Highway Safety Data: Costs, Quality, and Strategies for Improvement

Research Report

PUBLICATION NO. FHWA-RD-96-191

JANUARY 1998

U.S. Department of Transportation
Federal Highway Administration

Research and Development
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

This study was conducted by the Northwestern University Traffic Institute, Evanston, IL, under Federal Highway Administration (FHWA) contract no. DTFH61-91-C-00051. The report identifies, in detail, the issues and costs related to collecting and managing highway safety data and proposes ways to resolve them.

Two related reports, completed under the same contract, are also available. One entitled *Highway Safety Data: Costs, Quality, and Strategies for Improvement, Executive Summary* (FHWA-RD-96-027) provides a summary of the costs and quality of highway safety data and the strategies to improve the data. The other, *Highway Safety Data: Costs, Quality, and Strategies for Improvement, Final Report* (FHWA-RD-96-192), provides details on data collection methodologies and how to package them effectively to meet the needs of individual agencies.

Copies of these reports will be available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161. A limited number of copies will be available from the R&T Report Center, HRD-11, FHWA, 9701 Philadelphia Court, Unit Q, Lanham, MD 20706. The phone number for the R&T Report Center is (301) 577-0818.

A. George Ostensen
Director
Office of Safety and Traffic
Operations Research and Development

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-RD-96-191		

4. Title and Subtitle

HIGHWAY SAFETY DATA:
COSTS, QUALITY, AND STRATEGIES FOR IMPROVEMENT
RESEARCH REPORT

5. Report Date

January 1998

6. Performing Organization Code

7. Author(s)

Ronald C. Pfefer, Richard A. Raub, Roy E. Lucke

8. Performing Organization Report No.

9. Performing Organization Name and Address

Northwestern University Traffic Institute
405 Church Street
Evanston, Illinois 60208

10. Work Unit No. (TRAIS)

3A3B

11. Contract or Grant No.

DTFH61-91-C-00051

12. Sponsoring Agency Name and Address

Office of Safety and Traffic Operations R&D
Federal Highway Administration
6300 Georgetown Pike
McLean, Virginia 22101-2296

13. Type of Report and Period Covered

Research Report
December 1991-January 1996

14. Sponsoring Agency Code

15. Supplementary Notes

Contracting Officer's Technical Representative (COTR): Jeff Paniati, HSR-20

16. Abstract

The goal of this project was to analyze the collection and management of highway safety data by identifying issues and costs, and proposing means of resolving those issues and reducing the costs. Initial emphasis addressed known elements of the highway safety system with emphasis on collecting and storing relevant data. Sources included the literature and what the States, local agencies, and researchers throughout the Nation were doing to improve data handling for the three largest sources: traffic crash, roadway inventory, and crash injury (emergency medical services and trauma). Although a review of uses was important, a major investigation remained outside the scope of the project. Visits to various providers of data throughout the United States helped identify and classify issues along with exemplary practices. A total of 41 issues were listed which affected collecting, storing, and managing traffic crash, roadway inventory and medical data. The most important issue is that of quality, with data accuracy the most critical. Lack of coverage is becoming an increasing problem which affects information used to recommend countermeasures.

A major thrust of the research was directed toward identifying the costs of collecting, reporting, and managing safety data. Lack of cost data or lack of applicability of most data collected for roadway inventory and crash injury preclude an extension of the cost model beyond that of crash reporting. The three processes, collecting, reporting, and managing crash data, are estimated to cost 19.20 dollars per crash report filed based on personnel and equipment costs but disregarding "sunk costs." Additionally the report estimates a range of costs by severity of the crash, number of vehicles involved, and region of the country. These ranges were found to be significantly less than those estimated by others.

Finally a set of strategies was identified along with goals to be met. The research team identified 23 strategies which were capable of being introduced without requiring substantial additional effort. A number of strategies were taken from exemplary practices discovered at the State and local level. Each of these strategies was evaluated as it related to meeting the goals and objectives, and to reducing costs of operation. The report concludes with discussion of how strategies may be grouped and introduced as a package. All of the recommended strategies will provide both short- and long term benefits.

17. Key Words	18. Distribution Statement
Costs, safety data, strategies, crash data, roadway inventory, medical data	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	203	

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

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

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.71	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C + 32	Fahrenheit temperature	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

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

(Revised September 1993)

Highway Safety Data:
Costs, Quality, and Strategies for Improvement
Research Report

Table of Contents

Table of Contents (continued)

Table of Contents (continued)

Table of Contents (continued)

Table of Contents (continued)

Table of Contents (continued)

List of Figures

List of Tables

List of Tables (continued)

List of Tables (continued)

Acknowledgments

Many persons who regularly handle crash, roadway inventory and medical data at the State and local level provided their cooperation, time, and data in order to prepare this report. To attempt to name all of the contributors would risk missing names, particularly in agencies where many members of sections and departments worked jointly with the project team, but for which names were not always available. However, specific attention is given to the highway safety data units or divisions in California, New Jersey, Ohio, Pennsylvania, Virginia, and Wisconsin for pulling crash reports and providing final crash data records for analysis of quality of those reports.

Managers of seven Federal data bases gave time to the team for interviews regarding the operation of the systems. They also provided documentation used as part of the review of these efforts.

Additionally, the project team acknowledges the graciousness and willingness to cooperate shown by local and State police agencies and their management throughout the United States as the project members attempted to gather timing data for vehicular crashes. A number of these agencies also contributed self-timing reports prepared by their officers.

The Northwestern University Transportation Library provide substantial assistance in uncovering and providing literature in the field. This literature search resulted in an annotated bibliography published separately, along with references found in this report. Hema Ramachandran, Dorothy Ramm, and Mary McReadie provided most of the references assisted by student workers from the University.

The late James O'Day not only served as a consultant for the project, but gave insight into the many issues affecting the collection and management of crash safety and roadway inventory data. Jim also reviewed much of the work leading to the establishment of goals and objectives and in the evaluation of strategies. These reviews helped ensure that the analyses were clear and remained focused.

Finally, Chong Qing Wu, as masters degree student working for the Traffic Institute helped build the annotated bibliography. He also was responsible for analyzing and describing the various crash reports from the States.

Executive Summary

The goal of analyzing the collection and management of highway safety data was to identify issues and costs, and to propose means of resolving those issues and reducing the costs. This research report summarizes the findings and proposes strategies designed to improve both collection and management of these data. It is a synthesis of numerous reports produced on various aspects of the research over the 3-year period of investigation.

Initially, the research team addressed known elements of the highway safety system with emphasis on collecting and storing relevant data. Concentration was on what the literature had described and what the States, local agencies, and researchers throughout the Nation were doing to improve the handling of safety data from the three largest sources: traffic crash, roadway inventory, and crash injury (emergency medical services and trauma). Early in the project, a decision had been made to limit the scope of investigation to collection and storage. Some review of uses for the data occurred, particularly as these uses affected how safety data were collected and maintained. However, a major investigation of uses and users remained outside the project scope.

As part of the initial assessment, members of the team visited various sites throughout the country. During these visits the team viewed the processes involved in collecting, reporting, and managing data. They also timed police officers in an attempt to separate the times required to manage the traffic-crash scene from those needed to collect and report crash safety data. Although some discussions did involve the collection of emergency medical services and trauma data, the team discovered that the current state of work in this area was very limited. Little could be determined related to the costs and issues of handling these data. Therefore, this source is treated to a lesser extent in this report than issues and costs related to handling data from vehicular crash and the roadway.

Although the uses of highway safety data were not addressed in detail, major functions of use were summarized. This summary shows how these functions use the data related to: case management, problem identification, countermeasure selection, program evaluation, and administration. The elements of use are important in helping classify the issues and subsequent strategies for improvement.

The visits to the sites also helped identify and classify issues. They identified exemplary practices which later served as one basis for designing strategies for improvement. Additional information related to issues, particularly those of quality came from the review of literature. A total of 41 issues were listed which affected collecting, storing, and managing traffic crash, roadway inventory and medical data.

Substantial attention is given to the issues related to quality. The literature has addressed these issues with emphasis on coverage and accuracy. Even without agencies reducing the number of crashes investigated, there already exists serious problems with a failure to report many of the crashes that occur. Although most fatal crashes may be reported, more than 60 percent of those involving only property damage may be missed. This percentage is climbing as States increase thresholds for reporting, and agencies reduce the number of incidents in which they are willing to prepare a report.

Accuracy of data is most critical in properly locating crashes to allow highway engineers to use the reports for improving design. However, review of reports submitted by some of the agencies visited during the course of the project showed that accuracy (and completeness) can have significant gaps. As a result, many of these reports may have limited value in their use for improving highway safety.

A major thrust of the research was directed toward identifying the costs of collecting, reporting, and managing safety data. Because information was available or could be gathered relatively easily for handling traffic crash data, the development of a cost model was restricted to this source. The model was not extended to collection and management of data for roadway inventory or emergency medical services. In the case of the roadway inventory, much of the effort is undertaken for highway maintenance and construction. Distinguishing the aspects related solely to safety could not be performed easily. Additionally, only a small amount of the data collected in the roadway inventory applies to highway safety.

Identifying costs for collecting and managing crash injury data faces the same problems as for roadway inventory. In addition, no mechanism is available for centralized handling of the data. Each source operates independently; gathering costs from even a sample of the sources would have been a prohibitive task.

In order to measure the costs of collecting and reporting crash data, an attempt was made to measure the time required by those responsible for obtaining them. The team attempted to time the work of police officers at the scene. Because few crashes could be observed, several agencies also provided time on a self-reporting form. From these data, the research team was able to generate an estimate of time which they believed represented at least the minimal amount. Using the estimates provide a basis for estimating the costs of collecting and reporting data. Added were costs of their management. These latter costs came from budgets provided by agencies which the team visited.

All three processes, collecting, reporting, and managing crash data are estimated to cost 19.20 dollars per crash report filed. Using the times gathered along with estimates of changes in activity based upon the severity of a crash and the number of vehicles involved, the report is able to estimate a range of costs by severity of the crash, number of vehicles involved, and region of the country. These ranges were found to be significantly less than those estimated by others.

In order to help evaluate strategies for improving the process of handling the data, the report defines a set of goals and objectives, along with principles which govern implementation of the strategies. Five goals and sub-goals are presented:

1. Minimize costs collecting and managing the data;

2. Maximize the safety in collecting the data,
   a. maximize effectiveness of the operation at the scene;
   b. Minimize the exposure of involved parties;

3. Minimize effects on the organization;

4. Maximize quality of the data for users, and

5. Maximize suitability of the data for the users.

Established for each of the goals and sub-goals are objectives along with measures of effectiveness. A set of principles also was identified. These prescribe expectations affecting the system, persons who work in the system, and data from the system.

As a prelude to identifying strategies, the report describes practices currently undertaken by states and local agencies which attempt to resolve various issues. Many of the practices are unique and have proven effective, although few have been evaluated formally.

Finally, the research team identified 23 strategies which were capable of being introduced without requiring substantial additional effort. Each of these strategies was evaluated as it related to meeting the goals and objectives, and to reducing costs of operation. Effort was made to limit strategies to those which were non-technological in nature, particularly as they applied to collecting and managing highway crash data; however, technologies were considered where appropriate. The description and evaluation of strategies displays 11 strategies which should reduce the costs of collecting and managing crash data; nine are considered to have a relatively high cost-effectiveness. For the roadway inventory, three strategies were considered to have relatively high cost-effectiveness. These strategies were technical in nature. Seven strategies, also technical in nature, were recommended for collecting and managing emergency medical data.

Finally, the report concludes with discussion of how strategies may be grouped and introduced as a package. In many cases, strategies do not stand alone; they can be combined and introduced as a group. Regardless of the method used for introduction, many of the strategies are considered important enough to be implemented now. Even though some are technical in nature, their application is possible in the current environment and their costs are reasonable. All of the recommended strategies will provide both short- and long-term benefits.

Highway Safety Data:
Costs, Quality, and Strategies for Improvement
Research Report

1. Introduction

Scope of the Project

The analysis of highway-safety data addressed issues related to collecting and storing these data along with strategies to address the issues. Four objectives guided this analysis:

1. Determine the costs currently associated with collecting and managing highway-safety data with emphasis on traffic crash reporting.

2. Determine the quality of data collected for three sources: traffic crashes, roadway inventory, and emergency medical services provided to crash victims.

3. Identify issues which affect the collection, storage, and use of data from the three sources.

4. Recommend strategies for reducing costs, improving quality, and resolving other issues which affect the collection, management, and subsequent use of the data.

Three key data sources, crash, roadway-inventory, and medical files, are used by the safety community to help select countermeasures, and to serve as a basis for evaluation of safety programs and applications. The data form an integral part of the safety management system (SMS) concept commissioned in the Intermodal Surface Transportation Efficiency Act (ISTEA). SMS provides assistance in making cost-effective decisions.

The research team focused upon collection and management of the data. How users employed the data were outside the scope of the project except as how users might affect collection and management of those data.

This document provides a summary of the research methods used. It describes what the research team learned and provides descriptions of strategies which will provide a better product. Included in the activities were the following:

- Initially defining the problem based on the literature and contacts with State highway-safety personnel to determine what individual States had been doing.
- Designing methods to collect additional information particularly through visits to various agencies selected because of practices which appeared noteworthy.
- Developing a cost model.

- Identifying users and uses of highway-safety data.
- Identifying issues associated with the use of data and a set of strategies designed to address these issues.
- Assessing the strategies to establish relative cost-effectiveness.

The last two activities relate to strategies which provide the guidelines for improving the highway-safety system. Implementation will be of value to the managers of the systems.

The Highway-Safety System

Managers of the highway system will use an information system which supports managerial activities and decision making. The decisions can be divided into four primary areas:

1. Problem identification.
2. Countermeasure selection.
3. Program implementation and operation.
4. Effectiveness evaluation.

The manager selects guidelines, policies, programs, etc. to improve the operation of the highway system. While some of these are directed at the highway-management system itself, most become inputs to the highway system, which in turn produce mobility, risk, and other socioeconomic impacts.

The focus of this project is the safety-information system which assists decision makers. Safety data are collected, in part, because of the desire to provide decision makers with information regarding past conditions in the highway system. Because the purpose of the information is to support decision-making, those data should be selected and defined specifically to serve that purpose.

Data regarding traffic crashes also are collected and processed for use by the social system for related activities, e.g., criminal- and civil-court actions, and for other administrative needs remotely connected to highway-safety functions. Because the scope of this project is intended to include only data required for safety analyses, a careful definition is provided for elements to be included within the rubric of safety data.

Highway-Safety Data

Safety data for highway crashes derives primarily from police crash reports. However, a number of other agencies or individuals obtain or provide safety data in the course of performing functions not directed toward the four areas of activity listed above. These occur at four levels within the system; Federal, State, local, and private. The prime sources are shown in table 1. With the exception of traffic law enforcement, a preponderance of the data available is used for purposes other than highway safety.

Addressing all of the sources identified in the table is not feasible. The primary focus selected for the project was the police crash report (PCR) produced by traffic law enforcement. However, the use of roadway-inventory data from engineering also was selected to be addressed, with emphasis on road attributes. Finally, medical sources were considered, where feasible, primarily to evaluate them for acquisition of data on injury severity and driver condition prior to the crash.

Table 1
Sources of Highway-Safety Data

1. Traffic law enforcement (TLE);
2. Engineering;
3. Medical;
4. Motor vehicle administration (MVA)
5. Automobile insurance industry;
6. Adjudication;
7. Education;
8. Highway-safety agencies;
9. Involved persons; and
10. Involved private companies/agencies.

The crash report is used for two primary purposes:

1. Ease management.
2. Safety analyses.

Case management involves use of the PCR for dealing with the individual occurrence, as may be required for law enforcement, adjudication, sanction and insurance purposes. It may also include uses by governmental agencies for such things as rectifying government-property damage, or ensuring financial responsibility. These uses are often considered the primary reason that law-enforcement personnel are used to investigate crashes. Some of the objectives of case management include creating a safer highway. However, they do not specifically relate to highway-safety analyses.

In general, safety analyses are performed on sets of records from PCR's, or aggregations derived from sets of PCR's. Safety analyses may address individual crash reports, but this is usually in the context of trying to establish underlying patterns. Examples of data that are useful for safety analyses include the driver's intended maneuver, the location of the crash, the manner of collision, and the contributing circumstances.

Differentiation between the two functions of the data when costs are estimated for collection and management is important. The costs associated with case management should not be assigned to safety-data collection and management. Where the same data are used for both purposes, a logical rule is needed for handling the associated costs. This is discussed further under the section on cost.

Because the agencies listed in table 1 (with the exception of the highway-safety agencies) are collecting the data of interest for principal reasons other than safety analysis, the costs associated with them are not to be considered safety-data costs. That is, the costs experienced by these other agencies in collecting their data would be experienced whether or not the safety analyses were being performed. However, if the data from these sources are to be used in safety-data analyses, the cost of acquiring those data from the alternative sources should be included in cost estimates.

Organization of the Report

This report contains 8 chapters. The next chapter identifies processes employed for collecting and analyzing data for the project. Chapter 3 addresses uses and users of highway-safety data. Collection and analysis of safety-data are described in the fourth chapter. Costs of collecting these data form the basis for chapter 5. Goals, issues, and noteworthy practices appear in chapter 6. The seventh chapter presents strategies and analysis of their cost-effectiveness and ties together the previous discussion. Finally, chapter 8 provides a summary of the report.

2. Processes Employed to Collect and Analyze Data for the Study

Initial Work

The study set out to determine what was known related to costs and quality of highway-safety data, then to fill key gaps in the knowledge. For this purpose, initial work was based upon:

1. Literature reviews.
2. Responses received from an initial inquiry to the Governor's Highway Safety Representative of each State.
3. Pilot studies conducted in one State.
4. Previous knowledge of the project team.

As a result, the team discovered a large body of knowledge regarding the functions and issues related to the collection and management of crash data, but limited published material regarding the other two sources, roadway-inventory and emergency medical services data. The literature describing costs of obtaining and managing traffic-safety data was almost non-existent. The initial work produced:

1. A synthesis of the literature.
2. A plan for field data-collection.
3. A framework for analyses.
4. Preliminary models for estimating the cost of collection and management of crash data.

Data Collection

Selection of Sites for Visiting

An integral part of the project was visiting traffic-safety agencies in several States to document processes and issues related to crash data, roadway inventory, and emergency medical services (EMS) and medical (trauma) data management. Sites selected were considered representative of the Nation. Selection was based upon:

1. The response to the initial letter of inquiry sent to each State.
2. Specific attributes of interest regarding the safety-data systems in each State.
3. Geographic distribution.
4. The existence of urban centers in the State with major highway-safety system operations.
5. Opportunity for scheduling some visits in the course of other project-team travel.

A letter was sent to the safety-data agency director in each State. This letter requested information about their processing of safety reports, copies of crash reports and manuals associated with them, and contacts for safety-data reporting, roadway-inventory, and medical data. Thirty six States responded. Even though they had not responded to the initial survey, the States of Michigan, New Mexico, and Utah were added because they maintained exemplary practices which the project team believed should be studied, or they were members of the Highway Safety Information System (HSIS) operated by the U.S. Department of Transportation, Federal Highway Administration (FHWA).

From a matrix of attributes about each State, decisions were made as to what sites should be visited. Data about attributes were obtained through telephone interviews of safety-data base managers or State traffic engineers. Supplementing the interviews was information gathered from a number of sources including personal contacts. The States also were organized by region - northeast, south, north central, and west. The list of States selected as sites appears as table 2.

Table 2
List of States Selected for Visits

New Jersey	Michigan
Pennsylvania	Wisconsin
Virginia	Illinois
Florida	Utah
Ohio	California

Characteristics of the selected States appear in table 3 (on the following pages). States A, B, G, I, and J served as primary ones. The work done at secondary sites was a sub-set of the comprehensive package. Most often, the work at secondary sites focussed upon documenting the cost and procedure for managing PCR data. At the five primary sites the research team performed a comprehensive set of activities, including:

1. Direct measurement of police officer's PCR completion-time.
2. Documentation of PCR data-management costs and procedures.
3. Documentation of roadway-inventory data collection and management procedures.
4. Documentation of medical data collection and management procedures.
5. Conduct of focus groups with users of highway-safety data, regarding uses of the data.
6. Documentation of noteworthy practices.
7. Acquisition of selected crash-reports and crash-files for further assessment.

The results of the visits are summarized in several reports, the most comprehensive of which contained a review of safety-data management, roadway inventory, costs, and exemplary practices. This was the *State Data Book* submitted separately to the FHWA with a copy also maintained at the Transportation Library, Northwestern University.

PCR Completion-Time

Limited data available in the literature regarding the time that a police officer, or other type of data collector, spent at traffic-crash scene, and how much of that time was spent upon safety-data collection. The project team devised a simple form to measure time an officer spent in safety-data collection along with other activities at a crash scene. Direct measurement confirmed an anticipated low productivity for obtaining direct measures of PCR completion-times. Therefore, the team devised a self-reporting instrument for officers to complete for the project. While this approach had less validity than direct measurement, the potential sample sizes and cost-effectiveness was considered superior. Adjustment of self-reporting by officers to account for inaccuracies, by relating it to direct observations, proved more difficult than anticipated.

Data-Management Costs and Procedures

Similar procedures were used to document data-collection and management procedures for PCR's, roadway inventories, and medical sources. In each case principal managers of the data at the State level were interviewed regarding the process. Representatives from local agencies, which represented data collectors, also were interviewed. Furthermore, in the case of PCR management, the interviewees were asked to provide as detailed a set of information as feasible regarding the costs associated with the operation. In many cases, more than one State agency was involved in managing PCR data. In such cases, each agency's representative(s) was interviewed. In almost every case, the project team observed the process, as well as collecting manuals, diagrams, and other materials which documented the procedures in more detail than could be derived from the interview.

Conduct of Focus Groups with Users

During visits to primary State-sites, at least one focus group session was held with users of highway-safety data. The groups averaged about five users, some had twice the average. Although the selection of users was left to the ultimate judgement and willingness of the local host, the groups were generally well rounded with local and State users from a variety of disciplines, and even participation of representatives of a few private groups. The group first was asked to describe their agency's functions and what safety analyses they performed. One of the project team members serving as moderator, selected several examples of safety analyses and requested members of the group describe how the prepared and used the analyses in detail. The group also was asked to comment on the quality of the data, using a set of definitions for quality that were provided by the group leader. Finally, they were asked to describe any analyses they were unable to do because of lack of data, or accessibility to the data. The summaries of these sessions provide insights for the assessment of data quality, and was used for identifying issues and designing strategies.

Documentation of Noteworthy Practices

During the course of interviews and focus-group sessions, as the team members identified some practice being used that was considered unique or exemplary, they would pursue documentation of it. A general framework was developed for the interviewer to follow, but the interviewees were given complete latitude to describe the practice in a manner they saw fit.

Table 3
Attributes of State Systems Studied

States Selected

Attribute	A	B	C	D	E	F	G	H	I	J
Total Crashes Procsd.	137,822	122,887	50,660	113,913	356,458	280,258[1]	364,848[2]	410,980	196,176	491,000
Fatal Crashes	579	760	269	1,385	1,303	966	1,290	1,249	2.218	3,707
Injury Crases	40,792	50,865	22,490	86,310	119,509	98,512	97,478	103,246	125,338	215,000
Prpty. Damage Only	96,450	71,262	27,901	46,218	235,646	180,780	265,130	306,485	68,620	272,000
Reporting Threshold	$500	$1000	$750	Injury or tow[3]	$150	$500	$400	$300	Injury and fatal[4]	$500
Single Form Used?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes[5]	Yes	No
Motorist Report?	No	No	No	Yes (rarely used)	No	No	No	Yes	No	No
Short Form?	No	No	No	No	No	No	No	Yes	Yes	No
Type of Form	Traditional[6]	Template	Template	Traditional	Traditional	Traditional	OMR	Template	Traditional	Traditional & Computer based
Principal Agency Managing	Transportation	Motor Vehicles	Public Safety,[7] Transptn.	Transportation	Highway Safety	Transportation[8]	State Police	Transportation	Highway Safety	Highway Patrol
Other Agencies Managing	none	State Police (commercial vehicles)	(see above)	none	none	Motor Vehicle Services, Highway Safety	none	Secretary of State, State Police	Transportation, Health & Rehabilitive Services	Motor Vehicles, Transportation
Manner of Data Entry	One-step - full entry	Two steps - skeleton, full entry	One-step - full entry	Two steps - skeleton, full entry	Two steps - skeleton, full entry	One-step - full entry	One-step - scan then keyed	One-step - full entry	One-step - full entry	One-step - full entry

Acquisition of Selected Crash Reports

During the course of the interview with the primary manager of PCR data, a request was made that the project be provided with copies of between 25 and 50 crash reports as they existed prior to any handling. In addition, a printed copy of the records created from the selected set of crash reports was requested. The selection of the set of crash reports was left to the agency, to avoid creating undue burden by imposing a randomized process. The agency knew, however, that the purpose of the request was so that the project team could analyze the quality of the original PCR and the degree to which the management process improved the quality. Review of reports received suggested that agencies selected reports from a recently arrived batch without any further selectivity applied; therefore, the sample appeared to be representative of that agency's handling of crash reports in general.

Interviews of Managers of Federal Data Bases

The cost and quality of highway-safety data bases cannot be fully addressed without considering the files operated by Federal agencies. Seven were chosen as the most relevant to this project. The manager of each was interviewed regarding data collection and management of the systems, as well as associated. The managers also provided materials about their respective data bases.

Analyses

A series of analyses were performed on the data collected during the site visits, as well as that acquired from other sources. These include:

1. Tabulation and development of descriptive statistics for PCR-completion times.

2. Derivation of total and unit costs for PCR data-management.

3. Assessment of crash-report data quality by analyzing copies of PCR's as submitted by law enforcement agencies, and comparing them with the final data base record created from each PCR.

4. Derivation of estimates of national frequencies of highway crashes, by type, using the General Estimates System (GES), maintained by NHTSA.

5. Derivation of estimates of average law-enforcement personnel rates, using a data base available through the Law Enforcement Management and Administration Statistics (LEMAS) data base, maintained by the Bureau of Justice Statistics of the U.S. Department of Justice.

6. Summary of costs and attributes of the selected Federal data bases.

Table 3
Attributes of State Systems Studied (continued)

States Selected

Attribute	A	B	C	D	E	F	G	H	I	J
Manner of Editing	Online, Batch	Online, Batch	Online	Pre-edit, Online, Batch	Pre-edit, Online, Batch	Manual, Batch	Online, Batch	Online, Batch	Batch	Batch
Manner of Crash-Data Coding	Reliance on officer coding	Reliance on officer coding	Reliance on officer coding	Maximum degree of interpretation	Reliance on officer coding	Reliance on officer coding	Reliance on officer coding	Reliance on officer coding	Reliance on officer coding	Moderate degree of interpretation
Manner of Location Coding	Manual (prison system)	semi-automated, (computer codes)	Manual	Automated	Automated (tape transfer)	Manual (prison system)	Automated	Manual	Manual	Manual
Type of Database	Relational	Relational	Flat file	Relational	Relational	Flat file[9]	Flat file	Flat file	Flat file	Flat file
DB Linkages Used for Data Entry	Veh & DL, EMS (planned)	Roadway inventory	none	Veh & DL, Rdwy inventory	Tape transfer	none	none	none	none	none
Data Mgmt Cost per PCR per Year	$6.30	$9.50	$20.10	$8.50	$7.04	$7.10	$2.40	$2.40	$5.80	$5.80

[1] latest available statistics are for 1988

[2] 1991 data

[3] tow is required because of crash damage

[4] accepts police report of PDO on short form, but not motorist's

[5] slight variation of form for one city

[6] changing to optical mark recognition in 1994

[7] both the Departments of Public Safety and Transportation maintain crash safety data bases; Public Safety codes data first and concentrates on driver and vehicle data primarily for financial responsibility purposes

[8] all reports go to Division of Motor Vehicles first; then are distributed

[9] changing to relational

The voluminous notes and materials collected for the project have been synthesized to a great extent into a set of summary materials. These were distributed to the FHWA as a site data book.[1]

[1] A copy also is available at the Northwestern University Transportation Library.

