Does progress meet expectations?
- Is it costing as much as anticipated?
- If everything had to be done over again, what would remain the same and what would be changed?

According to the Procurement Executives' Association (63), performance measurement allows an agency to assess its emphasis area goals and objectives. Performance measurement reveals which safety strategies are working and which are not and provides information on the strengths and weaknesses of different strategies. Performance measurement can also provide the necessary intelligence that the SPL and

Operations Manager require to identify and initiate steps to improve performance. This intelligence helps to prevent limited resources from being allocated to ineffective and inefficient strategies. It is more desirable to detect low performance early in the implementation of a strategy, while there is still an opportunity to do something about it, than after completion of a strategy. High levels of performance can be useful for media relations and can improve public accountability.

Performance measurement should not be an end in itself, but a tool for more effective management. To make effective use of performance measurement results, the Task Teams, supported by the RAE group, should report information in a manner that can be clearly understood by the SPL and Operations Manager.

A key deliverable of the performance measurement is a report that provides intelligence to decision makers in order to allow them to answer the questions previously listed. More specifically, the performance report should

- Identify which strategies or combinations of strategies are working (impact),
- Determine the effectiveness of each strategy or combination of strategies in reaching the emphasis area goals (impact),
- Identify areas of strengths and weaknesses in each strategy or combination of strategies (impact and process),
- Identify barriers to the successful execution of each strategy or combination of strategies (process),
- Monitor and track the efficiency of resources used during the execution of each strategy or combination of strategies (process),
- Assess the performance of stakeholder agencies (process), and
- Formulate and initiate appropriate steps to correct areas of low performance (impact and process).

A performance report could also serve as a means of informing all stakeholder agencies and their staff of how each strategy is progressing. This will allow stakeholder agencies to assess their own performance against predetermined performance measures and against the performances of other stakeholder agencies. Task Team members, coordinated by the Operations Manager, would develop this performance report during ISMPProcess Step 6. A performance report should also include the information presented in Table D-18.

### D3.2 DATA COLLECTION PROCEDURES

The procedures involved in collecting data to quantify safety are considerably different depending upon whether one is

**TABLE D-18 Description of requirements for a performance report**  
(from the National Audit Office [64])

Requirement	Description
<b>Data management procedures</b>	Explain how performance information was collected, validated, processed, and analyzed. These explanations will provide the reliability of the reported performance.
<b>Performance information</b>	The performance information should be presented in a manner that can be clearly understood by those who will assess the performance report (e.g., SPL and Operations Manager). A common practice is to include diagrams and tables to clearly show performance against targets.  Performance information can also be presented visually over time in a graph. Trends and patterns are much more obvious to the human eye when viewed graphically. Ideally, such a graph should contain not just the current status information but also the original estimate of expected progress, thereby allowing the viewer to make immediate comparisons.
<b>Analyze performance</b>	Present possible explanations for significant variations between performance achieved and performance expected. Explanatory information should assist readers in understanding the level of performance achieved and evaluating the underlying factors that affected reported performance.

measuring process or impact performance. Some examples of both process and impact measures are given in Table D-19. In addition, Figure D-16 depicts the relationships between various activities and results for an emphasis area aimed at reducing speed-related crashes and crash severities.

### D3.2.1 Process Performance Data Collection

A key deliverable of ISMProcess Step 5 is a detailed action plan that describes the activities required to effectively execute a strategy or a combination of strategies. Each activity in the detailed action plan includes a time schedule with start and end dates, allocation of resources, assigned responsibilities for activity execution, and goals for each activity. Process performance measurement requires that the following questions be answered:

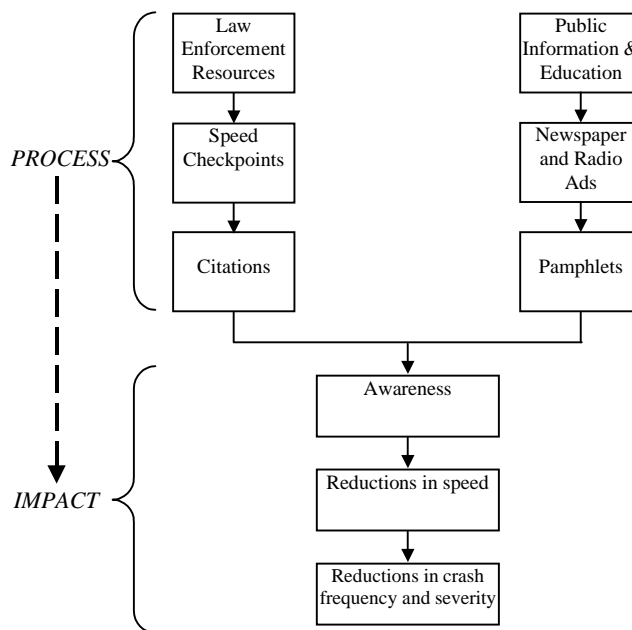
- Was the activity implemented? Was the activity completed? What factors influenced implementation or completion of the activity?

**TABLE D-19 Examples of performance measures**

Type of Performance	Metrics
<b>Process</b>	Citations issued Number of arrests Number of drivers stopped Number of TV advertisements aired Number of intersections converted to roundabouts Miles of rumble strip implemented
<b>Impact</b>	Reduction in crashes Reduction in crash severity Reduction on offense rates Reduction in recidivism rates Improvement in public awareness

- How many and what type of resources were used to implement or complete the activity, and at what cost?
- Who was involved in the implementation of the activity? Was leadership for the implementation adequate?
- What were the results of the activity?

One way to gather this information from the variety of people involved in a given activity is through feedback forms. NHTSA (65) recommends that the people who are responsible for completing feedback forms be involved in the form design. Members of the Task Team responsible for the strat-



*Figure D-16. Strategies aimed at reducing crash frequency and severity due to excessive speeds.*

egy (and related activities) should lead development of feedback forms. The feedback forms should be developed during ISMProcess Step 5, in conjunction with the development of detailed action plans.

A user-centered design process should be in place with the following three emphases:

- Early focus on users and their tasks,
- Empirical measurement, and
- Iterative design.

In other words, users of forms should be involved early in the design process and not after the forms have been produced. Formal usability testing should be conducted by first testing the forms in the field with real users performing their daily tasks. Finally, expect an iterative cycle of design and testing to finalize the feedback forms. Ideally, basic training should be provided to staff who will be completing the feedback forms. This training will ensure that forms are completed in a correct and consistent manner and enhance the quality of data obtained from the forms.

In addition to feedback forms, there are other ways to collect process performance measures:

- Schedule information to be updated through electronic or paper Gantt charts.
- Certain measures (e.g., staff hours, citations issued) can be obtained from existing personnel and financial and administrative management systems.
- Other measures can be collected through the systematic screening of data sources, such as newspaper archives, TV station records, and official records.

### D3.2.2 Impact Performance Data Collection

There are a number of different methods for collecting data to measure *impact* performance indicators. The following five methods are detailed in this document:

- Surveys,
- Psychometric and performance tests,
- Focus groups,
- Field studies, and
- Archival data.

#### D3.2.2.1 Surveys

One method of data collection, perhaps the most common, is the survey. Surveys gather information about the opinions, attitudes, knowledge, behaviors, and intentions of people. In highway safety, surveys are typically used to determine awareness of safety issues, including knowledge of current strategies. Creating surveys that obtain reliable information regarding road user awareness and knowledge is a complex topic that is best undertaken with the assistance and guidance of specialists in this field.

The use of surveys requires consideration of

- Bias,
- Survey instrument selection and design,
- Survey methodology, and
- Quality control.

Reliable survey data depends on

- Selecting representative survey respondents,
- Having a sufficiently large sample size, and
- Obtaining high response rates (63).

**Bias.** Bias may result if the survey is given to respondents who are not part of the target group or if the survey is conducted at times and locations that do not correspond with the characteristics of the target group/crashes. For example, a survey with participants selected only from a household phonebook may underrepresent college students that live in dormitories and depend upon cell phones (65).

**Survey instrument selection and design.** There are various ways of conducting surveys. People of interest can be surveyed through an oral interview or a written questionnaire. Some methods for administering surveys are

- Face-to-face personal interviews at home, at the office, or elsewhere;
- Mailed questionnaires;
- Phone interviews; and
- Written or verbal surveys at roadside checkpoints.

The method selected will depend on the research questions that need to be addressed and on the sample size or population that needs to be surveyed (which will affect the cost of conducting the survey). Often, available resources predetermine the method, sample size, and timeframe for the survey. In general, mail surveys are the least expensive method but can result in the lowest response rate. Face-to-face interviews can be the most expensive method of survey but usually garner a high response rate. Incentives can be employed to encourage higher response rates for mail surveys, as well as for other survey methods.

Surveys should be brief, with only the very basic information requested to measure the desired performance measure. The Procurement Executives' Association points out that surveys should be easy to complete in the shortest period of time (63). In addition, questions should be simple, direct, grouped by topic, and not open-ended. As stated previously, professionals who specialize in designing surveys should be actively involved in the survey process.

Incorporating background questions (e.g., gender, age, place of residence) will assist in a more detailed analysis of the results. For example, differences in categories of respondents' backgrounds may explain differences in responses to survey questions.



The Procurement Executives' Association recommends that a "comments" section be added at the very end of the survey for obtaining respondents' general comments, impressions, and recommendations (63). A "comments" section often contributes to higher response rates.

**Survey methodology.** It is important that survey participants be selected randomly and that they represent the target group. Nonrandom sampling or nonrepresentative sampling could introduce bias, as discussed previously. Stratified random sampling techniques can be used to ensure that the sample represents the target group on important variables such as gender, age, and ethnicity.

Surveys can gather information about

- A census of drivers of interest (e.g., all repeat traffic offenders in the jurisdiction);
- A representative, randomly selected, or stratified sample drawn from the population of interest (randomly assigned to complete a new educational course and all those assigned to no treatment);
- A cluster or group (e.g., all repeat traffic offenders within a certain county or city); and/or
- A quota sample of people in the population of interest (e.g., one-third of all repeat traffic offenders in the jurisdiction).

The sampling method will depend on the nature of the program being evaluated and the budget available.

Surveys for purposes of evaluation—for example, in a before-and-after study design—can be repeated with the same participants or repeated with a different representative sample. In these cases, the survey procedures and the survey questions must be standardized to ensure that the findings reflect real change before and after, rather than the outcome of different procedures and questions. This standardization includes interviewer training and quality monitoring in the case of oral interviews.

The primary limitations of surveys are the respondents' truthfulness, their ability to accurately recall past behaviors and events, and the inclination to respond in a socially desirable manner. Procedures can be embedded in questionnaires to minimize these limitations (e.g., truthfulness scales).

When response rates are low, a nonresponse bias becomes an issue to consider. Nonrespondents can be recontacted and asked to indicate the reason(s) they did not respond to the survey. Demographic and other characteristics of respondents and nonrespondents can also be analyzed to determine the extent of bias and to ascertain whether findings need to be adjusted to account for major biases.

**Quality control.** The survey instruments designed for adequate data collection are as important as the act of collection itself. Armstrong et al. (66) have some instructions for quality control that apply in many traffic safety situations:

- **Relevance:** The surveys should be designed to accommodate all relevant information, but *only* relevant information. If some information is required only in special circumstances, an efficient skipping pattern around "not applicable" items must be part of the design.
- **Completeness:** The survey, when completed, should have no missing data except for clearly "not applicable" items. This may require training of the people charged with completing the form (e.g., police, interviewers) and is especially important in the design of self-completed forms (i.e., surveys completed by the participants themselves). Pilot testing of newly designed survey forms is essential to identify problems at an early stage.
- **Accuracy:** A number of techniques are available to search for errors in the data collection; again, the training and supervision of data collectors will help to increase the accuracy of the collected data. The less time spent processing data prior to analysis, the more likely the data are to contain errors. A system that facilitates collecting data accurately in the first place will repay dividends. Nevertheless, time must be spent checking for the inevitable errors that do occur before the data can be released for information and analysis.
- **Brevity:** Care is needed to ensure that collected data are detailed enough to be useful, flexible enough to be adapted to a variety of purposes, and still as brief as possible. For example, it is more appropriate to record a driver's age at a roadside checkpoint than to categorize the driver into one of several age groups. The age categories may be sufficient to answer one set of questions, but another set of questions may require more detail. In some circumstances, it is more sensible to ask for the driver's birth date.

Quality control can also be enhanced by the use of computers and automated data collection devices, which serve to

- Reduce the time and cost of completing forms;
- Improve data entry and completeness by edit and consistency checks; and
- Add flexibility in terms of adding, deleting, or revising data elements.

Data collection staff should be adequately trained, with the help of clear and concise guidelines and processes on how to capture data and ensure the accuracy of data.

### D3.2.2.2 Psychometric and Performance Tests

The second method of data collection discussed here is psychometric and performance tests. These tests can be used to identify and assess problem drivers. Information derived from these standardized tests on the characteristics of problem drivers and the reasons for their driving errors can be

used to improve driver training and licensing programs. This information may also indicate the effectiveness of relevant strategies in changing knowledge, attitudes, skills, behaviors, and intentions of the targeted driving population (e.g., repeat traffic offenders, young or novice drivers).

Numerous psychometric instruments have been used to assess social, psychological, and behavioral factors associated with risky and/or problem driving. In the field of road safety, psychometric tests have been used to measure personality, attitude, alcohol abuse, sensation seeking, health compromising, and risky driving behaviors. Basic psychometric principles include reliability, validity, and norms.

Performance tests such as off-road and on-road skill tests have been used for research. These performance tests have also been used operationally by licensing agencies to ascertain whether drivers are competent and whether drivers have achieved the minimum standard of safe driving. Computer-based performance tests, such as a hazard perception test, have been applied in research and licensing contexts. Such tests must also meet standards of reliability and validity.

### D3.2.3 Focus Groups

The third method of data collection is the use of focus groups. Leader-directed group discussions can be used to obtain information on attitudes, perceptions, knowledge, behaviors, and intentions of small, well-defined, and representative samples drawn from targeted populations. The focus group technique, which has been used predominantly for consumer and marketing research, has also been effectively applied in the field of road safety.

The primary advantage of focus groups is that the issue under discussion can be covered comprehensively in a systematic and standardized manner. Focus groups lend themselves to more in-depth probing than is possible in individual interviews and surveys and, consequently, to a better understanding of the reasons behind attitudes and behaviors. Participants can be probed to determine why attitudes and/or behaviors changed the reasons they have been resistant to change and what is needed to produce change.

Group dynamics play an important role in such sessions, so group leaders should be properly trained in this technique. A limitation of focus groups is that the number of participants is relatively small. Therefore, even if the participants represent the target population, the results from one or more sessions may not generalize well to the larger population.

### D3.2.4 Field Studies

Field studies are the fourth method of data collection discussed here. They involve the systematic observation of road user behaviors on the highway. For evaluation purposes, road users are observed before, during, and after the implementation of a strategy to determine if the strategies have changed

the behaviors of the road users as predicted. For example, field observations can detect the presence or absence of a behavior or record some measurement of a condition (e.g., vehicle speeds, vehicle headways, or driver BAC).

Field studies are appropriate when existing (archival) data sources and other data-gathering methods, such as surveys, are unable to provide valid and reliable measures of the behavior change under investigation. Given that many strategies are intended to modify behavior on the highways as a precursor to crash reduction, real-world observation is a very appropriate data-gathering method.

To conduct a valid field study to measure an impact performance indicator, the following issues need to be considered:

- What procedure to use for collecting the observations,
- What to observe,
- Where and when to make the observations, and
- Sample size and/or observation duration.

#### D3.2.4.1 Procedures for Field Data Collection

The procedures to be followed during a field observation depend on the type of safety behavior that has to be observed. There are two main categories of safety behavior:

- Behavior that can be observed remotely and unobtrusively. Examples of behaviors included in this category are
  - Vehicle speeds,
  - Vehicle headways,
  - Certain vehicle deficiencies,
  - Vehicle loading (if using high weigh-in-motion devices),
  - Vehicle encroachments, and
  - Failure to obey traffic control devices (e.g., failing to stop at a red signal).
- Behavior that can be observed only by stopping the drivers, occupants, or pedestrians. Examples of behaviors included in this category are
  - The influence of alcohol or drugs,
  - Certain vehicle deficiencies, and
  - Driver and vehicle documentation (e.g., driver licenses, permits).

During the second type of observations, a brief survey questionnaire can be administered to the person to collect demographic information, trip purpose, and exposure-related information, as well as other information pertinent to the objectives of the evaluation.

Where possible, as traffic volumes allow, all vehicles passing the observation point should be observed. At higher volumes, only a sample of vehicles should be observed.

Vehicles should be selected randomly. For example, decide prior to beginning the survey that every  $n$ th vehicle will be observed. It is important not to deviate from this decision during the survey.

At the point of observation, there should be no outside factors (e.g., roadway or environmental characteristics) that could influence the behavior of the driver who is the subject of observation. For example, vehicle speeds should not be influenced by the presence of intersections, steep gradients, and weather conditions. Typically, when conducting speed surveys, the observer and the equipment should not be conspicuous to drivers.

**What to observe.** What to observe during a field observation is determined by the definition of the performance measure. Two example performance measures and related field observations are shown in Table D-20.

It is often not so simple when the performance measure relates to vehicle maneuvers or traffic conflicts. These situations often require the subjective interpretation of the observer—when a vehicle maneuver is a conflict and when it is not. It is possible that different observers may interpret vehicle behaviors in a different way. Field observations aimed at these types of performance measures require well-trained observers.

**Where and when to make the observations.** If the countermeasure strategy is site specific (e.g., intersection improvements), then field observations should be conducted at these locations.

If the strategy aims to have an areawide effect (e.g., reduce drunk driving or increase seat belt wearing rates), then field observations should be conducted at a representative number of locations within the target area. The combined results from these selected locations should give a representative indication of the magnitude of the performance indicator for the whole target area.

The number of locations depends on

- The size of the area,
- The desired level of accuracy,
- The cost of field observations, and
- The availability of resources.

To eliminate the potential for bias, it is recommended that locations for such field observations be chosen randomly. It is important that the same locations used for the collection of baseline information prior to strategy implementation be used for observations after strategy implementation. Traffic and environmental factors may change from location to location, at different times of the day, and on different days of the

week. Therefore, it is strongly recommended that observations be conducted under the same conditions (time, day, weather) during both the before and after observations.

When deciding when to conduct field surveys, it is important to consider the target periods. For example, the activities of a strategy aimed at curbing DUI will focus on times when DUI behavior and crashes are most prevalent (i.e., at night). Field observations to determine changes in the drinking rates should also be conducted during this time period.

**Sample size (observation duration).** A key issue when performing field observations is the required sample size. The larger the available sample, the higher the accuracy and reliability of the resultant measurements and inferences made from the data.

Typically, two types of data are collected during a field observation:

- *Continuous data:* This type of data is typically collected when a behavior is measured (e.g., vehicle speeds, vehicle headways, gap acceptance, or BAC levels). This type of data can provide *means, standard deviations, medians, and percentiles*. The data can also provide proportions (i.e., the proportion of sample sizes that exceed a certain threshold).
- *Binary data:* A yes or no answer to a question (e.g., is the occupant wearing a seat belt, or are the vehicle headlights working?). This type of data can be converted into proportions (e.g., the proportion of drivers that are wearing a seat belt).

A sufficient sample size is required to make meaningful inferences about statistical indicators such as means and proportions. For example, determining whether these measures have decreased or increased significantly between two periods and by how much must be based on an adequate number of samples. The formulas to determine sample sizes to make inferences about means and proportions are found in every standard statistics textbook.

**D3.2.5 Archival Data**

The fifth and final data collection method included in this report is archival data. Strategy evaluations can rely extensively on existing data sources. For example, police crash reports can provide a rich source of information on road users, vehicles, infrastructure, and the crash itself. The use of

**TABLE D-20** Example performance measures and the related field observations

Performance measure	What to observe
Proportion of drivers with BAC > 0.08	The BAC reading of the driver
Proportion of drivers not obeying stop signs	Whether a driver approaching a stop sign stops at the stop bar

such existing data sources avoids the high costs that can often be associated with collecting data.

However, existing data are generated principally for administrative, legal, or accounting purposes and not necessarily for road safety research and aggregate data analyses. For this reason, such available data may not conform to the exact requirements for the evaluation. The quality and completeness of information from existing data sources, as well as any reporting biases, need to be carefully assessed before using such available.

Appendix B2 discusses best practices for database management and the following archival data sources:

- State crash archives,
- Department of motor vehicles files,
- Highway crash databases,
- Vehicle registration files,
- Ambulance service and prehospital care data sources,
- Health care and morbidity data,
- Trauma centers, and
- Federal crash archives.

Single data sources often lack comprehensive information, but when merged with other data sources they can provide a rich source of shared information. Indeed, the value of integrating crash data with data from related sources has been recognized for some time. As early as the mid-1970s, Campbell (67) commented, “in addition to the process for crash reporting and compiling, a good system should have the means for linkage and match up with other files.” According to Erlander (68), “The system must be capable of coordination between different data sources so that data from different sources can be put together to give a total picture.” Unfortunately, for most states, political and practical difficulties have resulted in a lack of access to critical data and an inability to integrate data from multiple databases.

Considerable benefits can be achieved from an integrated data system. Such a system reduces overlap, duplication, processing, and storage requirements. In addition, an integrated data system enables the sharing of information among agencies and is vital for cost-effective safety efforts. Having timely and comprehensive information available to traffic planners and investigators and researchers can facilitate problem solving, the development of safety strategies, and their evaluation and improvement.

### **D3.3 OVERVIEW OF THE PROCESS AND IMPACT PERFORMANCE EVALUATION**

Within the context of the ISMProcess, a *process* performance evaluation is important for the following reasons:

- A detailed action plan has been formulated to achieve emphasis area goals after skilled and experienced persons collectively decided on the process measures and tactics

required. It should be determined whether or not the strategies were implemented according to the detailed action plan.

- An incorrect or inaccurate understanding of the relationship between levels of implementation and process performance measures could create future problems. An accurate and correct level of understanding of this relationship will facilitate future strategies and action plans.
- It is important to monitor the usage of resources to ensure that they are used efficiently and will be sufficient to complete the implementation of the action plan.
- Process performance measures could reveal gaps in the way that safety is managed or significant barriers and shortcomings in the ISMProcess implementation. In addition, performance measures may reveal threats in the external and internal environments; these threats could hinder the successful execution of strategies. Process performance measurement allows for the early identification of areas of low performance and gives management the opportunity to take corrective action.

*Impact* performance evaluation seeks to answer the question, “Is a strategy or combination of strategies effective in achieving the emphasis area’s goals and objectives?”

NHTSA (65) identifies two types of impact performance measures that can be used to conduct impact performance evaluations:

- Crash-based performance measures (e.g., a reduction in crash numbers, crash rates, and crash severity) and
- Intermediate (or proxy) performance measures (e.g., a reduction in speeding, increased seat belt use, or reduced roadside encroachment).

NHTSA (65) states that it is not always possible to determine the safety effect of a strategy at a local or even at a state level using crash-based performance measures for several reasons:

- Crashes are relatively rare, and random events and crash frequencies may fluctuate from time to time, even in the absence of any countermeasures. Since the number of crashes that occur at a location or within a community may be small, data may need to be aggregated over years to obtain a sufficient sample size. Crashes are subject to large stochastic components and may require large sample sizes. Obtaining sufficiently large sample sizes could take a long time (e.g., 35 years), by the end of which the strategy might be over. It is then too late to use the results of such an evaluation to improve the performance of the strategy.
- Crashes are influenced by a variety of factors. These factors (or variables) could include
  - Introduction of highway safety measures,
  - Economic conditions (e.g., unemployment),

- Exposure (e.g., the amount of driving done by drivers within a jurisdiction),
- Changes in driver profiles,
- Changes in population and/or vehicle numbers, and
- Changes in roadway and traffic standards.

NHTSA (65) concludes that, by using intermediate (or proxy) measures instead of using changes in fatality and injury levels, the cost of an evaluation can often be significantly reduced. For these reasons, the use of intermediate measures has been advocated. An example of the considerations involved in using intermediate measures is given in Figure D-17. This figure, taken from Dinh-Zarr et al. (69), presents a schematic of the relationship among strategy types, proxy measures, and crashes for seat belt use strategies.

In Figure D-17, a possible proxy measure for a seat belt program might be seat belt use. According to Dinh-Zarr et al. (69), increasing and maintaining high levels of seat belt use is essential. In all types of crashes, seat belts are approximately 45% effective in reducing fatalities in passenger cars and 60% effective in light trucks. It is estimated that seat belts reduce the risk of serious injury to head, chest, and extremities by 50–83%.

Dinh-Zarr et al. (69) identified 46 studies of the effectiveness of seat belt laws and found that all these studies reported a beneficial effect on safety. Dinh-Zarr et al. identified

- 13 studies examining the effectiveness of primary enforcement laws that found that greater benefits are associated with primary laws compared with secondary laws and
- 18 studies on enhanced enforcement programs that specifically target seat belt use.

The evidence found by Dinh-Zarr et al. indicates that enhanced enforcement programs are associated with an increase in seat belt use and a decrease in injuries. This review of seat belt studies would be helpful for a jurisdiction to esti-

mate the benefits (crash reductions) of strategies without performing extensive data collection. Similarly, one could probably infer crash reduction from strategies aimed at reducing BAC levels without actually measuring crash rates.

Previous research has already established and confirmed the extent and the nature of the relationship between driver behavior (as measured by speeds, BAC levels, seat belt wearing, etc.) and safety (as measured by crash rates, frequencies, and severities). NHTSA (65) recommends that some of the stumbling blocks related to the complexity, cost, and effort required for crash-based performance measures be overcome by focusing on measures of driver behavior and awareness (intermediate, or proxy, measures).

There are, however, many proxy measures (e.g., attitudes, safe driving knowledge, and road test scores) that cannot be relied upon as alternatives to crash rates (70). It is still important to include these proxy measures for two reasons:

- Most strategies, explicitly or implicitly, have targets that are viewed as necessary precursors to achieving safety goals and objectives (e.g., improving driving skills to produce safer drivers, which should result in fewer crashes). The relationship between intermediate objectives/measures (e.g., increased driving skills) and the crash-based objectives/criterion (e.g., crash rates) needs to be understood.
- Focusing only on the crash-based criterion (i.e., reduced crash involvement) provides little insight as to why a program failed to reduce crashes. By better understanding the “change agent,” one can make better decisions to eliminate or improve existing strategies and create new strategies.

NHTSA (65) identifies two different kinds of intermediate impact performance measures:

- Indicators of road user behaviour, including
  - Seat belt wearing rates,
  - Speeding rates,
  - Red light running rates,
  - Intoxication rates, and
  - Roadside encroachment rates.
- Indicators of awareness and public support, including
  - Awareness of safety strategies and initiatives,
  - Support for safety strategies and initiatives,
  - Knowledge of safety legislation, and
  - Driver attitudes (e.g., speeding).

### D3.4 PROCESS PERFORMANCE REPORTING REQUIREMENTS

#### D3.4.1 Process Performance Reporting

A guiding principle for process performance reports is to provide sufficient information for top management to make

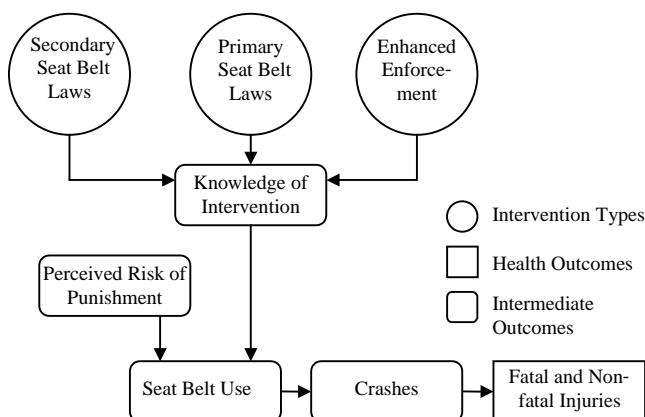


Figure D-17. Logic framework for seat belt use strategies (from Dinh-Zarr et al. [69]).

informed and intelligent assessments and decisions about the performance of the implementation of a strategy or combination of strategies.

To assess the performance of individual strategies, the process performance information could be given separately for each strategy or combined to assess the performance of a combination of strategies. To assess the performance of stakeholder agencies, the process performance information could be given separately for each agency.

Process performance reports should compare the implementation of the program with what was scheduled in the detailed action plan (as formulated by the multidisciplinary Task Team). This includes the following information for each scheduled activity:

- Activity,
- Start date/time,
- Duration,
- Completion date/time, and
- Degree of completion.

For activities that are still in progress at the time of preparing the performance report, the percentage of the task that is completed should be estimated and compared with the schedule. It is useful to illustrate actual progress versus scheduled progress using a Gantt chart (Figure D-18).

Progress performance reports should also compare implementation cost with the corresponding budget for the reporting period. Figure D-18, Table D-21, Figure D-19, and Figure D-20 indicate typical information that will allow the SPL and Operations Manager to track and monitor the use of resources against what was budgeted. This information can be used to determine whether the program will remain within budget. If the program is expected to exceed the budget, the progress information will provide the SPL with an opportunity to obtain additional resources or make modifications to the strategy.

#### D3.4.2 Impact Performance Reporting

During the formulation of the emphasis area goals, objectives, and strategies (ISMProcess Steps 2 and 3), a number of impact performance measures are identified. For each performance measure, an evaluation design outlines the data requirements and the methodologies that will be used to analyze these data. A schedule to collect and analyze the required data for the different performance measures is developed. This information is finalized in the detailed action plan (ISMProcess Step 5).

For each crash-based or proxy impact performance measure of a strategy, the impact performance report should include the following:

- **Problem Statement:** Present the problem (performance measure) to be studied.

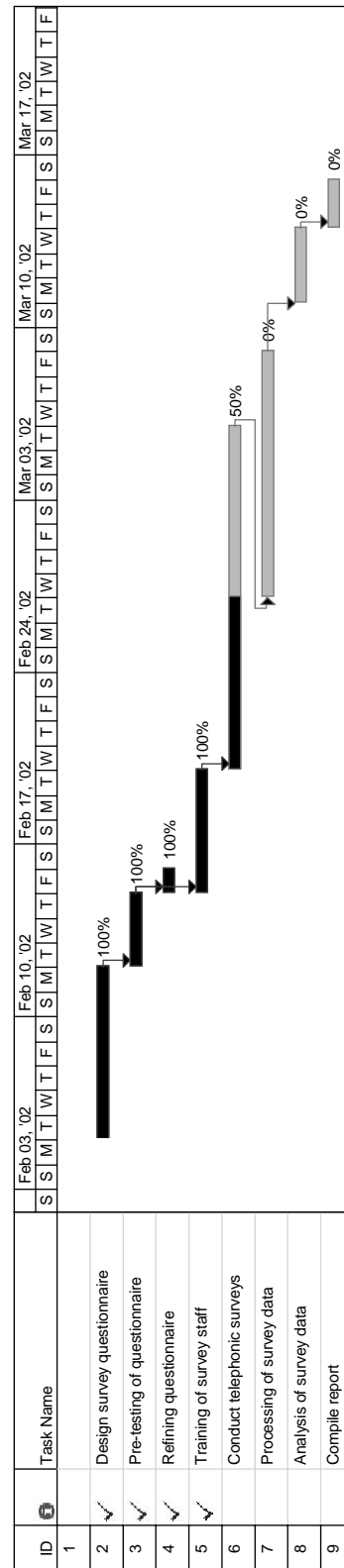


Figure D-18. Example of a Gantt chart.

**TABLE D-21 Monitoring of expenses versus progress example**

Description	Planned	Actual
Total funds expended	\$200,000	\$120,000
Total funds expended (%) of total budget	50%	30%
Percent complete	40%	25%

- **Participants:** Name the participants involved in the study. Describe who collects the data, who does the analyses, and so forth.
- **Data Collection Procedures:** Describe the materials and procedures used to collect data. Provide enough information so that the study can be replicated.
- **Data Analysis Methodology:** Describe the evaluation design and methodologies used to validate, process, and analyze the performance data. List all assumptions.
- **Results:** Present the results of the data in a clear and organized fashion. Keep in mind the target audience—the SPL and Operations Manager.
- **Discussion:** Interpret the results.
- **Limitations:** Discuss factors that affect the results. Discuss potential confounding factors that could or could not be accounted for and their potential impact.
- **Conclusions:** Summarize all of the above sections.

All impact performance measures of a given strategy must then be combined to determine the overall impact of the strategy. The purpose of the impact performance report is to answer the question “Is a strategy or combination of strategies effective in reaching the emphasis area goals and objectives?”

### D3.5 ASSESSING AND IMPROVING PERFORMANCE

The results from performance measurement must be effectively used to improve the ISMProcess. Performance results can be used to determine gaps between strategic goals, objec-

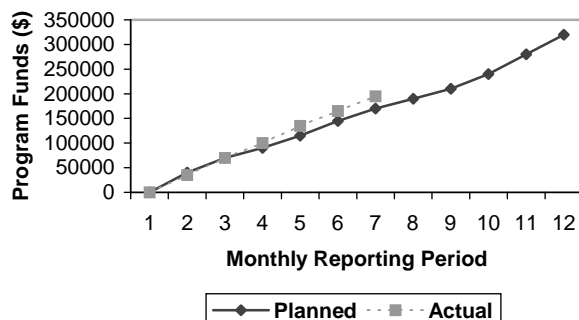


Figure D-19. Monitoring performance with respect to project expenses example.

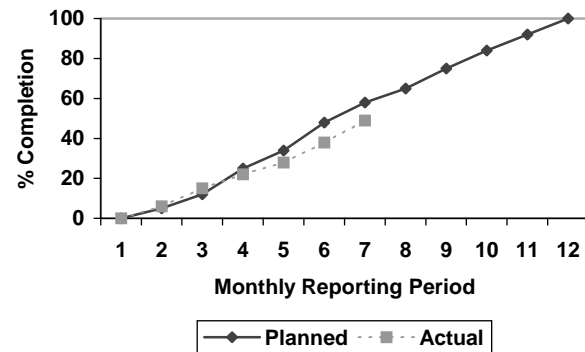


Figure D-20. Monitoring performance with respect to project progress example.

tives, and actual achievement. When a gap between current results and desired goals is identified, there is an opportunity to improve the process or strategy.

Reduced effectiveness could result from a variety of factors. A strategy may be working less efficiently, and therefore less effectively, because of the following:

- Resource cuts or shifting priorities (e.g., shift of enforcement to security from apprehension of impaired drivers);
- A change in the characteristics of the problem due to past success of the strategy (e.g., social drinkers view drunk driving as socially unacceptable, resulting in hard-core drinking drivers accounting for much of the alcohol-related crash problem); or
- External factors such as demographic shifts, social changes, economic cycles, infrastructure, institutional factors, political changes, and new road safety initiatives (e.g., there will be more young and aging drivers over the next few decades and, consequently, increases in their involvement in crashes).

Low performance could indicate the presence of one or more of the following barriers:

- **The action plan has not been implemented as scheduled.** Any unplanned deviation from the schedule could compromise the effectiveness of the emphasis area strategies.
- **Incorrect selection of activities to achieve goals and objectives of selected strategies.** This barrier could indicate a lack of understanding of how to address the emphasis area, possibly due to a lack of research or previous experience related to the particular emphasis area. It could also indicate the involvement of insufficiently skilled or inexperienced staff in the planning process.
- **Incorrect or inaccurate scheduling of activities considering the temporal and geo-spatial characteristics of the target crashes.** Activities may have been scheduled at times and locations that do not reach the target

group. For example, sobriety checkpoints scheduled during the daytime in quiet residential areas will not be as effective as those at night near entertainment locations.

- **Insufficient coordination among activities to produce synergistic effects.** Certain activities need to be scheduled in a way to support and enhance each other. A typical example is the requirement to combine public information and education activities and law enforcement activities to create a bigger impact on driver awareness and behavior.
- **Incorrect or inappropriate levels of implementation.** If not enough of an activity has been performed, the performance measure may be unaffected. This barrier could indicate a lack of understanding of the relationships between levels of implementation, process, and impact performance measures.
- **Ineffective use of resources, resulting in low productivity levels.** This barrier could indicate the insufficient use of technology. For example, it could indicate the need and justification to move away from manual law enforcement to automated law enforcement techniques where appropriate.

A key responsibility of the SPL and the Operations Manager is to assess and use the feedback (performance reports) given by the Task Teams. The role and function of the SPL and Operations Manager require that they understand and make corrective actions to eliminate areas of low performance.

The SPL and Operations Manager can make the transition from assessment to performance management in the following manner:

1. Reanalyze or initiate further analysis of data to enhance the understanding of the problem. The findings from this additional analysis need to be translated into modifications for strategies and action plans.
2. Redefine the goals, objectives, and performance measure criteria if it is evident that they are unrealistic given the current constraints and limitations of the ISMS system. The goals, objectives, and performance measures should be realistic and achievable.
3. Redesign, modify, or adjust the strategies and/or the action plan where low performance has been reported. The changes could be major (e.g., cancel certain strategies or add strategies) or minor (e.g., reschedule activities).
4. Change the levels of implementation and/or take measures to improve productivity. This step may require additional funds or the reallocation of existing funds.

Good feedback and communication from the Task Teams to the Operations Manager, SPL, and stakeholder agencies are vital to managing performance. Informing all stakeholders about the overall progress of a strategy keeps the stakeholders motivated and committed to the objectives and vision. Infor-

mation on corrective steps should be relayed to the relevant stakeholders who implement the decisions and recommendations of the SPL and the Operations Manager. Typically, requirements with respect to the modification of strategies and action plans will be communicated to the multidisciplinary Task Teams for action and implementation.