3. Uses and Users of Highway-Safety Data and Issues Affecting the Uses

A Framework for Users and Uses

A variety of public and private entities use or potentially would use highway-safety data if it were available in a convenient form. Each of these entities has a set of objectives. To achieve the objectives, the people in the entity require information and carry out activities based on the information derived from the data. The hierarchy of organizing information is shown in figure 1.

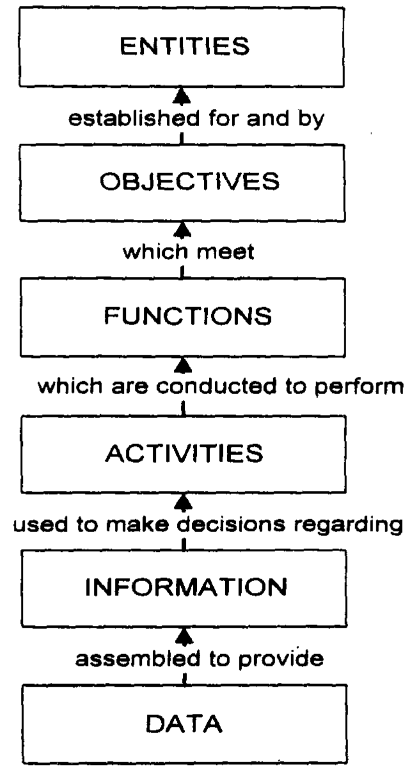


Figure 1
Flow of Information to Users

As an example, the objectives of a state department of transportation include the design of safe highways. The highway engineering function, i.e., planning, design, construction, operation and maintenance of highways is required to achieve this objective. The activities include the development of design standards that reflect safety considerations. To arrive at those safe standards requires information on the relative safety of different highway geometrics. That information is generated by combining crash and roadway data with other data, and performing the appropriate statistical analyses.

Affecting the users and uses are issues related to collecting and managing data from the three sources: police crash report, roadway-inventory, and crash injury reporting. During site visits and resulting from discussions with collectors, managers, and users of the data, the project team identified a number of important issues. These issues are addressed at the end of this chapter. Chapter 6 identifies noteworthy practices that have been taken at State and local levels to help address many of the issues.

Traffic-Safety Data Uses

Listed in table 4 are the functions that the highway-safety data users perform which are likely to require the use of safety data. This list is not exhaustive, and the entities' reliance on safety data for their functions ranges from minimal to extensive. The activities which are conducted by these data users can be grouped into five basic types, as they relate to the use of traffic-safety data:

Table 4
Users and Uses of Highway-Safety Data

Data User	Functions of this User
Law Enforcement Agencies	Enforcement of traffic laws (at all levels of government);
Traffic Safety Administration	Developing annual highway-safety plans including program funding;
Highway and Public Works Departments	Planning, design, construction, operation and maintenance of highways;
Motor Vehicle Administrators	Licensing, registration and control of drivers and vehicles for highway use;
Medical Providers (EMS Systems/Hospitals)	Management and treatment of highway trauma including emergency response and plans for patient receipt;
Adjudication System	Administration of law to highway operation;
Legislators/Regulators	Enactment of laws and regulations to develop and control the highway system;
Insurers	Provision of insurance for road users;
Schools (Public and Private)	Training highway users;
Commercial Road Users	Commercial transport of goods, and public transport of people;
Vehicle Manufacturers	Design, manufacture and repair of vehicles; and
Independent Research	Examine causes of, and potential remedies for, crashes but not as a part of other entities here.
Other Interest Groups	Representation of special interests related to highway safety (National Safety Council, American Automobile Association, Motor Vehicle Manufacturers Association, Mothers Against Drunk Driving).

1. Case Management - The handling of individuals, vehicles, and other property associated with a single crash.
2. Problem Identification - Using safety information to determine the existence and nature of safety-related problems.
3. Countermeasure Selection - Using safety information to guide choices actions.

4. Program Evaluation - Using safety information as part of the information set to determine the effectiveness of an action or program.
5. Administrative - Using safety information to support planning and operational decisions of the entity.

Activities for Which Highway-Safety Data Are Potentially Useful

Combining the functions for highway-safety data users and related types of activities provides more specific examples of what occurs within each of the entities. Table 5 summarizes uses. Again, the focus is on any use of safety data rather than on the extent of that use. Quantifying use is difficult; further functions of the various entities frequently overlap or are otherwise replicated at different levels of government.

Not all entities will conduct all activities in performing various functions nor will different entities conduct similar activities even though they may be performing the same function. In general, the safety-management, legislative, and research functions may carry out activities which address any aspect of other functions.

Project Scope

This project focused upon the middle three activities described above: problem identification, countermeasure selection, and program evaluation. Case management and administrative activities, while arguably related to safety analyses are addressed minimally. The scope also was limited to the data collection and management aspects of the process. Additionally, this project has focused upon three data sources out of the many which can provide data related to safety needs:

1. The crash report, produced primarily by law enforcement agencies.

2. Roadway inventories produced primarily by highway engineering agencies.

3. Medical crash data produced primarily within the trauma care system (limited to driver condition prior to the crash, and severity of injuries resulting from the crash).

Examples of Data Uses

During the site visits conducted as part of this project, a variety of uses and users of highway-safety data were discussed and observed. Users represented most of the entities listed above. Uses relied upon data from the three primary data sources listed.
The number of potential uses of highway-safety data is quite large along with the number of data user types. This project addressed how innovative approaches to the collection and processing of highway-safety data facilitated use of the data. It would be very difficult to assess how even a single new collection or processing strategy would affect all uses and

users of the data. When a set of such strategies is considered, assessing them across all uses is clearly not possible. Instead, a set of scenarios was developed as described in chapter 7 of this report. Evaluation of the strategies was based on these scenarios.

Table 5
Use of Highway-Safety Data: Functions and Activities

Function	Case Management	Problem Identification	Countermeasure Selection	Program Evaluation	Administrative
Enforcement of Traffic Laws	a. Track prosecution of violators. b. Perform crash analysis and reconstruction. c. Provide court liaison and testimony. d. Request driver re-evaluations to propose license cancellations or revocations.	a. Analyze attributes of crashes and violators (for selective enforcement). b. Identify problem drivers, locations, environments, laws and regulations.	Choose from among alternative selective enforcement, traffic direction and control, public information, education, emergency medical service and/or escort programs and legislative or regulatory actions to impact suspected problems.	Determine appropriateness of terminating, modifying, or expanding existing programs.	a. Staff Resource Allocation - Determine the amount and schedule of personnel to perform duties and carry out programs. b. Implement selected programs. c. Support tort liability defense.
Administration of Highway Safety Programs		Analyze attributes of crashes and violators, and identify problem vehicles, drivers, locations, highway features, laws and regulations.	Evaluate and/or select among administrative, highway engineering, enforcement, public information and education, medical, legislative, regulative and/or vehicle design strategies to correct suspected problems.	Determine appropriateness of adding, terminating, modifying and/or expanding existing programs.	
Planning, Design, Construction, Operation and Maintenance of Highways	a. Identify high-hazard locations. b. Determine needs for engineering action at specific locations (spots, segments and areas).	Evaluate and select among proposed alternative engineering actions, including alternative facility and network designs.	Evaluate the safety impact of changes in standards and practices, as well as recently implemented highway facilities and controls.		Support defense of tort liability actions.
Licensing, Registration and Control of Drivers and Vehicles	a. Provide driver and vehicle data to legal system. b. Conduct Hearings (where authorized). c. Examine and license drivers (e.g., use of "point" systems for violations and crashes). d. Assure financial responsibility. e. Track prosecutions. f. Provide court liaison and testimony.	Identify problem drivers (by attributes such as age, area of residence, type of vehicle driven, etc.).	Evaluate and select among alternative driver control and improvement actions (e.g., warnings, training, suspended license, revoked license, alternative sanctions), legislative and regulatory changes, enforcement activities and/or public information and education programs.	Evaluate the effectiveness of programs undertaken to determine if they should be modified, extended, or terminated.	

Table 5
Use of Highway-Safety Data: Functions and Activities (continued)

Function	Case Management	Problem Identification	Countermeasure Selection	Program Evaluation	Administrative
Management and Treatment of Highway Trauma	Evaluation and treatment of injuries to persons involved in traffic crashes, relating crash specifics to injury specifics.	a. Identify emergency response and trauma care system inadequacies and/or failures (e.g., response time, quality of diagnosis, quality of treatment, effectiveness of treatment, number of personnel, type and amount of facilities, and amount of associated costs). b. Identify vehicle, highway and human attributes which contribute to severity of injury, and degree of rehabilitation required.	Identify and evaluate alternative means for overcoming identified problems.	Evaluate the effectiveness of programs undertaken, to determine if they should be modified, extended, or terminated.	
Administration of Law Applied to Highway Operation	Track and manage prosecutions and dispositions, as well as establish sanctions.	Evaluate operations and monitor performance of judiciary and other officers of the court.	Identify and evaluate actions to overcome identified problems (e.g., alternative sanctions, administrative adjudication, judicial and/or attorney training).	Evaluate the effectiveness of programs undertaken, to determine if they should be modified, extended, or terminated.	Employ and schedule personnel and provide facilities to process cases in an efficient and just manner.
Enactment of Laws and Regulations to Develop and Control the Highway System		Identify policies, programs, regulations, legislation, funding levels, vehicles, and elements of the highway system which present hazards to the motoring public.	Identify and evaluate legislative and regulatory strategies to overcome the identified problems.	Evaluate the effectiveness of legislative or regulatory actions to determine if they should be modified, terminated, or maintained.	
Provision of Insurance for Highway Users	Claims settlement and legal defense.	Identify vehicular and highway system problems which result in claims being filed, as well as identifying problem drivers.	Identify and evaluate alternative means for reducing claims frequency and amounts.	Evaluate the effectiveness of programs undertaken, to determine if they should be modified, extended, or terminated.	Perform actuarial functions.

Table 5
Use of Highway-Safety Data: Functions and Activities (continued)

Function	Case Management	Problem Identification	Countermeasure Selection	Program Evaluation	Administrative
Training Highway Users		Identify patterns of driver attributes particularly associated with crashes.	Identify and evaluate alternative training content and methods to affect driver, pedestrian and bicyclist attributes. The focus will be on those attributes associated with crashes which, when corrected, have potential for reduction of crash frequency and severity.	a. Assess the effectiveness of driver training programs in enhancing highway safety. b. Evaluate the effectiveness of public information and education activities conducted in the school system.	
Transport of Goods and Public Transport of People	Claims settlement and legal defense.	Identification by fleet managers (e.g., of large trucking firms, transit agencies, taxicab companies) of vehicular and driver problems which contribute to crashes and/or increase injury severity.	Identification and evaluation of fleet managers of alternative means of reducing crash frequency, severity and costs.	Evaluation by fleet managers of the effectiveness of programs undertaken, to determine if they should be modified, extended, or terminated.	Assess actual and potential liability for fleet.
Design Manufacture and Repair of Vehicles	Claims settlement and legal defense.	Identify vehicular problems which contribute to crashes and/or increase injury severity.	Identify and evaluate alternative vehicle design, manufacturing and/or repair actions for reducing crash frequency, severity and costs.	Evaluate the effectiveness of designs, practices and programs undertaken to determine if they should be modified, extended, or terminated.	Assess actual and potential liability for design and manufacturing.
Independent Research	While much highway safety research takes place either within the other entities described above, or under contract to them, there is additional research conducted outside of the highway safety management system. Much of this research takes place at colleges and universities. Often the purpose of this research is to test a hypothesis and not necessarily to support management decisions. As such, the activities of this kind of research do not fit into the "case management, problem identification, etc." structure since the research is not necessarily oriented to a specific function.				
Representation of Other Interests Related to Highway Safety	Public and private interest groups exist for almost every aspect of highway safety. With regard to safety data, they may be considered similar to the highway safety management function. Current areas of emphasis for these groups include DUI, occupant restraint use, records systems, and vehicle design.				

A Taxonomy of Safety Data

Police crash reports. To facilitate discussion and analyze the nature of the data that are applicable to safety analyses, a taxonomy of data was created for the crash report which is usually completed by investigating police officers. This appears as table 6.

Table 6
Taxonomy of Traffic-Crash Data:
Categories and Types

		A	B	C	D	E
DATA CATEGORY		INVOLVED PERSONS	INVOLVED TRAFFIC UNITS	ENVIRONMENT	GENERAL CRASH DESCRIPTION	ADMINISTRATIVE
	1	Identifiers	Identifiers	Road Attributes	Location	Report No.
DATA	2	Injury	Damage	Weather	Date/Time	Jurisdiction
TYPE	3	Condition	Condition	Visibility	Collision Type	Agency
	4	Tests	Disposition	Damage	Facts	Collector Id.
	5	Treatment	Maneuver	Traffic Control	Accident Type	Investigation
	6	Attributes	Attributes			Enforcement
	7					Processing
	8					Medical

Many data elements are collected about a traffic crash as a basis for both case management and safety analyses. For example, the police crash report (PCR) contains information regarding personal identifiers, vehicle identifiers, and agency identifiers which are needed to provide:

- Insurance companies with information about the parties and vehicles involved.
- Courts with similar information.
- Law enforcement administrators with operational data.
- Linkages with other relevant data bases.

While some aspects of case management may be linked to highway-safety efforts, the relationship is minor when compared to the other types of safety analyses listed above. Examples of data that are useful for safety analyses include the driver's intended maneuver, the location of the crash, the manner of collision and the contributing circumstances.

Roadway inventories. The taxonomy adopted for roadway-inventory data is shown in table 7. Unlike the use of elements from the PCR by multiple users, those elements which appear on the roadway inventory are used almost exclusively by the highway engineering community.

Emergency medical services. The reporting of data from emergency medical services and medical trauma as related to vehicular crashes is not standardized. Further, consideration of medical data was limited to "driver condition prior to the crash" and "injury severity" (generally through use of the KABCO scale). Issues and potential strategies which apply to medical data collection are discussed in chapter 7.

Table 7
Taxonomy of Roadway-Inventory Data:
Categories and Types

DATA CATEGORY		A	B	C	D	E	F	G
		LOCATION	ALIGN-MENT	CROSS-SECTION	PAVEMENT	TRAFFIC	ADMINISTRATIVE	.MISCELLAN-EOUS
DATA TYPE	1	Govern. Unit	Horizontal	Lanes	Type	Volume	Dates	Special Features
	2	Referen. System	Vertical	Shoulders	Condition	Composition	Jurisdictions	Structures
	3	Route		Roadside		Controls	Linkages	
	4	Intersection		Median		Restrictions	Year Built	
	5	Direction		Appurtenances		Signals	Modifications	
	6			Clearances				

Data Needs for Safety Analysts

A brief general summary of data needs of highway-safety analysts appears in table 8. It provides an overview of needs across all highway-safety functions for the data types identified above. Each of the three major data sources selected for study in this project is addressed.

Classification of Issues

During the course of identifying how highway-safety data were gathered and managed, the project team documented a number of issues and noteworthy practices at sites throughout the Nation. This chapter briefly describes important issues. Framework for the discussion is that portion of the highway-safety information-system related to collecting and managing data. The data-collection phase can be further divided into collecting data at the scene and away from the scene. The issues relate to the three sources of data under discussion: crash, roadway inventory, medical. Table 9 shows nine possible classifications for the issues based on phase and component. These classifications will serve as the framework for subsequent discussion.

Data collection is the phase of obtaining and recording data about the crash or roadway which will become part of the highway-safety information-system. Data about a vehicular crash may be obtained both at the scene and away from it. The latter activity is used to add facts that the investigators were not able to obtain at the scene. Much of the collection of trauma

Table 8
Data Needs of the Highway-Safety Analysts

Source of Data	Data Types
Crash Report	1. Involved Persons - There is limited interest in specific identifiers of involved persons, other than for providing linkages. The primary interest is in the following data types: - Injury - Condition - Tests - Treatment - Attributes 2. Involved Traffic Units - The nature of interest is similar as for involved persons. The primary interest is in the following types of data: - Damage - Condition - Maneuver - Attributes 3. Environment - Interest is potentially high in the all of the data types included in this data category. The need of any specific analysis will depend upon its objective and scope. 4. General Crash Description - Interest is potentially high in all of the data types included in this data category. The need of any specific analysis will depend upon its objective and scope. 5. Administrative - The data of interest here are primarily for the purposes of data linkage and, possibly, jurisdictional variations.

Table 8
Data Needs of the Highway-Safety Analysts (continued)

Source of Data	Data Types
Roadway Inventories	1. Location - Location data are primarily used to provide linkage to crash data for merging data elements for analysis. 2. Alignment, Cross Section and Pavement - alignment, cross section and pavement data are primarily used by engineers for analysis of specific locations, determination of design standards, and construction needs assessments. Insurance industry users could have an interest in these data in a manner similar to training agencies and citizen groups seeking to do problem identification, or develop public education and information programs and materials. 3. Traffic - Traffic volume and composition data are primarily used in performing highway engineering and general safety management functions, to provide an exposure base for safety analyses. Traffic control data are used primarily by highway engineers. 4. Administrative - Data on date of construction, as well as modifications, can, be useful in trend analyses. Otherwise, administrative data generally are not used for safety analyses, but are employed by the data base administrator. 5. Miscellaneous - Miscellaneous data include special features or structural information that will be of interest to the same potential users, as may employ alignment and other geometric data in their analyses.
Medical Reports (EMS, Medical Examiner, Hospital and Trauma Records)	1. Driver Condition Prior - Driver condition is not readily diagnosed by a police officer or other non-medical data collectors at the site of a crash. The medical diagnoses performed by the trauma care system provide the potential for better quality data in this critical area than can be obtained elsewhere. 2. Injury Severity - Although many analysts need only a severity classification of fatal, personal injury and property damage crashes, the injury descriptions obtained from crash reports are not sufficiently accurate or precise for some analyses. The more informed and trained observers in the trauma care system may provide a useful alternative source for any user interested in details on injury severity.

and related medical data, e.g., treatment, admittance, and outcome, are collected at the hospital or other treatment sites which are away from the scene.

Data management is the process of entering data into manual and automated files, and maintaining those files. Where automated systems are used, data are entered into a data base through keying, scanning, or image processing. What is stored should be available for those who require information about highway safety.

Table 9
Classification Framework for Handling Highway-Safety Data

		Sources of Highway-Safety Data		
		Traffic Crashes*	Medical*	Roadway Inventory
Data Collection	A. Collecting			
	B. Reporting			
Data Management	C. Data entry and storage			

* can be "on-scene" or "away-from-scene"

The issues affect costs, quality, safety, and protection of privacy, often in combination. Resolution of issues may involve expenditures or redirection of processes. However, it often faces a conflict between the provision of safety data and management of the incident, which includes any case deriving from that incident. The discussion below highlights issues which the project team noted during their study of processes at sites nationally.

Issues Affecting Collecting and Reporting Data

Traffic crashes. After the responding officer has arrived, assessed the scene, and taken steps to stabilize the incident, he begins to collect and record data about the crash. These data will be used for case management and for crash reporting. In addition, the officer may leave the scene to collect additional data. The most common case is when the officer goes to the hospital to collect more information about injuries.

As the data are collected, the officer records them. The standard form is the police crash report (PCR) which contains information about the crash and serves for both case management and safety data. The officer handling the crash, support officers, or special investigators, such as the accident reconstruction officer, may prepare supplementary reports. Generally, these reports are used for managing the case rather than for safety data. An exception would be the Fatal Accident Reporting System (FARS) analysts at the State safety agency who might used the supplemental reports for additional data required by the FARS data base. A summary of issues related to collecting safety-data from crashes appears in table 10. Table 11 summarizes issues related to reporting those data.

Roadway inventory. Much of the data used by the roadway inventory is collected on-site and reflects the quality of that collection. Additional data come from construction and maintenance records. Generally these data are submitted to the State department of transportation where they are recorded on the roadway-inventory file.

Table 10
Issues Related to Collecting Crash Data

Collecting Data - at the scene of the crash 1. Potential conflict among officer roles at a crash scene may degrade the quality of safety-data collected (roles noted in priority order): a. determine needs for further response b. stabilize scene to prevent further harm c. traffic control d. manage operations at the scene e. collect data for managing the case; begin in-depth investigation f. enforce traffic and other laws g. collect crash safety-data 2. Perceived need by officer to establish fault, for insurance and enforcement purposes, may bias the collection and interpretation of data 3. Inconsistent coverage of vehicular crashes yields biased samples and includes - state and locally mandated reporting thresholds eliminate crashes, and - lack of police resources and motorists not reporting minor crashes to the police. 4. Incorrect use of resources may reduce quality of data collected resulting from: - lack of clearly defined roles for police at the scene, and - distractions while collecting and recording data. - multiple data gathering without proper communications of findings 5. Inaccuracy and imprecision in locating the crash, positions of vehicles and persons involved in the crash, and crash debris reduces usefulness of data for later analyses. 6. Collection of data *perceived by the officer* as having limited use will reduce the incentive to provide quality data. 7. Danger to officer at the scene reduces incentive to complete a thorough investigation of crash. 8. Inadequate initial and refresher training results in low-quality data gathering, as well as - limited feedback received by investigator, and - infrequent opportunities to investigate crashes 9. Adequate tools or means for collecting some of the data are not available, resulting in low quality including inaccurate or incomplete data. 10. Disturbance or destruction of the scene by traffic, fire, rescue, tow vehicle operators, as well as less well trained police officers.
Collecting Data - away from the scene) 11. Needs for gathering additional data from vehicles and persons may affect completeness of crash report, particularly when data are required for case management. 12. Sources of data may not be readily accessible to the investigating officer because - extent of injuries and medical treatment are not available; - vehicles taken to different locations, and - reluctance of witnesses to be interviewed

Tools used for collection range from handwritten forms to automated processes. Even with an automated process, there may not be verification built in during the collection of the data. One of the most common forms is use of photologging whereby roadway data are captured on videotape. These tapes may or may not be converted into a data base. Table 12 addresses issues related to collecting and reporting roadway-inventory data.

Table 11
Issues Related to Reporting Crash Data

1. Requiring officers to make judgements on matters for which they do not have adequate training will result in low-quality data, for example:
   - estimating speed, extent of damage, extent of injuries, and
   - determination of fault.
2. Care given to completing police crash reports often is directly related to the severity of the crash, resulting in inconsistent reporting.
3. Inadequate initial and refresher training results in low-quality data recording.
4. Poorly defined and inappropriate data elements and codes contribute to incomplete and inaccurate recording.
5. Even when comprehensive sets of codes are available for a given element, officers tend to select and use only a few.
6. Lack of codes or elements for specific needs results in loss of data.
7. Size and complexity of crash reports appear directly related to the likelihood of incomplete or inaccurate responses.
8. Reporting data which are also available from other sources or multiple recording of the same data increases the likelihood of inconsistency among data sources.
9. Completing report away from scene results in problems with accuracy and completeness because of failure to collect or forgetting data.
10. Review and quality control varies dependent upon attention and priority given to crash reporting by agencies.

Table 12
Issues Related to Collecting and Reporting Roadway-Inventory Data

Collecting Data

1. Lack of automated equipment to assist with collecting roadway data may result in inaccuracy, imprecision, and incompleteness in data collection.
2. Longitudinal geometrics (horizontal and vertical curvature) are not collected and will affect relating the crash to roadway characteristics
3. Roadway features may not be located accurately because of lack of proper measuring tools or assistance in measuring.

Reporting data

1. Limited feedback regarding accuracy of recording reduces the likelihood of collecting high-quality data.
2. Procedures for recording data are not available during construction and maintenance of roadways which affect quality.

Medical information. Data related to injuries, initial handling of the victims, and treatment are recorded at the scene by emergency medical services (EMS) personnel, and at treatment sites (trauma centers, hospitals, and other medical facilities) by other medical personnel. The data can provide substantial information related to the potential causes of injuries, the treatment given, severity, and the prognosis for recovery.

An over-riding issue affecting both EMS and trauma data is the lack of standardized methods for recording them. Once recorded, there are few centralized repositories which can be used by the safety-data community. Table 13 depicts issues with collecting and reporting data both at the scene and away from the scene, generally the hospital. In addition to the issues which specifically relate to handling of trauma data by EMS and hospitals, some of the issues related to collection and reporting of crash data also apply to medical information.

Table 13
Issues Related to Collecting and Reporting Medical Data

Collecting Data - at the scene of the crash 1. EMS personnel may not have information about the events leading up to the crash, deformations to involved vehicles, and location of persons in the vehicles limits later diagnosis of injuries. 2. Unavailability of correct names or accurate identifiers of injured persons precludes later attempts to match injured persons with a crash report. 3. Not all injured persons are seen or handled by EMS personnel.
Collecting Data - away from the scene 4. Inability to link patient with crash precludes associating injury data and treatment with crash facts. 5. Confidentiality of medical records precludes associating injury data and treatment with crash reports. 6. Classification of injuries is related to need for treatment and potential threat to life; it can not be related to the cause or outcomes, e.g. permanent disability. 7. Not all injuries are seen by medical units that might keep records of trauma-induced injuries.
Reporting Data 1. Forms are not standardized, both at the EMS and treatment levels. 2. EMS reports may be completed after delivery of the injured to a medical installation and rely on the memory of EMS personnel for accuracy and completeness. 3. Reports may not be prepared for persons not transported by EMS personnel causing loss of data regarding injuries. 4. Multiple responding EMS agencies each may record differing data about persons treated in a specific crash.

Issues Affecting Managing the Data

Data management is comprised of two parts: entry and storage. Data entry involves the act of preparing data contained on reports for subsequent storage. The activity may involve microfilming or imaging, manual editing, distribution to various processing units, and the actual keying or scanning of the data themselves. Not all reports are entered into data bases, or if they are, the data bases may not be centralized. Medical data, particularly EMS forms, rarely are captured on a system.

During storage, data are maintained for later recovery and use. Data may be stored in paper files, as images on microfilm and computers, and digitally. The computer format may depend upon out-dated software or up-to-date, data base management systems. They may be linked to images and to data residing in other data bases. Some States come close to having a distributed data base system, at least at the State level. The crash safety-data may be linked to driver, vehicle, roadway-inventory, and medical files. However, each file is maintained separately, and duplicate fields exist. No states currently integrate local data bases with the State files.

The most important issue affecting the quality and availability of data is costs. Upgrading systems, including the development of distributed systems is expensive. Few States have committed resources to upgrade their highway-safety management systems.

Traffic crashes. Data storage for traffic crash reports ranges from paper files at many police agencies to linked data base systems in some States. Some State and local agencies also are introducing imaging systems to store reports.

Once completed and reviewed at the local level, the PCR is sent to the State safety-agency for processing. An exception is reports which do not meet the reporting threshold established by the State. Some agencies also will record crash data on a local data base, normally from the police copy of the PCR. Only in Ohio (of those sites studied) does the local agency enter data directly to the State safety data base.

After receiving the report, all States key data into a safety data base. More than one agency may key data from the same report, some successively, others on a parallel path. Data may be entered at one time, or in two steps with some form of skeleton file being produced first. During entry, most States provide some form of manual review prior to keying the data. Many States have online editing systems which require correction of specific fields before the entry is accepted. They often use some form of batch editing to make corrections.