### D3.6 GENERAL ISSUES IN EVALUATION DESIGN

#### D3.6.1 Regression-to-Mean

Many studies attempt to evaluate the effect of a countermeasure on subsequent crash rates through simple, single-group, before-and-after evaluation designs. Such studies can never be considered definitive, and they are even more flawed when applied to entities (e.g., road sections, inter-sections, or drivers) that have been selected based on some statistical deviancy, such as inflated crash rates. Such studies are inevitably subject to natural regression artifact (regression-to-mean).

The regression-to-mean phenomenon causes an entity with high crash frequencies in one period to have lower frequencies in the following period. Similarly, entities with low crash rates will tend to have higher crash rates in the subsequent period.

The reasons underlying the regression-to-mean phenomenon are technical, but they derive from measurement error theory, which dictates that over the long run, good luck and bad luck will tend to average out. Entities with numerous crashes, some due to chance events, are unlikely to continue to have the same degree of bad luck over the next year. Conversely, some of the entities with exceptional crash records have just been lucky, and this same degree of good luck will not continue at the same rate.

The degree of regression-to-mean can be extreme, as indicated in Table D-22, which is from a California study (71). Table D-22 relates to a program in which drivers with a high conviction count in previous years were randomly assigned to a treatment (individual hearing) or control group.

In Table D-22, the group receiving treatment appears to have an enormous improvement in the subsequent crash and citation rates due to the treatment. However, note that the control group, which did not receive the driver improvement treatment, exhibited an almost identical reduction. Rather than 65–71% reductions in subsequent crash and citation rates, the proper comparison (treated versus control) indicates that the treatment produced no positive effect on crash rates (2.5% more crashes than control) and only a very small reduction in citation rates (3% fewer citations than control).

Although much of the improvement in the treated group reflects regression-to-mean, which can be adjusted for mathematically, other change agents operate to confound interpretation in absence of a control group. For example, the passage of time, maturation, and life changes can affect improvement.



**TABLE D-22** Records of treated and untreated negligent drivers 1 year before and 1 year after treatment

	Mean Crashes		Mean Citations	
	Treatment	Control	Treatment	Control
Pretreatment	0.568	0.602	3.82	3.87
Posttreatment	0.199	0.196	1.09	1.22
Reduction	65.0%	67.4%	71.5%	68.5%

Drivers are also subject to court fines, increased insurance rates, and numerous other contingencies that vary over time. Since the control group is also subject to these same time-dependent change agents, the control group provides a baseline for truly assessing the magnitude of improvement that can be attributed to treatment.

The regression-to-mean difficulty is perhaps more prevalent in the evaluation of improvements to highway locations (intersections and highway segments), since these locations are often selected for safety improvements on the basis of their poor safety performance. When performing a longitudinal before-and-after evaluation study, the regression-to-mean effect must be accounted for.

### D3.6.2 General versus Specific Deterrence

Deterrence theorists such as Ross (72) distinguish between the general deterrent effect of a strategy or treatment and the specific effects of a treatment.

General deterrence relates to the impact of a new law, policy, or treatment in deterring or influencing a large population of entities, often over an entire jurisdiction (e.g., treatments aimed at speeding, drunk driving, seat belt usage, red light running, or tailgating). Specific deterrence is concerned with the impact of a treatment or strategy that has been applied to a specific group of entities that has been deemed in need of remedial safety treatment.

The distinction between general and specific deterrence is critical because the two types of deterrence involve very different evaluation designs. For example, general deterrence evaluations are often amenable to interrupted time series designs, often using ARIMA models, while specific deterrence evaluations require the use of control groups and an assumption that the control groups do not benefit from general deterrence effects. These evaluation designs are discussed later in this appendix.

### D3.6.3 Role of Exposure

Exposure to crashes and traffic citations is influenced by the amount of driving, time of day, and places where people drive. When the amount of miles driven by individuals is known, it is possible to convert crash and citation frequen-

cies to a rate (incidents per million miles) or to use mileage as a covariate in computing mileage-adjusted rates. Unfortunately, evaluation studies require data at the individual driver level, and mileage information for individual drivers does not routinely exist in any state or federal database. Therefore, mileage information for individual drivers must be obtained on a selected basis as part of each evaluation study through questionnaires and interviews (73–76). The only other option is to use some type of induced exposure model (77, 78).

A more fundamental approach is to assess when exposure variables are essential to the validity of the study. When assignment to treatment is random, exposure ceases to be a problem (confounder) because the random assignment process ensures that the groups are initially equivalent in terms of all variables, including driving mileage. Even if subsequent exposure was reduced by the treatment and, as a result, produced a reduction in crash rate, this reduction would still be a causal safety effect—one mediated by exposure reduction. In fact, the objective of some countermeasures, such as license restriction and license suspension, is to reduce exposure.

Where assignment to treatment groups is not random, possible exposure differences become critical. Any differences in subsequent crash rates could be a function of preexisting differences in the amount and type of driving.

Crash rates are often used as a measure of safety at intersections and road segments. The inclusion of exposure in estimating safety is required to equalize for differences in the intensity of use in order to make comparisons more meaningful.

The use of crash rates to make comparisons requires that the relationship between crashes and exposure be linear. In other words, the crash rate should be independent of the level of exposure ( $E$ ).

According to Hauer (21), choosing a measure of exposure that is not linearly related to the crash measure could have the following consequences:

- When the relationship is not linear, the crash rate will change as the amount of traffic (exposure) changes, even if there was no strategy and the road remained the same. It is possible for the crash rate to decrease even as the facility becomes less safe. It is even possible when two facilities are compared with each other for the safer facility to have a higher crash rate than the other facility.

- When evaluating the effectiveness of a remedial measure, a nonlinear safety performance function could cause the effectiveness to be over- or underestimated.

#### D3.6.4 Quality of Data

Another issue in evaluation design is the quality of data collected. Performance measurement should be based on good information. In addition, performance data should be of a sufficient quality to support decision making and to report performance reliability. High-quality data have the following characteristics:

- Valid: Adequately represent actual performance.
- Comprehensive: Include sufficient data.
- Accurate: Are free from significant error.
- Consistent: Are collected using the same procedures and definitions across collection points and time periods.

Good information can be expensive to collect. Data collection procedures should not be an undue burden that impacts an agency's ability to perform its duties. Attempts to reduce the cost of collection could compromise the quality of data collected. There needs to be an appropriate balance between the quality of data and the cost of collection. In order to collect reliable data without placing an undue burden on cost, the National Audit Office (64) recommends the following steps for agencies:

1. Clearly define the performance measurement criteria.
2. Formally define the minimum criteria for good performance data.
3. Involve experts with domain knowledge early in the process.
4. Establish accountability for collecting and reporting performance data.
5. Take advantage of existing data sources and processes.
6. Enlist managerial support, a key component in ensuring good performance data.

Maximizing the use of data from existing management information systems can help to minimize the cost of collecting performance data. Where performance data are already used as part of routine supervision and monitoring, these existing systems could also be used to reduce the cost of data collection.

Many management information systems within highway transportation agencies already collect reliable data that could be useful for measuring performance indicators. Examples include

- Police crash records,
- Department of motor vehicle's driver and vehicle records,
- Law enforcement traffic citation systems, and
- Roadway inventory systems.

In addition, the quality of the data may be evaluated at a more technical level in terms of the fundamental characteristics of the data, which are discussed in the following section.

### D3.7 TECHNICAL ISSUES IN STUDY DESIGN

This section defines several terms that are used to address different pitfalls in the design of the evaluation of a safety strategy. Definitions are provided here and may also be found in NCHRP 20-45, "Scientific Approaches for Transportation Research," Volume 1, available on-line at <http://traffic.ce.gatech.edu/nchrp2045> and as a CD-ROM, CRP-CD-22.

#### D3.7.1 Causation

Computing statistical associations between two or more variables (e.g., excessive speeding and the number of fatalities) in the form of correlation coefficients, odd ratios, regression coefficients, and tests of significance is a relatively straightforward process. Although causation results in a significant association between an independent variable (a safety strategy) and dependent impact measure (a safety effect), associations frequently appear without any causative relationship because of extraneous variables.

For example, an epidemiological study matching cell phone use records to police reports of motor vehicle crashes may establish a correlation between the two, but such a study alone would not prove causation. Extraneous variables are factors that correlate with both the independent and the dependent variable of interest, thereby introducing a correlation between cell phone use records and police records of motor vehicle crashes that is independent of any causal association.

#### D3.7.2 Measurement

Measurement plays a key role when assessing the performance of a strategy or a combination of strategies. The concepts of systematic error, random error (variance), validity, and reliability are fundamental to the development of impact performance measures.

An impact performance measure is usually a proportion, mean, or rate computed over a large sample of entities (drivers, intersections, etc.) that are distinguished based on whether or not they have received a given strategy. The estimation of such parameters involves a number of considerations, including the concepts of precision and bias, which are in turn influenced by sampling error and measurement error. These errors may be random or systematic. These concepts are briefly discussed in following sections.

For the association between the safety of an entity and the entity's characteristics to be viewed as causal, it is necessary, as stated in the previous discussion on causation, that the relationship not be confounded by extraneous variables.

### D3.7.3 Precision and Accuracy

Precision is the degree of variability in the parameter estimate (e.g., mean crash rate of drivers aged 21–25) produced by a given sampling plan and sample size. As sample sizes increase, the variability (standard error) of the estimate usually decreases. The degree of precision is often represented by confidence intervals, and precise parameter estimates have small confidence intervals.

Accuracy refers to whether repeated samples of a given size ultimately converge to produce an unbiased estimate of the true mean. In other words, accuracy refers to the degree to which repeated samplings will be distributed symmetrically around the “true” mean.

A sampling plan can have excellent precision but low accuracy or vice versa. High-quality data should have both good precision and high accuracy. This combination of good precision and high accuracy normally requires the use of large representative samples and properly specified statistical models.

### D3.7.4 Validity, Reliability, and Sources of Error

The concepts of validity and reliability are frequently used to measure performance and skill. The definitions and formula for different types of test reliability and validity are detailed in a manual published by the American Association of Motor Vehicle Administrators (79). More general definitions can be found in a text on test construction and psychometric theory, such as *The Dependability of Behavioral Measurements: Theory and Generalizability for Scores and Profiles* (80).

#### D3.7.4.1 Validity

Validity refers to the extent to which a test or performance indicator measures what it intends to measure. Validity also refers to the ability of an observed measurement to capture the true value of the characteristic of interest. For example, is distance traveled over time a valid measure of exposure to crash involvement, given the variability in speeds of travel? Are nighttime crashes a valid indicator of a strategy or a combination of strategies aimed at reducing drunk driving? Campbell and Stanley (81) distinguish between two types of research design validity: internal and external (general).

Validity also refers to the ability to generalize the evaluation findings to a larger target population and operational environment. If the evaluation is based on highly nonrepresentative samples or uses the countermeasure in a way that is different than how the countermeasure is to be implemented operationally, the results of an internally valid study might not be generalizable, in which case the study would have little or no external validity.

For example, assume that the state of California was evaluating the effectiveness of vehicle impound strategies and

conducted internally valid studies in three cities. If it were found that the strategies reduced repeat-offense rates by 12%, 15%, and 17% in the three cities, could one generalize these estimates to an ISHSPlan implemented in the entire state? Although one might not be confident as to the repeatability of these results, it is likely that a statistically significant effect would occur at the state level.

There are situations where generalization can become problematic. Using the same impound strategy as an illustration, there might be considerably more doubt in generalizing the findings in California to a different state or province.

#### D3.7.4.2 Reliability

Reliability refers to the likelihood of getting similar results from repeated samples. If the same variable is measured multiple times and the result each time is the same, then the measure is considered highly reliable. A low reliability means that the results are likely to look quite different, even though the same variable has been measured each time. Reliability has important consequences for the validity of a test or measure, since validity cannot exceed reliability. A test or measure with zero reliability necessarily has zero validity.

For example, crashes are not really direct measures of driver behavior, but are consequences of multiple interacting variables of roadway, vehicle, and random external variables. As such, driving behavior itself may be quite reliable but not highly correlated with actual crash involvement (71, 82). The large measurement error inherent in using crash rates renders the crash rates susceptible to large regression artifacts (regression-to-mean). Regression-to-mean will most likely occur whenever individuals or groups of drivers are selected based on crash rate deviancy.

Although limited as measures of individual crash propensity, crash rates can have acceptable reliability when computed for aggregate driver categories (e.g., age, gender, and prior traffic conviction frequency). One can therefore make reliable actuarial estimates of a group’s crash expectancy even though one cannot accurately predict which individuals within that group will have crashes.

#### D3.7.4.3 Sources of Error

**Random error.** Variation that can be attributed to chance alone is called random error. Random errors may be due to subtle variations in the observer, circumstances, instruments being used, observed person, road location, or vehicle. Such errors are equally likely to over- or underestimate the true value. Measurements with a high degree of random error are said to lack precision. Achieving a high level of precision, in the face of high variability, is difficult in practice. It requires an increase in the number of observations and reduced sources of variability.

**Systematic error.** Systematic error results in consistent over- or underestimation relative to the true value, resulting in a biased estimate of the truth. Sources of systematic bias can be classified as one of two groups: selection bias or information bias.

Selection bias involves problems in the choice and recruitment of study subjects (people, vehicles, road sites, crashes for detailed investigation, etc.). This bias is particularly an issue when performance information is collected from surveys and when respondents are not selected randomly. This bias is evident when entities such as road locations or drivers are selected for further treatment on the basis of a high crash count, leading to the regression-to-mean phenomenon.

Information bias involves problems in the way the data about each subject are obtained. Common causes of information bias relating to measurements obtained from crash data are

- The underreporting of crashes,
- Inaccurate and incorrect crash information, and
- Missing crash information.

Both selection and information bias will affect the final measurements, including their reliability.

The underreporting of crashes can create significant information bias if analysis is based solely on police crash report data. According to a study done by Hauer and Hakkert (83) among 18 reporting authorities in the United States, Canada, the Netherlands, and Germany, the degree of crash underreporting is substantial and differs widely from one authority to another. The authors estimated that fatalities seem to be known to an accuracy of  $\pm 5\%$ . It was found that about 20% of injuries that require hospitalization are underreported (i.e., inaccurate severity or not reported), and only about 50% of all injuries sustained in motor vehicle crashes are reported to the police.

Hauer and Hakkert (83) state that the probability of reporting an injury to young children is 20–30%, and for people over 60, the probability is 70%. James (84) explains that the low reporting rate for children is related to the type of crash in which they are most likely to be involved in, namely bicycle crashes.

Hauer and Hakkert (83) found that the probability of reporting an injury is largest for the driver, less for the passenger, and even less for nonoccupants. Elvik and Mysen (85) confirmed that reporting levels tend to be higher for occupants and lowest for cyclists. This was the pattern for all 13 countries considered in their investigation (85). Elvik and Mysen found that the reporting of single-vehicle bicycle crashes is particularly low—less than 10% in all the countries studied.

Hauer and Hakkert (83) argue that most of what is said about road safety is based on crashes that have been reported and not on estimates of what actually occurred. This makes the safety problem appear to be smaller than it really is and confuses changes in safety with changes in crash reporting

trends. According to Hauer and Hakkert, if the inclination to report a crash is constant from period to period and between sites, comparisons on safety on the basis of reported crashes are legitimate. However, Hauer and Hakkert argue that this assumption is unrealistic. Several factors influence the probability of a crash being reported—factors that can change over time and from location to location.

### D3.8 EVALUATION DESIGNS AND ANALYSIS METHODS

The purpose of this section to

- Present an overview of the different key concepts and characteristics of evaluation design and study approaches,
- Explain the relevance of these characteristics to highway safety, and
- Provide illustrative examples of the ideas used in practice.

The reader is assumed to have at least a first-year-college-level knowledge of research design and statistics. For more details and elaboration, the reader is referred to the references and, particularly, to “Scientific Approaches for Transportation Research” (available on-line at <http://traffic.ce.gatech.edu/nchrp2045> and as a CD-ROM, CRP-CD-22), which is a comprehensive document. This section may be viewed as a supplement to “Scientific Approaches for Transportation Research” and not as a replacement or substitute. In addition, there is considerable overlap between the evaluation designs and analysis methods described in this section and those designs and methods discussed in Appendix D2 when evaluating new, innovative, or existing strategies for which there is otherwise insufficient information to estimate their effect.

The traditional textbook explanation of the scientific method is approximately as follows:

1. Form a hypothesis,
2. Collect data,
3. Analyze the data, and
4. Report conclusions.

Road safety study design is not that simple. While the first step of forming a hypothesis conforms to the traditional scientific method, the steps that follow and their order is less clear. The hypothesis is basically a question, any safety question, where the answer is unknown or unclear. The question may be, “How many lives would be saved if rumble strips were used on all roads in the state?” Stating this question as a hypothesis might yield, “Rumble strips will reduce fatalities by 4% in the state.” The type of question or hypothesis helps determine what type of analysis needs to be conducted. However, the question “Do rumble strips reduce fatalities?” may lead to very different types of analysis. Therefore, researchers may need to evaluate whether their hypothesis is appropriate.

One misconception based on the textbook-based scientific method is that first data are collected and then the analysis method is chosen. Ideally, the evaluation method would be chosen before the data are collected. Knowing *how* one is going to analyze data determines which data need to be collected. For pilot and evaluation studies, the original proposal for the study should always include details regarding the analysis. If the analysis method has not been decided upon before the data have been collected, then there is a very high possibility that the data collected will fail to adequately address the hypothesis. Therefore, while the analysis can only be conducted after the data have been collected, the analysis method should be selected before the collection of data.

For nonpilot studies (i.e., studies that are retrospective using historical data), the situation is obviously somewhat different in that one does not have the luxury of determining which data were collected in the past. In order to conduct the type of analysis needed, the historical data may need to be transformed, adapted, or selectively filtered to fit within the constraints of the necessary statistical test. Necessity is based on requirements needed to test the hypothesis.

The purpose of evaluation is to determine what effect a strategy has on safety. In order to be successful, an evaluation study must be designed to eliminate, or at least render improbable, alternative explanations for any observed changes. This task is accomplished by controlling other influencing factors so that their effects are either minimized or accounted for. Using a control group is one methodology for factoring out influencing factors. For example, Hauer (37, 86) discusses the study to determine the effect of the Reduced Impaired Driving Everywhere (RIDE) program on DUI injuries and fatalities. Many factors and strategies can influence alcohol-related injuries. Therefore, the change in alcohol-related injuries and fatalities after the implementation of the RIDE program could be due to a variety of factors. The RIDE program was not implemented in every police district, and districts without the program could act as control groups. By comparing the number of alcohol-related injuries and fatalities in districts with and without the RIDE program, it is possible to determine what reduction, if any, is due to the RIDE program.