In addition to the computer data base, local agencies and the States keep images of the documents on file. Most organizations keep a paper file. Many also will microfilm the crash reports. A few agencies are beginning to use imaging systems to maintain copies of the documents. The issues which affect management of crash data appear in table 14.

Roadway inventory. Most roadway inventories are not linked to the crash safety data bases. As a result, information about locations on the roadway where crashes are occurring have to be matched manually to summaries from the crash safety data base. Where locations of the crashes are not accurately coded, a problem in many States, the usefulness of the crash data is limited. Issues unique to roadway inventories appear in table 15.

Medical information. Most data related to medical information from both EMS and trauma organizations are not maintained centrally. Several States have been taking steps to centralize these data. However, because of privacy laws and the threats of litigation, data may not be shared even though they are available from a central source. The CODES project currently being funded by the National Highway Traffic Safety Administration is seeking ways to associate data with crash reports and to make them available for users. Many of the issues related to traffic crash reports also are related to medical data. The issues which affect the management of medical data are shown in table 16.

Table 14
Data Entry and Storage: Crash Data

1. Local agencies enter data into their files prior to transmitting reports to the State, thereby increasing delay in availability at State level.
2. Maintenance of paper files can reduce flexibility in access and use of local data.
3. The number of steps required for handling reports increases delay and likelihood of errors.
4. Separation of duties for recording data allows differing levels of quality and timeliness at the State level.
5. Multiple databases, particularly local and central, increase delays in data available centrally and increases handling of data.
6. Use of outdated database management systems reduce the likelihood of making changes and usually preclude linkages. They also can restrict the number of years of history maintained.
7. Editing is performed without access to original facts which limits ability to provide accurate and precise data.
8. Lack of automated assistance increases time required to process reports particularly for location coding.
9. A long period from crash occurrence until data are available from central source delays availability to users.
10. Paper files are the primary source of crash narrative and diagram limiting general availability of these two elements to of users who need them for specific crashes.
11. Variable confidentiality laws reduce availability of data to users.
12. State level processing not integrated with local level, reducing feedback needed for improving quality and quality assurance.
13. Change in forms without adequate review of entire process results in loss of data or quality of the data.
14. Lack of direct access by users reduces feedback needed to address quality as well as usefulness of data being recorded.

Table 15
Data Entry and Storage: Roadway-Inventory Data

1. Not linked to state safety data base reduces precision of information to that base.
2. Roadway inventory does not reflect maintenance and construction changes reducing accuracy of inventory.
3. Crashes are not part of inventory thus reducing ability to access one source in order to determine relationship of crashes and roadway elements.
4. Lack of longitudinal geometrics (horizontal and vertical curvature) results in less complete information relating the roadway to the crash.
5. Accuracy and precision are variable, limiting usefulness of determining role roadway may have played in crash.
6. Lack of inventory for classes of highways reduces coverage.
7. Identification of roadway features imprecise because of measuring schemes.
8. Long intervals between times of updating the inventory will not reflect changes which may be related to increases or decreases in traffic crashes.

Table 16
Data Entry and Storage: Medical Data

1. Confidentiality of medical records limits the association of injury data and treatment with types of crashes.
2. Lack of data bases prevent single access to data and unavailability of data from many sources.
3. Medical data not available until after discharge or admitting creates long delays in availability of data.

4. Collecting and Managing Safety Data

Overview

This chapter briefly addresses the aspects of quality as it affects safety data. The issues of quality are important in the subsequent derivation of costs discussed in the next chapter. After a discussion of quality, the chapter concludes with a discussion of the processes related to collecting and managing highway-safety data. It focuses on the main sources of data: crash, roadway-inventory, and medical (emergency medical services and trauma injury).

Quality of Safety Data

Dimensions of Quality

Quality is difficult to quantify. It has several dimensions. Within each dimension, the quality rating of a given set of data may vary according to the user and need for information. For example, one user may need substantial precision while another needs approximations. The coverage of the data needed may be wide ranging or very narrow. As a result an almost infinite number of quality requirements arise from users (see also O'Day 1992). For the purposes of this project, six "dimensions of quality" are used:

1. **Coverage**: Proportion of the universe sampled and representativeness of the sample.

2. **Accuracy**: Degree to which the crash report is correct, both in terms of what the report requires of the collector, and what the collector reports; includes retention and translation of the report in processing, both in terms of data entry and preserving the edited record in an unadulterated form.

3. **Completeness**: Data which are missing either because they are not collected, reported, or entered into the system, or which are inaccessible for reasons of confidentiality or "ownership."

4. **Precision**: Degree of detail.

5. **Consistency**: Among observers and collectors, and over time.

6. **Timeliness**: How soon the data are available after the event.

Discussion following will addresses issues of quality as it relates to the police crash report (PCR), roadway inventory and medical data with concentration on the PCR. Many of the issues related to quality appearing in literature also were raised by users interviewed during the course of the project. A final section under this part presents an analysis of crash reports obtained from sites visited.

Previous Studies of PCR Data Quality

Coverage, accuracy, and consistency are the dimensions which have attracted the most attention in the literature. Precision, completeness, and timeliness tend to be most explicitly dealt when a review is made of a specific traffic-records system.

Coverage. The degree to which the entire universe of traffic crashes is captured by the system is a function of several factors. Reporting levels are established by agencies so that low-severity crashes are not required to be reported by the police. The trend is to reduce the number of cases which require completion of the PCR. These reporting thresholds vary considerably across the States, as well as across agencies within a State, and are based on estimated costs to repair, injury severity, and combinations.

At one level involved parties go to the nearest police station to have an officer complete a "desk" report. Even with at-scene reporting, the number of elements reported will vary depending upon severity of the crash along with its likelihood of requiring further case management. Furthermore, regardless of laws and policies which require reporting, the motoring public often does not comply.

The degree of coverage, whatever it may be, seems to be declining as available resources within the law enforcement community decline. A recent report on the situation in one State [University of Texas at Austin 1984] expresses what is generally believed across the U.S.

> "The level of reporting ... is becoming an increasingly serious problem. The result of these movements has been a decrease in the amount of police-reported accidents and an increase in the amount of driver-only reported accidents. This situation has important implications...
> - Many accidents go unreported;
> - Driver-reported accidents may not be as reliable as police reports; and,
> - Systematic biases are introduced into the data, if drivers report certain kinds of accidents, while police report others."

A recent Transportation Research Board (TRB) report [TRB 1987] concludes that "most States do not have a good estimate of the number of missing accident reports, but it is generally agreed that large numbers of property-damage-only accidents and many minor injury accidents are not reported." This is considered by some analysts of the field of traffic safety to be the most fundamental issue to be faced with numerous studies and reviews of crash coverage in the literature [Hauer and Hakkert 1989, also Hutchinson and Kennedy 1964; Morrison and

Kjellstrom 1987; Harris 1990; Chapman 1983; Turner - National Safety Council 1989; and O'Day 1992].

An overview of 18 studies conducted through 1985, [Hauer and Hakkert 1988] concluded that in the U.S. a range from 35 to 55 percent of all crashes are reported by the police. Fatalities can be estimated to an accuracy of +/- 5 percent using reports from the police, while only 80 percent of the injuries requiring hospitalization and 50 percent of all injuries are reported. Equally important, the authors conclude that there are inconsistencies in coverage according to crash attributes, e.g., whether the injured party was a driver, passenger or non-occupant of a vehicle, and the number of vehicles involved. In spite of the important impact of coverage issues on the accuracy of analyses, State traffic records assessments being conducted by NHTSA [Brown June 1992, November 1992, and July 1993] do not explicitly deal with this issue.

Accuracy. The 1987 TRB report expresses the opinion that there is "widespread concern about the difficulty of the job that reporting police officers have. ... Errors are common in reported data such as occupant restraint usage, injury, severity, location references and vehicle identification numbers." The results of a survey based upon 35 respondents from 100 randomly-selected cities [Urban Transportation Monitor 1990] indicated that 71 percent believed the accuracy of their available crash data to be "good," while 26 percent rated it "reasonable" and only 3 percent found it to be "poor." On the other hand, a mail survey of the highway authorities in Great Britain [Ibrahim and Silcock 1992] resulted in conclusions that law enforcement agencies expended significant time on resolving officer reporting-errors. Location referencing, narratives, vehicle identifiers, road classification, and attendant circumstances were the most problematic.

The difficulty in assessing the accuracy of crash data is the availability of objective criteria, especially for the data collection. Judgments about accuracy of a crash report are difficult because those conducting the assessment were not present when the officer completed the PCR and have no other source of information about the crash. In some cases, it is possible to obtain data on the crash from other sources; however, the relative accuracy of the alternative source is not objectively known either.

An attempt to verify accuracy appeared in a study [Shinar, Treat, and McDonald 1983] which used a sample of 214 crash reports prepared by trained accident-investigation specialists matched to State and local PCR's. Using statistical measures of agreement and disagreement between the two sources of information on the crash, the authors reported that the "police are highly reliable in observing the correct location and date and may be considered to be sufficiently reliable in noting the day of week, number of drivers, passengers and vehicles involved in each accident." On the other hand, reporting of ambient weather, road and light conditions was moderately inaccurate, while the poorest agreement occurred with police notation of vertical curvature. PCR's were also identified as inadequate for reporting crash severity, primarily because approximately 15 percent of the crashes classified as property-damage-only (PDO) had personal injuries.

For traffic unit variables, i.e. vehicle and driver data, there was a greater difficulty arriving at valid conclusions because of the significant amount of missing data discovered. The conclusion, in general was that "the use of police data for evaluation the frequency and type of vehicle defects in accidents is very questionable."

Finally, the analysis of causes attributed to the crash on the PCR yielded the conclusion that "the police in fact are conservative in their attribution of causes since the probability of false alarms is extremely low for all causal factors ..." More detailed analysis resulted in the conclusion that the police are best at judging direct human-contributions, e.g., failure to yield, while the indirect human contributions, e.g., alcohol and fatigue, as well as the environmental contributors to the crash, are the ones that the police are least adept at identifying.

The findings of Shinar were echoed by Howard [1979] in South Australia. Inconsistency and very low accuracy also were found by Hutchinson and Kennedy [1964], Hargroves and Hargroves [1981], and Stein [1993].

Hall [1984] attributed inaccuracies in coding highway factors as causes to police bias which looks for driver errors. He is supported by others [Baker 1983; Hutchinson 1987]. The study also cites the need to replace the collection of many data elements by police with linkage to related archival files which contain driver, vehicle and roadway data in a more accurate and precise form.

Very limited information is available on how data-management affects accuracy either through correcting inaccuracies or by introducing new ones. A limited review of coded data was conducted in one State [Roy Jorgenson Associates, Inc. 1986] using fifty fatal-crash reports from 1985 found at least one error in every field captured by the traffic records system. Most errors resulted from inaccurate location coding; review of randomly selected fatal-crash reports showed differences from reported to actual location from 0.032 km to 2.17 km (average of 0.804 km) [see also Hall 1984].

One State [Mounce, Delucia, and Hilger 1991] recently conducted an evaluation of the quality of data which comes from its reporting system. This was done by identifying projects which used the crash records in the process of making a decision on investments in improvements at specific sites. The project team compared the contents of the automated record with the original PCR. No serious problems were found with the accuracy or completeness of the crash data file "with regard to seven variables typically used in problem identification and funding allocation decisions." This suggests that the translation of the report to the record does not suffer significantly in quality. It does not address the accuracy of the original PCR. Even so, the authors found errors ranging from 3.3 to 4.2 percent.

Completeness and Precision. The results of a survey based upon 35 respondents from 100 randomly selected cities [Urban Transportation Monitor 1990] indicated that 57 percent believed the level of detail of their available accident data to be "good," while 40 percent rated it "reasonable" and only 3 percent found it to be "poor." As cited in Shinar, the PCR's tend to be incomplete in reporting of vehicle and driver causal factors. A study of 1,088 PCR's, for crashes in Illinois in the early 1960's [Hutchinson and Kennedy] found 18.3 percent of the reports were not completely filled out. Although the State of Illinois depends upon the collision diagram for coders to enter certain data into the system, 15.3 percent of this sample did not include a collision diagram.

A study in another State [Hargroves and Hargroves 1981] found an overall omission rate on its PCR's of 1 percent of the data elements. They translated this into "one out of four accident reports contains one or more omissions." However, the omissions might be centered on a few elements of the form and may not significantly affect the quality of safety-data analyses. The series of State interviews [Hughes, *et al.* 1993] identified significant occurrence of incomplete reporting of Vehicle Identification Numbers (VIN's), crash diagrams, and narratives.

With regard to precision, there is a need to balance between report designs which require overly specific responses and ones that offer only vague options. The former can lead to excessively long lists of choices and inappropriate response options. The latter can result in no useful information. The State study cited above [Hargroves and Hargroves] concluded, on the basis of analysis of a representative sample of 10,000 PCR's, that "in many cases, vague or poorly defined response options were the most commonly used items for the data element."

Consistency. The set of critical automated reporting elements (CADRE) was developed to help standardize reporting of specific elements. A study of the crash reports used in 21 States (Wu 1993) concluded that most States do not contain the CADRE elements. Where they do, the codes adopted for the element were not completely in agreement with those recommended in the CADRE system. He also concluded that there was little consistency between the States. Even where common variables appeared for crash facts, the codes used for them were widely divergent.

While little else is available regarding this in the literature, one study [Gerrard and Mosher 1960] compared PCR's in the State of California with alternative archival sources of crash-related data, documented the variation between PCRs submitted by the State police and those submitted by city police departments. The data elements of interest on this study were those that could be derived for traffic units from DMV files, weather related data, and ambient light conditions. In almost all cases, there was an overlap of confidence intervals for State and city agency PCR's. The intervals for cities tended to be larger, possibly associated with the significantly smaller sample-sizes of the city samples.

Timeliness. The timeliness of safety data, i.e. its currency, is a subject of concern among safety analysts, but not one on which there is complete agreement. The results of a survey based upon 35 respondents from 100 randomly selected cities [Urban Transportation Monitor 1990] indicated that only 38 percent considered the timeliness of their available crash data to be "good," while 38 percent rated it "reasonable" and 24 percent found it to be "poor." A study by Hughes, *et al.* showed data to be available to end-users from 25 to 210 days after the crash. In one State [Brown November 1992] a two-year backlog of crash reports for data entry existed. Added to this delay was a reduction in completeness and coverage for those data eventually entered.

The varying needs of the users result in a range of currency required from the data. Unfortunately, some State systems have a combination of limited resources and inefficient processes which slow the availability of data. One State's review [Kelsh, Reitzler, and Rauth 1984] identified crash data which were as much as 3 months old as a major operational deficiency. Another State system study [Illinois Department of Transportation 1988] identified problems with processing which seemed to be the major reason for a delay of 6 months before crash data were available to safety analysts. Efforts have been made to demonstrate the application of new technology, use of streamlined processes, and injection of additional staff to reduce the time until the data are accessible to analysts [Lucke and Stenzel 1981].

User-Identified Quality Issues

As part of the focus groups conducted for this project in several of the States visited, users were asked to comment upon the quality of the data which they derived from safety-data sources. Table 17 summarizes comments and gives a "flavor" of the responses received. Appendix A contains a more complete set of those comments. In general, the comments are those made in a particular State, or city, so they are not representative of all the States visited. In a few cases, the comments were repeated among multiple groups.

Assessment of State Crash Files

Selection of crash reports. Rare opportunities are available to place the crash data base of a State under detailed analysis to assess the overall quality of what resides in the file. Most States do not do this as a regular, or even infrequent, practice. There are exceptions to this general state-of-practice. Some States conduct regular quality-assurance reviews. However, the results of these generally are not available in a form useful for research such as occurred with this project. The team, therefore, sought to conduct limited analyses of its own on crash-report data where a State was willing to cooperate and expend some resources to provide the data. In addition, the review of quality HSIS [Council 1991] is used to supplement findings summarized by the project team.

The team arranged to acquire a small set of PCR's from each of five States. They were selected and copied at-random (but without specific instructions on how to achieve randomization) upon receipt at the submitting agency, but before any processing was started. The reports then were compared to the final data base record.

Table 17
Comments from Users Regarding Quality of Crash Data

Coverage	• Reporting of PDO crashes varies among jurisdictions; local thresholds vary from the State threshold. • Pedestrian, bicycle, and hit and run crashes are not well-reported.
Accuracy	• Officers tend to be accurate on reporting the "when, here, and who" aspects of the crash, but less accurate with the "how and why" aspects. They also do not report roadway related elements accurately. • On scannable reports (OMR), officers mark the wrong "bubbles." • Use of "other" and "unknown" occur too frequently.
Completeness	• Seating positions and occupant injuries often are missing. • Reports lack details on contributing circumstances for roadway elements. • Pedestrian age and drinking, and helmet usage is not determined.
Precision	• Precision is highly variable, and the narratives and diagrams tend to be oriented to establishing fault, rather than recording facts. • Officers often provide insufficient information to locate crashes precisely; they frequently use "other" for coding elements. • Some codes for contributing circumstances are imprecise, e.g., "careless driving" or "speed too fast for conditions." • Roadway elements are reported imprecisely.
Timeliness	• Crash and citation data are available barely in time for preparation of the State Highway Safety Plan each year. • Need for timeliness varies significantly among users. • Agencies create their own files because of lack of timely data.
Consistency	• Reporting thresholds and injury severity is not consistently reported. • Reporting of contributing circumstances increases with crash severity. • Narratives and coded elements often do not agree, particularly for driver condition.

A total of 246 reports were received from the 5 States as shown in table 18. Of these, one half represented reports of injury-producing crashes, the other half for PDO. Because the relationship between PDO and injury crashes is not of the magnitude of 2:1 as exists nationally, the results of this review can not be taken inferentially. Finally, of the crash reports received, 14 from 1 State contained collisions with deer. The special manner in which these crashes are reported resulted in removing them from the sample.

From the remaining 232 reports, 16 elements were selected for analysis. Six represented descriptors of the crash; 10 covered the vehicle, highway, and driver. The data elements for safety-belt usage, speed limit, and estimated speed were added because the literature has identified these as problematic. Criteria included that the elements are common to most States,

and that the elements were important descriptors of the crash. Not all States reported the same elements. A list of the selected elements is shown as table 19.

Table 18
Crash Reports Submitted for Review

Type of Crash	STATE					ALL STATES
	A *	B	E	I	J	
Hit & Run	0	0	7	3	2	12
Single vehicle - PDO	3	3	9	1	3	19
Injury	2	7	11	8	15	43
Multi-vehicle - PDO	12	15	44	5	9	85
Injury	5	4	22	22	20	73
All Crashes - PDO	15	18	60	9	14	116
Injury	7	11	33	30	35	116
TOTAL	22	29	93	39	49	232

* excludes 14 reports involving a vehicle and a deer (all reports missed one or more elements)

From the remaining 232 reports, 16 elements were selected for analysis. Six represented descriptors of the crash; 10 covered the vehicle, highway, and driver. The data elements for safety-belt usage, speed limit, and estimated speed were added because the literature has identified these as problematic. Criteria included that the elements are common to most States, and that the elements were important descriptors of the crash. Not all States reported the same elements. A list of the selected elements is shown as table 19.

Table 19
Elements Selected for Review

1. Crash a. Manner of collision b. Severity of the crash c. Location relative to road d. Condition of the road e. Time of day vs light conditions f. Most harmful event	2. Vehicle, Highway and Driver a. Safety equipment used b. Driver condition c. Severity of injury (driver) d. Driver contributing factor e. Vehicle contributing factor f. Highway contributing factor g. Vehicle type h. Estimated vehicle speed I. Posted speed limit j. Maneuver preceding collision

Although a review of 232 crashes provides a limited basis for conclusions, it does highlight issues. The most important ones are discussed in greater detail. In general, the findings suggest that as many as 6 percent of all elements coded may be in error.

Errors detected - general. The 232 police crash-reports had 301 "errors" among the 15 elements. These errors included both misreported and missing data. Table 20 and figure 2 present a summary of the findings, showing the variation in errors among the States. Overall, missing codes accounted for more than one half of the errors. These incomplete elements occurred on more than one-third the reports analyzed. Of the driver and vehicle elements studied, no errors appeared in the reports from one State; another had errors in 30.6 percent of the elements.

Degree of correction of the errors was variable among the States. None of the errors detected in crash reports from state B were corrected in the data base records. On the other hand, State I detected and corrected 91 percent; however, 20 of the 21 corrections applied to improperly labeled "direction of travel."

Although a "total of errors" is shown in table 20, it is descriptive only. Each State describes and codes elements slightly differently. Errors occurring in one State, e.g., the "manner of collision," might not occur in another because of these differences. This highlights the inconsistency in definitions used for data elements.

Sources of error. Errors found can be divided into several sources. The most important were those appearing because:

- Directions for completion appeared *unclear.*
- Data are *not available* or the element *may not appear important.*
- *Failure to record* the data properly.

The most common example of *unclear* needs was where the officer left "contributing cause" blank more than 50 percent of the time. As noted above, the officer appeared to interpret the instructions to mean "complete this only if the element is 'at fault'."

Missing "posted" and "estimated speeds" are examples of data which are not recorded because they are *not available* or *may not appear important.* When speeds were entered, they were not consistent with apparent damage done, e.g. speed of "5 mph" when both vehicles appeared to be total losses.

The final case of error stems from *failure to record* data properly. The most frequent example of this failure was miscoding the "location of the crash relative to the road" which occurred in

Table 20
Elements Reviewed and Errors Identified

NUMBER OF ELEMENTS REVIEWED	STATE A	B	C	D	E	ALL STATES	
Crash	5	6	6	4	6	5.4	(avg)
Vehicle/Driver	9	4	9	8	8	7.6	(avg)
ELEMENTS IN ERRORS OR MISSING							
Crash - Number							
PDO	18	8	19	0	1	46	
Injury	10	4	9	0	6	29	
All Crashes	28	12	28	0	7	75	
- Percent *							
PDO	24.0%	7.4%	5.3%	0.0%	1.2%	7.3%	
Injury	28.6%	6.1%	4.5%	0.0%	2.9%	4.6%	
All Crashes	25.5%	6.9%	5.0%	0.0%	2.4%	6.0%	
Vehicle/Driver - Number							
PDO	77	6	60	12	9	164	
Injury	20	4	14	11	13	62	
All Crashes	97	10	74	23	22	226	
- Percent *							
PDO	30.6%	4.5%	6.1%	10.0%	4.0%	10.1%	
Injury	17.1%	6.7%	2.7%	2.6%	2.8%	4.2%	
All Crashes	26.3%	5.2%	5.0%	4.2%	3.2%	7.3%	

* percent is based on average errors per report divided by number elements reviewed

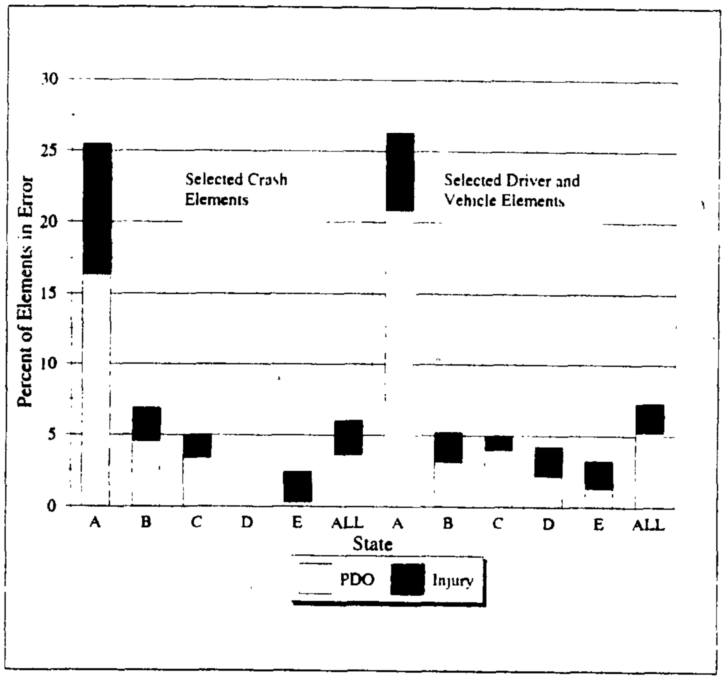


Figure 2
Errors on Crash Reports

States A, B and E. Generally, if the crash occurred at mid-block, it was coded as such even though the narrative clearly indicated that the crash was "intersection or driveway related" and codes for "intersection-related" and "driveway" were available. In other cases, officers mismatched time of day and the light condition, and recorded vans as cars or trucks even though the code for "van" was available. Data entry personnel rarely corrected these mistakes.

Quality of State crash files from HSIS. The Highway Safety Information System (HSIS) is sponsored by the Federal Highway Administration, and operated under a contract with the University of North Carolina, Highway Safety Research Center (HSRC). It currently maintains crash and roadway-inventory files created from "raw" files supplied by five States: Illinois, Michigan, Minnesota, Ohio, and Utah. The HSRC team reviews the data that are received from the State before the files are re-structured for HSIS purposes [Council 1991].

In general, their reviews show that each State's files are relatively complete and accurate. Specific problems that were identified varied from State-to-State, many of which focused on variables considered non-essential for the expected uses of HSIS. Variables with significantly-high incompleteness included:

1. Contributing factor, or cause of crash.
2. Driver physical condition.
3. Driver/occupant ejection.
4. Safety equipment (one State did not enter this into the file because it was considered too unreliable).

Inconsistencies occurred in the distributions of coded values for some of the key variables. These appear to be caused in many cases by changes in the design of the form. Both the addition and deletion of variables caused inconsistencies in reporting, when comparing the years before and after a change. Alteration in the code structure also caused inconsistencies in reporting, either because officers apparently continued to use old codes or misunderstood the new ones. In a few cases, the codes themselves created problems, especially when they were ambiguous or the alternatives were not mutually exclusive.

Analysis of univariate distributions suggested several areas of inaccuracies. Most consistently cited problem variables were those dealing with driver physical condition (generally related to DUI), safety equipment usage (generally related to indication of seat-belt use), and classification of the crash, i.e., as rear-end, head-on, right angle, etc.

Summary and conclusions from the error analysis. The analysis indicates that errors are occurring at a rate as high as 50 percent for some elements and approximately 6 percent of all elements studied. These errors often arise from missing data and are rarely being corrected. Even those States which pride themselves on substantial manual pre-editing, as well as automated editing both online and in batch mode, did not correct the errors identified from this cursory analysis. With perhaps the exception of errors in manner of collision and location relative to the road, one might argue that others are not critical to analyses of highway safety-data. However, that argument then questions the need for the element in the data base.

Putting the argument for "need" aside, this analysis confirms that errors of both omission and inaccurate recording occur frequently on police crash reports. The number of errors and their source vary from State to State. No patterns could be discerned regarding elements more or less likely to be incorrect. Most significant of the errors, however, was the frequency with which elements were not coded. The other important finding was that during data entry, many of the errors are not corrected even where missing values occurred. Most States are pressed to complete the recording of data in a timely manner and would not have the personnel and time available for such a thorough review. Because a completely error-free file is not feasible, the collectors, managers and users are faced with making decisions about tradeoffs between quality of data and other considerations.

Collecting and Managing Traffic Crash Data

Initial Handling of the Crash

The police crash report (PCR) represents the core source of safety data. It is the focus for this project on the cost of data collection. Figure 3 presents a diagram showing the process of collecting and managing PCR data. Issues related to costs and quality are presented following each section discussing aspects of collecting and managing traffic-crash data.

Before any police involvement can occur relative to a crash, they must be made aware of its existence. In most cases, a report of a crash is made to a public safety dispatch operation by a citizen; in rare instances, the police may find the crash while on patrol.