A variety of research designs and techniques have been developed specifically to assist in the design of evaluation of strategies or treatments (81, 86, 87). Evaluations fall into two broad categories: randomized experimental designs and quasi-experimental designs. (Note: some material in the next paragraphs has been also incorporated into Appendix D2, where similar evaluations are applied to new or innovative measures.)

### D3.8.1 Randomized Experimental Design

In an evaluation design, the key method for deciding which entities (e.g., drivers, road sites, and vehicles) should be included in a pilot study is random selection. Entities are ran-

domly selected for the application of a countermeasure. Examples of countermeasures include formal driver training or red light cameras. A group of randomly selected entities receives no countermeasures and is the *control* group, the group used for comparison purposes. The other randomly selected entities receive the countermeasure and are the *treatment* group.

Random assignment reduces the possibility that a confounding effect may somehow influence the results of the study. A confounding effect is defined as a variable in a statistical model that is correlated with a variable that is not included in a model. For example, confounding may result in an improvement being attributed to a strategy when, in fact, the improvement is due to a variable unrelated to the strategy. In a well-designed evaluation study, all extraneous variables that might confound interpretation of the study results are eliminated or controlled. Two examples of the confounding effect from highway safety literature are the following:

- When behind-the-wheel driver training was first introduced in high schools, initial studies found that teenagers who completed training had much lower crash rates than those who did not. This statistical association was hailed as evidence for the safety benefits of driver training. But it was subsequently demonstrated that those who chose to enroll in driver training classes differed on numerous other characteristics (attitude, grade point average, etc.) and virtually all of these personal characteristics were associated with more favorable crash and citation rates independent of any beneficial effects of driver training. Biases of this magnitude are particularly strong in any evaluation where entry into a safety strategy is voluntary or based on self-selection. Subsequent better-controlled evaluations of driver training programs have failed to produce unequivocal evidence of any safety value (88).
- The evaluations of the effectiveness of court adjudication sanctions (fine, jail, traffic violator schools, alcohol treatment programs) are often corrupted by confounding factors. The procedures used by judges in determining a driver's sanction are not random. Even when done as part of a formal evaluation project, courts rarely follow an established random assignment protocol. Consequently, the resulting sanction groups will often differ on numerous characteristics as a function of the judge's discretionary authority and subjective perception of which sanction is likely to be the most appropriate and effective for a given offender.

In a randomized study, statistical differences between the treatment and control groups are assumed to be due to the effects of the strategy. A carefully designed evaluation study will determine two things about a strategy:

- The *magnitude and direction* of effect of the strategy (How large is the effect, is it in the hypothesized direction?) and

- Whether the effect of the strategy is statistically *significant* (Is the difference real and repeatable or a random fluctuation?).

The smaller the magnitude of an effect is, the greater the amount of effort is that is required in the evaluation study to conclude that an effect is significant. A greater effort implies that a greater amount of data needs to be collected to measure a smaller effect. As previously discussed, collecting more data is more expensive in terms of budget and resources. One method of obtaining more data at reduced cost is to use proxy measures (i.e., measures that assess related or highly correlated characteristics to the variable of interest). For example, the number of conflicts at an intersection is highly correlated to the number of crashes at an intersection. Since there are many more conflicts per hour than crashes, much more data can be collected by counting conflicts than by counting crashes.

### D3.8.2 Quasi-Experimental Designs and Observational Studies

Controlled evaluation using random assignment is not always possible or appropriate. For example, it is not possible to randomly assign people to certain ages or sexes or to assign the historical number of crashes to different locations. In addition, in many cases it may be unethical to arbitrarily apply countermeasures that are believed to reduce safety or to withhold countermeasures that are believed to improve safety from a control group. In such cases, randomized experimental designs are either not possible or not appropriate.

When randomization is not feasible, one could consider quasi-experimental alternatives. Campbell and Stanley (81) and Cook and Campbell (87) have written extensively on evaluation designs, primarily as they relate to evaluating educational programs. Campbell and his colleagues do not actually advocate quasi-experimental designs; in fact, they point out that quasi-experimental designs frequently lead to biased results. The authors recommend that randomized designs be used as extensively as possible and quasi-experimental designs be used only if there are no other options.

However, because of the difficulties involved in adequately randomizing all factors of interest in traffic safety, experimental designs are infrequently applied to evaluate countermeasures and strategies. Quasi-experimental designs are more often used with efforts to minimize potential biases. A highway safety-oriented paper by Peck discusses using quasi-experiments to evaluate the specific effects of DUI sanctions and DUI treatment programs (89). By far, the most common quasi-experimental designs used in evaluating safety treatments are variants of the nonequivalent control or *cohort* study (also known as a weakly randomized treatment-control design). Cook and Campbell (87) describe the variants and conditions under which these designs can provide valid and relatively definite results.

### D3.8.3 Before-and-After Comparison Studies

In a recent book, Hauer (37) has developed much of the methodology in this area of safety analysis. The book identifies the special problems created by the peculiarities of crash and related data. Hauer presents the latest methods to account for these problems in observational before-and-after studies. Before-and-after refers to a class of studies where measurements are taken before the implementation of a strategy and then after the implementation. Before-and-after studies encompass all methodologies by which one may study the safety effect of an implemented strategy on a group of entities (road sections, intersections, drivers, vehicles, neighborhoods, etc.) (37).

Persaud (11) states that some or all of the observed changes in safety following a treatment can be due to factors other than the treatment and need to be separated from the treatment effect. These factors include traffic volume changes, trends in crash occurrence, and random fluctuation in crash counts.

The observed change in safety following a treatment is the sum of the following four effects:

- **Treatment effect:** The change in the level of safety was caused by the influence of a strategy. The estimation of this effect is the primary objective of a strategy evaluation study.
- **Exposure effect:** A change in exposure is likely to have a proportional affect on crash frequency. Between the before and after periods, an increase or decrease in the exposure (e.g., traffic volumes) may occur. An increase in traffic volumes after the implementation of a strategy could by itself lead to an increase in crashes. Not accounting for the exposure effect could result in the underestimation of the treatment effect whenever traffic volumes have increased, and vice versa. A countermeasure can directly influence exposure—for example, an extended learner period reduces exposure for young drivers by limiting their driving habits and, consequently, may reduce the number of young drivers involved in crashes. Thus, the effect on crashes is a direct result of the exposure effect of the countermeasure.
- **Confounding effect:** Many possible causal factors are difficult to identify and measure. Between the before and after periods, there are changes in the traffic composition, driver composition (e.g., more older drivers), law enforcement activity, legislation, crash-reporting practices, and so forth.
- **Random effect:** Variation in data can often be attributed to chance alone.

Crash counts in a time period are random variables that have a Poisson-like distribution around a long-term mean (90). This long-term mean is referred to as the “true” level of safety. It is unlikely that an observed value will be equal to the long-term true mean because of the discrete and random

nature of crash data. The difference between the true mean and the observed value is referred to as “random error.” Remedial measures are normally applied to crash sites with a high historical crash frequency. This application introduces selection bias or the regression-to-mean effect (90). In the after period, the count will generally decrease, even in the absence of any remedial measure, thereby creating a false sense of success.

According to Persaud (11), there are two distinct methodologies to conduct before-and-after studies:

- Conventional before-and-after comparisons and
- The empirical Bayesian procedure.

The above methodologies have different results when eliminating the exposure, confounding, and random effects from the observed change after treatment.

#### *D3.8.3.1 Conventional Before-and-After Comparison*

The simplest, but least powerful, of all evaluation designs is the before-and-after study with no comparison group. This method involves a comparison of the dependent measure before and after the introduction of a strategy to determine the extent of change. For example, the average crash frequency among a target group in the 5-year period before the strategy could be compared statistically with the average crash frequency in the 5 years following the introduction of the strategy. If the strategy had the anticipated effect, the crash frequency should be smaller after the strategy than it was before.

Although the before-and-after study with no comparison groups provides an indication of the extent of change over the period in question, it does not provide compelling evidence that the change was directly a result of the strategy. There are numerous explanations other than the strategy that could account for any observed change. For this reason, other, more powerful techniques must be employed when possible.

One of the simplest, most common means for improving the design is to include a control group of entities that were not exposed to the strategy. The more comparable the characteristics of the treatment and control groups, the greater the validity of the results. The addition of a control group increases the strength of the before-and-after study. However, strong conclusions cannot be made about the treatment effect, since there are still numerous other potential explanations for any observed changes in safety.

An elaborate statistical treatise on the conventional before-and-after methodology is a draft FHWA report by Griffin and Flowers (91). The report is intended to be a manual that documents and discusses the following six different statistical methods. These methods may be used to determine whether and to what degree selected highway projects are reducing crashes:

- Simple before-and-after design,
- Multiple before-and-after design,
- Simple before-and-after design with yoked comparison,
- Multiple before-and-after design with yoked comparison,
- Simple before-and-after design with yoked comparison and check for comparability, and
- Multiple before-and-after design with yoked comparison and check for comparability.

“Multiple” designs combine information from a series of treatments to produce a more stable estimate of treatment effect.

Yoked comparisons are characterized by three measures: time per treatment site, before and after at the treatment site, and before and after at a comparison site. These three measures control for extraneous factors such as changes in traffic conditions, reporting thresholds, and other factors known and unknown.

The comparability check is to ensure that crash trends in the comparison group mirror those in the treatment group in each of the before and after periods. For example, if crashes are rising at 5% per year in the treatment group during the before period, then one should expect crashes to rise by 5% per year in the comparison group during the after period.

Hauer’s book (37), the report by Griffin and Flowers (91), and other prominent sources such as Pendleton (15) emphasize the problem of regression-to-mean. Hauer states that using a comparison group to control for regression-to-mean is problematic, since sites must be matched on crash frequency in order to control for changes in safety due to a random fluctuations in crash counts. For example, if a treatment site had five crashes of the target type before implementation, the matched comparison site should also have had five crashes in the same period. Given this substantial data requirement, the empirical Bayes approach is preferred over conventional before-and-after designs, as acknowledged by Griffin and Flowers (91), for situations “where regression-to-mean might be at play.”

#### *D3.8.3.2 The Empirical Bayes Approach*

The empirical Bayes approach overcomes the difficulties in selecting a proper comparison group by not relying on one. In particular, it overcomes the difficulty of finding comparison sites with the same number of crashes when safety is a consideration in selecting sites for treatment. The approach also facilitates the proper accommodation of traffic volume changes and time trends in a jurisdiction’s crash experience.

The objective is to estimate the number of crashes expected in the after period at the treatment site had there been no treatment. The treatment effect is the difference between this estimate and the number of crashes actually recorded after treatment. The number of crashes expected in the after period had there been no treatment is a weighted average of the number of

crashes in the before period and the number of crashes expected on sites with similar traffic and physical characteristics.

To estimate the weights and the number of crashes expected on sites with similar traffic and physical characteristics, a *reference* group of sites similar to the treated site is used for extrapolation as described by Pendleton (15). Where sufficient data are available, a multivariate model, or safety performance function, that relates crash experience to traffic and physical characteristics of sites in the reference group is calibrated and used to estimate the weights and the number of crashes expected for the reference group.

Hauer (37) refers to this as the “multivariate EB [empirical Bayes] method,” while Pendleton (15) calls it the “EB method with covariates.” The approach is preferred over conventional before-and-after approaches that directly estimate the reference group crash experience. However, there are two drawbacks with the empirical Bayes method:

- Suitable reference population data for calibrating the models are rare in practice.
- The task of calibrating a multivariate model can be challenging, even for those with substantial statistical knowledge.

To overcome these drawbacks, some analysts seek to adapt models developed by others for reference populations similar to those of interest. To this end, there is considerable research underway to develop a comprehensive suite of models for a variety of reference populations.

The main obstacle to applying the empirical Bayes approach, according to Persaud (11), is that the methodology, though conceptually simple, can be cumbersome to apply, especially for analysts without the required background in statistics. Even the provision of software packages such as FHWA’s BEATS (Bayesian Estimation of Accidents in Transportation Studies) (92) has done little to help in this regard. As Pendleton (15) notes, “the issue of who should use this complex methodology requires careful consideration,” and the version of BEATS existing at the time “requires additional effort to even be usable by the statistically sophisticated researcher.”

### D3.8.3.3 Cohort Studies

A cohort study is a study in which entities that presently have a certain condition and/or receive a particular treatment are followed over time and compared with another group that is not affected by the condition under investigation. In the simplest variant of this study, which is the one most relevant to road safety evaluations, one group (e.g., drivers) receives a treatment (e.g., driver training). A comparison group is identified as one that does not receive the treatment. This creates a difficulty because assignment to treatment and comparison groups is not random (the groups are usually differ-

ent in a variety of characteristics, some of which affect safety performance).

For example, Sebastian (39) evaluated the safety effectiveness of different left-turn phasing operations at signalized intersections by dividing intersections into three different groups depending on the left-turn phasing arrangement—permitted, protected/permitted, and protected only. The average left-turn crash rates for these three groups were compared with each other. However, the selected intersections varied in a number of characteristics that could affect safety performance. For example, most approaches without a dedicated left-turn lane operated with permitted phasing.

When assignment of treatments or exposures is not possible, the investigator must adopt a more passive role, usually by conducting an observational study. Sometimes it is possible to identify two otherwise comparable groups, or cohorts, that differ by some variable of interest (e.g., exposure to a particular program or a particular engineering feature) and follow them over time to assess differences in consequences (e.g., their crash record). After controlling for measured confounding variables, the study subjects can be compared even though they have not been formally randomized as in an experiment. For example, one can use administrative databases of traffic convictions or crash records to obtain direct estimates of risk of conviction or crash.

The cohort design is among the most common designs used in the evaluation of highway safety strategies. Even studies conducted many years ago have often employed sophisticated statistical analyses. Though not an example of an evaluation study, a good example of a cohort design is provided by Brezina (38), whose study examined a sample of approximately 49,000 Ontario drivers over a 39-month period in the late 1960s. Brezina examined patterns of crash and conviction occurrence associated with sex, age, degree of experience, and size of community. Patterns were identified using life table analyses for first crash, multiple regression to predict crashes, and measures of association between convictions and crashes.

The cohort design is commonly used in the evaluation of safety strategies targeted at road locations. In this context, they are also referred to as “cross-section” or “with and without” studies. For these applications, the differences in crash experience between two cohort groups is usually partly due to factors other than the measure being evaluated, and these factors cannot be controlled for.

Differences in traffic volume and other characteristics can be controlled for through the use of a regression model in which crashes are first related in a regression equation to a variety of highway features, including traffic volume. The safety effect of making a change in one or more variables can then be estimated using the equation to calculate the resulting change in crashes. Council and Stewart (40) evaluated the safety effects of converting rural two-lane roadways to four lanes based on regression equations relating crashes to the annual average daily traffic (AADT) for roads with these two types of cross sections. This can be tricky to interpret, because



it may not be possible to statistically control all possible confounders such as speed, sight distance, and geometry. This difficulty is the reason why these studies may lead to counter-intuitive conclusions (e.g., that seemingly “good” measures such as left-turn lanes or illumination are bad for safety).

The studies to examine the effectiveness of graduated licensing programs (GLPs) include a number of cohort studies. For example, a cohort of new drivers licensed before a GLP is introduced may be selected to compare with a cohort of new drivers licensed after the GLP is in place. These designs have been used to evaluate GLPs in Ontario (41), in Florida (42), and in Nova Scotia (43). Many, if not all, of these studies have relied exclusively on administrative databases to identify suitable new drivers (i.e., eligible for the GLP in their jurisdiction, whether subject to it or not) and to examine their crash frequencies as recorded in their driver records for a defined period of time.

Numerous factors or conditions favor a cohort study involving nonrandomly assigned cohort groups. Inferences from nonrandom studies become less vulnerable to validity threats if some of the following conditions exist:

- Some degree of active control in assignment of subjects,
- Moderate or negligible self-selection in determining treatment group membership,
- A small degree of bias on the covariates,
- A treatment effect (differences in adjusted means) that is much larger than the bias,
- A covariate pool that reflects most of the important factors known to be related to outcome,
- Minimal measurement error on the covariates, and
- An unadjusted treatment effect that is statistically significant and opposite to the direction predicted by the observed bias.

Cohort studies can also be enhanced by use of a control variable (89). A control variable is a measure that reflects exposure but is expected to be uninfluenced or less influenced by treatment. For example, Peck (89) cites two California studies in which interpretation of the effects of DUI countermeasures were enhanced by comparing effects on alcohol-related crashes with crash rates for non-alcohol-related crashes. Griffin and Flowers (91) present a number of before-and-after evaluations on the effects of continuous shoulder rumble strips on on-the-road crashes. Each design uses day-time crash frequency as control measure based on the logical premise that rumble strips would have a much greater effect on nighttime crashes.

#### *D3.8.3.4 Case Control Studies*

Case control studies compare entities that already have a certain condition with entities that do not. A cohort study may not be feasible if entities are exposed to different risks of crash or injury. Or the outcome of interest may be so rare

that large numbers must be followed for extended periods to accumulate a reasonable number of outcome events. In these circumstances, the case control design may be an effective alternative.

In general, case control studies require much smaller sample sizes than do other designs; this is important when data must be collected for each individual involved in the study, such as the amount and type of driving done by a group of drivers. The case control design identifies subjects who have the outcome of interest, such as crashes (cases), and compares them with a group that has not experienced this outcome (controls) with respect to factors that might change the risk of the outcome.

For example, one could look at intersections with crashes involving elderly pedestrians and compare the characteristics of these intersections with intersections that have not experienced such crashes. Analyses might find a number of differences between cases and controls, so that a simple cause-and-effect relationship that implicates one factor is not clear. Statistical analyses that are able to examine the association with one factor while controlling for other factors, such as multiple logistic regression, are commonly used to clarify these relationships.

Haddon et al. (44) provide a good example of a case control study, though not an example of an evaluation study. Haddon et al. carried out a case control study in New York City to examine factors that might be associated with fatal injuries to adult pedestrians. The cases were adults hit by motor vehicles at a known site and time in Manhattan who died of their injuries and were autopsied. For each case, a team of investigators returned to the site of the crash at the same time of the day and day of the week to recruit four pedestrians of the same gender as the case. Sites varied from outside Grand Central Station in the evening rush hour to a street in Harlem at 2:00 a.m. Each control pedestrian was asked to provide information that had been obtained for the case in the investigation of the fatality (e.g., age, race, place of residence, travelling alone or in company, and origin and destination of the trip). At the end of the interview, a breath sample was collected to assess BAC, to compare with BAC obtained at autopsy from each case.

Matching controls to cases on gender and site was not sufficient. The controls were much younger than the cases, and age was likely to confound other factors. The statistical analyses available to Haddon et al. when these data were collected were very different from what is possible today. Nevertheless, the authors were able to demonstrate a strong association with alcohol after stratifying for age in the comparisons.

Despite their utility when the outcome of interest is rare (e.g., fatal crashes), case control studies are not as heavily used in traffic research as might be expected.