Police officers initially arriving at crash scenes have other duties to carry out prior to, as well as in conjunction with, crash-data collection. The first is to stabilize the scene. This may require controlling traffic, providing medical assistance to the injured, and summoning others to assist with or take over those tasks.

After carrying out this initial assessment, the officer must determine how (or if) the crash is to be reported. If the damage is below the jurisdiction's reporting threshold, the officer will advise the involved parties as to their responsibilities. In other instances, the first arriving officer may determine that specialists will be needed because of the circumstances of the particular crash. They also may do preliminary data collection, or may elect to leave all such activity to the specialists.

Even if a report will be needed, the initial responding officer may continue to manage the scene including controlling the flow of other responders, traffic, and bystanders. He also will assist and complete administrative activities such as advising involved parties of their rights and responsibilities. Assignment of a civilian investigator can start the process of data collection earlier.

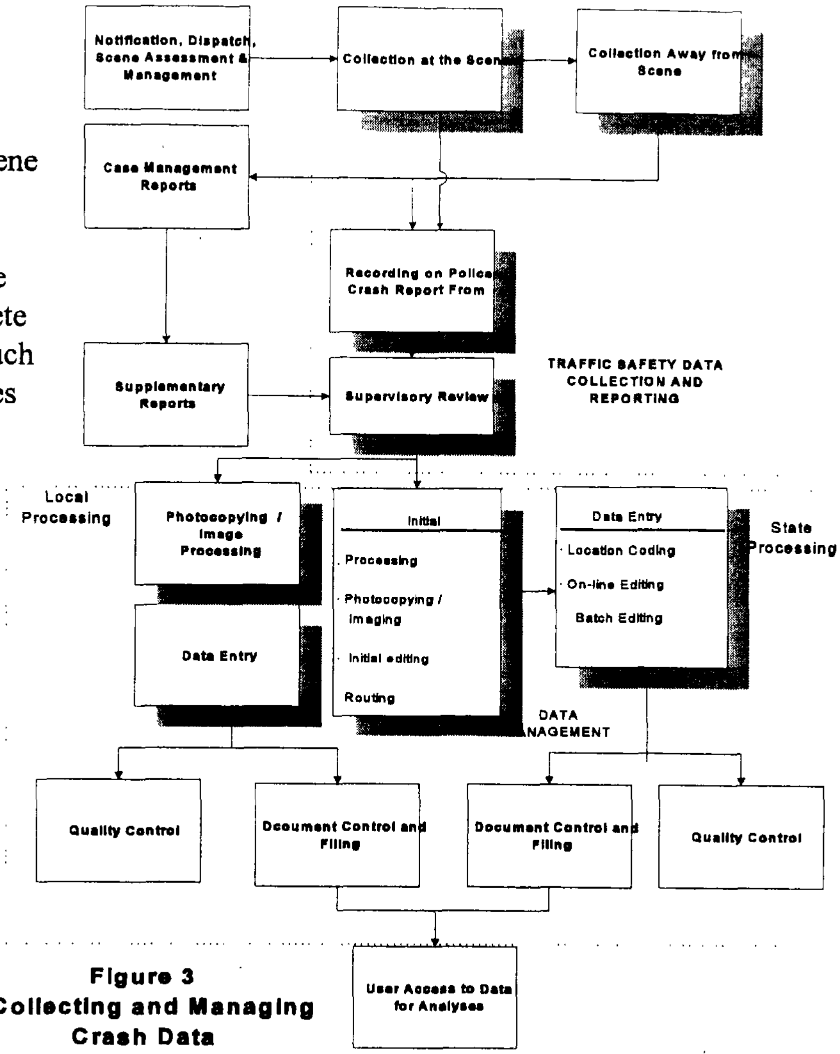


Figure 3
Collecting and Managing Crash Data

The majority of the non-data collection activities take place prior to most data collection and reporting, but some are concurrent. An officer who is directing traffic can observe facts about the accident scene and vehicles that will be recorded later. Similarly, discussions with EMS and tow personnel can provide information on persons or vehicle damage that will be recorded on the report form later.

Cost issues. Costs involve personal and equipment needed for dispatch and response.[1] Most of the costs derive from personnel time along with some attributed to use of equipment. In general, however, these activities are not attributable to safety-data collection and recording.

---

[1] Note: In many jurisdictions minor crashes are reported at police facilities rather than in the field. These instances usually involve a desk officer or clerk filling out the crash report form based on information provided by involved parties. For these cases, costs are limited to the time needed to ascertain basic crash facts and fill out the report form.

Quality issues. Because management of the scene does not involve data collection, quality is not a direct issue here. However, in managing the scene, the officers involved must often use judgment as to the importance of preserving marks, debris, and position of vehicles and persons involved in the crash. Also, the degree to which the officer is prepared and able to make mental notes about the crash scene while assessing and managing it, will increase the likelihood of quick and accurate reporting of the crash. Rapid response is important to arrive before potentially useful items of evidence can be lost or destroyed.

Data Collection at the Scene

Drivers and others persons. Information from and about people is an important and time-consuming part of data collection. It involves acquiring and recording elements which fall into two categories: general descriptors and crash-related facts. General descriptors include: personal identifiers, addresses, and individual attributes, e.g., age, physical characteristics, and sex. Crash-related facts include any information about the drivers, occupants, and other involved persons. They will include determining facts and obtaining opinions regarding events and actions of involved parties. Also included is initial determination of injuries, although the reporting officer may change the level of severity after a later visit to the hospital.

Cost issues. In a complex crash, the officer may take a significant amount of time to locate the involved parties, including witnesses. The recording of names, addresses, telephone numbers, and other identifiers consumes a major part of the time spent upon data collection. Most of these are not considered safety data. However, completion of a narrative, and to a lesser extent, a diagram, is usually based upon information gleaned from talking with involved parties. Some State-level data management depend upon the narrative and diagram to provide the data. The officer also may perform tests and prepare citations. While these are data which may be useful in a safety-information system, the tests are being performed for law enforcement purposes. Only the cost of acquiring test results and entering them into the safety data base should be counted.

Quality issues. The most common problem is inaccurate coding of injury. A crash may be considered severe if a person is transported by ambulance even where only a "complaint of injury" occurs. In addition, the narrative and diagram often lack accuracy and precision because the officer sees their use in a different light (case management) than does the safety community.

Involved traffic units. Data elements are needed for each involved vehicle. As with information about involved people, the traffic unit elements can be separated into two categories, general descriptors and crash-related facts. The general descriptors include attributes about the vehicle; crash-related facts refer to damage, pre-crash elements, and maneuvers. The process involves recording data from a registration document and the vehicle itself, in addition to information obtained from drivers, passengers, and witnesses.

Cost issues. The major effort involved in recording these data is associated with identifying the attributes of the vehicle and describing damage. The officer may be asked to record make, model, vehicle type, color, registration data and the Vehicle Identification Number (VIN). Damage and vehicle condition descriptions must be entered based upon the officer's inspection of the involved vehicles. All this effort can use a significant amount of time.

Quality issues. Incorrect recording of the VIN is a significant problem. Officers are also not adequately trained to make good judgments on how to classify damage to the vehicles or identify equipment violations. Further, most PCR's are not well formatted to capture equipment defects or information about the vehicle and its damage. An exception is the federally mandated commercial-vehicle reporting standards, developed by the National Governor's Association [NGA Elements], which are beginning to appear as States up-date their crash reports.

Roadway environment. Data elements are collected on both permanent and temporary environmental conditions. This includes both highway features and weather-related conditions. Many of the roadway items are observable directly by the reporting officer while approaching the scene, and immediately upon arrival. An item, such as the functioning of traffic control devices, usually must be actively sought after initial handling of the crash.

Cost issues. Most of the data are usually entered by checking appropriate boxes. This aspect of the data collection usually does not involve a significant amount of officer time.

Quality issues. There is evidence that officers do not do an accurate job of recording the basic roadway data. Furthermore, the data requested on the PCR is generally highly qualitative and not of sufficient detail to support many analyses.

Observe, note, and measure elements related to the crash. In some cases, the officer may need to note information that will be useful for later crash reconstruction or independent determination of the sequence of events of the crash for identifying and locating: types of debris, tire and other marks, e.g. gouges, and pools or trails of vehicular fluids. The officer or specialist may need to take pictures of the traffic units involved, and measure their positions, as well as the positions of debris and markings. Often these data are taken only when a serious injury has occurred or other potential need for accident reconstruction may be present. The proper collection of measurements and photographs often requires special training; therefore, it may require a specialist to be at the scene of the crash.

Cost issues. Most of the data are acquired when criminal or civil litigation is expected. The data rarely are added to the safety data base. Therefore, the costs should not be applied to collecting safety data.

Quality issues. Use of a specially trained individual to identify and interpret the marks and debris associated with a given crash usually enhances quality as opposed to those obtained by a patrol officer.

Post-scene activities. For many traffic crashes, the reporting officer completes all tasks and forms prior to leaving the scene. He may also complete reports at a nearby off-road location. In some cases, information will be collected away from the crash scene or through return visits to the scene.

The most common need for off-scene data collection arises when crash-involved persons are taken to hospitals before the police can take statements from them. In this case, the officer will travel to the hospital, but may try to talk to the person by telephone. For serious crashes, the police may take witness (and suspected-offender) statements in a controlled situation, such as a police station, rather than at the scene. When detailed examinations of traffic units are needed, these too will be done at off-scene locations, such as a salvage yard.

Although increasingly uncommon, some jurisdictions require that all reports be typewritten. Therefore, an officer would complete only field notes at the scene and then type the report at a later time. It is more common for officers to recopy (or type) the accident report form for serious accidents where litigation is likely.

Cost issues. Significant amounts of officer time can be devoted to post-scene data collection and reporting. Most of the costs can be ascribed to case management rather than handling safety data.

Quality issues. Other than taking statements at a hospital, post-scene data collection is usually carried out by a specialist. The quality of reporting is high.

Law enforcement agency review and processing. The officer will submit completed PCR's to the immediate supervisor at the end of the shift, or at some other specified interval. The immediate supervisor and then higher level management may review the PCR for completeness, consistency, accuracy, etc., for both case management and safety-data collection. Errors in the PCR may be modified by the reviewer (with or without consultation with the submitting officer) or returned to the submitting officer for additions or corrections. Review practices vary from almost non-existent (especially in larger urban police agencies where no traffic unit exists) to careful scrutiny and quality control especially where an exemplary traffic unit is in operation.

Cost issues. While the unit cost of personnel time to perform the review functions is usually higher than that of the officer, the amount of time involved per PCR is small, so that the added cost is a small proportion of the total.

Quality issues. The care given to quality of safety-data collection at this stage of the operation can potentially have a major impact on the overall quality of the data in the information system. This is not just because of the importance of correcting errors at a point nearest to the source of the report, it also creates a message for the officer about the departmental interest in quality for this aspect of the officer's job.

Circumstances affecting data collection time

Number of involved vehicles and persons. The single variable that appears to have the greatest impact on the time expended for safety-data collection and reporting is the number of vehicles involved in the crash. Each additional vehicle involved requires the investigating officer to record another set of vehicle descriptors. Increasing the number of involved persons, such as occupants, also increases reporting time.

Type of Vehicles. The involvement of a commercial vehicle can significantly increase the time spent on the scene. However, data obtained generally are not safety related.

Severity. As the severity of the crash increase, particularly for fatal crashes, officer time involved in reporting a crash can increase dramatically. Most officers recognize the increasing probability of litigation as crash severity increases and will more carefully record more detail in their investigation.

Hit-and-run crashes. Crashes where one or more drivers leave the scene can result in major time expenditures for police officers. As with severe crashes, time spent at the scene generally is attributable to case management rather than collection of safety data.

Other circumstances. In addition to the circumstances specifically described above, which can significantly affect crash reporting time, other spurious conditions that can also affect officer time. An example of this is adverse weather conditions. In some cases, police stop investigating all but the most serious crashes.

Local Processing of Crash Data

Originating law enforcement agency. Processing by the law enforcement agency originating the PCR may involve data entry, copying, or image processing of the form both for distribution to, or access by, other agencies, and for transmittal to other agencies, both at the State and local level. Data are then stored and maintained for the use of the agency or other units of local government. Hard copy of the form may be maintained for a period then microfilmed. At least one State has begun to use local agencies for direct entry of data into the State's data base to eliminate duplication of entry at the local and State levels. If data entry occurs, generally there will be editing at the local level.

Processing by others. Even though a local enforcement agency may maintain an automated data base of the PCR's it originates, other local agencies such as engineering often will maintain their own file based on the PCR's. This usually occurs because the system operated by the enforcement agency is not readily accessible to the other local agency and may not readily support the different use intended by the other agency.

Although costs occur in managing crash data at the local level, they duplicate those efforts performed at the state level. Therefore, the costs are not included in the model.

Processing by the State

Initial State processing. Copies of the PCR are transmitted usually via US mail to the State agency having official responsibility for maintaining PCR files. Receiving clerks usually will sort the PCR's into groups to fit the nature of the state-agency's process. Initial processing will include matching of PCR's with other sources of information on the crash, primarily citizen reports where they are required by the State. Scannable forms require significant attention to assure that the form is not bent, wrinkle, stapled, etc., to prevent jams to the scanner or a high proportions of rejections. Computer-developed forms may have been uploaded, or may arrive on disk. In some States, local communities already are widely using computer-based forms and compatible data base systems. Unfortunately, some States are not able to accept the electronic forms, thus requiring the originating agencies to submit hard copy of the electronic report.

The initial processing also can include manual editing of the PCR's. Some States will assign skilled personnel to conduct a thorough review of the form at this stage with corrections made based upon evidence gleaned from the entire report.

Forms management also usually occurs. This can include establishment of batches, with associated batch controls, photocopying or image processing. The point at which the PCR is imaged, if at all, tends to occur either at the outset of the process or at its completion.

> Cost issues. Most obvious of the cost issues is the duplication of data management that occurs at the State and local levels. The lack of systems integration results in redundant processes being performed. Also, where significant editing is involved at this stage it is labor-intensive.
>
> Quality issues. The quality of data entering a crash data base primarily is affected by the editing and correction performed at this stage. Manual editing can be successful because the reviewer has access to the narrative and diagram to assess the accuracy of other entries which is not possible yet with automated editing routines. At least one State has found that the use of local data entry for state databases has improved quality because the entry operator has ready access to the originating officer and does not hesitate to interact.
>
> The point at which imaging is done can affect quality of information, when the PCR is later referenced, as is often done for site-specific analyses. If the report is imaged before editing and correction, a person reviewing the image may not see the results of the corrections. Where corrected, or supplemental, reports are filed, most systems have procedures to link the corrected report to the original one.

State data entry and edit. Many variations exist among the States of how this part of the process operates. Some States will create an initial or skeletal record completing it at a later point of the process. Others separate the coding and entry activities, so that each is performed by a different group; some have only one point in the process at which coding and entry occur assigning the task to one person. In a few cases different agencies will do coding and entry of different parts of the form into the same data base. For example, one State system creates an initial entry when the motor vehicle administration agency enters driver and vehicle data. The form then goes to the department of transportation (DOT) for location coding and entry of other crash facts. One State has a control group through which each PCR must pass at each stage of processing. The report moves between groups of one kind or another at least seven times while being processed.

Manual location-coding is often done by a separate group because of the special knowledge and skills required. Location coding is most often done by the group in the DOT most directly responsible for maintaining a highway inventory. At least two States use prison labor for location coding. Some States are performing location coding assisted by automated systems.

The manner for coding of crash facts is highly varied, and depends somewhat upon the design of the PCR. Some States will employ highly skilled coders who will interpret the narrative and diagram, along with other data on the form, to establish codes for many elements not appearing on the form. Many States are tending to move away from this, using a PCR designed to have the officer do the coding on the report. A scannable or optical mark sensitive (OMR) form can eliminate the need for some manual data entry, but still involves significant entry of identifiers and similar items that can not use OMR. While one State is using optical character recognition (OCR) and OMR for its statewide EMS run sheet, this is in the experimental stage.

Some States are linked in real time with files from their motor vehicle administration agency. These will serve as sources for entering data about persons and vehicles. Use of this linkage reduces the number of entries made from the PCR. In a few States, linkage is two ways. The crash record links to data from driver and roadway files, and provides a record of the crash to these two files.

A few States provide no verification at data entry. On the other hand, many use both an on-line edit system and a batch edit, the latter being performed at the point when the data base is to be updated.

The federally mandated Fatal Accident Reporting System (FARS) operation usually requires special handling of fatal crashes. Data entry for fatal crashes ranges from uncoordinated operations to a FARS unit which handles all data entry.

Cost issues. A host of cost issues are associated with data management. Most systems involve more than one agency in the process. In the case of location coding, some States were moving crash reports back and forth between agencies. The inefficiencies of the movement and attendant controls are clear. There are similar problems with multiple data-entry points. On the other hand, linkages with driver and vehicle data bases creates efficiencies in data entry.

Manual location coding is another labor intensive effort. Several States have, or are automating this. However, where automation has been implemented, many rejected reports occur because of inadequate description of the location by the reporting officer.

The cost of developing and maintaining a modern data base management system for a crash data base can be significant. The States visited that had modern systems, experienced costs of several million dollars for development and implementation. The cost-effectiveness of these systems is accepted, but not well documented. As personal computers, and associated software are adopted, the development costs of entry and edit routines may be significantly reduced.

The use of automated forms can significantly reduce costs. OCR and OMR forms may involve special handling before being scanned. On the other hand, they result in the need for significantly less data-entry staff compared with traditional forms. OMR forms will continue to require special handling that use of OCR can eliminate.

Quality issues. While much can be done to improve quality during data collection, the processing aspect of the operation is the key point for maintaining quality control. Important aspects of quality control can occur during manual editing, coding, data entry, and post entry. A variety of manual and automated techniques are used. Most States employ on-line and batch edit routines. These routines seem to be the single-most quality-assurance technique of any employed by the States; however, most of the local agencies with whom the project team spoke, were not familiar with what was being checked. A few States share the error reports with the submitting agencies either directly or through periodic visits.

Once errors have been identified, they must be corrected properly to make the effort effective. In some States, the corrections were often based upon coder opinion; other States were able to be contact the originating agency. However, limited resources sometimes prevented States from pursuing corrections to data elements not considered "critical." A couple of States do have aggressive quality-assurance programs in effect.

As noted, local data-entry is believed to increase quality. Use of computer-based reporting allows on-line edits to occur in the field. The officer is then forced to make a correct entry to the extent that the computer program is able to build in safeguards.

Another aspect of data quality is lack of timeliness. A fairly common problem among States is the extended time required from the moment of the crash to the point when the data are available to users in the system. Delays may occur because of personnel shortages, or the process, or both. In one State, the processing for the financial responsibility requirements had to be completed before entry could proceed greatly delaying progress. In another State, limited staff and an inefficient handling system with multiple changes of hands resulted in significant backlogs.

Use of automated data-entry methods (e.g., OCR, OMR, and computer disk), while offering potential cost savings, can have a negative impact on quality. If the PCR is subject to less manual or automated review and edit, the likelihood of errors getting into the record can increase. This may also be true, to a lesser extent, as more reliance is made on officer self-coding.

Storage and maintenance. The final step in the safety-data management involves the conversion of the entered data into system files for maintenance and processing. If a record has passed the required batch editing, it will be formatted and added to the data base. The rest of this process involves ongoing maintenance of both the files and the system within which the files are housed. Most States that were visited maintain 3 to 4 years of data on the computer system. Some maintained data on-line for longer periods. Almost all maintained data in tape format for at least 10 years.

Cost issues. A cost consideration for this aspect of the process involves the amount of data stored and the length of time it is stored and maintained. While these costs are small compared to other aspects of the process, they are often under scrutiny by the management information systems (MIS) operations involved. Further, the cost of maintaining a data base can be affected by its design. Modern relational data bases are generally more flexible than the older flat-file designs. Furthermore, software development for the newer data bases can be done more efficiently. Similarly, personal computer and workstation systems can be more easily modified than mainframe systems.

Quality issues. The primary quality issue apparent here is the accessibility of the data. As the data are stored and maintained, the manner of storage affects accessibility, and therefore, the timeliness with which the data are available.

Data-Quality Assessment

While quality control procedures may be employed during data collection and management, some states have identified a need to monitor the effectiveness of those procedures and make appropriate changes to overcome problems noted. This quality assurance is not widely performed among the States visited. However, one State had a comprehensive program which included a regular analysis of the crash file to determine the integrity of the data being stored in the system. The results of the analyses allowed the state to provide feedback to those completing the crash reports, as well as to plan changes in the editing and correction procedures. The effectiveness of the feedback has not been evaluated.

Cost issues. The State performing the quality-assurance program has a dedicated data-integrity team consisting of three programmers and an assistant. This group produces error-checking reports on a monthly, semi-annual, and annual basis. This effort, however, can be costly.

Quality issues. The use of data-quality assessments at the back end of the data management has proven useful. A balance must be found between stringent control of quality and timeliness of data entry.

Roadway Inventory Data

Review of Roadway Inventory Processes

While the primary focus of project was on the police crash report, significant attention also was given to roadway inventory data as a source for highway-safety analyses. Interviews were conducted with managers of roadway inventory systems in several States visited as part of project. The purpose was to document the process for collecting and managing the data, as well as to identify data-quality issues. The process that was documented is shown in a generic form in figure 4, parts of which are discussed in more detail below.

Data that often appear as part of a roadway inventory system may be divided into roadway-specific measurements and attributes. The extent of the network for which an inventory was maintained varied considerably from State-to-State. Most focused attention on the State's Interstate and primary highways. Some States maintained two levels of detail, thus including secondary roads but with a reduced data-set. All States visited included the FHWA Highway Performance Monitoring System segments within their inventory activities, but with special handling to meet Federal requirements.

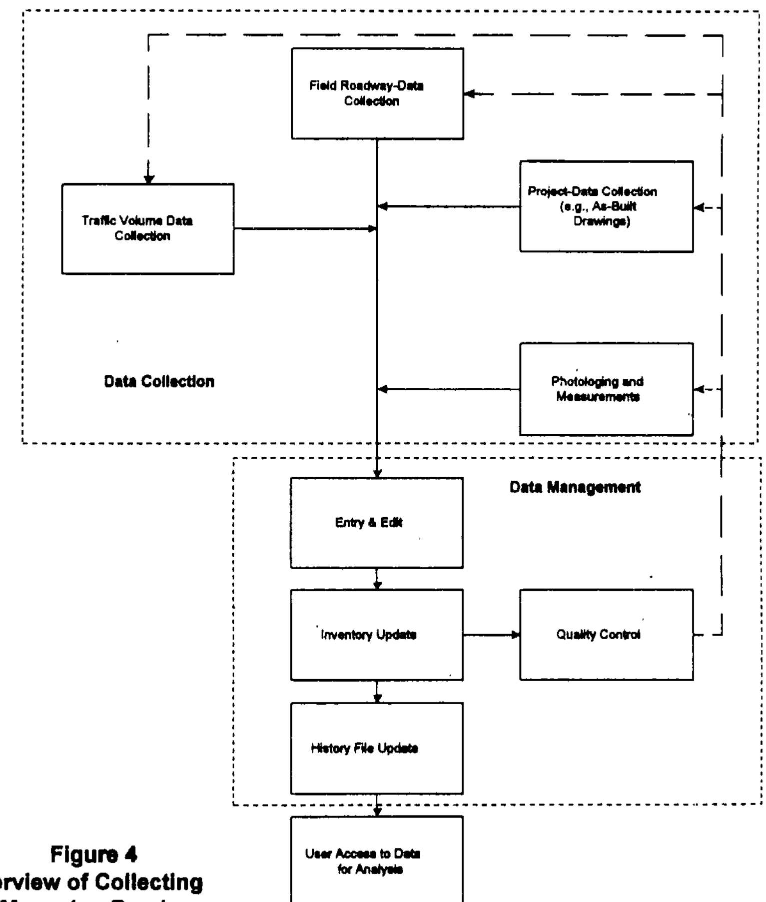


Figure 4
Overview of Collecting and Managing Roadway Inventories

Data Collection

Most of the States obtained roadway-specific data from DOT district personnel which submitted inventory systems. Once the file had been established, changes were made by "exception." That is, the district (and to a limited extent the central headquarters) are responsible for monitoring construction, maintenance, and repair work to identify where changes are being planned or made to the network that would require the inventory to be updated. This often involved reviewing permit applications, work orders, as-built plans, construction drawings, and newspaper articles.

Two States studied, however, operate a regular field data-collection program for updating their roadway inventory. Crews in a vehicle are equipped with measuring devices and a portable computer. In one State, a single crew covers 9,000 Interstate and principal highways on a 4-year cycle. In the other State, vans are available in each district for district crews to cover 41,000 miles of State-maintained roads on a 5-year cycle. The computer is used to refer to the existing roadway inventory file for comparison of what is seen in the field. The field crew will enter changes as they move along the road, stopping as necessary to make measurements. The results of the field work are stored on disks which are uploaded to the mainframe system and analyzed. In one of the States, if there are major deviations from past records, a crew will be sent out from the central office to validate the record.

Many of the States use photologs. Two that were visited do not have a photolog program primarily because of the cost involved. Older systems are being converted from film to video disc. One State was using video cameras in the field. Another State included an instrumentation package in their photolog vehicle, so that quantitative data could be recorded on the image. These included altitude and attitude of the vehicle.

Quality issues. The specific methods and procedures used have a varied impact upon the quality, especially accuracy and completeness. Those States that depended upon district offices almost exclusively, not only for the data but for quality control, seemed to have the greatest problem with maintaining a quality data base.

In general, photolog operations seem to be handled independently of the roadway inventory. Using photolog data as a source for inventory updates does not seem to be widely practiced, but does exist. The integration of photolog activities and roadway inventory could result in improved quality.

The roadway inventory data elements that are collected vary somewhat from State-to-State, but not significantly, unless special-purpose files are created. The collection of data regarding longitudinal geometrics, i.e., horizontal and vertical alignment, is the content-area of greatest weakness in most systems. Many do not capture any longitudinal geometrics. Some will use construction plans and as-built drawings. One State is experimenting with the use of data from gyroscopes and altimeters in an instrumented photolog-vehicle to derive curvature and cross slope. Most State systems focus only on primary State highways. This narrow coverage limits the usefulness of the system for comprehensive safety-analyses.

The currency and nature of traffic data is another matter of concern. While some States have fully integrated traffic volumes into the inventory system, others have weakly linked data which are derived more from an ad-hoc effort than from a comprehensive statewide-effort

The State systems studied varied significantly in the types of roadways covered. Most State systems focus only on primary State highways. This narrow coverage limits the usefulness of the system for comprehensive safety-analyses.

Data Preparation, Edit, and Entry

Several of the States that were visited, process inventory-update forms submitted by district offices to a central office. The process at the central office includes reviewing the data in the forms, making corrections as possible, and contacting the submitter to resolve problems. It also involves coding the data for entry. In some cases, the district is responsible for data preparation and entry. Few States have sophisticated editing programs.

One State used the photolog as a source of some inventory data. They overlay grids on the photolog images to take the desired measurements.

> Quality issues. Data receive combinations of manual and automated edits. These are critical to the quality of the file. For instance, remote entry from inexperienced, infrequent users can create errors which must be trapped. Review at the central office by experienced personnel would seem to be an important aspect of quality control. Leaving review and edit at the district level can lead to inconsistencies between districts, unless there is a strong on-line edit involved, or the central office is actively monitoring the submittal of the various districts.

File Linkages and Updating

When records are updated, or new ones entered, they generally are placed in a temporary file for a later batch-update. In several cases, the batch update also includes edit checks. In such cases, rejected records are listed with the errors and the central office works to resolve these. In one State, if the remote entry meets the edit criteria, the record is immediately updated. Another State requires any deviation greater than 6.1 m on the location of an item already in the record to be checked in the field.

Several States download traffic and crash data from other data bases into the roadway inventory. Linkage is either accomplished electronically or by tape transfer. Traffic data primarily are limited to ADT. These data are usually derived from the State's regular traffic census, generally conducted by the State highway planning agency and residing in one of its highway-needs data bases. In some of the States, data from the roadway inventory also is downloaded to the crash file.

> Quality issues. When data are acquired via linkage the quality of those data are often not subject to controls of the acquiring agency. When an interdependency exists, an organizational link is needed to allow for resolution of problems and quality assurance.
>
> The proliferation of "specialized" inventory files, e.g., guardrails, traffic control device, can be symptomatic of outdated data base systems. Lack of flexibility to modify data base structure, coupled with difficulty the end-users may be having accessing the data, lead to the development of independent data bases. This practice can lead to inadequate or inconsistent quality controls,

inefficiencies, and restricted benefits being derived, although modern data base management and access systems have overcome this problem.

Assessment of the Quality of Roadway-Inventory Systems

The site visits made as part of this project did not include specific activities to measure the quality of roadway inventory systems. However, a good source of insight on this is available through the HSIS operated under contract by the Federal Highway Administration. The user manuals produced for the system include an assessment of the files in the system which include roadway-inventory data from five States. The five States are not representative of the status of systems in the U.S. Rather, they tend to represent the exemplary systems, especially from the perspective of having good-quality data, which allows linkages to be made between crash and roadway data.

The analyses of the roadway inventory files done for HSIS, were limited to what could be derived from the files themselves rather than direct observations and measurements. However, the analysis showed these states' inventory systems were well-maintained and generally of high quality for the data represented.

Although direct measurement of accuracy and consistency was not usually reported in the HSIS manuals, the general level of accuracy was. In general, a high degree of accuracy was found when distributions of variables were compared for related variables within a State, and the same variables between States. On the other hand, one State (not in HSIS) that was visited by the project team, had done its own study of data accuracy and found that almost half the data elements in the inventory fields were in error. This had led the State to institute a comprehensive quality assurance program which had reduced errors to below 10 percent.

Precision was not directly addressed except for location coding. Features having some length sometimes were not coded to beginning and end points. Midpoints were use instead, rendering precise relationships, e.g., between crashes and roadway curvature, difficult if not impossible to perform. In some State files, interchanges were not adequately accounted for making it extremely difficult to identify them for focused analyses.

Most States believe their files to be fairly complete. Instances were found, however, where a field was left uncoded (blank) because the feature was not present, instead of indicating lack of presence with a zero. In some other cases, blank fields appeared where they had specific meanings, even though they were not meant to give that meaning. In a few cases and with a few variables, incompleteness appeared to be a problem. In one State, speed limit was not coded for 41 percent of the mileage. In another State almost none of the records had an entry in the field for design speed. Where fields existed for longitudinal data, most of the values were missing. In one State, horizontal and vertical curve data appeared only if it was considered sub-standard. Files from another State showed some lengths along links longer than the links themselves.

Coverage is generally limited to State-maintained highways, principally the high-level ones. Mileage for which a full-featured inventory is maintained generally ranges from 9,000 to 15,000. However, one State keeps comprehensive inventories on 41,000 miles of state-maintained roads. Many of the States keep limited inventory data on secondary and municipal segments as well. As the requirements of recent federal legislation take affect, this probably will change.

In summary, roadway-inventory systems seem to be given more attention in the last few years, prompted by the antiquity of the current systems, the advances of technology, and the requirements of recent Federal legislation. Major improvements are being considered in several of the States contacted. The data maintained on these systems were found to be adequate for many safety analysis purposes, with the following limitations:

1. Data collected from a widely-divergent group of district personnel.
2. Limited quality-control.
3. Limited content, either because key longitudinal-variables were not included, or the values of some of the variables had not been entered in most of the records.
4. Lack of timely data in the file, and The file may not be up-to-date.
5. A lack of history of highway changes which allows multiple-year analysis of safety issues.

Crash-Related Medical Data

From a societal perspective, obtaining accurate information regarding injuries is among the more important data collection tasks in the investigation and reporting of traffic crashes. Most significant costs associated with crashes are medical expenses and lost productivity due to crash-incurred injuries.

The medical data referenced above are *post-crash* data. Another, often overlooked aspect of crash-related medical data is the medical condition of crash-involved parties *prior to* the crash. (In this context, the term "pre-crash medical condition" is used in a general sense to refer to the complete intrinsic physical, mental and emotional state of the involved parties just prior to the crash, not just those parties' basic physical situation.) When assessing the role of medical data in traffic crashes, both pre-crash and crash-produced conditions need to be included.

Data Collection and Management

Pre-crash medical data. For almost all circumstances, there is only a single source of readily accessible pre-crash medical condition data, the police crash report. In some States the only pre-crash medical data collected relate to driver/pedestrian sobriety; others, however, ask the reporting officer to indicate pre-crash medical condition more generally. These crash report forms usually have an area marked "driver condition" where the officer has a set of listed conditions that can be indicated such as shown in table 21.

Table 21 Pre-Crash Physiological Elements
• physical disabilities • previous injuries • fatigue/sleepiness • emotional upset • age • illness • alcohol or drug *presence* • alcohol or drug *impairment*

Often, impairment as defined legally, i.e. driving under the influence of alcohol or other drugs, is recorded in a separate area from "medical condition," or sometimes that information is recorded in two places, one focusing on condition and the other on legal impairment. The collection of pre-crash medical information rests almost entirely on the police officer who completes the crash report.

Terms often are not well defined, and the officer given no instruction as to whether these conditions are to be recorded *any time they are present* or only *when they are factors in the crash.* For example, a person could be missing a limb and be stopped legally at a traffic signal when another vehicle crashes into them. That person would clearly have a disability, but clearly that disability was not a factor in this crash scenario. Based on what the project team learned during interviews, most officers would enter the "driver disability" condition for this scenario. Also, officers are seldom given instructions as to how to record situations where more than more than one condition is present.

Some pre-crash medical data may be located in hospital records systems. For some emergency department treatments, and in almost all cases if a patient is to be admitted to the hospital, toxicology screening will be conducted. While notations as to drug or alcohol presence (and possibly other conditions as well) may be noted on the patient record, there is little evidence that such information is routinely or systematically included in any hospital data bases.

Management of pre-crash driver medical information data rests mostly at the State level. Because the pre-crash condition is usually an objective item, data entry usually is straightforward and little interpretation of this item is possible.

Injury severity. In almost all instances, medical data relating to crash severity comes from the police crash report form. Generally, the reporting officer assigns an injury category to each involved party and the crash is categorized based on the most severe injury recorded. The most commonly used crash-injury reporting scheme (recommended by both ANSI and CADRE) uses five injury categories as defined in table 22.[2]

---

[2] Injury codes are rarely revised in either State or local data bases with the exception of an involved party's death. If a party dies in a specified time after the crash, usually 30 days, the crash record will be revised to indicate the fatality.

Some States go beyond reporting overall injury severity and require officers to indicate body regions that are injured and separately indicate injury severity for each body area. Other States classify injuries only as minor or severe.

In most cases, the reporting officer will classify injuries based on his interpretation of what the injury looks like or on an involved party's claim of a non-visible injury. The "look" of an injury can often be quite deceiving, easily resulting in mis-classification.

Table 22
Crash Injury Reporting

K -	Dead at the scene or before the report was completed.
A -	Incapacitating injury
B -	Non-incapacitating visible injury
C -	Reported non-visible injury
0 -	No apparent injuries

Police officers often will ask medical personnel for their opinion as to injury severity. Most often, these medical personnel will be emergency medical technicians (EMT's) called to the scene of the crash. The police officer's seeking of advice as to medical condition is informal and usually consists of simply asking how severe the injuries appear. If the police officer goes to the hospital to interview the crash-injured party, a similar question may be posed to doctors and nurses there. The final decision as to classifying injury severity is, in any case, the police officer's.

A flow chart depicting the collection, recording and storage of all data from the medical community can be found in figure 5 Unlike pre-crash medical information, however, there are other sources of crash severity information. All medical personnel who treat crash-involved parties keep records of some nature. However, these records are usually for case-specific purposes and the data contained therein are not usually accessible by other entities.

The greatest potential source of crash injury information is probably through emergency medical service (EMS) providers. They usually will check and create a record for all crash-involved parties at the crash scene, whether they are claiming an injury then or not. However, most EMS records do not address overall injury-severity and consist mostly of narratives describing injury or illness. Relatively little objective information on injuries is collected by most EMS systems. Further, most States do not maintain a centralized EMS data base.

Hospital emergency departments do collect and record more information on injury severity, but collect it only on persons who come or are brought to their facilities. Most hospitals have records systems that describe injuries and outcomes in detail, but for patient confidentiality reasons, these records are almost impossible for other entities to access.

Collection and management of injury-severity information by the medical community is for purposes of case management and billings. Little aggregate information comes from the medical community addressing crash injuries in any epidemiological manner other than occasional special studies.

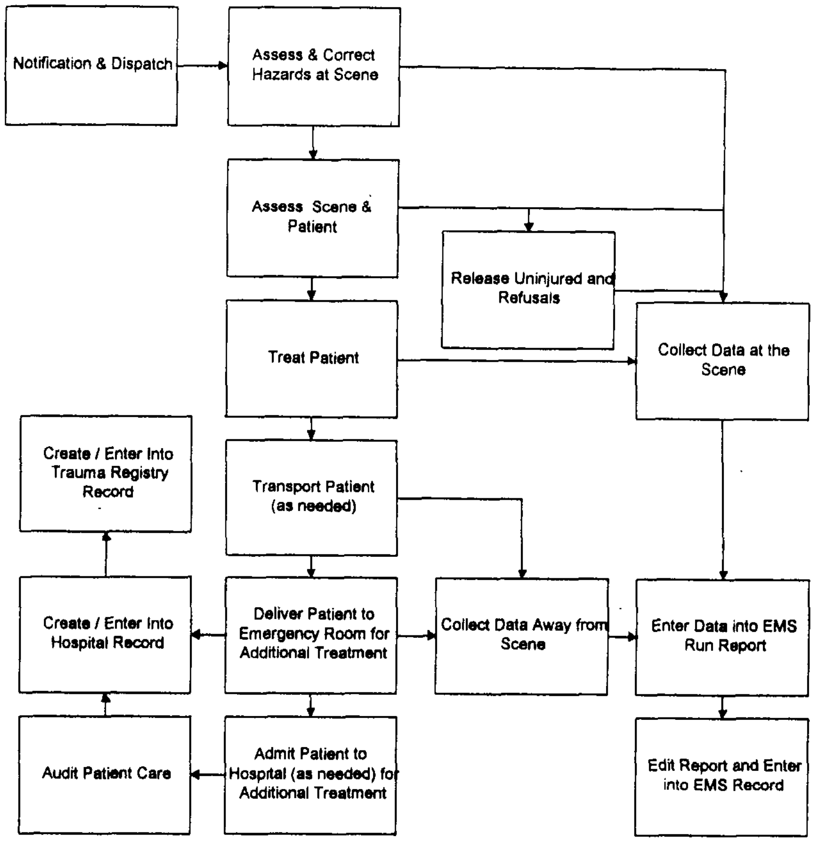


Figure 5
Overview of Collecting and Managing Medical Data

Quality of Medical Data

Pre-Crash medical data.

When assessing the quality of pre-crash medical data by the six dimensions of quality already defined (accuracy, precision, completeness, coverage, timeliness and consistency), some of the dimensions are less relevant than others. It is not possible to assign values of coverage and timeliness to pre-crash medical data because pre-crash medical information is a field on the police crash report form and has the same degree of coverage and timeliness of other data on the report form.

Accuracy is arguably the most important dimension of quality because it is vital to know if pre-crash medical condition was a factor in the crash. In most crashes, the information recorded for this category will be self-reported by the involved party and subject to bias, e.g. fatigue is believed to be a substantially under-reported crash factor.

Completeness is, in most States, a simple matter to assess. Either the "condition box" is completed or it is not. If a reporting officer failed to complete this section and the report reviewer failed to notice the omission, it is probable that "condition" would be either a missing field or interpreted as "unknown" in the data record. Assessment of HSIS State files indicated that this is a significant problem [Council 1991].

Precision of pre-crash medical condition data is, in most states, a function of the number of categories available to describe condition on the police crash report. Further, officers are not trained sufficiently to provide precise diagnoses of injuries.

Consistency of individual officers their interpretation of pre-crash medical condition is usually acceptable, but there is no simple method for assessing consistency across officers or agencies. Providing clear definitions of the condition options and following up with officer training and examples is the recommended approach for ensuring consistency.

Injury severity. Researchers believe that injury severity information is among the least reliable data items found on the PCR, e.g. O'Day 1991. The reason for this is simple. Police are being asked to make assessments in areas where they have little if any training, and to record those assessments before the full extent of injuries are known. Furthermore the definitions used for injury severity levels are often misunderstood not only by the reporting officer, but by users of the data as well. Some states use less precise words such as "serious" or "severe" rather than "incapacitating" for the most severe non-fatal injury classification.

Some agencies also assume that if persons have to be removed from the crash scene by ambulance, they are incapacitated. In an era of "precautionary medicine," people whose injuries prove to be of no medical consequence are often transported to hospitals be ambulance "just to be sure." Also, some injuries such as head and facial lacerations can look much more serious than they are. Conversely, some very serious injuries, e.g. internal injuries may not be visible. This emphasizes the problem with the traditional "KABCO" coding system. In general however, the police probably err on the side of overstating injury severity. This conclusion is based on project team experience, and views expressed by agency personnel during field site visits.

Data collection by EMT's usually does not call for objective categorization of injuries like the PCR. Typically, the EMT will record, in addition to personal identifiers and relevant medical history, patient vital signs, symptoms, complaints, and treatment given. Much of this recording is narrative in nature and is not readily transferable to objective coding.

The most comprehensive system examined in the project, Milwaukee County, Wisconsin, does objectively code involved body parts for each injured person and a coded description of the injury. Even there, however, no specific indication of overall patient injury is provided.

When crash-injured parties are taken to a hospital, specific recording of injury severity usually is done. Medical information usually is available in two hospital-oriented records systems. The first is the hospital trauma registry. All crash victims treated and admitted into the hospital are entered into the trauma registry. The registry will contain specific information on all injuries, and an Abbreviated Injury Scale (AIS) score for each injury. The AIS indicates the degree to which an injury threatens a person's life, and, to a lesser extent, degree of incapacitation.

The other hospital-oriented system for determining injury severity is the so-called discharge code. This is a code, usually assigned by a medical records technician who reviews the patient's record, that indicates very specific injuries, treatments, and long-term prognoses. The primary purpose of the discharge code is for medical insurance purposes. Reimbursement for treatment is based upon the discharge code-defined nature of the injury.

Coverage. Most crash-injured parties are probably never treated for that injury by a physician. Those treated are, to some extent, a self-selected group. While, the most severely injured are usually transported to medical facilities, in other cases only those who want to go are taken (or later go on their own). Several previous studies, e.g., Hauer and Hakkert [1988] have shown that police reports sometimes cannot be found for people treated at hospitals for crash-related injuries. Similarly, hospital reports cannot be found for people listed on the police crash report as having been transported to the hospital.

Accuracy. This element is the most important dimension of quality. Appropriate classification of crashes by severity is important for assessing societal costs of those crashes. For severity based on the PCR, the information recorded for this category will be the officer's best assessment of what he sees and hears. The reporting officer will almost always ask if the involved parties are injured in any way and, for potentially serious injuries, ask medical personnel for guidance as to classification.

Some studies have compared police officer classification of injury severity with hospital severity information. In a recent study completed by the Northwestern University Traffic Institute [Orsay and Lucke 1993], a population of crash-injured drivers who were admitted to the hospital for their injuries was examined. Admission to the hospital is a good criterion for a severe injury since it implies a threat to life, incapacitation, or both. Of this population, the police assigned the most serious injury category to 66 percent of the sample. Seventeen percent were reported to have a moderate injury, 13 percent a possible injury, and 4 percent no injury. Other similar studies [States 1990; Barancik and Fife 1985] have shown comparable (or worse) results.

Although medical mis-diagnoses are not unheard of, assessment of injury severity can be assumed to be accurate by medical professionals. The exception to this might be EMTs, but they rarely classify injuries by severity. Their focus is on field identification of medical problems and providing appropriate field treatment.

Completeness. Most States require the officer to indicate overall injury severity for the crash (based on the most severe injury to any party) and an injury code for all drivers passengers, and pedestrians. (Some States require passengers to be listed on the report form only if they are injured.) In these cases, either the appropriate sections are complete or they are not. Based upon conversations with personnel at sites visited, classification of the crash based on severity is a data category that is unlikely to be omitted. Also, as discussed earlier, medical records from any source are usually complete in terms of describing injuries, but those records are sometimes still not sufficient for severity classification purposes.

Precision. The preciseness of severity data from police crash reports is, in most States, a function of the number of categories available to describe injuries. These categories range from description of overall injury as minor or severe to describing severity on a five or more element scale for each involved body part.

Both trauma and discharge codes are precise. Trauma codes have a potential bias in that their focus is describing threat to life. People can have injuries that are permanently incapacitating but pose little threat to life.

EMS records systems. as mentioned before, generally are not precise in terms of classifying severity. Often, the only objective classification is a trauma score, which, like the hospital trauma registry AIS score, focuses on threat to life.

Consistency. Individual officers are probably consistent with their interpretation of injury severity. However, inconsistencies may occur among officers, agencies and regions of the country. Clear, consistent definitions and training remains the recommended approach for reducing inconsistency. Medical professional must be assumed to be consistent with their classification of injury severity.

Timeliness. Timeliness and coverage are issues when using other sources to acquire severity data. For data based upon hospital discharge codes, the patient must be discharged from the hospital. For most cases, this is not a problem since most treatment is only an emergency department visit or brief hospital stay. Prolonged hospital stays, however, would clearly affect the timely availability of this data item. Timeliness of data from EMS and trauma registry systems would depend upon the nature of those systems and how quickly field data were entered and output reports prepared.

Costs of Collecting Medical Data

The cost of collecting pre-crash medical data is negligible. Providing that information almost always consists of only filling out one of the several boxes on the police crash report form. From the police perspective, the cost of collecting severity information is not much different from collecting pre-crash medical information. It is also a matter of filling in the appropriate areas on the crash report form.

Occasionally, the police officer will seek advice from the medical community in assessing severity. Police officers will often ask EMS personnel how serious a person's injuries appear to be. If the officer needs to go to the hospital, the officer will ask doctors or nurses about the involved person's condition. It would be unusual for a police officer to go to the hospital just to check on condition, however. Interviewing the person regarding the crash is usually the purpose of going to the hospital.

From the medical perspective, collecting and recording injury-severity information is done for medical purposes. As a result there is no cost that should be assigned to crash-data collection. Costs may be incurred when attempts are made to link medical data with crash records.

At present there are no known systems where any medical data are automatically linked to crash data. Where such linkages have been attempted, they have been for special studies. The most ambitious of these efforts is the CODES project currently being conducted by the National Highway Traffic Safety Administration. These projects are examining approaches for linking crash-report data bases primarily with hospital discharge data bases using probabilistic software to link both files. Each CODES study exceeds several hundred thousands of dollars.

Summary of Cost and Quality Issues

Most of the time associated with collecting data from traffic crashes are associated with managing the cases and providing information which can assist insurance companies with liability issues. A small portion of time is spent collecting safety data. As a result, most of the costs incurred by police and others are not applicable to safety aspects.

The time spent managing data were identified and are used to help develop costs. However, a critical element in the time spent relies on how agencies handle the crash reports after they receive them. Some systems are relatively streamlined and depend heavily upon computer assistance. Others, still require multiple manual processing and can be costly to operate.

No attempt was made to identify costs for collecting roadway-inventory and crash-injury data. Although some of the data from both sources are included with the safety data base, most of the data collected are for purposes specific to each source. A number of issues related to quality also affect the usefulness of the roadway inventory.

More critical to the use of data is their quality. Some of the States collecting data have taken steps to help improve the quality of data collected and entered into their systems. Most of this effort has emanated from the agencies responsible for managing the data. However, because most of the quality rely upon the officer collecting data at the scene, there remains a high degree of variability in the quality of those data. Furthermore, some agencies are attempting to enhance the quality of their data through the use local data entry and linked systems. More importantly, however, is that issues of quality including accuracy, precision, consistency, and completeness remain to be resolved.

5. The Cost of Collecting and Managing Crash Data

Computed costs for collecting and managing safety data provide a basis for cost-effectiveness evaluations and for establishing policy on sharing costs. Assessment of cost-effectiveness can be used to establish policies, review alternative strategies, and evaluate programs for improving data collection and management. This chapter summarizes the development of and results from a model to estimate the cost of collecting and managing data from police crash reports. Specific objectives of this effort include:

1. Estimating the unit cost of collecting and managing safety data associated with a crash, and

2. Estimating national costs associated with safety-data collection and management.

Assumptions and Methods

Assumptions

Two basic questions apply when determining whether a cost should be allocated to safety-data collection and management:

1. Is the cost associated with safety data, and (given a "yes" answer to the first question)

2. If safety-data collection were not being done, would the activity occur anyway?

As the research uncovered, many of the costs are related to case management not the handling of safety data, and therefore the answer to the first question is "no." A "yes" answer to the second question indicates that the costs associated with the activity are "sunk costs." Even when safety data were collected, the activity would have been performed regardless; therefore, the costs would have been incurred.

Instead of a strict accounting of costs allocated over multiple operations in an organization, an alternative approach was used. It estimates the personnel time involved in the collection and management of safety data and calculates costs using salary, employee-benefit and indirect-cost rates. While this avoids a costly accounting study, it can introduce errors because the indirect-cost rate can include agency expenditures not all applicable to safety data but which are not feasible to separate, e.g. the cost of a crime laboratory. A rate was derived which excludes, to the extent possible, these potential inaccuracies.

The approach being taken for this study also has implications for how to treat costs associated with the alternative safety-data sources to the PCR. Using the rule of sunk costs just enumerated, expenses related to collecting and managing data for an alternate source are not included unless those data were collected specifically for that source. Furthermore, costs for all sources must be reduced to the extent that fees are collected for the data, e.g. copies of crash reports.

Applicable Cost Studies from Literature

The highway-safety literature has little to offer on the subject of estimating the cost of safety-data collection and management. Initial contact with individual states regarding cost experience also was not productive. One state conducted an analysis of costs associated with processing crash data (Kelsh, Heitzler, and Rauth 1984). The report contained a significant discussion regarding the duplication in the handling of crash reports. Six different groups within the state managed the reports, often duplicating the processes of each other. Costs were estimated for each of these units of government. The results are shown in table 23.

Table 23
Previous Estimates of Crash Data Collection and Management Costs
Comparative Summary of Annual Crash Data
Processing Costs for All Agencies (Virginia)

Agency	Personnel	Computer	Indirect	Misc.	Total	Unit Cost
VDH&T	$362,500	$11,500	$29,000		$403,000	$2.94
DMV	251,200	25,300	20,100	2,500	299,100	2.18
DSP	254,900	13,200	20,300		288,400	2.11
VHTRC	25,800	16,900	2,100		44,800	0.33
VDTS	14,300		1,100		15,400	0.11
OPTS	6,700		500		7,200	0.05
TOTAL	$915,400	$66,900	$73,100	$2,500	$1,057,900	$7.72

Adapted from Kelsh, Heitzler & Rauth (1984)
VDH&T = Virginia Department of Highways & Transportation
DMV = Department of Motor Vehicles
DSP = Department of State Police
VHTRC = Virginia Highway and Transportation Research Council
VDTS = Virginia Department of Transportation Safety
OPTS = Office of Pupil Transportation Safety

Other state studies have identified similar redundancies in data management. A CCSRS design study conducted for the State of Illinois (Illinois Department of Transportation 1988) recommended eliminating redundant entry of crash data by three different agencies at the state level, plus additional handling at the local level. A recent paper (Bozack 1991) summarizes a similar problem in the state of Maryland, where nine different agencies were managing data from the PCR.

Miller, *et al.* (1991) includes estimates of the cost of police response to traffic crashes and compares them with an earlier NHTSA (1983) estimate. These are shown in table 24. The values derived by Miller were based upon at-scene time estimates provided from a survey of five urban police departments and five state police agencies, each having an automated dispatch operation. The estimates are based upon total reported time the officers logged on the crash. The purpose for their estimate did not lead them to further disaggregation or refinements.

Table 24
Previous Estimates of Crash Data Collection and Management Costs

Severity	Miller (1991)	NHTSA (1983)
PDO	$9	$11
Minor	$59	$55
Moderate	$74	$77
Serious	$81	$111
Severe	$88	$154
Critical	$94	$185
Fatal	$185	$185

1988 Dollars Adapted from Miller, *et al.* (1991)

Because the police officer is the primary safety-data collector, national estimates are feasible. Police salary data are available from The Law Enforcement Management and Administration (LEMAS) statistical data base operated as part of the Bureau of Justice Statistics of the U.S. Department of Justice. These data are collected in a national survey of a representative sample of law enforcement agencies, including data from state and local police departments. National sources of salaries, benefit rates and indirect cost rates for civilian personnel involved in safety-data processing were not been found. This information was acquired during visits to state agencies to document PCR management practices.

Analysis-Units

When calculating the unit cost of safety-data collection, the desired approach would be to compute the cost for each element. However, the unavailability of sufficiently accurate and detailed cost data makes such an approach unrealistic. The approach used in this study was to address the PCR as a whole. Some attempts were made to separate costs into important groups of data elements as well.

Costs of data collection and management associated with a PCR will vary depending upon a number of factors. The major cost component for data collection is the police officer's time. The discussion below documents how this can vary depending upon:

1. The severity of the crash;
2. The number of traffic units involved;
3. Whether the crash is in an urban or rural area, and
4. Region of the country.

The classifications used to estimate costs appear in table 25. The six basic cases represent severity and number of units involved. Additional variations result from regional factors and area-type. Cost of collecting data was derived taking each of these variables into consideration.

Table 25
Classifying Units of Analysis for Cost Estimation

Severity	Roadway Location
Property-Damage-Only	Rural
Minor & Moderate	Urban & Suburban
Severe & Fatal	
	Region
Number of Traffic Units	Northeast
Single	North Central
Multiple	South
	West

For managing data, the cost was calculated in a less desegregate manner. There is much less variation represented by the factors. Furthermore, the nature of the available PCR management cost-data limits flexibility for analysis. The specific manner in which the costs are calculated is outlined follows.

Cost Model

Data Collection

Figure 6 presents the formulas applied to calculating the costs of collecting and reporting crash data for any analysis-unit, using the PCR. Although the model allows for four types of personnel (patrol officer - PO, crash investigator - AI, support officer - SO, and supervisor - SU), data from LEMAS only classifies salaries for a sergeant and a patrol officer. The specialist (AI) and support officer (SO) are have the same costs as a patrol officer (generally their rank); these are treated together for the remainder of the discussion.

Because the equations are applicable to any analysis-unit, costs for any portion of the activity can be estimated. Needed in such a case are the time estimates for the selected portion of the PCR for the each type of personnel involved.