Unlike some cohort studies, case control studies rarely rely solely on the data already available in administrative databases. Such databases may provide part of the data needed, but usually involve additional data collection from the cases and controls. For example, it has been suggested that case control

design could be very useful in some aspects of the evaluation of graduated licensing, especially when factors that are not part of what is routinely available on all drivers are to be examined, such as a driver's employment status or annual distance driven. However, case control studies are less useful for evaluating the effects of treatment programs because the strategies of interest would often not be present, at least at sufficient sample sizes, in the pre-outcome histories of the subjects.

#### D3.8.3.5 Interrupted Time-Series analysis

Time-series analysis is used to develop forecast models for time-indexed data; in other words, each data point is associated with a time-stamp. The following example serves to illustrate how an interrupted time-series design was used to evaluate the effect of a sobriety checkpoint program.

Lacey et al. (93) used an interrupted time-series design to measure the effectiveness of Tennessee's statewide sobriety checkpoint program called *Checkpoint Tennessee*. The dependent variable and measure of effectiveness used in the study was "drunk driving fatal crashes." A drunk driving fatal crash was defined as a fatal crash in which one of the involved drivers had a BAC of 0.10%. Data for the study were obtained from NHTSA's Fatality Analysis Reporting System (FARS).

Two techniques were used to account for confounding factors:

- A model of drunk driving fatal crashes in five states surrounding Tennessee (Kentucky, Georgia, Alabama, Mississippi, and Louisiana) was developed using the same procedures to see if an effect occurred coincident with Tennessee strategy. Such an effect might indicate

a regional or, possibly, a national factor not associated with the strategy.

- All fatal crashes were also included as an explanatory variable in the model for Tennessee and the model for the five surrounding states.

To account for a possible lag between the time the program was started and the time an impact occurred, the analysis was conducted using different start dates.

The ARIMA analysis method developed by Box and Jenkins in the 1970s, and incorporated in the SAS statistical package as PROC ARIMA, was used to analyze the time-series data. Lacey et al. (93) found that the best results were obtained when "all drunk driving crashes" was the dependent variable. *The model showed that the strategy variable (Checkpoint Tennessee) had a significant effect, amounting to a reduction of about nine drunk-driving fatal crashes per month (t ratio = -7.06).* This was a 20.4% reduction over the projected number of drunk driving fatal crashes that would have occurred with no safety strategy.

The results are depicted graphically in Figure D-21. The heavy line (labeled "model, program") represents the ARIMA time-series model fitted to the actual data. The light line (labeled "model, no program") shows the series after the start of the checkpoint program had there been no program.

According to McBurney (48), an interrupted time-series design is a "research design that allows the same group to be compared over time by considering the trend of the data before-and-after experimental manipulation." The strategy interrupts sequential measurements of the performance measure (dependent variable). Differences in the data series before the strategy is implemented are examined to determine if the strategy caused the dependent variable to change.

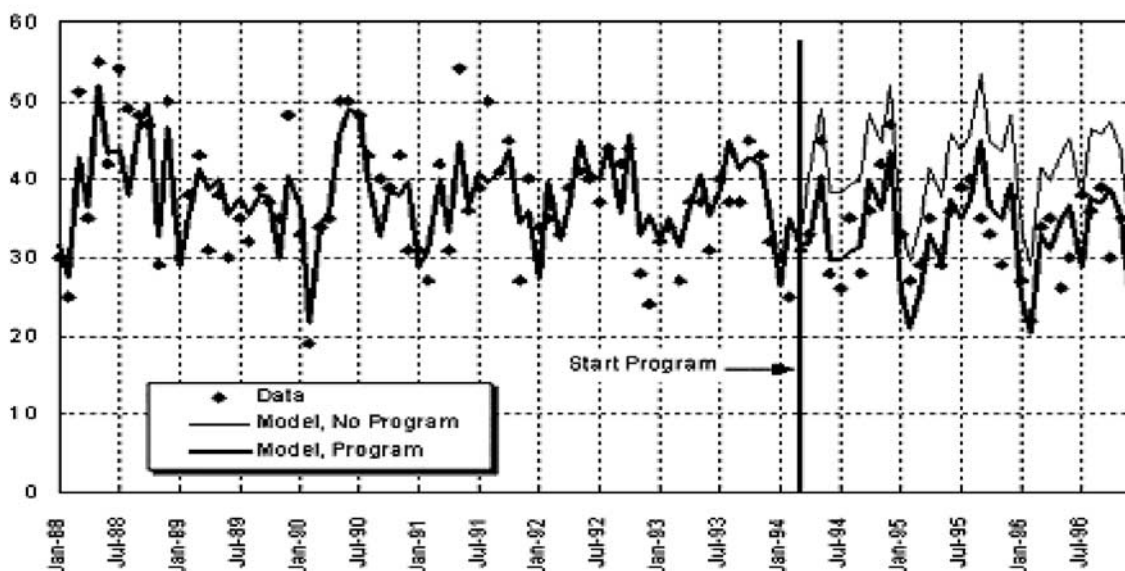


Figure D-21. ARIMA model of drunk driving fatal crashes in Tennessee, ALL fatal crashes as an input (1988–1996).

The interrupted time-series design is one of the most powerful evaluation designs and has been widely used in the traffic safety literature to assess the impact of a variety of strategies, including changes in drinking and driving legislation (94) and GLPs (43).

The interrupted time-series approach requires a relatively long series of equally spaced observations both before and after a strategy. Monthly summaries of traffic crashes are often used for this purpose. The analysis involves examining the data statistically for the presence of long-term trends and cycles in the series and determining whether these trends and cycles changed at the time the strategy was introduced.

According to an online tutorial by Polson (95), the influence of extraneous (confounding) factors is the greatest threat to internal validity in an interrupted time-series design. One way to account for confounding factors is to include a control condition. For example, time-series data from a control group can be assessed (a design called *Interrupted Time Series with a Nonequivalent No-Treatment Control Group Series*; see Cook and Campbell [96]). Another possibility would be to collect time-series data for another dependent measure that should not be affected by the treatment variable but could be affected by likely confounding variables (a design called *Interrupted Time Series with Nonequivalent Dependent Variables*).

A frequently used approach in evaluating DUI laws is to compare monthly pre- versus postcrash rates involving alcohol (or late-night crashes as a surrogate) with a comparison series involving non-DUI crashes (or daytime crashes). These designs are often easy to execute, assuming that the necessary monthly or weekly historical crash data exist and are adequate for detecting moderate-to-large strategy effects. However, the designs often lack sensitivity for detecting small strategy effects.

The design must effectively rule out most alternative explanations for any observed changes in the dependent measure, thereby enhancing the validity of the inference that the changes are attributable to the strategy.

#### D3.8.3.6 Regression Discontinuity Method

The regression discontinuity method sometimes involves a comparison of the trends in the annual series of the outcome measure before and after the strategy. In essence, this method is a time-series analysis using annual data. It can be used to determine the extent to which the prevailing trends in the dependent measure were altered subsequent to the introduction of the new strategy. The analysis involves calculating the regression line for the dependent measure over time with the inclusion of a parameter to represent the strategy (i.e., the implementation of the new strategy [97]).

This method can be used with any outcome measure for which there is a number of equally spaced data points (e.g., annually) before and after the strategy. The regression discontinuity method can rule out the possibility that any apparent change in the outcome measure was the result of a preexisting trend. The regression discontinuity method can also be made

more powerful through the addition of a control group with data over the same period of time.

Polson (95) summarized the regression discontinuity method as follows:

- The performance metric is measured before the strategy.
- The strategy is applied.
- The performance metric is measured after the strategy.
- A regression line is calculated separately for the performance measures before and after the strategy.
- The regression lines are compared to determine the effect of the strategy on the performance measure.

Beirness et al. (98) used regression discontinuity procedures in an evaluation of the vehicle seizure/impoundment (VSI) and administrative license suspension (ALS) programs implemented in Manitoba in late 1989. The authors examined the annual number of fatally injured drivers with positive BACs in Manitoba in the 5 years before and the 6 years after the implementation of the ALS and VSI programs. Linear trends were plotted for the number of drinking driving fatalities before and after the introduction of the two strategies. The authors found that prior to 1990, the linear trend was rising, but following implementation of the ALS and VSI programs, the trend was flat. Applying the regression discontinuity method revealed a strategy effect that approached significance ( $p < 0.08$ ). This suggested a disruption in the series corresponding to the implementation of ALS and VSI programs. In the comparison province of Saskatchewan, the trend lines revealed that prior to 1990, the number of drinking-driver fatalities was decreasing; since 1990, the trend continued downward, but at a much slower rate. Regression discontinuity analysis revealed no significant disruption in the series between 1989 and 1990 ( $p > 0.1$ ). Based on these findings, the authors concluded that

regression discontinuity analysis revealed an effect of the strategy in Manitoba of marginal significance, but a similar analysis of data from Saskatchewan failed to find a significant change in drinking-driver fatalities corresponding to the time of the introduction of ALS and VSI in Manitoba (98).

Although the previous types of studies have focused on time-series applications, regression discontinuity can be used in any study where there is a strong covariate and subjects have been assigned to treatment based on their scores on the covariate.

### D3.9 PRACTICAL APPLICATIONS FOR THE DEVELOPMENT OF PROCESS AND IMPACT MEASURES AND THE DESIGN OF RELEVANT EVALUATION METHODS

#### D3.9.1 Instituting Graduated Licensing for Young Drivers

Graduated driver licensing for young drivers is one of AASHTO's 22 key emphasis areas. Considering graduated

driver licensing for young drivers is appropriate from two perspectives:

- The elevated crash risk of young, inexperienced drivers is well documented in the United States and elsewhere. Accordingly, young drivers are an identifiable emphasis area that needs to be addressed in all jurisdictions
- Driver licensing, and in particular a system of graduated licensing, is a promising solution to reduce the crash involvement of young drivers, since
  - The program directly addresses the two factors that give rise to the higher crash risk of young drivers— inexperience and immaturity—by ensuring that beginners acquire their initial on-road driving experience under low-risk conditions and by delaying full licensure until they are older and
  - A growing body of research demonstrates that graduated licensing is an effective safety measure (all evaluations conducted to date have reported positive safety benefits).

None of the evaluation studies of graduated licensing conducted to date, however, illustrate all aspects of the evaluation design and system measurement process in highway safety. Taken together, these studies can be used to demonstrate how the process can and should operate. The studies that are referenced in this section include evaluations of GLPs in

- Nova Scotia (35, 43, 99),
- Michigan (100, 101), and
- North Carolina (102).

#### *D3.9.1.1 Problem Identification and Solutions*

As mentioned above, the young driver crash problem is well documented, as numerous age-based studies have found that young drivers have the highest per-driver and per-distance crash rates. Typically, however, concern about the problem results from a tragic crash in which a teenage driver is at fault and several innocent victims are killed. Intense media attention to the crash raises public awareness about the problem and a demand for political action.

The government agency responsible for road safety typically reacts to this call for action by commissioning a study to examine the contemporary magnitude, characteristics, and causes of the young driver crash problem. Per-capita and/or per-driver age-based crash rates (e.g., the number of 16-year-old driver crashes per 10,000 licensed drivers compared with the crash rate for older drivers) are examined, confirming the elevated crash risk of young drivers in that jurisdiction. Based on such research, the government agency considers possible solutions that have been suggested for reducing the risk of crash among young people. Graduated driver licensing is among the solutions and has become increasingly popular over the past decade. This is illustrated by the fact that more

than 40 states have adopted some version of graduated driver licensing.

#### *D3.9.1.2 Process Measures*

The effectiveness of a GLP can be examined in operational or administrative terms (i.e., how well has it been working). This examination can encompass, very broadly, all aspects of the management and delivery of the program. Process measures for GLPs include

- Number of drivers licensed in each stage of the program,
- Number of drivers who take and successfully complete driver training,
- Length of time that novices remain in each stage of the program,
- Number of license test attempts and failures at each stage of the program,
- Number of police charges for violating the conditions of the program, and
- Number of drivers who graduate to a full license.

#### *D3.9.1.3 Impact Measures*

**Intermediate, or proxy, measures.** Several intermediate or intervening factors that can impact the effectiveness of a GLP in achieving its safety goals have been examined in recent studies. These factors include the following:

- *Knowledge about the program*—to be effective, parents and teens need to understand how the program works, and they must be aware of the penalties for noncompliance.
- *Support for the program*—to be effective, the program requires the support of parents and their teens because a lack of support can result in noncompliance.
- *Type and duration of practice driving*—to be effective, novices need to experience a wide range of different traffic situations under the extended period of supervised practice.
- *Compliance with the restrictions*—to be effective, teenagers need to comply with the various restrictions and conditions placed on them.

The evaluation design for the proxy measure follows.

Questionnaire-based, self-report surveys have been employed to obtain information on the above issues. For example, Mayhew et al. (99) conducted telephone interviews with 450 teenagers, ages 16–18, and 500 parents, with teenagers ages 16–18, in the Nova Scotia GLP to obtain information on their knowledge about the program, their attitudes toward it, and the level of support for and compliance with its restrictions. Waller et al. (101) surveyed 814 parents of young drivers in Michigan to determine how the parents were dealing with the new GLP and how acceptable the parents found it.

In the Michigan program, a responsible adult is required to certify that a young driver had received extended (at least 50 hours) supervised practices. Such extended supervised practice is a critical aspect of an effective GLP. The Waller study surveyed parents of young drivers who had completed the supervised driving requirement. The survey was administered at licensing offices when young drivers accompanied by a parent (or other certified supervisor) were seeking Level 2 licensure. The survey requested information about

- The supervisor's relationship to the driver;
- What, if any, assistance was offered in the driving supervision task, and whether the help was used;
- The type and duration of supervision; and
- What limits the parents placed on the student's driving beyond the state's requirements.

The issue of compliance with the restrictions imposed by GLPs is also being addressed in several ongoing studies with driver record data (i.e., examining information on the number of GLPs and any convictions or violations of novice drivers in the program).

**Crash-based impact measures.** The aim of a GLP is to reduce the incidence of crashes for the primary target group (i.e., young drivers). The evaluation design for this measurement follows.

A variety of research designs and techniques have been used to evaluate the safety impact of GLPs. These studies have employed quasi-experimental research designs because graduated licensing is typically applied to all young novices obtaining a license, thus making random assignment to the old or new licensing programs impossible. A few studies have used interrupted time-series analysis—one of the most powerful quasi-experimental evaluation designs—to examine the monthly number of crashes of the primary target group for extended periods before and after the program was implemented (99). The primary advantage of time-series analysis is that it rules out the possibility that any observed short-term changes could be the result of long-term trends or cycles in the data series and not of the program itself. Such studies typically test for both an abrupt, permanent change and a gradual, permanent change following the introduction of graduated licensing. The design effectively rules out most alternative explanations for any observed changes in the outcome (dependent) measure, thereby enhancing the inference that the changes are attributable only to the program.

Most studies, however, used before-and-after comparisons with controls to assess changes in crashes among the primary target group (43, 100, 102). This approach involves comparing the prevalence of crashes among the primary target group prior to the introduction of the program with the crash experience of these groups after the program was implemented. Typically, the year the program was implemented is omitted from the comparisons because of transitional changes in licensing

(e.g., a rush to get licensed just prior to implementation to avoid the new requirements).

Most of the evaluations compare per-capita crash rates in the before and after periods to control for fluctuations in population among the primary target group. A few studies also compare per-driver crash rates to control for changes in licensing, which could account for changes in the frequency of crashes in the target group.

Internal and external control groups not exposed to the new program have been used in most of these studies to control for the effects of other confounding factors or events that could influence the prevalence of crashes (e.g., an economic recession that would impact both young and older drivers alike). Internal controls typically include an older age group of drivers; external controls include similar age groups in other jurisdictions that did not implement a GLP.

Studies, particularly of U.S. programs, typically examine the impact of graduated driver licensing on the crashes of teen drivers (e.g., 16 and 17 years old), the primary target group for the program. The programs that have been evaluated in Canada apply to all novices, not just young ones, so that the evaluations have examined the impact on all novice drivers as well.

The types of outcome measures examined in these studies have varied considerably. However, most studies have examined the effect of the program on all motor vehicle crashes as well as crashes that involve injury. A few studies have also included a measure of nonfatal injury crashes as well as alcohol-related crashes (e.g., crashes in which the police reported the driver as having been drinking or a "surrogate" measure such as single-vehicle crashes of male drivers that occurred at night).

Investigations completed to date have focused on the overall short-term impact of the GLPs. These studies provide few insights into the reasons why graduated licensing is effective (e.g., whether it leads to safer driving or whether the benefits are attributable to reductions in the amount being driven, especially during the extended "learner stage" of supervised driving). Ongoing studies of GLPs are attempting to address these issues. For this purpose, novice driver record data containing information on start and end dates for each license stage are being analyzed to examine, for example,

- The transitional and long-term effects of the implementation of the program on licensing (e.g., rate of licensing just prior to, and subsequent to, the date of implementation and age at which novices decide to become licensed—some may wait until they are 18 to avoid the system);
- The relative contribution of each stage of the program (i.e., learner and intermediate levels to the overall safety effect); and
- The long-term impact of the program into full licensing on crashes.

### D3.9.2 Aggressive Driving

The goal of this emphasis area is to reduce crashes that are attributable to aggressive driving. The AASHTO Implementation Guide on aggressive driving collisions identifies two basic objectives and five strategies for addressing the issue of aggressive driving (Table D-23). Evaluation designs for one of the objectives and related strategies are described briefly in the following sections.

#### D3.9.2.1 Deter Aggressive Driving

##### **Educate and impose sanctions against repeat offenders.**

The target of educating and imposing sanctions against repeat offenders is drivers showing a history of violations that indicate a pattern of aggressive driving. Driving records can be used to identify aggressive drivers (i.e., high-frequency violators), with some further classification or assessment system developed and applied to distinguish between aggressive and nonaggressive violations.

Education for repeat offenders has been identified for impacting the behaviors of drivers with a history of aggressive driving. This strategy is an educational anger management class that uses a facilitation model to modify aggressive driving behaviors of repeat offenders. Repeat offenders can be induced to take the course either by court order or by a motor vehicle department allowing attendance as a substitute for a license action.

*Process Measures.* The purpose of the process evaluation is to determine if the new educational course is implemented as intended and is operating in an efficient and effective manner. The measures of process evaluation can be developed in parallel with the development of the operational definition of aggressive driving (i.e., the process or assessment procedures for identifying aggressive drivers for the strategy) and the course curriculum as well as teaching strategies. Information that needs to be collected could include the following:

- Number of judges, other court personnel, and/or department of motor vehicles staff trained in the procedures and requirements for enrolling drivers in the program;
- Number of instructors trained in the delivery of the course;

- Number of judges and/or department of motor vehicles staff and percentage of all trained judges and/or department of motor vehicles staff that actually refers repeat offenders to the class;
- Number of aggressive drivers identified by the courts and/or department of motor vehicles for remediation (i.e., how well is the random assignment working);
- Number of aggressive drivers and percentage of all eligible offenders that elect to attend the class as a substitute for a license action;
- Number and percentage of aggressive drivers assigned to the course who successfully complete it; and
- Revenue collected from course attendees.