Data Management

The approach to estimating data-management costs is somewhat different than used for data collection. The estimate starts with the total annual duty time for the involved staff and uses an estimate of the percent of the person's annual time which is allocable to safety-data management to arrive at an annual time allocable to the process. This approach reflects the most feasible format for measuring safety-data involvement for each type of activity. Once unit

Costs for Collecting Safety Data

Cost per Unit

$$C_u^c = \Sigma(S_i/DT_i)*I_{P,i}*T_i^c*(1+f_{B,i}+f_{O,i}) \quad (1)$$

for i = PO, AI, SO, and SU

where:

- C = cost
- u = unit
- c = safety data collection
- S = annual salary (dollars per year);
- DT = annual duty time (minutes per year);
- I_P = percent of analysis units involved in;
- T = time spent on data collection (minutes);
- f_B = fringe benefit rate (pct of salary/100);
- f_O = indirect cost rate (pct of salary/100);
- PO = Patrol Officer;
- AI = Accident Investigation Specialist;
- SO = Support Officer;
- SU = Supervisor.

and for S, where:

$$S_i = [S_{i/r}\,(1+g)^{(yr - ye)}]*f_R \quad (2)$$

where:

- $S_{i/r}$ = salary from reference data;
- g = inflation (growth) rate;
- y_r = year of salary reference data;
- y_e = year of cost estimates;
- f_R = regional salary factor.

Total Costs

$$C_t^c = N*C_{uc} \quad (3)$$

where:

- t = total;
- N = frequency of use of analysis units;

Figure 6
Costs for Collecting Safety Data

estimates have been derived, they can be applied in a reverse mode based on the estimated number of PCR's affected.

In addition to personnel costs, there are often significant computer system development and maintenance costs. Although hardware and software costs should be included, the proportion of system used in many states usually is so small that it is negligible, unless the agency operating the computer system has a charge-back system. However, at least one state visited by the project team used dedicated hardware for its PCR management. Equations 6 through 9 in figure 7 which is the cost model for data management were developed for use when hardware and related data processing is dedicated to safety-data management.

Determination of Values for Independent Variables - Data Collection

Values were determined for each of the independent variables, based upon information available in the literature, and that collected during site visits to selected states. In some cases, the values are considered appropriate and well documented. Some values were arrived at using assumptions or extrapolations. The discussion below presents the values used and outlines their bases.

Salaries and benefits. Salaries and benefits were determined for police personnel from published data based upon the results of national surveys contained in LEMAS. Benefits

Costs for Data Management

Total PCR Management Personnel Cost

$$C_t^m = \Sigma(n_i * S_i^m * DT_i^m * p_i)\ (1+f_B+f_O) \qquad (4)$$

for i = each clerical and supervisor classification involved

where:

- S_i^m = hourly salary;
- DT_i^m = annual paid-hours;
- m = safety-data management;
- n = number of persons of classification "i";
- p = percent of annual hours of personnel time associated with managing safety data.

Estimating Unit Personnel Cost

$$C_u^m = C_t^m/R \qquad (5)$$

where:

R = the number of PCR's managed

and:

$$CS = D + M \qquad (6)$$

where:

$$D = (D'*CR)-(V*SF) \qquad (7)$$

where:

- M = annual maintenance and operating cost
- D'= the lump sum development cost
- V = salvage value of the system
- CR = capital recovery factor $= \{g(1+g)^L\}/\{(1+g)^L - 1\}$ (8)
- SF = sinking fund factor= $g/\{(1+g)^L - 1\}$ (9)
- L = life of the system development investment (years)

Figure 7
Costs for Data Management

include pensions, health, disability and life insurance, social security or other mandatory retirement plans, workman's compensation, and unemployment compensation. Minimum and maximum salaries, for both sergeants and patrol officers, are available for 1989 (LEMAS 1990) as part of a data set created from the survey. These data, and the benefit rate (expressed as a percent of salary), were averaged. The results are shown in table 26 and figure 8 for each of the regions to be used for the analysis-units.

Indirect rates. The LEMAS data-set does not report indirect costs for agencies. Overhead, or indirect costs, preferably would be estimated directly from a model which accounts for all elements of this type of burden. In one study, Kelsh, Heitzler, and Rauth (1986) arbitrarily decided to use an indirect cost factor of 10 percent of salaries, but this was for data processing. Based on estimates on the various components of overhead, e.g. equipment, administration, and housing, the project team estimated indirect costs at 90 percent of direct costs.

Annual duty time. Because benefit rates include benefit-time off, the full duty year is used for the calculations in the cost formulas. The LEMAS-based reference (Hoetner 1990) indicates a consistent use of the forty-hour week for law enforcement agencies. This translates into 2,080 hours or 124,800 minutes per year. However, actual on-duty hours average 1760 per year; therefore, the costs must be increased by 18.2 percent to represent actual time available. All values used for calculating costs appear in table 27.

Table 26
U. S. Law Enforcement
1990 Officer Average Salaries and Benefits

REGION	PATROL OFFICER MIN.	MAX.	MID-POINT
Northeast	$21,804	$25,583	$23,700
South	$16,963	$20,451	$18,700
Midwest	$20,192	$23,488	$21,800
West	$21,796	$26,134	$24,000
All Regions	$19,598	$23,258	$21,400

REGION	SERGEANT MIN.	MAX.	MID-POINT
Northeast	$31,536	$33,259	$32,400
South	$21,906	$25,969	$23,900
Midwest	$27,033	$29,224	$28,100
West	$28,488	$33,351	$30,900
All Regions	$26,308	$29,626	$28,000

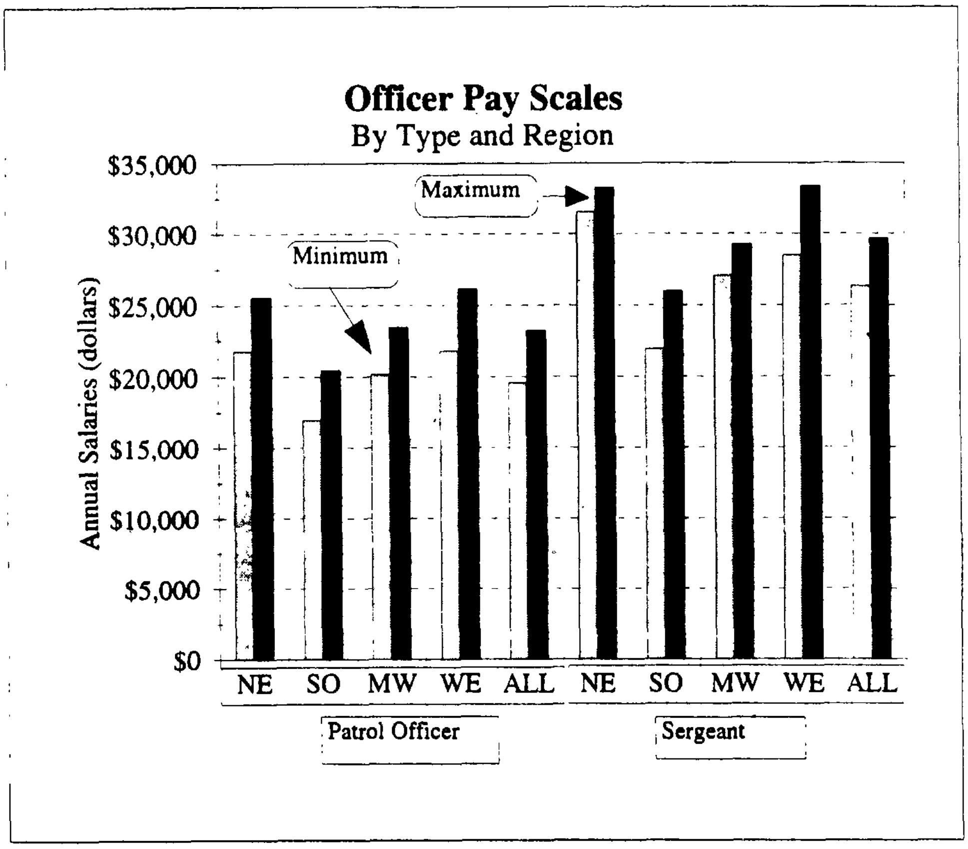


Figure 8
Officer Pay Scales
Source: Law Enforcement and Management Administration Statistics

Table 27
Cost Factors for Base-Case Calculations

Variable	Region: Base Estimate	North East	Mid-West	South	West
Inflation	4.0%	4.0%	4.0%	4.0%	4.0%
Officer Benefits	18.2%	19.3%	18.1%	17.9%	17.0%
PCR Management Benefits	18.0%	18.0%	18.0%	18.0%	18.0%
Officer Duty Time-Off Factor	18.2%	18.2%	18.2%	18.2%	18.2%
PCR Mgt. Personnel Time-Off Factor	18.0%	18.0%	18.0%	18.0%	18.0%
Law Enforcement Agency Indirect Cost	90%	90%	90%	90%	90%
PCR Mgt. Agency Indirect Cost	40%	40%	40%	40%	40%
Reference Year	1990	1990	1990	1990	1990
Officer Annual Hours	2080	2080	2080	2080	2080
Annual Duty Hours	1760	1760	1760	1760	1760
Patrol Officer - cost per minute	$0.42	$0.47	$0.43	$0.37	$0.47
Supervisor (sergeant) - cost per minute	$0.55	$0.64	$0.55	$0.47	$0.61

Personnel costs used. Table 27 also displays the resulting total personnel costs for officers involved in crash data-collection. It combines the salary with benefit and indirect-cost burdens. It also includes the effect of benefit-time off. The values are shown in units of dollars-per-minute so that they can be readily applied to the estimates of data-collection time, which were estimated in minutes.

Crash-data collection time - the literature. As has been discussed previously in this report, report, police officers assigned to investigate traffic accidents spend time on many activities which do not directly involve safety-data collection. "Other" activities involve managing the crash scene, taking enforcement actions against crash-involved traffic law violators, or attending other administrative tasks, e.g., explaining motorist report forms to involved parties. Obtaining the actual time spent collecting safety data proved difficult.

Three studies were located that attempt to address time spent by police officers on the accident investigating and reporting (Howard, Young, and Ellis 1979; Raub and Ferguson 1986; and Stenzel and Lucke 1991). While all are well-done studies and have value, each falls short of what was needed for the current effort. The most significant difficulties encountered in these reports is that they were concerned with officer time spent on the entire AI (accident investigation) process, not just safety-data collection. Furthermore, the studies focused on somewhat narrow aspects of quality.

The most comprehensive of the reports was produced by Howard, Young, and Ellis (1979), for the state of South Australia. This was the only effort found that attempted to time specific tasks completed by the police in investigating and reporting crashes. While considerable analysis of these timings was conducted for their study, it had limited use for this effort. The first limitation is that all South Australia police officers ride in two-officer patrol cars. As a result, the Howard study had to account for the time expended by both officers. In most cases, one of the officers appeared to do little that was directly applicable to the reporting process, so their time was attributed to "making observation," thus inflating that category.

A second difficulty with the report arises because the survey team members who did the timings were rarely at the accident scene either prior to, or arriving with, the investigating officers. The team usually left the scene before all officer activities were completed. The inability of the team members to be with officers throughout the process raises questions as to the accuracy of the time intervals indicated on any given activity. Additionally, they timed only data collection, not its recording.

In their report, Raub and Ferguson (1986) were endeavoring to show the *total* time spent by Illinois State Police officers on crash-related activities, and then extrapolate to the total cost of crash investigation and reporting. The officer-time reporting system used by Raub had one important limitation for making accurate time estimates, however. The times reported appear accurate if the officer completed all crash-related activities as a continuous process. If the officer logged off an crash, performed other duties (including general patrol) and then returned to complete the PCR, the delayed time could not be captured and associated with the original event.

Raub and Ferguson's study has two limitations relating to this project. The first is the problem of capturing discontinuous time spent on an crash, as described above. The estimates of time spent on crash reporting must therefore be considered conservative since there are no doubt many instances of officers interrupting their time on completing a report. Second is the absence of any differentiation between time expended on safety-data collection and time spent on all other accident-related activities such as scene management and taking enforcement actions. The available data provide only time assigned to overall investigation and reporting, not to any subcategories within that task.

A project was conducted for the National Highway Traffic Safety Administration (Stenzel and Lucke 1991) to develop a model for estimating the number of officers needed for an agency's patrol function. To carry out the computations for the model, agencies need to determine their average total time expended on investigating and reporting crashes. For the thirty-two state police districts in eight states that participated in the field test of the model, an average of 3.1 hours was expended per crash with a range of 0.83 to 5.83 hours. These figures included travel times (which were lengthy for some of the rural districts) and all other crash-related activities. There was no effort made to focus on safety-data collection times.

Field data collection of times

Initial investigating officer. The literature provides limited information on the time police spend handling a vehicular crash. Where estimates of overall times have appeared none has disaggregated the time into components, such as:

1. Gathering and recording safety data;
2. Gathering information for case management, and
3. Managing the scene.

This section summarizes the work done on this project to obtain time estimates for components. The initial plan for obtaining times for reporting crashes involved observing and timing officers as they handled a crash. The components measured during the observations are shown in table 28. A test performed in the Rockford, Illinois Police Department, showed the process to be workable. The observations not only allowed team members to measure times taken, but also to observe how quality of data collection might be affected during different parts of the process.

Table 28
Components of Observations

1. Classification by Severity
   a. severe and fatal (K and A-level injury)
   b. minor and moderate (B- and C-level injury)
   c. property damage only (0)
2. Components of Reporting Time
   a. inspecting damage
   b. investigating the scene
   c. interviewing principals
   d. interviewing witnesses
   e. recording crash facts
   f. recording person data
   g. recording vehicle data
   h. preparing narrative
   i. preparing diagram
3. Time Handling the Crash
   a. travel time
   b. at-scene time
   c. other reporting time

Given the success of the pilot test, direct observations were made as various members of the research team visited sites for other project-purposes. Opportunities were taken to ride with police officers in the expectation of gathering additional data. Overall, the team observed 31 crashes in four states; 22 involved property damage only, and 9 involved personal injury. The observations were primarily made for municipal law enforcement agencies. A strong likelihood exists of the observer having an effect on the officer's performance; however, no better means was identified, however, for directly collecting this information.

Early in the study, the team discovered that few crashes could be directly observed using the method adopted. Crashes did not occur with sufficient frequency to use project personnel productively. In an effort to obtain enough data to address at least some of the range of variables listed above, a self-reporting form was prepared and given to cooperating police agencies for their officers to complete. Over a two-month period, the participating departments completed these forms for 274 crashes.

Local agencies did not provide the level of cooperation for reporting times as did the state police. Only six percent of the self-reported cases were from local police departments. The remainder came from two mid-western state police agencies.

Self-reported times were subject to even more measurement problems than direct observations. However, these were considered the best available alternative. The project team believed that between the two sources reasonable estimates could be derived.

Analysis of the data gathered has provided a base of time which can be used for computing the costs associated with police handling of vehicular crashes. However, the data were not sufficient to address many of the dimensions originally identified. Moreover, differences between observed and self-reported times were substantial, making estimates tenuous.

After reviewing the observations and self-reporting forms, several components of the data-collection process were chosen for analytical purposes. In accordance with definitions adopted for analysis-units for the project, crash severity was divided into three classes: "severe" (fatal and A-level injury), "minor and moderate" (B and C-level), and "property damage only" (PDO). Times for collecting and recording data were classified into:

1. Gathering the data, and
2. Completing the report (components comprising: objective portion, narrative, and diagram).

Finally, other recorded times also allowed computation of travel to, and time spent at, the scene. The classifications of data arrived at, therefore, is shown in table 29.

Table 29
Final Classifications for Recording Times

1. Severity of the Crash
   a. severe
   b. other injury
   c. property damage only
2. Time - Safety Data
   a. gathering data
   b. completing objective portion of form
   c. narrative
   d. diagram
3. Time - Handling Crash
   a. travel to scene
   b. on-scene time

Both self-reporting and on-site observations were analyzed. Self-reporting produced times for 274 crashes (some reports did not capture all elements); on-site observations provided coverage for 31 crashes. The times for collecting and recording safety data were substantially different depending upon the source. Compared to on-site observations, times reported by officers were approximately four times higher. This relationship, as shown in table 30, held regardless of the severity of the crash. Most of the differences arose in the component "gathering data." Police reported times ranged from 31 minutes for PDO crashes to 95 minutes for serious-injury ones. On-site observation showed a range between 8 and 28 minutes.

Table 30
Time Spent Reporting Vehicular Crashes
Summary of Median Self-Reported and Observed Times - in Minutes

SEVERITY OF THE CRASH

	PROPERTY DAMAGE ONLY			MINOR INJURY			SEVERE INJURY		
	Self-Reported	Observed	Estimated	Self-Reported	Observed	Estimated	Self-Reported	Observed	Estimated
n	204	22		55	*8*		24	**2**	
MEDIAN	54	22	22	70	*25*	29	127	**39**	52

VEHICLES INVOLVED

	ONE			MULTIPLE		
	Self-Reported	Observed	Estimated	Self-Reported	Observed	Estimated
n	142	*1*	143	127	31	158
MEDIAN	50	*27*	17	65	19	29

GEOGRAPHIC LOCATION

	URBAN			RURAL		
	Self-Reported	Observed	All Urban	Self-Reported	Observed	All Rural
n	131	30	161	142	*2*	144
MEDIAN	48	19	22	65	*18*	30

numbers in *italics* indicate n is too small for median values to be reliable

The differences between reported and observed arose in part because reporting officers were not able to separate time carefully; whereas, the observer was able to distinguish among various activities that the investigating officer was doing. Further, officers made a mental estimate of the time required, rather than using a stop watch or other timing device, and the self-reporting forms were usually completed after the crash reporting was finished.

Finally,the average self-reported values for time were skewed by several high values which were considered unreasonable, and therefore likely to be in error. An attempt to use a median value from both self-reported and observed times also was unacceptably high because most of the times came from self-reporting. To provide times for computing costs, a compromise was made. The time to collect and record safety data for PDO crashes was taken from the median value for observed cases. The median value for each level of severity was adjusted based on differences found in self-reported times among the three levels, PDO, minor injury, and severe injury. These values are the basis of table 31. The table also shows the median times for the analysis-units which were derived from proportioning the median times within each cell in accordance with the variations among self-reported times.

Table 31
Time Spent Reporting Vehicular Crashes
(Combined Self-Reported and Observed Times - in Minutes)
Showing Severity, Location, and Number of Vehicles

	P.D.O. (M=22 min)		MINOR INJURY (M=29 min)		SEVERE INJURY (M=29 min)	
	URBAN (M=22 min)	RURAL (M=30min)	URBAN (M=22 min)	RURAL (M=30min)	URBAN (M=22 min)	RURAL (M=30min)
SINGLE (M=17 min)	12	16	15	21	27	38
MULTIPLE (M=29 min)	16	22	21	29	37	52

M - median values

Involvement of Other Police in Crash Investigation. The above time estimates reflect only the time expended by the primary investigating-officer on the various cases. In many instances, however, assisting officers and supervisors may appear at the crash scene and participate in safety-data collection. Also, in many law enforcement agencies, crash investigation specialists will report some of the crashes while others will be reported by general patrol officers. No references were found in any literature that addressed time expended by support officers at crash scenes or the proportion of crashes reported by AI specialists versus patrol generalists. Times for supporting officers could not be obtained.

In most instances, more than one police unit will respond to a reported traffic crash, even if only one is actually assigned. From experience, the police know that crash scenes are often confused and complex, and assistance with restoring the scene is usually welcome. These back-up officers often take over direction of traffic around the crash and work with the investigating officer in calling for tow trucks and other needed assistance.

Occasionally support officers also will record names and addresses of passengers and witnesses, and assist with supplementary report forms. The investigating officer still will transcribe the personal information onto the PCR. For crashes where scene measurement is needed, an assisting officer commonly will help with the tape measure or with recording measurements. However, rarely does more than one officer complete the PCR.

As crash severity increases, not only does the likelihood of having support officers increase, but it is also more likely that patrol supervisors will come to the crash scene. While supervisors will less often play an active role in managing the crash, they can and will fill the same roles as the support-officers. The amounts of time spent by support officers and supervisors on safety-data collection are estimated based on the ride-along, interviews with police officers, and the experience of the project team, but not direct or indirect measurement.

Finally, even where agencies assign specially trained crash investigators instead of general patrol, these specialists often are compensated at the same rate as general patrol officers. Therefore, no differences salaries or benefits were ascribed to these specialists. Table 32 shows the components used to establish the costs of collecting and reporting data at a traffic crash.

The percent of crashes handled by patrol, support, and supervisors appear in table 32(a). Values shown in table 32(b) provide estimates of the percent of effort that each of the four groups of officers are expected to give to collecting and recording safety data from any one crash. Finally, using as a base the median times established and shown previously, a set of times can be estimated for each group of officers within the classifications of analysis-units as shown in table 32(c).

Unit cost estimates. Multiplying the estimated time spent at the scene from table 32(c) by the unit-costs per officer per minute as originally shown in table 27 yields a set of costs to collect and record data for the police crash report (PCR). The results of these computations appear in table 33. Costs range from 7.00 dollars per PCR for a PDO in a urban/suburban area to 35.00 dollars per PCR for a multiple-vehicle serious injury/fatality crash in an rural area. The values are, for the most part, significantly less than estimated in previous studies (Miller, *et al.* 1991, and NHTSA 1983) as previously shown. However, the values for the previous studies were based on time at the scene, not reporting crashes. When applying costs against observed and self-timing values for travel and "at-scene" times, the costs for handling crashes does come closer to the other two studies.

Determination of Values for Independent Variables for Data Management

Methodology

One of the principal activities during visits by project-team members to the selected states, involved documenting the costs associated with managing crash data. To avoid burdening the personnel providing information, the team accepted costs as provided by the host agency validated only as possible during the course of walking through the processing of the PCR.

Early in the process, the project team discovered substantial variation in practice among the states visited. Furthermore, often two, and sometimes three, agencies were involved in PCR management. For example, location coding was often done within the DOT, but the DOT might not maintain the data. When DOT maintained the data, location coding might be done in a different office of the agency. This made it infeasible to take a single approach to estimating the cost of data management.

Table 32
Crashes Handled and Time Spent

(a) Percent of Crashes (Analysis-Units) Handled by Collector

			Percent of Crashes Using			
				Principal Collector		
No. Traffic Units	Locale	Severity	Supervisor	AI Specialist	Patrol Officer	Support Officer
Single Vehicle	Rural	PDO	5%	2%	98%	0%
		M/M PI	20%	10%	90%	5%
		SPI & F	75%	50%	50%	25%
	Urban and Suburban	PDO	10%	5%	95%	0%
		M/M PI	50%	50%	50%	15%
		SPI & F	95%	75%	25%	35%
Multiple Vehicle	Rural	PDO	5%	2%	98%	0%
		M/M PI	20%	10%	90%	5%
		SPI & F	75%	50%	50%	25%
	Urban and Suburban	PDO	10%	5%	95%	0%
		M/M PI	50%	50%	50%	15%
		SPI & F	95%	75%	25%	35%

(b) Percent of Time Spent Collecting Data at a Crash

			Percent of Time at the Crash		
				^Principal Coll.	
No. Traffic Units	Locale	Severity	Supervisor	Patrol / AI Specialist	Support Officer
Single Vehicle	Rural	PDO	10%	100%	25%
		M/M PI	25%	100%	50%
		SPI & F	25%	100%	50%
	Urban and Suburban	PDO	10%	100%	25%
		M/M PI	25%	100%	50%
		SPI & F	25%	100%	50%
Multiple Vehicle	Rural	PDO	10%	100%	25%
		M/M PI	25%	100%	50%
		SPI & F	25%	100%	50%
	Urban and Suburban	PDO	10%	100%	25%
		M/M PI	25%	100%	50%
		SPI & F	25%	100%	50%

(c) Estimates of Crash Data Collection Time for all Personnel (minutes)

			Time of Involved Personnel (Minutes)		
				Principal Coll.	
No. Traffic Units	Locale	Severity	Supervisor	Patrol / AI Specialist	Support Officer
Single Vehicle	Rural	PDO	2	22	5
		M/M PI	7	29	14
		SPI & F	13	52	26
	Urban and Suburban	PDO	2	16	4
		M/M PI	5	21	10
		SPI & F	9	37	18
Multiple Vehicle	Rural	PDO	3	26	7
		M/M PI	9	35	17
		SPI & F	15	62	31
	Urban and Suburban	PDO	2	19	5
		M/M PI	6	25	12
		SPI & F	11	44	22

PDO = Property Damage Only
M/M PI = Minor and Moderate Personal Injury
SPI & F = Serious Personal Injury and Fatal

Table 33
Unit Costs Adjusted for Proportion of Cases Handled

No. Traffic Units	Region	Property Damage Only		Minor & Moderate		Severe & Fatal	
		Rural	Urban & Suburban	Rural	Urban & Suburban	Rural	Urban & Suburban
Single Vehicle	Northeast	$10.26	$7.41	$14.69	$12.00	$33.29	$25.90
	Midwest	$10.26	$7.39	$14.57	$11.78	$32.47	$25.15
	South	$8.10	$5.84	$11.55	$9.38	$25.95	$20.15
	West	$10.39	$7.49	$14.83	$12.05	$33.34	$25.90
	All	$9.27	$6.68	$13.23	$10.77	$29.82	$23.17
Multiple Vehicle	Northeast	$12.32	$8.89	$17.63	$14.40	$39.94	$31.08
	Midwest	$12.31	$8.87	$17.49	$14.14	$38.96	$30.19
	South	$9.71	$7.01	$13.86	$11.26	$31.14	$24.18
	West	$12.47	$8.99	$17.79	$14.46	$40.01	$31.08
	All	$11.12	$8.02	$15.88	$12.93	$35.78	$27.80

- Includes involvement of support officers, supervisors, and allocation to AI specialists
- Based on 1989 salary and benefit values from LEMAS database

In some cases, it was possible to determine a fairly accurate count of full-time equivalent (FTE) staff, by personnel classification, associated with each of the major steps in the management process. In such cases, the agencies were able to supply salary rates by classification, in some form, as well as a factor used to estimate employee benefits as a percentage of salary. In other cases the agency was able to give a total budget for the section managing the PCR's, but could not give any disaggregate information. Further, employee benefit-factors and indirect (overhead) costs were not known, nor could be found. Because some operations used the same personnel to perform several functions during PCR management, the teams could not disaggregate the costs of the different activities. However, a cost estimate was made for each state, using what was available and making assumptions where necessary to complete the analysis.

Estimates of the Data Management Costs

Table 34 and figure 9 provide a summary of the estimated costs for PCR management. The estimates are disaggregated to show data-management, entry, and edit activity costs separately from location-coding, as well as those associated with the development, implementation and maintenance of hardware and software systems.

The volume of PCR's managed ranged from as few as 15,000 to almost 500,000. Some states had recently revised, or even re-written their computer processing systems while others were laboring with systems that had remained stagnant for ten or more years. Each state

Table 34
Summary of PCR Management Cost Estimates for Study Sites

State	Data Mgmt. Entry & Edit	Locn. Codg.	Sub-Total	Hardwr/ Softwr Systems	Total
A (1)	$3.59	$0.07	$3.66	$2.88	$6.54
B (2)	$2.14	$2.39	$4.53	$5.25	$9.78
C (3)	$16.90		$16.90	$3.85	$20.76
D (2)	$6.85		$6.85	$1.89	$8.75
E (2)	$4.35		$4.35	$2.94	$7.29
F (4)	$4.81		$4.81	$2.54	$7.35
G	$1.79		$1.79	$0.73	$2.53
H	$1.55	$0.76	$2.31	$0.19	$2.49
I	$3.47	$1.05	$4.51	$1.51	$6.02
J	$4.97	$0.62	$5.59	$0.37	$5.96

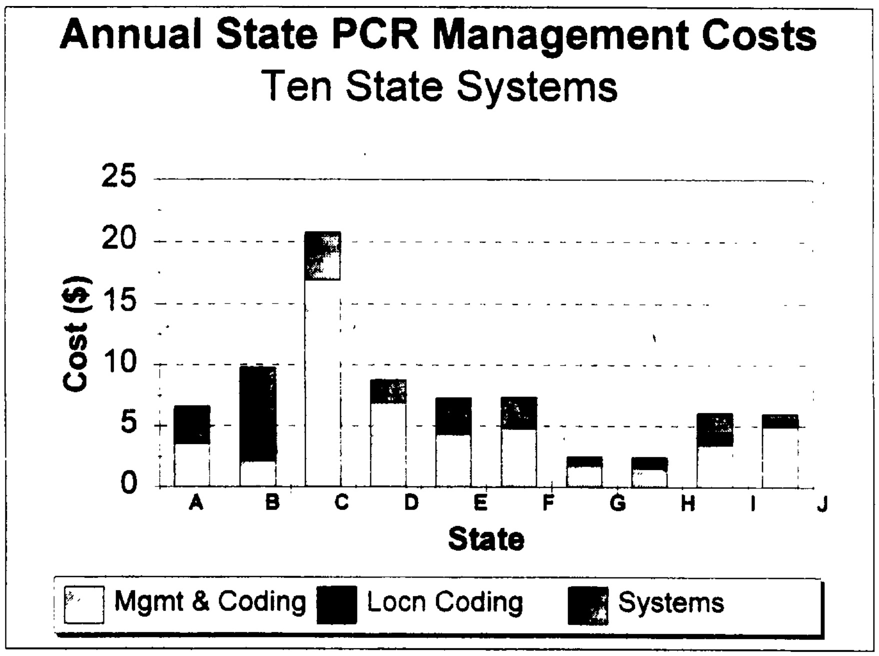


Figure 9
Annual State PCR Management Costs

determined the information that would be captured in the file from their PCR and the types and severity of crashes that would be reported. Each also decided which crashes would be location coded, and how much supporting information from other computer databases would be accessed and added to the basic PCR data. The variety of functions performed by the states in this study, the range of sophistication of their computer systems, and the variations in PCR volume, make it difficult to compare one state to another. However, several observations and conclusions can be made that pertain to all the states studied.

Most noticeable in similarity between the states was the high proportion that manual procedures were of the total management cost. In one state, 94 percent of the cost of processing PCR data was associated with manual handling of documents and entry of data into computer systems. The lowest percents (low 60's) were found in states which had recently upgraded their computer systems and purchased dedicated computer hardware to improve handling of PCR data.

The two states with the lowest cost per PCR (2.44 and 2.42 dollars) had volumes of 350,000 and 400,000 respectively. The high volume in one of these states provided the incentive to automate their PCR processing by developing a scannable form. The other state had a relatively old computerized database-management system which meant its annual automation-costs were low. However, it also used a single set of individuals to perform most of the tasks of editing, coding and entry, thus minimizing personnel paper-handling.

In contrast to these two states, one state which processed almost 500,000 PCR's in 1992 had an annual cost of 5.67 dollars per PCR processed. This state has not yet chosen to use online computer systems for processing the data. Instead, it uses a batch entry and editing system with individuals who specialize in just one aspect of the process. In this state, the paper PCR document must be circulated between various specialist-groups and agencies, and is monitored by production-control employees to insure PCRs are not delayed, mis-routed, or lost.

Some state included in this study extract pertinent information from other databases through computerized linkages. These data, such as driver name and address, or roadway features, can then be automatically included with in the PCR record. The linkages can be relatively expensive to develop, but eliminate manual keying efforts when implemented. These linkages may reduce the time required by an officer to complete a PCR in the field by reducing the number of data items that must be captured. However, a few data elements are gathered in this manner, and cost savings are therefore not likely to be significant. The major benefit may be in the accuracy of the data.

Certain conclusions can be made concerning the factors that affect the overall cost of processing PCR data:

1. Economies of scale can occur with high-volume PCR processing. A computer system costs essentially the same to develop whether it is used to process 1,000 or 100,000

PCR's. Other economies of scale can come into play. The three states with the highest PCR volume were also the states with the lowest PCR processing costs.

2. Efficiencies can be achieved by minimizing the handling of the PCR, whether it is for coding, data entry, editing, or control. Fewer steps for the PCR improve efficiency of the process.

3. Minimizing the data-entry tasks will reduce the costs which represent the majority of those currently experienced. Scannable forms are just one example. Use of computers to help collect data will result in electronic fields which can be directly dumped to a computer without keyed entry. The relative efficiency of this latter strategy depends upon how well the computer prevents errors entering the record in the field.

Table 34 shown previously provides an average cost for processing a PCR in the states visited of 7.50 dollars per year. This includes data handling, entry, edit, location coding, and hardware and software systems costs derived from a highly variable set of 10 individual values. The limited sample prevents any statistical testing (as demonstrated by the standard deviation of 5.00 dollars). However, the states were chosen because they employed differing processes appear to represent a range of conditions expected across all 50 states. In the absence of any reason to do otherwise, the unit costs derived here will be used without adjustment to make total unit-cost estimates and national cost estimates.

Estimates of National Frequencies of PCR's

A 1989 file of the NASS-GES was used to gain an understanding of the distribution of crashes by severity. The number of crashes were estimated here for the United States as a whole, and for four separate geographical regions as defined in the NASS program. They serve as a basis for estimating the total time spent by police officers in crash investigation. Tables of expected error in these estimates are presented in the National Accident Sampling System, General Estimates System Technical Note, 1988 to 1990. Briefly, the probable error ranges from 11 percent for the smallest estimate to 6.9 percent for the largest, although there are potential bias errors in the these data that are larger than either of these.

Blincoe (1990) has shown that GES shows about 15 percent fewer police-reported crashes overall than counted by the FHWA annual statistics report. Because the FHWA report comes from data contributed directly from the various states' crash files, it is likely to be more complete. Because of the lack of further information regarding variation, regional variations have not been adjusted.

A second kind of missing data is represented by cases that are investigated (taking up police officer time) but are not formally reported (and thus do not appear in the FHWA or GES statistics). Many derive from crashes involving only property damage which is below the reporting threshold; however, a number of vehicular crashes involve injuries. Work on

CODES has been uncovering more than a few hospital admissions for which no police report can be found.

As a result of these other considerations, estimates presented in table 35 are likely to be somewhat understating the actual experience. The impact of these potential variances on costs can be measured through sensitivity testing.

Concern exists about injury distributions because states in the southern region exhibit a higher than average proportion of "A" injuries (O'Day 1993). The sum of serious and fatal injury crashes (K and A) for the south region (207,932) divided by the sum of all injury crashes for the south region (2,236,960) is 0.0925. For the remainder (the other three regions) of the country the ratio is 0.045, or one-half that of the south. This occurs because the definition used for "A" injury in many of the southern states is considerably more liberal than average. For this study, the severe and fatal injury crashes in the south were adjusted to reflect the percentages elsewhere increasing the number of minor and moderate injuries in the south region compared to the NASS tables. These revised values also are included in table 35.

Estimates of National Costs of Safety for Data Collection and Management

The unit costs described and presented above were applied to the frequencies for each analysis-unit, providing estimates of safety-data collection and management costs for PCRs associated with each analysis-unit. Resulting regional and national estimates of safety-data collection and management costs are shown in table 36 for the values previously derived and presented. This is referred to as the "base condition." The total national cost for safety-data collection and management was estimated at 122 million dollars. Figure 10 demonstrates that safety-data management constitutes approximately 40 percent of the total. The overall

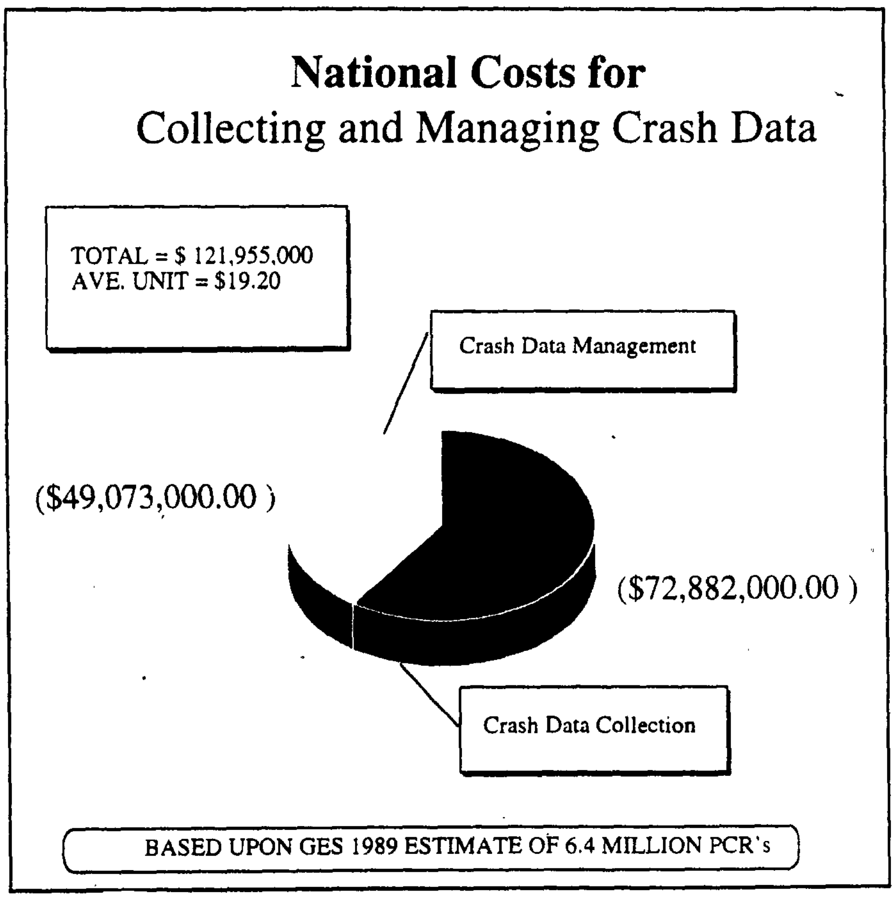


Figure 10
National Costs for Collecting and Managing Crash Data

Table 35
1989 National Crash Frequency Estimates Used for Calculations

No. Traffic Units	Region	Property Damage Only		Minor & Moderate		Severe & Fatal		Sub Total		
		Rural	Urban & Suburban	Rural	Urban & Suburban	Rural	Urban & Suburban	Rural	Urban & Suburban	Grand Total
Single Vehicle Crashes	Northeast	162,661	55,042	74,782	45,577	14,648	10,002	252,091	110,621	362,712
	Midwest	381,106	135,738	110,484	41,444	29,256	10,128	520,846	187,310	708,156
	South	225,704	118,951	117,691	59,600	30,536	12,508	373,931	191,059	564,990
	West	81,320	68,227	48,497	47,214	17,470	11,963	147,287	127,404	274,691
	Sub Total	850,791	377,958	351,454	193,835	91,910	44,601	1,294,155	616,394	1,910,549
Multiple Vehicle Crashes	Northeast	278,778	208,339	126,441	90,854	12,137	5,883	417,356	305,076	722,432
	Midwest	574,764	452,877	177,069	128,062	29,515	13,320	781,348	594,259	1,375,607
	South	540,086	586,008	240,764	238,190	38,213	22,709	819,063	846,907	1,665,970
	West	144,427	299,818	64,147	146,099	11,293	20,409	219,867	466,326	686,193
	Sub Total	1,538,055	1,547,042	608,421	603,205	91,158	62,321	2,237,634	2,212,568	4,450,202
	Total	2,388,846	1,925,000	959,875	797,040	183,068	106,922	3,531,789	2,828,962	6,360,751

Based upon NASS General Estimates System from 1989

(Note: crash data from 1994 show a decrease in 1.4% for reported crashes, from 6,360,751 to 6,273,900) unit cost for safety-data collection and management associated with a PCR is estimated as 19.20 dollars for the base case.[1]

Another way of examining the costs is by number of vehicles involved and the severity of the crashes. Figure 11 demonstrates that single vehicle crashes constitute slightly less than a third of the national cost. National costs generally decline inversely with severity of the crash, primarily in proportion to their relative frequency of occurrence.

The values shown in the exhibits referenced above were based on times derived from the limited observations and adjusted using self-reported data. This adjustment was done because the analyses of self-reported times suggested values which appeared unrealistically high. Had the self-reported times been used, the costs of collection and management would have been approximately 35.70 dollars instead of 19.20 dollars per crash reported.

Sensitivity Analyses

A number of assumptions made in developing the values for the cost model as well as in constructing the model itself. The confidence in the estimate of unit and total costs is limited by these potential sources of error. An estimate of a national value is intended only to be an order-of-magnitude evaluation. However, the cost model also is intended for assessing the impact of alternative strategies for improvement. Therefore, of use is how the end result of the model (the measure of which is taken as the average unit-cost) may vary with variations of the inputs. This can provide a perspective on the size of the order-of-magnitude, as well as the ability of the model to detect changes that may result from various strategies. To accomplish this, several key variables were altered from the base condition to determine the effect on the average unit cost. The results are discussed below.

The cost of officer time collecting crash data predominates in the total. The primary components of that cost include; salary, benefits, indirect cost, and estimated time collecting crash data. Figure 12 shows the affect of varying collection time from the base condition. Because the field data showed significant differences between times determined from direct measurement and those self-reported by officers, the affect of increasing time estimates by as much as 100 percent are of interest. Given that data collection costs represent a large proportion quarters of the total, a 100 percent increase in data collection time will raise total costs 59 percent, resulting in an average unit cost of collection and management of slightly more than 31.00 dollars per PCR.

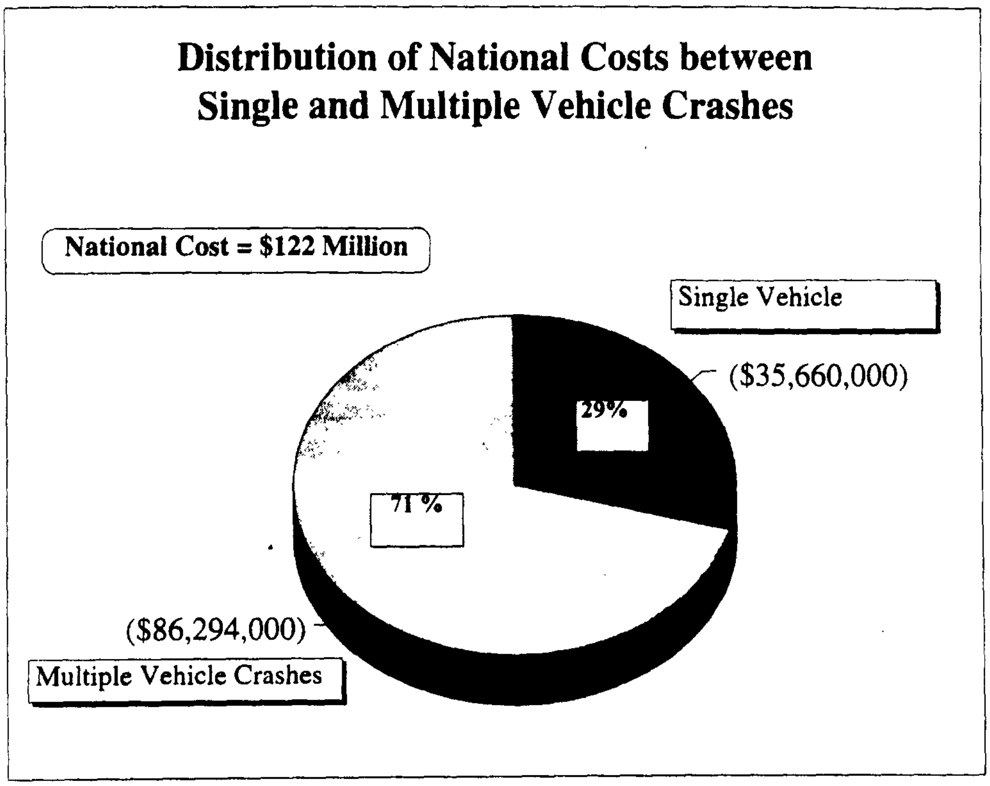
---

[1] This Research Report uses data based on salaries and costs derived in 1988. The authors elected not to change all the tables and figures to update the values to 1995. Had this step been done, based on a 40.7 percent change in the Consumer Price Index during the past seven years, the average cost of handling a crash would have increased from $19.20 to $27.00.

Table 36
Calculated Total National Safety Data Collection and Management Costs

Property Damage Only		Minor & Moderate		Severe & Fatal		Sub Total Rural and Urban		
Rural	Urban & Suburban	Rural	Urban & Suburban	Rural	Urban & Suburban	Rural	Urban & Suburban	Total Collection
$1,670,000	$408,000	$1,099,000	$547,000	$488,000	$259,000	$3,257,000	$1,214,000	$4,471,000
$3,908,000	$1,004,000	$1,610,000	$488,000	$950,000	$255,000	$6,468,000	$1,747,000	$8,215,000
$1,827,000	$695,000	$1,359,000	$559,000	$792,000	$252,000	$3,978,000	$1,506,000	$5,484,000
$753,000	$456,000	$642,000	$509,000	$521,000	$277,000	$1,916,000	$1,242,000	$3,158,000
$7,883,000	$2,526,000	$4,650,000	$2,088,000	$2,740,000	$1,033,000	$15,273,000	$5,647,000	$20,920,000
$3,434,000	$1,852,000	$2,229,000	$1,308,000	$485,000	$183,000	$6,148,000	$3,343,000	$9,491,000
$7,073,000	$4,018,000	$3,096,000	$1,810,000	$1,150,000	$402,000	$11,319,000	$6,230,000	$17,549,000
$5,247,000	$4,106,000	$3,336,000	$2,682,000	$1,190,000	$549,000	$9,773,000	$7,337,000	$17,110,000
$1,801,000	$2,696,000	$1,141,000	$2,113,000	$452,000	$634,000	$3,394,000	$5,443,000	$8,837,000
$17,101,000	$12,408,000	$9,661,000	$7,797,000	$3,262,000	$1,733,000	$30,024,000	$21,938,000	$51,962,000
$24,984,000	$14,934,000	$14,311,000	$9,885,000	$6,002,000	$2,766,000	$45,297,000	$27,585,000	$72,882,000
						AVERAGE UNIT COSTS		**$11.50**

No. Traffic Units	Region	Crash Data Management Cost					
		Data Mgmt. & Entry	Location Coding	Indirect Cost	Hrdwr. & Softwr.	Total Mgmt.	**GRAND TOTAL**
Single	Northeast	$1,102,000	$342,000	$578,000	$776,000	$2,798,000	**$7,269,000**
	Midwest	$2,151,000	$668,000	$1,128,000	$1,515,000	$5,462,000	**$13,677,000**
	South	$1,716,000	$533,000	$900,000	$1,209,000	$4,358,000	**$9,842,000**
	West	$835,000	$259,000	$438,000	$588,000	$2,120,000	**$5,278,000**
	All	$5,804,000	$1,804,000	$3,043,000	$4,089,000	$14,740,000	**$35,660,000**
Multiple	Northeast	$2,195,000	$682,000	$1,151,000	$1,546,000	$5,574,000	**$15,065,000**
	Midwest	$4,179,000	$1,299,000	$2,191,000	$2,944,000	$10,613,000	**$28,162,000**
	South	$5,061,000	$1,573,000	$2,654,000	$3,565,000	$12,853,000	**$29,963,000**
	West	$2,085,000	$648,000	$1,093,000	$1,468,000	$5,294,000	**$14,131,000**
	All	$13,520,000	$4,201,000	$7,088,000	$9,523,000	$34,332,000	**$86,294,000**
Grand Total		$19,324,000	$6,005,000	$10,132,000	$13,612,000	$49,073,000	**$121,955,000**
				AVERAGE UNIT COSTS		**$7.70**	**$19.20**



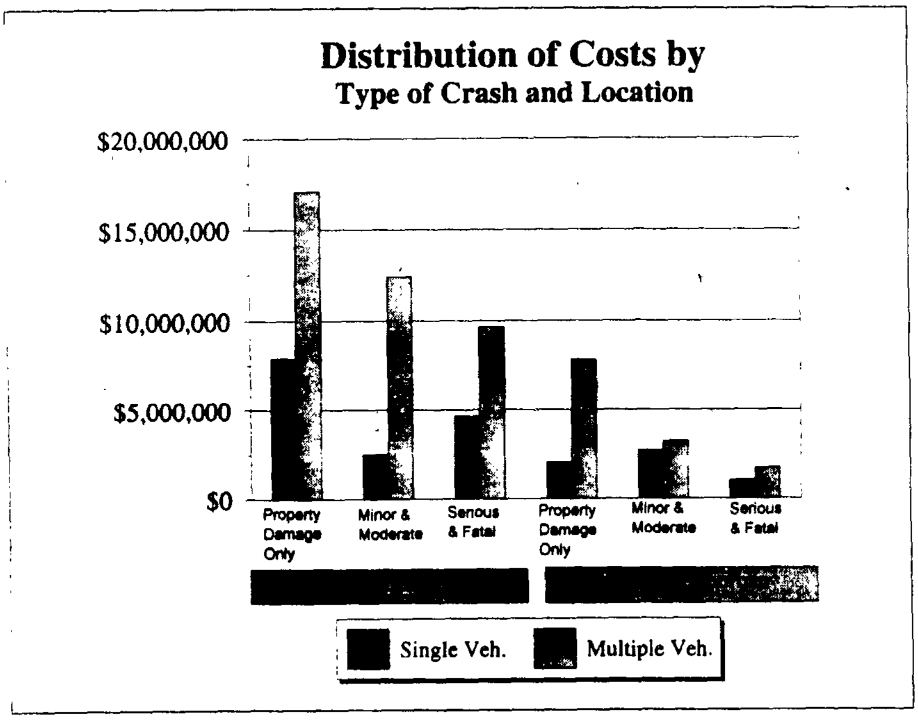


Figure 11
National Annual Cost of Crash Data Collection and Management

A similar difference was demonstrated when changing salaries and benefits for officers. Therefore, strategies that use lower-cost personnel such as non-sworn crash investigators could result in lower unit costs, even though these person may take the same amount of time to collect and record data.

Because data-management costs comprise about approximately 40 percent of the total cost, changes in these costs will have a lesser effect. Figure 13 confirms that a 50 percent change in data management costs would change total unit cost by about 20 percent.

Other sensitivity analyses were performed. They displayed smaller, but in some cases still substantial, changes from the base conditions, e.g. increasing indirect costs by 50 percent increased operating costs by 16 percent. Sensitivity analyses showed that assumptions made for costs can play an important role. Without firm values for many of these assumptions arguments for significantly higher costs of handling police crash reports are reasonable. However, for Chapter 7 which is devoted to assessing strategies, the base condition of 19.20 dollars per crash is used. Of this amount, 7.50 represents the costs for processing the police crash report. The remaining 11.70 dollars is the cost for collecting and reporting the crash data themselves.

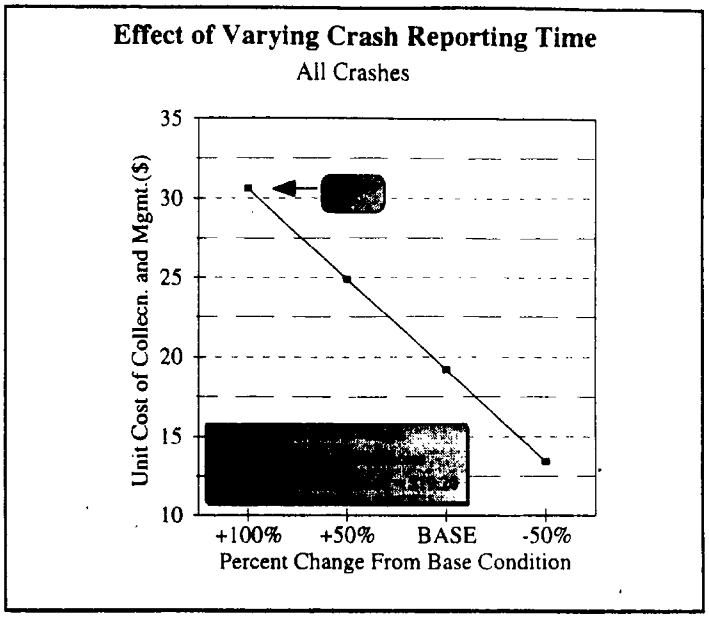


Figure 12
Effect of Varying Crash Reporting Time

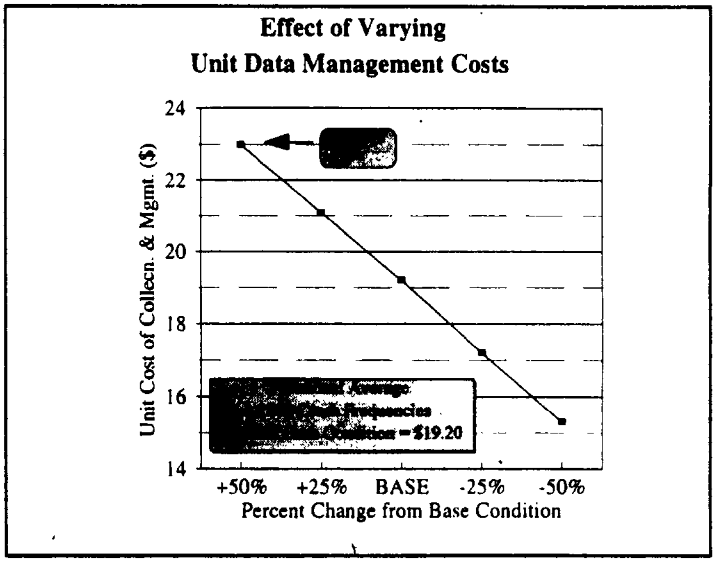


Figure 13
Effect of Varying Unit Data Management Costs

6. Goals, Objectives, and Practices Affecting Highway Safety Data

In order to help resolve issues and address costs of handling traffic safety data, a set of goals and objectives has been formulated. In addition, on-site visits and discussions with members of the safety community have identified noteworthy practices which can form a basis for the development of strategies designed to meet the goals and objectives. This chapter identifies the mission, goals, and objectives which are crucial to improving the system, along with noteworthy practices which can meet the goals and objectives. The next chapter addresses strategies and evaluates those strategies based on cost-effectiveness measures.

Previous chapters have addressed issues and costs related to collecting and managing safety data. Also included in the discussion were issues relating to quality of the data and how the processes of collecting and managing data affect the quality. The noteworthy practices described in this chapter demonstrate how current efforts are attempting to resolve the issues.

Mission, Goals, Objectives, and Principles

Mission of the Highway Safety Data System

The highway-safety information system operates within a larger system context. It is a subsystem of the highway-safety decision-making system, which in turn is a subsystem in the highway transportation system. This hierarchy continues, but the safety system can be a major source of information in a broader decision-making context.

The mission of the highway-safety information system is to provide information in support of decision makers in the most useful manner. The functions which result from this mission include collection and management of highway safety data. Users should interact with the source of information through a decision-support system which can guide the user through interaction with the data.

Goals, Objectives, and Principles

Goals. For this project, five goals have been identified within the mission of the highway-safety information system. Each of these goals, in turn, has objectives which need to be met. The goals apply to each of the sources of data: crash, roadway inventory, and medical.

1. Minimize the costs of collecting and managing the data.

2. Maximize safety in the collection of data:
   a. maximize the effectiveness of operation at the scene.
   b. minimize the exposure of involved parties.

3. Minimize effects on organizations, institutions, and operating systems.

4. Maximize quality of data for uses and users

5. Maximize suitability of data for uses and users

Objectives identified for the goals. Goals tend to be general; objectives help define them more clearly. An objective is specific and suggests measurement of achievement. For each of these goals are a set of objectives which will be used for evaluating strategies.