In addition to the above program-related data, information can also be obtained from those agencies and persons involved in the implementation and delivery of the new educational program. Judges, department of motor vehicles staff, and trainers can be surveyed using a semistructured questionnaire to obtain their opinions, comments, and concerns about the delivery and operation of the program as well as suggestions that might improve the overall efficiency of the strategy.

Participants can be surveyed upon course completion to obtain their views and opinions about the value of the course, for themselves personally and for others attending, as well as their recommendations for improving program delivery and course content. If the program is voluntary, eligible offenders who accept other license action instead of the program can be interviewed to determine why they choose not to participate. Course participants who do not complete the program can be interviewed to determine the reasons for dropping out.

*Impact Measures—Specific Deterrence.* Educating and imposing sanctions against repeat offenders targets a specific group of repeat offenders (i.e., identifiable aggressive drivers). Accordingly, the impact evaluation can examine the specific deterrent effects of the new educational course—the impact of the program on persons directly affected by it. This impact could be measured by

- Statistically significant increases in knowledge and attitude improvement following course completion, as well as at 6- and 12-month follow-up periods;

**TABLE D-23 Objectives and strategies for aggressive driving from the AASHTO Implementation Guide**

Objective	Strategies
Deter aggressive driving in specific populations, including those with a history of such behavior, and at specific locations.	Target enforcement. Conduct educational and public information campaigns. Educate and impose sanctions against repeat offenders.
Improve the driving environment to eliminate or minimize the external “triggers” of aggressive driving.	Change or mitigate the effects of identified elements in the environment. Reduce nonrecurring delays and provide better information about these delays.

- Statistically significant decreases in self-reported intentions to drive aggressively in different traffic situations causing frustration;
- Statistically significant reductions in speeding violations and aggressive driving incidents in a 12-month period following course completion; and
- Statistically significant reductions in crash rate in a 12-month period following course completion.

The evaluation design is amenable to a randomized experimental design in which target offenders are randomly assigned by courts or (preferably) a motor vehicle department to one of several treatments, such as

- No treatment control,
- A simple warning letter,
- A traditional group meeting focusing on knowledge and attitudes using a lecture-based approach,
- The new educational course focusing on anger management using a facilitation model, and
- A brief individual hearing or counseling session in which the offender receives a license control action (e.g., 1-year probation with restrictions and a 360-day license suspension).

The treatment comparisons will ultimately depend on the existing driver improvement actions in the jurisdiction that develops and pilots the new educational course.

The randomized design has to be modified if the new program is voluntary and not court-ordered or mandated by the motor vehicle department (i.e., if eligible offenders are allowed to attend as a substitute for a license action). Eligible offenders who voluntarily take the course may differ from those who select the license action. Demographic and driver record information have to be examined to ensure that observed changes in outcome measures (e.g., subsequent moving violations and aggressive driving incidents) are only the result of course completion. Any significant, preexisting differences between treatment groups has to be controlled statistically.

Differences in subsequent 12-month driving records can be evaluated through logistic regression or analysis of covariance. Survival analysis can be used to examine, for the various groups, the probability of not violating or having a crash (i.e., “surviving”) at various time intervals after treatment (e.g., 3 months, 6 months, and 9 months). Examining the probability of not violating or having a crash over time helps to determine the extent to which the new education course prevented, reduced, or delayed repeat violations and crash involvement.

Improvements in knowledge, attitude, and behavioral intentions can be evaluated through changes in before-and-after test scores. Self-report questionnaires can be administered prior to the strategy and shortly following completion. To determine the duration of the effects on knowledge, attitudes, and self-

reported behaviors, the questionnaire can be administered again 6 months and 12 months following program completion. For this purpose, offenders can be required or induced to complete follow-up interviews (e.g., as a further condition for retaining the license).

*Impact Measures—General Deterrence.* A coordinated enforcement and public awareness campaign aimed at aggressive driving would be expected to have global effects on the prevalence of aggressive driving behaviors targeted as well as on crashes, particularly those that appear to indicate aggressive driving. Impact measures include statistically significant reductions in

- Aggressive driving in the targeted area, as measured by reductions in observed and/or self-reported aggressive behavior and citations over time;
- Overall crash rates over time; and
- Proportion of crashes attributable to aggressive driving.

*Evaluation Design.* Educating and imposing sanctions against repeat offenders is amenable to a quasi-experimental research design. Potential high-risk sites can be identified throughout one or more states. Some of these sites can become the target of the enhanced enforcement and public information campaign. A sample of similar sites from another area of the state can constitute the control sites. Relevant rates on the measures should be collected and tabulated, including crash rates in each group prior to initiating the campaign. Treatment X multiplied by interactions can be evaluated for evidence of significant effects on the criterion measures.

Baseline data on the incidence of aggressive driving, citations for aggressive driving, and crashes associated with aggressive driving can be obtained for the treatment and control sites prior to, during, and after the implementation of the coordinated enforcement and public information campaigns.

A variety of evaluation designs can be employed to measure changes in the prevalence of aggressive driving. Following are two examples:

- **Observational surveys.** Studies that involve observing the performance of drivers on the highway are frequently conducted to estimate the prevalence of a behavior such as speeding or running red lights. Observing traffic in treatment and control sites before, during, and after the coordinated campaign generates information on the number of drivers engaging in a predetermined set of behaviors thought to have relevance to aggressive driving. Traffic volumes at each site during the observation period should be recorded to standardize the counts of aggressive driving. Filming can be used to observe driving behavior.
- **Self-reported surveys.** Motorists in the treatment and control sites can simply be asked how frequently, and under what circumstances and conditions, they engage in aggressive driving behaviors and have personally experi-

enced such behaviors by others. Self-reported surveys administered before, during, and after the coordinated campaign provide a direct and efficient method to estimate changes in the prevalence of aggressive driving as well as ascertain the level of awareness about, and support for, the targeted enforcement and public information campaign. Moreover, this approach provides the opportunity to determine the motivations and situations that precipitated the aggressive behavior and to obtain demographic and psychological information on drivers who do and do not drive aggressively. This approach overcomes the major limitation of other research designs (e.g., observational surveys) and data sources (e.g., charge and crash data) that only provide information on the observed behavior and not the motives and intent of the behavior (e.g., observations by police or field staff of tailgating on the highway may reflect a lack of ability of the driver to judge a safe following distance or aggressive driving or something else).

*Target Enforcement and Conduct Educational and Public Information Campaigns.* The two strategies of “target enforcement” and “conduct educational and public information campaigns” are discussed together in this section because of the functional dependency between them. As observed in the AASHTO Strategic Highway Safety Plan Implementation Guides, a coordinated publicity campaign can be a great help to increase the driving public’s perceived likelihood of being apprehended for driving aggressively. Public information and awareness is an essential ingredient to the success of selective and increased enforcement. Conversely, public information designed to deter a problem such as aggressive driving is ineffective in the absence of an increased subjective perception of detection and negative contingencies. Therefore, the above two strategies have been linked and introduced in combination.

*Process Measures.* The process evaluation involves careful monitoring of the targeted enforcement and the public awareness campaign to ensure that the strategies are implemented in an efficient and effective manner and that the level of intensity is comparable across treatment sites and over the duration of the campaign. Process measures include an increase in the following:

- Police personnel allocated to the selected regions,
- Officer time on targeted enforcement efforts,
- The number of drivers cited or referred to courts or department of motor vehicles for sanctions or reexamination,
- The amount and number of TV and other media public information and education messages, and
- Public awareness of aggressive driving and support for the coordinated initiative.

Feedback forms administered before, during, and after the campaign can be used to determine the number of people exposed to, and aware of, the media messages, as well as to enhance enforcement on the highways. Focus groups can be used to develop the media messages and, subsequently, to determine reactions to the campaign. These focus groups should include motorists identified by means of self-reporting and driving history who do and do not drive aggressively.

*Impact Measures—Specific Deterrence.* The goal of public information campaigns is to impact the behavior of “populations” and, therefore, is not expected to have a specific deterrent effect (i.e., on a specific, targeted group of aggressive drivers).

*Impact Measures—General Deterrence.* Targeted enforcement targets a specific group of repeat offenders and does not include efforts to inform the public about the strategy. Accordingly, the strategy may not have any measurable global effects on aggressive drivers in the general population or on crashes that result from aggressive driving.

### **D3.9.3 Unlicensed, Suspended, and Revoked Driver Strategies**

The AASHTO Implementation Guide on unlicensed, suspended, and revoked driver accidents identifies 5 basic objectives and 10 strategies for addressing the issue of unlicensed drivers and drivers with suspended or revoked licenses (Table D-24). Potential performance measures and evaluation designs for the first four objectives and related strategies are described briefly in the following sections.

#### *D3.9.3.1 Apply Special Enforcement Practices*

**Increase enforcement in selected areas.** Increasing enforcement in selected areas targets times and places where U/S/R (unlicensed/suspended/revoked) drivers appear to be over-represented. Selected enforcement to apprehend U/S/R drivers should be conducted at these high-risk times and places.

The execution and evaluation of this strategy involves the following steps:

1. Identify databases that contain information on the incidence of U/S/R drivers by location and time. This step requires crash, police citation, and/or traffic conviction data that contain information on license status and the time and location of U/S/R driving incidents.
2. Develop a selective implementation enforcement protocol.
3. Determine the required number of sites.
4. Obtain cooperation from states and local authorities.



**TABLE D-24 Objectives and strategies for unlicensed, suspended, or revoked drivers from the AASHTO Implementation Guide**

Objective	Strategies
Apply special enforcement practices.	Increase enforcement in selected areas. Routinely link citations to driver record. Create and distribute “hot sheets.”
Restrict mobility through license plate modification or removal.	“Stripe” license plate. Impound license plate.
Restrict mobility through vehicle modification.	Immobilize/impound/seize vehicle. Install ignition interlock device (IID).
Restrict mobility through direct strategy with offender.	Monitor electronically. Incarcerate.
Eliminate need to drive.	Provide alternative transportation choice.

5. Develop a specific research design and data collection procedures.
6. Implement the strategy and obtain data on the specified process and outcome measures.

*Process Measures.* The purpose of the process evaluation is to determine if there has been an increase in enforcement activities aimed at U/S/R drivers and an increase in citations and conviction rates of U/S/R drivers.

Potential process performance measures for this strategy are

- Number of enforcement personnel assigned to the high-risk selected areas,
- Number of cited U/S/R drivers,
- Court conviction rate, and
- Increase in vehicle impoundment.

*Impact Measures—Specific Deterrence.* The purpose of the specific deterrence impact evaluation is to determine if there has been a reduction in the subsequent crash rates and conviction rates of convicted offenders.

Potential specific deterrence impact performance measures for the strategy of increasing selective enforcement are

- Subsequent 1-year rate of convicted offenders and
- Subsequent 1-year traffic citation rate of convicted offenders.

*Impact Measures—General Deterrence.* If the program has a general deterrent effect, one would expect to see

- A reduction in the incidence of cited offenders over time and
- A reduction in the proportion of crashes involving U/S/R.

*Evaluation Design.* The evaluation design is a prospective nonequivalent two-group cohort design with inclusion of available covariates to increase statistical precision and to reduce assignment bias. The covariates include age, gender,

prior driving record, and aggregate “ecologic covariates” based on the zip code of the driver’s residence (see DeYoung [103] for examples of how to compute socioeconomic indexes for zip codes by use of census data).

The comparison group should be selected from department of motor vehicles files and consist of suspended or revoked drivers who were not convicted of a suspended or revoked citation during the same time period. At a minimum, the comparison group should be matched to the experimental group on city of residence and, preferably, zip code of residence. Given the uniqueness of the criterion distribution and the small percentage of subjects expected to have more than one posttreatment incident, logistic regression and COX proportional survival analysis are the preferred statistical analysis methods. Since there is no way of identifying unlicensed drivers in most department of motor vehicles file systems, the specific analysis of effects would not normally include unlicensed drivers.

In some states, the available data may be amenable to an ARIMA time-series analysis if the necessary number of pre-strategy time points for crash rates exist, and the license status of the crash-involved drivers can be determined. Another technical contingency is whether the number of crashes in the experimental and control series is sufficient to meet certain assumptions required by ARIMA.

The existence of the necessary historical data is even more problematic for measuring the incidence of U/S/R driving. If an ARIMA approach cannot be justified, the only option is to perform a before-and-after analysis of annual rates of U/S/R driving convictions. Rates might be computed for an experimental and comparison area for, say, three single-year periods prior to the program and two single-year periods subsequent to the program. In this type of design, the unit of sampling might be number of zip codes or enforcement regions within the experimental and control group areas. Evidence of a treatment effect would be in testing the significance of the year (before and after) times the treatment (group) interaction using a log linear contingency table analysis. For crashes, another option is to use crashes not involving U/S/R drivers as an additional

comparison using one of the before-and-after designs with “yoked comparisons” contained in Griffin and Flowers (91). Some of these designs contain provisions for checking for biased covariates (“checks for comparability”) and allow for single versus multiple sites.

**Routinely link citations to driver record.** Routinely linking citations to driver records requires that all citations be regularly checked against driver records to determine license status. According to the AASHTO Strategic Highway Safety Plan Implementation Guides, this check should ideally occur at the time of apprehension.

*Process Measures.* The purpose of process evaluation is to determine the extent to which the strategy has increased the detection of U/S/R driving and the successful conviction of U/S/R-cited drivers.

Potential process performance measures for this strategy are

- Proportion of U/S/R drivers who are appropriately cited by police and
- Conviction rate in court on the cited charge.

Process evaluation will require the availability of reliable baseline data on the number and proportion of improperly licensed drivers that are detected on the basis of routine enforcement.

*Impact Measures—Specific Deterrence.* The purpose of the specific deterrence impact evaluation is to determine if there has been a reduction in the subsequent crash rates and conviction rates of convicted offenders.

Potential specific deterrence impact performance measures for this strategy are

- Subsequent 1-year crash rate of convicted offenders and
- Subsequent 1-year traffic citation rate of convicted offenders.

Execution and evaluation of this strategy requires a non-equivalent control group design in which the license check versus no check is based on the geographical area of the citing officer unless the status check can be randomized or “quasi-randomly assigned” within a selected enforcement region. Under this latter preferred design, police make status checks on certain days and not others—for example, even-versus odd-numbered days. Drivers who are appropriately cited as a result of the license status check constitute the experimental group, and drivers who avoided *not* being cited as result of the nonstatus check constitute the control group. This design has the advantage of creating equivalent groups, and the statistical analysis is a straightforward application of logistic regression or COX proportional survival analysis.

*Impact Measures—General Deterrence.* If the strategy has a general deterrent effect, one would expect to see

- A reduction in the incidence of cited offenders over time and
- A reduction in the proportion of crashes involving U/S/R drivers.

*Evaluation Design.* The design should involve similar options as described in the strategy to increase selective law enforcement. Note that if the randomized design is used to evaluate the specific effects, an evaluation of general deterrence cannot be based on that same data set or enforcement area.

#### D3.9.3.2 Restrict Mobility Through License Plate Modification or Removal

**“Stripe” license plate.** According to the AASHTO Strategic Highway Safety Plan Implementation Guides, the aim of license plate striping is to discourage unlicensed driving by canceling the vehicle registration on a vehicle operated by a U/S/R driver and covering the annual renewal sticker with a striped “zebra” sticker. A sticker that has been “zebra striped” is considered probable cause for an officer to stop a vehicle to check the license status of the driver.

*Process Measures.* The purpose of process evaluation is to determine to what extent eligible vehicles driven by improperly licensed or unlicensed drivers have been “zebra striped” and appropriately reported to the state licensing agency. Potential process performance measures for the strategy of striping license plates are

- The proportion of the eligible vehicles driven by an unlicensed driver that are striped and appropriately reported to the state licensing agency and
- No significant variation in percentage as a function of the demographics and characteristics of the offender unless authorized by the law.

*Impact Measures—Specific Deterrence.* The purpose of a specific deterrence impact evaluation is to determine if there has been a reduction in the crash and citation rates involving “zebra striped” vehicles and the drivers of such vehicles.

Potential specific deterrence impact measures for this strategy are

- Subsequent crash rate of striped plate vehicles and unlicensed offenders who were driving such vehicles and
- Subsequent citation rate of striped plate vehicles or unlicensed offenders who had committed an offense resulting in a plate striping.

*Impact Measure—General Deterrence.* The purpose of a general deterrence impact evaluation is to determine if there has been a decrease in the proportion of casualty crashes (fatal and injury) involving U/S/R drivers after the implementation of the strategy.

A potential general deterrence impact measure for this strategy is the proportion of casualty crashes (fatal and injury) involving U/S/R drivers following implementation of the law.

*Evaluation Design.* The logistical requirements, strategic considerations, and research design options are almost identical to the preceding two strategies with some modifications. In addition to identifying a control group of nonstriped plate subjects from the same period as the striping, a second control group of U/S/R violators should be identified from the state licensing agency's files for the year prior to implementation of the striping program. The use of a control from the prior year was used by DeYoung (103) in his evaluation of the specific effects of the California vehicle impound program because of concern over large selection biases influencing application of striping policy during the enforcement period. The use of two control groups permits an assessment of assignment bias and superior statistical adjustments. Differences in overall state crash and conviction rates associated with differing calendar year for measuring posttreatment outcomes can be controlled for statistically.

Interrupted time-series analyses using an ARIMA model should be used to determine if the monthly number of casualty crashes involving unlicensed drivers decreased following implementation of the strategy. Two levels of control series should be used:

- A matching state and
- The number of crashes involving only licensed drivers within each state.

If the license status of nonfatal crashes is not available from state files, the analyses can be limited to FARS data.

An alternative to comparing crash volumes of unlicensed and licensed drivers is to model the following ratio:

$$\frac{C_U}{C_L}$$

where

- $C_U$  = Crashes involving U/S/R drivers and  
 $C_L$  = Crashes involving only licensed drivers.

**Impound license plate.** The aim of this strategy is to discourage improperly licensed driving by impounding and destroying the license plate of the vehicle in which the driver is apprehended.

*Process Measures.* The purpose of a process evaluation is to determine the extent to which the license plates of eligible offenders have been confiscated and the extent to which drivers driving without license plates are apprehended and convicted. Potential process performance measures for this strategy are

- Number of offenders who have their license plate confiscated and
- Number of drivers driving a vehicle with a confiscated plate who are convicted of the offense.

*Impact Measures—Specific Deterrence.* The purpose of a specific deterrence impact evaluation is to determine if there has been a reduction in the crash and citation rates during the period of confiscation. Potential specific deterrence performance measures for this strategy are

- Number of crashes during the period of confiscation,
- Rate of DUI re-offense during the period of confiscation,
- Number of alcohol-related crashes during the period of confiscation, and
- Number of moving violation suspensions during the period of confiscation.

*Impact Measures—General Deterrence.* Impounding license plates would have a deterrent effect on offenders who would be exposed to the strategy should they drive illegally. Potential general deterrence performance measures for this strategy are

- Number of casualty (fatal and injury) or fatal crashes involving unlicensed drivers,
- Number of fatal had-been-drinking (HBD) crashes involving unlicensed and U/S/R drivers (note: this is limited to fatal crashes because alcohol designations are often not reliable for injury crashes), and
- Number of late-night (10 p.m. or later) casualty (fatal and injury) crashes involving U/S/R drivers (note: late-night crashes are frequently used as a surrogate for alcohol-related crashes).

*Evaluation Design.* The research design is the same as those used for striped plates except that two additional ARIMA models are involved for measuring the general deterrence effects on HBD and late-night crashes. The two types of crashes involve a comparison of two time series within the selected state. In the case of HBD-fatal, non-HBD-fatal can serve as a control series; for late-night crashes, daytime (6 a.m. to 7 p.m.) crashes can serve as a control series.

### D3.9.3.3 Restrict Mobility through Vehicle Modification

**Immobilize/impound/seize vehicle.** According to the AASHTO Strategic Highway Safety Plan Implementation Guides, the goal of this strategy is to impound or immobilize the vehicles of multiple offenders.

*Process Measures.* The aim of the process performance evaluation is to determine the extent to which the impound-

ment and immobilization of eligible vehicles have increased. Potential process performance measures for this strategy are

- Number of eligible vehicles that would be impounded, immobilized, or forfeited, depending on which options are available to the selected state, and
- Number of U/S/R drivers who become eligible to be licensed or reinstated and do so within 3 months of impoundment.

*Impact Measures—Specific Deterrence.* The purpose of specific deterrence impact evaluation is to determine whether the subsequent crash and citation rates of drivers decrease after the vehicles are impounded or immobilized. Potential specific deterrence performance measures for this strategy are

- The subsequent 1-year crash rate of the drivers of impounded vehicles,
- The subsequent 1-year rate of DUI offenses, and
- The subsequent 1-year rate of moving violation citations.

*Evaluation Design.* The evaluation design considerations for immobilizing, impounding, or seizing vehicles are similar, if not identical, to those of license plate confiscation. For examples of such designs, see DeYoung (103, 104).

An additional strategic consideration with vehicle impound and immobilization concerns the identification of the impounded vehicle sample. In California, for example, there is currently no centralized file of impounded or forfeited vehicles (103). This makes identification of any sample extremely difficult because of the need to access numerous local police databases, which also creates problems in terms of the reliability of even enumerating the total number of vehicles impounded. Any implementation and evaluation of this strategy needs to establish a centralized database at the outset.

*Impact Measures—General Deterrence.* Potential general deterrence performance measures for vehicle impound and immobilization are

- The number of casualty crashes by U/S/R drivers and
- The number of HBD fatal crashes involving U/S/R drivers following enactment of the program.

*Install Ignition Interlock Device.* According to the AASHTO Strategic Highway Safety Plan Implementation Guides, the goal of the ignition interlock strategy is to install an ignition interlock device into the offender's vehicle so that the driver can operate the vehicle only when sober. This strategy can reduce hazardous driving by relicensed U/S/R drivers.

*Process Measures.* Potential process performance measures for the ignition interlock strategy are

- Number of offenders with ignition interlocks installed and
- Rate of proper installation and compliance with state specifications and maintenance.

*Impact Measures—Specific Deterrence.* Potential specific deterrence performance measures for this strategy are

- Number of subsequent 1-year crashes caused by the targeted offenders,
- Number of 1-year subsequent HBD crashes, and
- DUI recidivism rate.

*Evaluation Design.* In addition to many of the logistical considerations outlined above, any evaluation would be influenced by whether the statute is mandatory or allows judicial discretion. Another consideration is the timing. Some ignition interlock requirements are applied at the time a driver qualifies for conditional reinstatement, whereas others are installed at the time of conviction for a DUI offense.

Assuming that a random assignment design is not feasible, the only option is some form of prospective nonequivalent control group design. The control group consists of offenders who, for some reason, are not required to install an interlock. Detailed covariate data have to be collected on both groups to reduce self-selection and other confounding variables. If this type of design is adopted, an additional control should be selected from DUI offenders convicted in the year prior to the ignition interlock law. This control could be less confounded by selection biases, but would be subject to potential biases emanating from the different time periods of the observations. Total population crash and conviction rates from the time periods could be used to assess this source of bias and provide a standardization referent for an adjusted logistic regression or log linear analysis.

*Impact Measures—General Deterrence.* The ignition interlock strategy is aimed only at drivers whose vehicles have been fitted with an ignition interlock and therefore does not involve a general deterrence evaluation.

#### *D3.9.3.4 Restrict Mobility Through Direct Strategy with Offender*

**Incarceration for DUI offenders.** According to the AASHTO Strategic Highway Safety Plan Implementation Guides, the goal of incarceration is to restrict the activities, including driving, of repeat offenders.

*Process Measures.* Potential process performance measures for this strategy are

- The proportion of eligible offenders assigned to one of the programs,

- The mean amount of time served, and
- The costs of the sanctions.

*Impact Measures—Specific Deterrence.* Potential specific deterrence performance measures for this strategy are

- Whether jail and house arrest groups have at least 20% fewer crashes than the community service group during the 12-month period following initiation of the sanction;
- Whether the jail and house arrest groups have at least 25% fewer HBD crashes than the community service group during the 12-month period following initiation of the sanction;
- Whether the jail and house arrest groups have at least 30% fewer DUI convictions than the community service group during the 12-month period following initiation of the sanction;
- Whether, when computed from the time the incarceration ends, the above reductions will reduce to 10%, 15%, and 20%; and
- Whether, compared with similar nonjailed offenders from a prior year, the jail and house arrest groups have 30% fewer crashes, 35% fewer HBD crashes, and 40% fewer DUI convictions during the 12 months following initiation of the sanction.

*Strategic Considerations.* House arrest (electric monitoring) and incarceration are used as sanctions for DUI offenses. In most states, incarceration is mandatory for repeat offenders. In some states, including California, brief jail is mandatory for first-time offenders. However, in California and numerous other states, some form of community service is frequently used in lieu of actual jail time, particularly for first-time offenders. It therefore seems logical to evaluate house arrest, jail, and community service as optional variants of a sanction intended to punish and, in the case of jail and house arrest, temporarily remove offenders' access to a vehicle. It appears feasible to evaluate the comparative effects of these sanctions in a randomized three-group design. Offenders deemed by the court to be appropriate for incarceration are randomly assigned to jail, house arrest, or community service.

The length of the sanctions are graduated by number of priors (first-time, repeat), and it might also be necessary to limit the target groups to first- and second-time offenders because of the reluctance to waive in-jail incarceration for third-time offenders. To make the sanction severity more equitable, the length of house arrest and community service can be longer than jail time.

If the randomized design listed above is not feasible, another option is to assign offenders discretionarily to the above sanctions based on the judgment of the court or a pre-sentence assessment formula. The amount of time of each sanction can be varied randomly within each sanction level.

Some offenders can receive only very brief exposures to serve as quasi-controls for those receiving the full length.

The only other option can be some form of nonequivalent control group design, but the selection biases are likely to be too extreme to allow for valid inferences. Note that neither of the above designs has a pure control group unless one chooses to view community service as a control. It therefore might be desirable to use as controls offenders from the previous year who qualified for jail but did not receive any.

*Evaluation Design.* A randomized design has been presented above. If not feasible, the only quasi-experimental option that seems feasible is to use offenders from the two previous years who meet the eligibility criteria. This option results in a difference in the calendar period for monitoring subsequent performance, but there are a number of methods for adjusting or standardizing the effects for time-related differences in statewide crash and conviction rates. In any event, this option makes the design a prospective nonequivalent control design with time of measurement as an additional covariate.

Given the 24-month follow period and the need to measure the effects from both the beginning and the end of the incarceration or house arrest, the COX proportional survival analysis with two treatment strategy points appears to be the best method of analysis. If a randomized design proves feasible, the method of statistical analysis is the same, but there is less concern over selection and other biases. Sample size considerations should be based on the same criteria for the strategy for increased selective enforcement.

*Impact Measures—General Deterrence.* Incarceration does not entail a general deterrent analysis. Many of the offenders may also be under license suspension and in alcohol treatment programs during the same period used to evaluate the effects of jail and house arrest.

#### D3.9.4 Head-On Crashes

The goal of the head-on crashes emphasis area is to reduce head-on crashes associated with highway (i.e., nonintersection) segments. A head-on crash occurs when a vehicle crosses the centerline or a median and crashes into a vehicle approaching from the opposite direction. Head-on crashes can be the result of inadvertent actions by the driver or deliberate actions of the driver (e.g., driving on the wrong side of the road). The AASHTO Strategic Highway Safety Plan Implementation Guides identify two objectives and seven strategies for addressing the issue of head-on crashes (Table D-25).

Evaluation designs for the first objective and related strategies are described briefly below. The strategy to install centerline rumble strips will be used to illustrate issues relating to the design and implementation of process and impact performance evaluations.

**TABLE D-25 Objectives and strategies for head-on crashes from the AASHTO Implementation Guide**

Objective	Strategies
Keep vehicles from encroaching into opposite lane.	Install centerline rumble strips for two-lane roads. Provide wider cross sections on two-lane roads. Provide center two-way, left-turn lanes for four- and two-lane roads. Reallocate total two-lane roadway width (lane and shoulder) to include a narrow “buffer median.”
Minimize the likelihood of crashing into an oncoming vehicle.	Use alternating passing lanes or four-lane sections at key locations. Install median barriers for narrow-width medians on multilane roads. Implement strategies over extended sections of highway or only to selected spot locations.

#### *D3.9.4.1 Keep Vehicles from Encroaching into Opposite Lane*

##### **Install centerline rumble strips for two-lane roads.**

The purpose of centerline rumble strips is to alert drivers who may inadvertently encroach into opposing lanes.

*Process Measures.* The purpose of a process evaluation would be to determine

- The number of road miles where centerline rumble strips have been installed,
- The number of hazardous locations where centerline rumble strips have been installed, and
- The number of vehicle miles of travel exposed to centerline rumble strips (requires information on traffic volumes and the length of treatment locations and be retrieved from an agency’s roadway inventory and traffic volume data files).

The information required to obtain the process performance measures should be collected from agencies on whose highways centerline rumble strips have been installed using formal reports.

*Impact Measures.* Two types of measures—proxy (intermediate) measures and crash-based impact measures—can be used to conduct an impact performance evaluation. It should be noted that in order to do a crash-based performance evaluation, a sufficient sample of crashes is required at the treatment locations. Head-on crashes are generally only a fraction of the total crashes at a location and can require long before and after periods for a sufficient sample of head-on crashes to come available. In cases like these, a proxy (intermediate) impact performance evaluation can provide some idea as to whether centerline rumble strips will be effective in reducing head-on crashes.

*Intermediate or Proxy Measures.* The goal of implementing centerline rumble strips is to reduce vehicle encroachments into the opposing lane. The purpose of a proxy (intermediate) performance evaluation is to measure the change in the centerline encroachment rate before and after the implementation of centerline rumble strips.

*Evaluation Design.* The evaluation will require the measurement of centerline encroachment rates using field observation studies at a number of preselected locations (e.g., locations that have previously experienced problems with centerline encroachments). The before and after encroachment rates will be compared to determine if there has been a significant reduction in the encroachment rate. Encroachment rates can be determined by counting the number of vehicles that encroach over the centerline at a location during a specified period of time and expressing this number as a proportion of the total number of vehicles that have traveled through the location during this time period.

In designing a field observation study like this one, consideration needs to be given to the required sample size to detect the desired change in the encroachment rate. Should the sample size be unrealistically large, consideration can be given to using automatic sensing devices to detect encroachments. It is important that the after study be conducted in a manner that will minimize the effect of changes in confounding factors (e.g., traffic volumes, traffic composition, and weather) between the before and after periods. This manner of conducting the study might require that the after study be conducted during the same time of day, by the same observers, during the same day of the week, and under similar weather and environmental conditions as during the before study.

The relationship between encroachment rates and traffic volumes can be nonlinear, and the field observations should be conducted under conditions of similar traffic volumes. The analysis of data will require the hypothesis testing about

two proportions and the construction of a confidence interval for the difference between the two proportions.

The effect of centerline rumble strips on safety at a location may be a function of a number of variables (e.g., traffic volumes, roadway width, travel speeds, type of terrain, and passing opportunities). These factors need to be accounted for in an evaluation to determine the safety effect of centerline rumble strips. This makes it very difficult to conduct a cohort study, which requires that, if possible, the comparison and treatment groups be similar in all respects except for the treatment (i.e., centerline rumble strips).

If safety is a consideration in determining the locations to implement centerline rumble strips, the assignment to the treatment group is nonrandom and the potential for regression-to-mean is introduced. Not accounting for the regression-to-mean bias could result in an overestimation of the safety effect.

The conventional before-and-after method is not recommended because it does not account for the regression-to-mean effect. The accuracy of the conventional before-and-after evaluation with a control group relies on the validity of the assumption of similarity between the treatment and control group sites. The validity of this assignment can be questionable. For example, to account for the regression-to-mean effect, the crash experience of the comparison group should be similar to that of the treatment group. Finding such a similarity can be difficult given that the common practice is to apply safety measures in a nonrandom manner to locations with the worst safety records.

To overcome the problems associated with the cohort study and the conventional before-and-after approach and to account for exposure, regression-to-mean, and confounding factors, the recommended approach is to use an observational before-and-after study using the empirical Bayes approach.

From data about unchanged locations and a consideration of existing models, predictive regression models are calibrated for crashes to traffic flow and geometry. This allows the empirical Bayes method to account for possible regression-to-mean effects. The same model can also be used to account for changes in exposure. A comparison group can be used to account for changes in weather, demographics, and other factors.

In selecting a comparison group, caution should be exercised to ensure that crashes did not “migrate” or “spill over” onto control group locations. It is possible, for example, that centerline rumble strips delay overtaking maneuvers until a section without centerline rumble strips is reached.

The empirical Bayes before-and-after method with a comparison group can be used to predict the expected number of crashes after the implementation of centerline rumble strips, given the unique traffic flow and geometric characteristics of the treatment locations. This prediction will be compared with the number of crashes in the after period to determine the safety effect of the centerline rumble strips. The variance of the safety effect and its confidence interval should be calculated to determine the accuracy of the safety estimate.

## Appendix D References

1. Cameron, M., Haworth, N., Oxley, J., Newstead, S., and Le, T., "Evaluation of Transport Accident Commission Road Safety Television Advertising." Vol. 52, Monash University Accident Research Centre (1993).
2. <http://care.cs.ua.edu> (2002).
3. Huang, H., Schneider, R., Zegeer, C., Khattak, A., and Lacy, J. K., "Analysis of Serious Crashes and Potential Countermeasures on North Carolina Highways." Washington, D.C., Transportation Research Board 81st Annual Meeting (2002).
4. Datta, T. K., Kar, K., and Schattler, K. L., "Statewide Traffic Crash Analysis Using GIS in Michigan." CD-ROM, Institute of Transportation Engineers, Nashville, Tennessee (2002).
5. Vitetta, B. A., and Abdek-Aty, M. A., "Using Induced Exposure to Investigate the Effect of Driver Factors in Traffic Safety." CD-ROM, London, Ontario, Third Transportation Specialty Conference of the Canadian Society for Civil Engineering (2000).
6. Zador, P. L., Krawchuk, S. A., and Voas, R. B., "Relative Risk of Fatal Crash Involvement by BAC, Age, and Gender." DOT-HS-809-050, U.S. Department of Transportation, National Highway Traffic Safety Administration (2002).
7. DeYoung, D., Garber, N., Peck, R. C., and Li, D., "Estimating the Exposure and Fatal Crash Rates of Suspended/Revoked and Unlicensed Drivers in California." *Accident Analysis and Prevention*, Vol. 29, No. 1 (1997) pp. 17–24.
8. Dissanayake, S., and Lu Jian, J., "Use of Induced Exposure Method to Study the Highway Crash Involvement of Driver Groups under Different Light Conditions." Third Transportation Specialty Conference of the Canadian Society for Civil Engineering (2000).
9. Zegeer, C. V., *NCHRP Synthesis of Highway Practice 91: Highway Accident Analysis System*, Transportation Research Board, National Research Council, Washington, D.C. (1982).
10. Zegeer, C. V., *NCHRP Synthesis of Highway Practice 128: Methods for Identifying Hazardous Highway Elements*, Transportation Research Board of the National Academies, Washington, D.C. (1986).
11. Persaud, B. N., *NCHRP Synthesis of Highway Practice 295: Statistical Methods in Highway Safety Analysis*, Transportation Research Board, National Research Council, Washington, D.C. (2001).
12. Flowers, R. J., and Griffin, L., "Development of a Plan for Identifying Highway Locations that May Be Overrepresented in Accident Frequency and/or Severity." Texas Department of Transportation, Traffic Engineering Section, Austin, Texas (1992).
13. Hauer, E., Lovell, J., and Geva, I., "Designing a Study to Estimate the Effect of Daytime Running Lights by Decision Analytic Methods." *Transport Canada*, Ottawa, Ontario (1984) pp. 1–92.
14. Hauer, E., "On the Estimation of the Expected Number of Accidents." *Accident Analysis and Prevention*, Vol. 18, No. 1 (1996) pp. 1–12.
15. Pendleton, O., "Evaluation of Accident Analysis Methodology." FHWA-RD-96-039, Federal Highway Administration, Washington, D.C. (1996).
16. Persaud, B. N., Lyon, C., and Nguyen, T., "Empirical Bayes Procedure for Ranking Sites for Safety Investigation by Potential for Safety Improvement." *Transportation Research Record 1665*, Transportation Research Board, National Research Council, Washington, D.C. (1997) pp. 7–12.
17. Persaud, B. N., Cook, W., and Kazakov, A., "Demonstration of New Approaches for Identifying Hazardous Locations and Prioritizing Safety Treatment." Lisbon, Portugal, Seventh International Conference: Traffic Safety on Two Continents (9-22-1997).
18. Bahar, G., Szymanski, E., and Persaud, B. N., "Safety and Operational Improvement Program for Halton Region." Canadian Multidisciplinary Road Safety Conference (CARSP), London, Ontario (2001).
19. Bahar, G., DiLorenzo, T., Munro, J., and Persaud, B. N., "Prioritization of Interchanges and Ramps Based on Potential for Safety Improvements." Institute for Transportation Engineers, Monterey, California (2001).
20. Hagle, J. L., Witkowski, J. M., Morris, C. N., Pendleton, O. J., and Quayle, E., "Bayesian Identification of Hazardous Locations." *Transportation Research Record 1185*, Transportation Research Board, National Research Council, Washington, D.C. (1988) pp. 24–36.
21. Hauer, E., "Exposure and Accident Rate." *Traffic Engineering and Control*, Vol. 36, No. 3 (1995) pp. 134–138.
22. U.S. Department of Transportation, Office of the Assistant Secretary for Transportation Policy, "Revision of Departmental Guidance on Treatment of the Value of Life and Injuries." [http://ostpxweb.dot.gov/VSL\\_background.htm](http://ostpxweb.dot.gov/VSL_background.htm) (2002).
23. "Federal Aide Policy Guide SubChapter J—Highway Safety Part 924—Highway Safety Improvement Program." <http://www.fhwa.dot.gov/legregs/directives/fapg/cfr0924.htm>, 23 CFR 924 OPI: HHS-11, U.S. Department of Transportation, Federal Highway Administration (1991).
24. Federal Highway Administration, "Highway Safety Improvement Program (HSIP)." FHWA-TS-81-218 (1981).
25. McFarland, W. F., Griffin, L., Rollins, J. B., Stockton, W. R., Phillips, D. T., and Dudek, C. L., "Assessment of Techniques for the Cost-Effectiveness of Highway Accident Countermeasures." FHWA-RD-79-53, Federal Highway Administration, Washington, D.C. (1979).
26. Federal Highway Administration, "Summary of Proceedings: Benefit/Cost Applications Conference Exploring the Application of Benefit/Cost Methodologies to Transportation Infrastructure Decision Making." FHWA-PL-96-014 (1995).
27. Briss, P. A., Zaza, S., Pappaioanou, M., Fielding, J., Wright-DeAgüero, L., Truman, B. I., Hopkins, D. P., Mullen, P. D., Thompson, R. S., Woolf, S. H., Carande-Kulis, V. G., Anderson, L., Hinman, A. R., McQueen, D. V., Teutsch, S. M., and Harris, J. R., "Developing an Evidence-Based Guide to Community Preventive Services—Methods." *American Journal of Preventive Medicine*, Vol. 18, No. Supplement 1 (2000) pp. 35–43.
28. Brown, D. B., "Safety Investment Allocation by Dynamic Programming." *AIIE Transactions*, Vol. 5 (1973) pp. 245–249.



29. Brown, D. B., Bulfin, R., and Deason, W., "Allocating Highway Safety Funds." *Transportation Research Record 1270*, Washington, D.C., Transportation Research Board, National Research Council (1990) pp. 85–88.
30. Chen, G., Wilson, J., Meckle, W., and Cooper, P., "Evaluation of Photo Radar Program in British Columbia." *Accident Analysis and Prevention*, Vol. 32 (2000) pp. 517–526.
31. <http://www.lindo.com/table/lingot.html>. (2002)
32. McGeehan, D., and Samuel, L. H., "Procedures for Prioritizing Road Improvements under the Statewide Highway Plan." Virginia Department of Transportation Research Council (1991).
33. Chowdhury, M., Garber, N., and Li, D., "Multiobjective Methodology for Highway Safety Resource Allocation." *ASCE Journal of Infrastructure Systems*, Vol. 6, No. 4 (2000) pp. 138–144.
34. Liu, C. C., and Chen, H., "Safety Resource Allocation Programs and Input Processor—Users Manual." FHWA-TS-88-20, Federal Highway Administration (1988).
35. "Safety Resource Allocation Programs: Implementation Technique." FHWA-TS-88-19, Federal Highway Administration (1988).
36. [www.cochrane.org](http://www.cochrane.org). (2002)
37. Hauer, E., "Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety." Oxford, United Kingdom, Pergamon Press, Elsevier Science Ltd. (1997).
38. Brezina, E. H., "Traffic Accidents and Offences: An Observational Study of the Ontario Driver Population." *Accident Analysis and Prevention*, Vol. 1 (1969) pp. 373–395.
39. Sebastian, K., "Collision Analysis of Left-Turn Maneuvers at Signalized Intersections." Transportation Research Board 78th Annual Meeting (1999).
40. Council, F. M., and Stewart, J. R., "Safety Effects of the Conversion of Rural Two-Lane to Four-Lane Roadways Based on Cross-Sectional Models." *Transportation Research Record 1665*, Transportation Research Board, National Research Council (1999) pp. 35–43.
41. Boase, P., and Tasca, L., "Graduated Licensing System Evaluation—Interim Report 98." Toronto, Ontario, Safety Policy Branch, Ministry of Transportation for Ontario (1998).
42. Ulmer, R. G., Preusser, D. F., Williams, A. F., Ferguson, S. A., and Armer, C. M., "Effect of Florida's Graduated Licensing Program on the Crashes of Teenage Drivers." *Accident Analysis and Prevention*, Vol. 32, No. 4 (2000) pp. 527–532.
43. Mayhew, D. R., Stein, H. S., and Des Groseilliers, M., "Impact of the Graduated Driver Licensing Program in Nova Scotia." *Journal of Crash Prevention and Injury Control*, Vol. 2, No. 3 (1999) pp. 179–192.
44. Haddon, W. J., Valine, P., McCarroll, J. R., and Umberger, C. J., "A Controlled Investigation of the Characteristics of Adult Pedestrians Fatally Injured by Motor Vehicles in Manhattan." *Journal of Chronic Diseases*, Vol. 14, No. 6 (2002) pp. 655–678.
45. Jones, I. S., and Stein, H. S., "Effect of Driver Hours-of-Service on Tractor-Trailer Crash Involvement." Insurance Institute for Highway Safety, Arlington, Virginia (1987).
46. Griffith, M. S., "Safety Evaluation of Rolled-In Continuous Shoulder Rumble Strips Installed on Freeways." *Transportation Research Record 1665*, Transportation Research Board, National Research Council (1999) pp. 28–34.
47. Persaud, B. N., Retting, R., Garder, P., and Lord, D., "Safety Effect of Roundabout Conversions in the United States: Empirical Bayes Observational Before-After Study." *Transportation Research Record 1751*, Transportation Research Board, National Research Council (2001).
48. McBurney, D. H., *Research Methods*. Third edition, Pacific Grove, California, Brooks/Cole Publishing Company (1994).
49. Voas, R. B., Tippet, A. S., and Fell, J., "The Relationship of Alcohol Safety Laws to Drinking Drivers in Fatal Crashes." *Accident Analysis and Prevention*, Vol. 32, No. 4 (2000) pp. 483–492.
50. Scopatz, R., "Methodological Study of Between-States Comparisons, with Particular Application to .08% BAC Law Evaluation." Transportation Research Board 77th Annual Meeting (1998).
51. Tarko, A., Aranky, S., and Sinha, K., "Methodological Considerations in the Development and Use of Crash Reduction Factors." Transportation Research Board 77th Annual Meeting (1998).
52. Marsh, W. C., *Educational Approaches to Driver Improvement: An Experimental Evaluation with Negligent Drivers*. CAL-DMV-RSS-78-66, California State Department of Motor Vehicles (1978).
53. Miller, T. R., "The Effectiveness Review Trials of Hercules and Some Economic Estimates for the Stables." *American Journal of Preventive Medicine*, Vol. 21, No. 4S (2001) pp. 9–12.
54. Retting, R., Nitzburg, M. S., Farmer, C. M., and Knoblauch, R. L., *Field Evaluation of Two Methods for Restricting Right Turn on Red to Promote Pedestrian Safety*. Insurance Institute for Highway Safety (2000).
55. Persaud, B. N., Retting, R., Garder, P., and Lord, D., "Observational Before-and-After Study on the Safety Effect of U.S. Roundabout Conversions Using the Empirical Bayes Method." Transportation Research Board 80th Annual Meeting (2001).
56. Hingson, R., Hereen, T., and Winter, M., "Lowering State Legal Blood Alcohol Limits to 0.08%: The Effect on Fatal Motor Vehicle Crashes." *American Journal of Public Health*, Vol. 86, No. 9 (1996) pp. 1297–1299.
57. Spiegelhalter, D. J., Myles, J. P., Jones, D. R., and Abrams, K. R., "Bayesian Methods in Health Technology Assessment: A Review." *Health Technology Assessment*, Vol. 4, No. 38 (2000) pp. 17–23.
58. Melcher, D. J., Dixon, K. K., Washington, S., and Wu, C., "Bayesian Analysis Development: The Feasibility of 'Subjective' Engineering Evaluations for Assessing Road Safety Improvements." TRB 01-0491, Washington, D.C., Transportation Research Board, National Research Council (2001).
59. Elvik, R., "Measuring the Quality of Road Safety Evaluation Studies: Mission Impossible?" Washington, D.C., Transportation Research Board 81st Annual Meeting. For a copy of the paper, email the author at [re@tio.no](mailto:re@tio.no) (2002).
60. Zaza, S., Carande-Kulis, V. G., Task Force on Community Preventive Services, Sleet, D. A., Sosin, D. M., Elder, R. W., Shults, R. A., Dinh-Zarr, T. B., Nichols, J. L., and Thompson, R. S., "Methods for Conducting Systematic Reviews of the Evidence and Economic Efficiency of Interventions to Reduce Injuries to Motor Vehicle Occupants." *American Journal of Preventive Medicine*, Vol. 21, No. 4S (2001) pp. 23–30.
61. Elvik, R., "The Safety Value of Guardrails and Crash Cushions: A Meta-Analysis of Evidence from Evaluation Studies."

- Accident Analysis and Prevention*, Vol. 27, No. 4 (1995) pp. 523–549.
62. Hagenzieker, M. P., Bijleveld, F. D., and Davidse, R. J., “Effects of Incentive Programs to Stimulate Safety Belt Use: A Meta-Analysis.” *Accident Analysis and Prevention*, Vol. 29, No. 6 (1997) pp. 759–777.
  63. Procurement Executives’ Association, *Guide to a Balanced Scorecard Performance Management Methodology*. U.S. Department of Commerce (1999).
  64. National Audit Office (NAO), “Good Practice in Performance Reporting in Executive Agencies and Non-Departmental Public Bodies.” Comptroller and Auditor General, London, United Kingdom (2000).
  65. “The Art of Appropriate Evaluation: A Guide for Highway Safety Program Managers.” DOT HS 808 894, Washington, D.C., National Highway Traffic Safety Administration (NHTSA) (1999).
  66. Armstrong, B. K., White, E., and Saracci, R., “Principles of Exposure Measurement in Epidemiology.” *Monographs on Epidemiology and Biostatistics*, Vol. 21, Oxford, United Kingdom, Oxford University Press (1994).
  67. Campbell, B. J., “Requirements of a Mass Accident Data System.” Canberra, Australia, Department of Transportation, Expert Group on Road Safety (1978) pp. 7–14.
  68. Erlander, S., “Suggestions for an Improved System for Statistical Data about Road Accidents and Road Traffic.” *Accident Analysis and Prevention*, Vol. 6, No. 2 (1974) pp. 125–139.
  69. Dinh-Zarr, T. B., Sleet, D. A., Shults, R. A., Zaza, S., Elder, R. W., Nichols, J. L., Thompson, R. S., Sosin, D. M., and Task Force on Community Preventive Services, “Reviews of Evidence Regarding Interventions to Increase the Use of Safety Belts.” *American Journal of Preventive Medicine*, Vol. 21, No. 4S (2001) pp. 48–65.
  70. Griffin, L., Powers, B., and Mullen, C., “Impediments to the Evaluation of Highway Safety Programs—Final Report.” MVMA-UNC-74-17, Arlington, Virginia, Motor Vehicle Manufacturers Association (MVMA) (1975).
  71. Peck, R. C., “Toward a Dynamic System of Driver Improvement Program Evaluation.” *Human Factors*, Vol. 18, No. 5 (1976) pp. 493–506.
  72. Ross, H. L., “Deterring the Drinking Driver—Legal Policy & Social Control.” Lexington, Massachusetts, Lexington Books, D.C. Heath & Company (1982).
  73. Burg, A., “The Effects of Exposure to Risk on Driving Record.” University of California, Los Angeles, Institute of Transportation and Traffic Engineering (1973).
  74. Harano, R. M., Peck, R. C., and McBride, R. S., “The Prediction of Accident Liability through Biographical Data and Psychometric Tests.” *Journal of Safety Research*, Vol. 7, No. 1 (1975) pp. 16–52.
  75. Harrington, D. M., “The Randomized Clinical Trial.” *Journal of American Statistical Association*, Vol. 95, No. 449 (2000) pp. 312–319.
  76. Peck, R. C., and Kuan, J., “A Statistical Model of Individual Accident Risk Prediction Using Driving Record, Territory and Other Biographical Factors.” *Accident Analysis and Prevention*, Vol. 15, No. 5 (1983) pp. 371–393.
  77. Banks, W., Shaffer, J., Masemore, W., Fisher, R., Schmidt, C., and Zlotowitz, H., “The Relationship between Previous-Driving Record and Driver Culpability in Fatal, Multiple-Vehicle Collisions.” *Accident Analysis and Prevention*, Vol. 9, No. 1 (1977) pp. 9–14.
  78. Robertson, M. D., and Drummer, O. H., “Responsibility Analysis: A Methodology to Study Effects of Drugs in Driving.” *Accident Analysis and Prevention*, Vol. 26, No. 2 (1994) pp. 243–247.
  79. American Association of Motor Vehicle Administrators, *AAMVA Guidelines for Knowledge & Skill Testing* (1999).
  80. Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N., *The Dependability of Behavioral Measurements: Theory and Generalizability for Scores and Profiles*. New York, New York, Wiley (1972).
  81. Campbell, B. J., and Stanley, J. C., “Experimental and Quasi-Experimental Designs for Research.” Chicago, Illinois, Rand McNally College Publishing Company (1963).
  82. Peck, R. C., “The Identification of Multiple Accident Correlates in High-Risk Drivers with Specific Emphasis on the Role of Age, Experience and Prior Traffic Violation Frequency.” *Alcohol, Drugs and Driving*, Vol. 9, No. 3-4 (1993) pp. 145–166.
  83. Hauer, E. and Hakkert, A. S., “Extent and Some Implications if Incomplete Accident Reporting.” *Transportation Research Record 1185*, Washington, D.C., Transportation Research Board, National Research Council (1988) pp. 1–10.
  84. James, H. F., “Under-Reporting of Road Traffic Accidents.” *Traffic Engineering and Control*, Vol. 32 (1991) pp. 574–583.
  85. Elvik, R., and Mysen, A. B., “Incomplete Accident Reporting: Meta-Analysis of Studies Made in 13 Countries.” *Transportation Research Record 1665*, Washington, D.C., Transportation Research Board, National Research Council (1999) pp. 133–139.
  86. Hauer, E., <http://traffic.ce.gatech.edu/nchrp2045/>. (2002)
  87. Cook, T. D., and Campbell, D. T., “The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings.” Chicago, Illinois, Rand McNally (1979).
  88. Peck, R. C., “The California Performance Evaluation Test—An Improved Method of Assessing Driver Competency.” *New to the Road: Reducing the Risk of Young Motorists*, Los Angeles, California, Youth Enhancement Service, University of California (1996) pp. 95–101.
  89. Peck, R. C., “Evaluating DUI Program Impact in Quasi Experiments.” *Alcohol, Drugs and Driving*, Vol. 10, No. 3-4 (1994) pp. 207–215.
  90. Abbess, C., Jarret, D., and Wright, C. C., “Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment, with Special Reference to the ‘Regression-to-Mean’ Effect.” *Traffic Engineering and Control*, Vol. 22, No. 10 (1981) pp. 535–542.
  91. Griffin, L., and Flowers, R. J., “A Discussion of Six Procedures for Evaluating Highway Safety Projects.” DTFH61-95-R00118, College Station, Texas, Texas Transportation Institute (1997).
  92. Pendleton, O., “A Users Manual for BEATS—Volume II.” FHWA-RD-91-014, Washington, D.C., Federal Highway Administration (1991).
  93. Lacey, J. H., Jones, R. K., Smith, R. G., and Mid-America Research Institute, “Evaluation of Checkpoint Tennessee: Tennessee’s Statewide Sobriety Checkpoint Program.” DTNH22-94-C-05064, Washington, D.C., National Highway Traffic Safety Administration (NHTSA) (1999).

94. Ross, H. L., "Law, Science and Accidents: The British Road Safety Act of 1967." *Journal of Legal Studies*, Vol. 2 (1973) pp. 1–78.
  95. Polson, "Experimental Designs." <http://psych.athabascau.ca/html/Design/> (2002).
  96. Cook, T. D., and Campbell, D. T., "Quasi-Experimentation: Design and Analysis Issues for Field Settings." Boston, Massachusetts, Houghton Mifflin Company (1979).
  97. Jude, K., "Estimating the Effects of Social Interventions." New York, New York, Cambridge University Press (1981).
  98. Beirness, D. J., Simpson, H. M., and Mayhew, D. R., "Evaluation of Administrative License Suspension and Vehicle Impoundment Programs in Manitoba." Ottawa, Ontario, *Transport Canada* (1997).
  99. Mayhew, D. R., Simpson, H. M., Ferguson, S. A., and Williams, A. F., "Graduated Licensing in Nova Scotia: A Survey of Teenagers and Parents." *Journal of Traffic Medicine*, Vol. 26, No. 1-2 (1998) pp. 37–44.
  100. Shope, J. T., Molnar, L. J., Elliot, M. R., and Waller, P. F., "Graduated Driver Licensing in Michigan: Early Impact on Motor Vehicle Crashes Among 16-Year-Old Drivers." *Journal of the American Medical Association (JAMA)*, Vol. 286, No. 13 (2001) pp. 1593–1598.
  101. Waller, P. F., Olk, M. L., and Shope, J. T., "Parental Views and Experience with Michigan's Graduated Driver Licensing Program." *Journal of Safety Research*, Vol. 31, No. 1 (2000) pp. 9–16.
  102. Foss, R., Feaganes, J. R., and Rodgman, E. A., "Initial Effects of Graduated Driver Licensing on 16-Year-Old Driver Crashes in North Carolina." *Journal of the American Medical Association (JAMA)*, Vol. 286, No. 13 (2001) pp. 1588–1592.
  103. DeYoung, D., "An Evaluation of the Specific Deterrent Effects of Vehicle Impoundment on Suspended, Revoked and Unlicensed Drivers in California." HS-8089 7272, Washington, D.C., National Highway Traffic Safety Administration (NHTSA) (1997).
  104. DeYoung, D., "An Evaluation of the Specific Deterrent Effects of Vehicle Impoundment on Suspended, Revoked and Unlicensed Drivers in California." *Accident Analysis and Prevention*, Vol. 31, No. 1 (1999) pp. 45–53.
-

Abbreviations used without definitions in TRB publications:

AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
IEEE	Institute of Electrical and Electronics Engineers
ITE	Institute of Transportation Engineers
NCHRP	National Cooperative Highway Research Program
NCTRP	National Cooperative Transit Research and Development Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
SAE	Society of Automotive Engineers
TCRP	Transit Cooperative Research Program
TRB	Transportation Research Board
U.S.DOT	United States Department of Transportation