Objectives considered applicable to the collection and management of highway safety-data appear in the list below. In the list, the goal appears in bold faced print and is described briefly. Also shown for each objective elements which can provide measures of effectiveness (MOE). Most of the objectives can be applied to all three sources of data upon which this project is focussing.

1) **Minimize the costs of collecting and managing the data**
Costs incurred for collecting, reporting, and handling the data will be minimized. Included are costs of people, capital, and operations.

   a. costs of personnel required
      - direct and indirect costs
      - time spent collecting
      - time spent managing
      - time spent editing and reviewing
   b. costs of capital
      - equipment
      - materials
   c. costs of operating the information system
      - development and replacement
      - maintenance

2) **Maximize safety in the collection of data**
The time taken to clear a crash, and to collect and record data, as well as the number of, and exposure to, persons involved is minimized. In addition, any process minimizes congestion.

   a) Maximize the effectiveness of operation at the scene.

      1. time for response of appropriate personnel and equipment
      2. congestion at the scene
         - provide positive guidance through and around the scene
         - quickly clear the scene
         - rapidly restore scene to normal traffic flow
      3. time required on-scene to collect data

b) **Minimize the exposure of involved parties.**

1. time subjects involved in or related to the crash remain at the scene
2. number of officers and related safety personnel responding
3. time safety personnel remain at the scene

3) **Minimize effects on organizations, institutions, and operating systems.**
Organizations prefer status quo. Changes to the safety information system should, therefore, have a limited effect on organizations and their operations. Included are changes to processes, staffing, structure, and interference with other organizations, as well as time required to implement the process.

a. changes to processing systems
   - new systems
   - revisions to current systems
b. changes in staffing required to operate the system
   - number or persons
   - diversity of personnel clarifications
c. changes to the organizational structure
d. interference with other organizations
e. the time required for implementation

4) **Maximize quality of data for uses and users**
Six dimensions of quality are needed to reflect the complex needs of users. The attributes of consistency and coverage are measured over time, and coverage is also measured over geographical space. Quality attributes are:

a. accuracy
b precision
c timeliness
d completeness
e. consistency
   - among collectors and recorders (each collector records the data in the same manner with the same definition)
   - over time (data are collected based on the same definition over time)
f. coverage
   - by classification (all classifications, e.g. severity of crash, are completely covered)
   - over time (the same frequency of coverage occurs over time)
   - geographically (the same frequency of coverage occurs regardless of location of that coverage, e.g. the same data are gathered for local streets and expressways)
   - temporally

5) **Maximize suitability of data for uses and users**
A maximum number of cases are made available to the widest body of users. These users have easy access to the data with a maximum of technical support.

a. the number of potential uses and users

b. number of cases (sample-size) available for analytical purposes
c. the ease of access to, and use of, data
d. access to technical support

Principles for collecting and managing data. Guiding the operation of the highway-safety information-system are general principles which are critical to the mission of acquiring appropriate data, managing them, and making them accessible to users. The principles apply to three dimensions of the safety-information system: the system itself, people, and data. Applicable principles appear below

1) The System

- involves all participants, collectors, reviewers, data managers, and users.
- provides ongoing monitoring, auditing, and evaluation, and feedback, as well as provisions for keeping it current.
- crosses organizational lines and interacts with multiple files and sets of data.
- subjects data to quality control as early in the collection and management process as possible.
- provides for entry of data into the system as close to the point of collection as possible.
- does not duplicate the collection of any data element within the system.
- maintains identity of the source of each data element and record (e.g., identifies if crash report is a citizen report, desk report or officer field investigated).

2) Persons Who Work in the System

- are well trained and regularly updated in all aspects of the system, but especially in the one for which they are responsible.
- are highly motivated, especially through job-related rewards.
- receive regular feedback on performance, as well as usefulness of their efforts.
- handle the minimum required data.
- use data-collection instruments which are easy to understand and manipulate.
- are well equipped.
- record the data using systems which require manual intervention only where automation cannot deal with it adequately, but which provide
  - verification;
  - checks for errors, and
  - corrections or avenues for correction.
- edit and correct the data based upon access to the original facts of the case.
- forward (transmit) findings immediately.

3) Data Available to the System

- reside in readily and easily accessible locations where each element resides in only one location, but is accessible to all others.
- are available in a manner which is transparent and irrelevant to the managers and users.
- reside in systems which are easy to maintain, modify, and link.
- have an adequate and consistent history.

- are overseen by an organization which is effective in performing its duties to maintain quality assurance for the collection, recording, analyses, entry, and maintenance of the data (generally the organization which has the greatest need for the data).

Noteworthy Practices

A number of state and local agencies throughout the nation are attempting to resolve issues with practices that the project team has considered noteworthy. These efforts affect the collection and management of crash safety-data. This section provides a summary of each practice.

Description of the practices was gained through on-site visits to 10 states and one city, as well as telephonic contact with another police agency. Table 37 presents a list, by title, of each of the 23 practices. It groups them according to broad categories. The list should not be considered inclusive; rather, it represents information learned as the project team made its visits. Those sites not visited could well have noteworthy practices which would be important in addressing the issues raised previously.

Traffic Crash, Data Collection and Management

State-level User Groups

1. Illinois Traffic Safety Information Systems Council (ITSISC)
   Agency. Illinois Department of Transportation (IDOT)

Overview. The IDOT has created ITSISC to coordinate the efforts to improve traffic-safety information systems in the state. This was done both in response to recommendations made by a NHTSA-sponsored, traffic-records assessment team, and the experience that IDOT had using a coordinating task force when recently re-designing the state crash report. Members represent a wide distribution of collectors, managers, and users of crash safety-data.

Cost considerations. The operation of the council is on a volunteer basis. Staff support for the council currently is provided by IDOT's Traffic Safety Division through its regular personnel.

Impacts. The committee represents a potential for affecting the quality and efficiency of the highway-safety information-system in Illinois. The framework for the strategic plan emphasizes the areas of systems integration, improved quality, easier access and use, and effective system administration.

Table 37
Noteworthy Practices

Category	Practice Title	State/Agency
Traffic Crash Data		
State-level User Groups	1. Illinois Traffic Safety Information Council (ITSISC)	IL
	2. Accident Records System Advisory Committee (ARSAC)	NJ
	3. Central Accident Processing System Committee (CAPS)	VA
Quality Assurance Programs	4. Data Element Sponsorship	FL
	5. Quality Assurance Program	WI
	6. Outreach Program	WI
Direct Communication with Local Agencies	7. Communicating with Local Data Collectors	OH
	8. Communicating with Local Data Collectors	PA
Scannable Forms	9. Scannable Police Crash Report Form	MI
	10. Scannable Police Crash Report Form	WI
Automated Systems	11. Vision 2000	CA
	12. Local Agency Data Entry	OH
	13. Linked Databases	PA
Non-sworn Crash Investigators	14. Use of Non-Sworn Crash Investigators: Community Service Officers	CO (Ft. Collins)
	15. Use of Non-Sworn Crash Investigators	FL(Ft. Lauderdale)
	16. Use of Non-Sworn Crash Investigators	SD (Rapid City)
Miscellaneous	17. State-level Interpretation of Crash Facts from PCR	PA
	18. Proactive Problem Identification	CA
Roadway Inventory		
Roadway Inventories - Automation and Upgrade	19. California Highway Inventory and Performance System	CA
	20. Automated Updating of Roadway Inventory	PA
	21. Automated Roadway Inventory Data Collection and Straight-Line Diagram	VA
Quality Assurance Programs	22. Quality Improvement Program: Roadway Inventory	FL
	23. Quality Assurance Program - Crash Location Coding	FL

2. Accident Records System Advisory Committee (ARSAC)
Agency. New Jersey Division of Highway Traffic Safety (DHTS)

Overview. To assist with upgrading their current safety database, New Jersey has formed ARSAC. In addition to the Division of Highway Traffic Safety as the lead agency, directors from 14 other state agencies are represented. The committee is responsible and has been given authority for making decisions regarding the collecting, recording, and use of traffic crash data.

Cost considerations. Only indirect costs are involved such as those related to the time spent at meetings and task forces. Some non-personnel services may be required, e.g. travel and printing, for members of the committee, but these should be minimal.

Impacts. ARSAC is a committee of users. It represents a unique opportunity to have a disparate body of persons deal with a common topic. The effect of such work should be to enhance the flow of data from the crash report, while at the same time recognizing the varied needs of the users. Reaching an understanding by which the Committee has authority to make decisions was an important element.

3. Central Accident Processing Committee System (CAPS Committee)
Agency. Commonwealth of Virginia, Department of Motor Vehicles (DMV), Department of Transportation (DOT), and Department of State Police (DSP)

Overview. When the current central accident processing system (CAPS), operated by the DMV, was being developed, a CAPS Committee was formed to oversee the system development. The committee included representatives of the DSP, DMV, and DOT The committee now plays a central role in maintaining the system and quality of data available from it.

Cost considerations. The cost of participation is borne by each agency for its personnel. Cost of committee operation was not estimated, but appears to be minimal. Costs of implementing committee actions is a separately borne, primarily by DMV as the agency responsible for CAPS.

Impacts.

a. The committee has fostered interagency communications.
b. Safety-problem identification has been facilitated and become more comprehensive since it has become a committee function.
c. The commonwealth has been more successful at receiving federal safety grants now that a more comprehensive approach is possible.
d. A better working relationship has resulted with both FHWA and NHTSA.

Quality Assurance Programs

4. Data Element Sponsorship
Agency. Florida Department of Transportation (DOT)

Overview. The Florida DOT, Transportation Statistics Office has been through a major redevelopment of their roadway inventory system. As part of the redevelopment an assessment was made of the data elements to be included in the file. A strategy was employed to place the inclusion of data elements into a "market" environment by requiring each data element to have at least one "sponsor." The sponsor(s) is to have responsibility for the definition, collection and quality control of the data element.

Cost considerations. Presumably the process eliminated unneeded data and saved money. It also spread some of the costs of data acquisition and management among the user agencies. The savings have not been documented.

Impacts. There was a significant reduction in the number of data elements maintained in the inventory.

5. Quality Assurance Program
   Agency. Wisconsin Department of Transportation (DOT), Traffic Accident Section(TAS)

Overview. The Traffic Accident Section of WisDOT developed a quality assurance plan which had two major components. The first was the establishment of a "quality-infrastructure." . This component focused on basic management and operational activities to provide an environment which would support efforts directed toward quality improvement. The second component was the identification of specific quality-assurance strategies, their prioritization and implementation. The result has been a multi-faceted approach, both within TAS and outside, for improving the quality of crash data available to users.

Cost considerations. The principle costs are personnel time. Most of the personnel time is required from the TAS. Data processing and personnel costs are explicitly provided for in the TAS budget. Costs for the quality improvement program are estimated at 10 percent of the overall costs of the division in which TAS operates. This translated to about $100,000 per year. However, the cross-training of personnel resulted in higher classifications for these employees because of their broader range of skills. The higher classifications translate into higher wage rates.

Impacts.

a. The office dealing with license suspensions has found that errors have been reduced to an insignificant level.
b. FARS files have been fully coordinated with the state file.
c. Cross training data-handling personnel has reduced the time required to respond to public inquiries, and improved efficiency of PCR handling.
d. The program momentum has created a sense that unless there are constant and continuing improvements being made to the system, there is something wrong.
e. The savings to the division within which TAS operates, just as a result of the quality improvement program in TAS alone, allowed the division to meet its entire "cutback requirements" that had been allocated to it as part of a budget reduction.

6. Outreach Program
   Agency. Wisconsin Department of Transportation (DOT), Traffic Accident Section(TAS)

Overview. Over the period of several years, the WisDOT, Traffic Accident Section has put in place several components of a program to reach out to those who provide and use traffic-crash data. Most of this activity has occurred since mid-1991. The purpose is to obtain the best quality data possible and facilitate its productive use. Those involved believe that the outreach effort has been effective in improving the motivation of officers to provide better-quality data when reporting crashes. Components currently in place include:

a. A newsletter;
b. On-site visits to local users and providers;
c. Provision of crash information to local communities;
d. Involvement of those outside the section in planning and decision making affecting the PCR, and
e. Officer recognition.

Cost considerations.
a. The per-issue cost for the newsletter is estimated to include about two-staff weeks plus the cost of printing.
b. The TAS data processing budget has been increased to account for additional requests from local communities. This removes another barrier to system use, because local agencies are no longer asked to pay for the computer work.

Impacts.
a. Officers express amazement that the crash data are being used for so many purposes other than insurance.
b. There has been a significant increase in requests from law-enforcement agencies for ad hoc reports.
c. Telephone inquiries have increased because direct personal contact with someone from TAS makes the inquirer less hesitant. Inquiries come from both officers who want to know how to complete a form correctly, as well as those in the agency who want information from the system.
d. Data-entry personnel interviewed indicated a noticeable improvement in report quality as the program was phased in. Evaluation has not been performed.

Direct Communications with Local Agencies

7. Communicating with Local Data Collectors
Agency. Ohio Division of Highway Safety (DHS)

Overview. An integral part of Ohio's highway-safety management is communication between the central agency and data collectors at the local level. This outreach is done formally through regularly scheduled training sessions, and informally through contact between data entry personnel and local agency personnel. Both sets of communications tend to reduce the traditional "distance" between local personnel including the data collector and the central, safety-data operation. Local personnel become part of the system.

Cost considerations.
a. Training requires travel and costs for conducting it.
b. Telephone costs can be high if many contacts are made between data entry personnel in Columbus and local agencies.
d. Time is taken from data entry jobs to make the contact. Each data entry person will key fewer reports.

Impacts. Reducing errors in collecting and recording crash data means that the data will be of higher quality. Information based on those data then should have a higher value. Given that communication increases the awareness of the need for quality, the impact of this outreach is positive.

8. Communication with Local Data Collectors
   Agency. Pennsylvania Department of Transportation (PennDOT)

Overview. Two staff members from the Division of Traffic Safety Data provide ongoing training for new officers and as a refresher for other officers throughout the state. Additionally, the trainers make special visits to agencies to help resolve issues that central staff have noted, e.g. miscoding locations or missing data. The division also produces a bi-monthly newsletter which is directed toward the reporting officer.

Cost considerations. Two staff members are needed. Additional costs are involved in preparing, printing, and mailing the newsletter. PennDOT does not keep data which would provide a basis for determining total costs. Such a basis would be useful, particularly coupled with a study of the effectiveness of the contact.

Impacts. On the presumption that trained reporting officers will produce more accurate and complete data, then the Pennsylvania traffic safety-database should contain higher quality statistics than are available from states not giving personal training and follow-up. Higher quality data would mean that studies of crashes using data from Pennsylvania might be better than those depending on data from other states. Regardless of potential use of data, the pursuit of training and the use of newsletters is linked closely to an over-riding concern on PennDOT management that they have high quality data.

Scannable Forms[1]

9. Scannable Police Crash Report Form
   Agency. Michigan State Police (MSP)

Overview. The MSP is responsible for collecting and processing police crash reports in the state. Because of down-sizing, they sought to eliminate staff interpretation of reports by using officer self-coding. The revised form and processing system incorporated OMR and was implemented in less than one year. Since introduction in 1992, numerous problems were encountered and re-solved, e.g., a need to reinstate the narrative and diagram. Reaction to the form and resulting data has been mixed, but significant improvements have overcome many of the earlier problems.

Cost considerations. Determining changes in costs resulting from the switch to a scannable form is difficult. The Office of Highway Safety Programs staffing was in transition, with reductions occurring throughout the organization. Comparing the cost of a process which maintains pace

[1] Although scannable forms using optical mark recognition (OMR) coding have been introduced in several states and appear to be serving as a practical means of reducing time required to process crash reports, they have not changed how data are collected and coded. As such, the use of such forms should be considered an interim step in enhancing the *collection* and *management* of traffic safety-data.

with the inflow of PCR's with one that did not is not appropriate. However, significant reductions occurred in staff required for data entry.

Impacts. Effects of the change in PCR format and processing were both primary, i.e. meeting the objectives, and concomitant resulting in unforeseen events. The changes resulted in:

a. The eight-month backlog of police crash reports initially was reduced to 26 days; however, further loss of staff has resulted in the backlog increasing to 7 weeks. Based upon a sample of 6200 forms analyzed in the early stages of implementation, about 15 percent required special handling. Additionally, a significant amount of information on the form still must be keyed into a file.
b. The Department of Transportation (MDOT) has not been able to use the data from the new forms because it did not alter its system to accept the new format.
c. An MDOT analysis of the data processed from the new form through November, 1992 indicated numerous problems with data quality, but the problems were not compared to those occurring with the previous form.
d. The revised processing at MSP has created a new quality problem. The records created from the scanning process and keying process are created independently and later are merged in a batch matching process on the central mainframe. The process results in a small portion of unmatched reports. Those that are not matched via automated means are not corrected later because of the press of other work.
e. The new PCR resulted in statewide uniformity of crash reports. It integrated the supplementary truck crash report, as well. The addition of codes on sequence of events and location relative to road offers an opportunity for improved analyses, once the officers are correctly reporting it.
f. The outreach program conducted by OHSP has seemed to be important for identifying mid-course corrections.
g. The change to the scannable crash report occurred within the context of personnel and organizational changes. These significant concomitant changes make it difficult to separate the role of the scannable form in the impacts noted.

10. Development of a Scannable Police Crash Report
Agency. Wisconsin Department of Transportation (DOT), Traffic Accident Section(TAS)

Overview. In March, 1992 WisDOT began developing a traffic crash report that could use OMR which would allow automated entry of many of the data elements. A task force was created to design the form and field test it. This was accompanied by a comprehensive implementation plan which had policy, training and legislative components. A WisDOT committee was also formed to coordinate the development of a database management system. The target implementation is January 1, 1994, with data being available from the system on-line two or three months later.

Cost considerations. The work is being conducted by a senior person working on special assignment, with the assistance of an engineering technician to do software work. Task-force time was also involved, but not measured. The field test will involve extra officer time, since the same crash will have to be reported on two different forms.

Impacts. The form had not been finally designed, or implemented, at the time of the site visit. Therefore, bottom-line impacts have not yet been experienced. However, at the time of review, the project team was not sure that WisDOT had conducted a thorough assessment of what Michigan had done. Such an assessment is critical if the costs of change along with problems encountered are to be kept to a minimum.

Automated Systems

11. Vision 2000 Plan for Automation
Agency. California Highway Patrol (CHP)

Overview. Based upon a strategic plan for a law enforcement information system developed by the State of California in 1990, CHP adopted a plan for automation, calling it "Vision 2000." The concept involved the use of mobile digital computers on which all forms required by the Patrol would be automated. Furthermore, the portable device would have capability to read magnetically striped driver licenses, and input global positioning system (GPS) data from a receiver connected to the device. Finally, the portable systems would use wireless communications and a statewide communications network to transmit the data collected.

The CHP has been aggressively pursuing the implementation of the concept, initially through a series of demonstration and pilot studies of portable computing devices. Lack of funding has constrained the rate of progress.

Cost considerations. CHP estimates that the complete effort will cost about 72 million dollars, plus require about 65 man-years over the next four years. Added would be allocations for non-personnel costs. Funding cutbacks in the state budget have slowed these efforts. Other portions of the system included:

a. A statewide automated citation system estimated to cost about 7.8 million dollars
b. A study of a statewide communications system at 500,000 dollars, with start-up estimated to be an additional 50 million dollars.
c. Approximately 3 million dollars needed to implement an automated vehicle locator system for CHP.

Impacts. The automated crash report is being used by about 1000 officers, and there is a backlog of demand for training to add users. However, the crash reporting system is currently unable to accept the electronic upload, so the form must be printed out in the current format of the crash report. Significant savings in officer completion time, compared to manual completion, have been reported but not documented.

Additionally, the state has been experimenting with automated preparation of traffic citations. The magnetically-striped drivers license along with input from the computer go into a hand-held device which generates the citation. The test resulted in a reduction in errors from 13 percent to 8 percent. This was estimated to save significant staff time (not documented). About 62 percent of the officers in the test area used the device by the end of the project. While automated transfer of the data saved some court time, the type of device and the trial nature of the software resulted in a slight increase in the time required for an officer to complete the citation, compared to handwritten ones. Problems also were experienced using low-band communications frequencies along with persons de-magnetizing the stripe.

12. Local Agency Data Entry - Crash Reports
Agency. Ohio Department of Highway Safety (ODHS)

Overview. Police agencies in Ohio can enter elements from the police crash report directly into the ODHS safety system. Development of a relational database for highway safety records helped make local data entry possible. Agencies have available all of the online editing features available at the central site. As communications links improve, more local entry will become available. ODHS management believes that decentralization has improved the quality of the data recorded.

Cost considerations.

a. Fewer central data-entry personnel are needed. Although, the workload falls on the local agency, because it is dispersed, most agencies can provide the entry as part of other tasks.
b. Data arrive corrected so less time and money are being spent by central staff making corrections. The number of errors discovered during batch editing appears to be fewer than with the other reports. Less time is spent in making corrections at the batch edit, but this value has not been documented.
c. Agencies have access to the data for reporting. ODHS does not have to incur costs associated with specially preparing and mailing reports to the agencies.
d. On the other hand, more computers are needed as well as increased costs are incurred for communications lines. However, as more databases are made available, e.g. roadway, emergency medical services, driver licensing, the costs of the computers and links will be spread over a wider user base.

Impacts. Errors in the records processed at the final editing have decreased. Local personnel have become active participants in coding highway crash data. They have become supportive of the needs of the ODHS, particularly in terms of quality. However, the department has not measured the differences between local data entry and central data entry; a study of the differences would be beneficial.

Another effect on quality is that when corrections are made, it is done closer to the source both in time and distance. Because data entry is occurring shortly after the PCR is taken, details remain fresh. Also, the person requesting corrections is not an individual in "Columbus," but generally someone the officer knows.

13. Linked Databases: Crash Records with Driver/Vehicle and Roadway Inventory Records
Agency. Pennsylvania Department of Transportation (PennDOT)

Overview. The PennDOT safety-database is linked to both driver and vehicle records, and the roadway inventory. This enables data management personnel, both from the input and user sides to access the databases for crash data. The linkages are built into the data-entry program. They provide facts about the driver based on driver license or name, and about the vehicle based on inspection registration number, owner name, registration plate, or VIN. The roadway inventory provides the sector and offset data for locating the crash, and makes available elements about the roadway, e.,g. alignment, roadway type, pavement type, etc. The roadway linkage also allows safety-data-entry personnel to locate crashes automatically. Linkages provide two-way handling of the data. Date, time, and severity of the crash, number of vehicles, and type automatically become part of the driver, vehicle, and roadway databases.

When producing reports, the safety-database manager can use elements from the driver file, e.g. citations or previous crashes. Driver licensing staff have crash information available directly. PennDOT engineers can obtain a range of roadway data for any specific crash. All of the linkages are transparent to the users.

Cost considerations. Once linkages among systems have been developed, the process of using the data from the other files occurs at a minimal direct cost. However, PennDOT has not taken advantage of using linked data to reduce costs of data collection. Police officers still record much of the data available elsewhere. The coders then enter the data and use linkage for validation only. The roadway inventory is the only file that provides data which are not available from the crash report to coders or analysts.

Impacts. By using linked data, PennDOT is able to help ensure high quality data, particularly as they relate to drivers, vehicles, and the roadway. Accurate data for the roadway is important to engineers and can help save time by providing summary information about the roadway and crashes automatically from the database without having to resort to individual crash reports which is done in other states. Additionally, researchers interested in the relationship between types of vehicles and crashes can obtain these data, generally with a greater degree of accuracy than had the state depended upon the police officer to record facts about the vehicle accurately at the scene. Finally, because the crash itself is recorded on the driver license and roadway inventory systems, users of these systems have timely information about crashes.

Non-Sworn Crash Investigators

14. Use of Non-sworn Crash Investigators: Community Service Officers
    Agency Fort Collins, CO

Overview. The Fort Collins Police Department uses their Community Service Officers (CSO) for traffic crash investigation. The CSO's work regular shifts and are assigned to all crashes occurring during their shift. They can cite drivers for violations of traffic laws but cannot make physical arrests. In addition to their handling of traffic crashes, they perform other traffic and police services including: vehicle investigation and taking reports of crimes where a suspect does not have to be interviewed. Their work on non-enforcement activity has freed officers in the department to handle those tasks which require the presence of a police officer.

Cost considerations. Pay for the Community Service Officer is set at 75 percent of the top pay for a patrol officer which currently is 39,000 dollars per year. The pay scale arose from a study done by the department in which they compared what the CSO was doing to similar positions in other departments. The lower pay also means that indirect costs such as FICA and pensions will be lower. Differential in pay, approximately 9000 dollars per officer, results in a savings of approximately 63,000 dollars annually as opposed to the salaries which would have to be paid for 7 police officers to replace the CSO's.

Other indirect savings exist. CSO's receive less training; they are available more hours for duty. Police now have more time for enforcement or investigation. Finally, the CSO becomes an expert in certain aspects of work such as traffic crash investigation. They can perform the task faster and are more likely to perform a quality job.

Impacts.

a. lower costs
b. persons devoted to traffic crash investigation who have the time to complete their work
c. expertise arising from repetition; higher quality investigations
d. freeing of police officers to concentrate on enforcement and patrol

15. Use of Non-Sworn Crash Investigators: Public Service Aids and Traffic Homicide Investigators
Agency. Ft Lauderdale (Florida) Police Department

Overview. The reporting of many traffic crashes by Ft. Lauderdale is performed by non-sworn Public Service Aids (PSA) rather than by sworn police officers, when available. Otherwise, motorcycle traffic enforcement officers will serve as the next available investigator. The PSA works daily, except between midnight and 6:00 am. An additional group of non-sworn homicide investigators handle all fatal or probable fatal crashes and will be called out as needed if such a crash occurs when none are on-duty. These non-sworn personnel go to the scene will take over the investigation at that point.

Cost considerations. The non-sworn investigators earn about two-thirds the salary of sworn officers with similar experience. These start at almost 15,000 dollars per year; sworn officers earn almost 23,000 dollars initially. Because the average non-sworn officer has much less experience than the average sworn officer, the average pay difference is almost double. Benefit and indirect costs also are lower.

Impacts.

a. The crash investigators are well integrated into the department
b. They are accepted by other employees and no "turf" problems have arisen.
c. Since the PSA's provide both a service at a reduced cost and serve as a potential cadre for sworn officers, their impact is positive.
d. They also handle more crashes than most patrol officers and can become quite proficient in completing crash reports.

16. Use of Non-Sworn Crash Investigators
Agency. Rapid City (South Dakota) Police Department

Overview. The Rapid City Police Department has employed non-sworn personnel as crash investigators for more than 15 years. They complete all investigations of crashes; however, they cannot write a citation or make an arrest. When not handling crashes, the investigators assist with a wide range of traffic duties including auxiliary patrol. Their ability to devote their time to traffic crashes has provided a well trained and dedicated set of investigators who, the department believes, provide quality investigative services.

Cost considerations. The crash investigator starts at 9.95 dollars per hour. After three years, they earn 11.70 dollars per hour. The department provides a uniform, vehicle, and a wide range of training. The investigators work approximately the same number of hours annually as sworn personnel and receive the same ongoing training. If the investigators earned as much as an officer, the department would disband the operation.

Two savings exist. First, the pay differential is approximately two dollars per hour or 4000 dollars annually. The four investigators represent more than a 16,000 dollar annual savings in salaries (additional savings arise from a lower amount of benefit costs).

The other savings is one of time. According to the divisional manager, the investigator is more proficient in handling a crash. Less time is spent at the scene and the data gathered more likely to be accurate and complete. This means less time is required for police who need to assist at the scene, as well as less supervisory time to review documents. Because of the expertise of the investigators, management also believes that team members spend less time in court and are more likely to obtain convictions. None of this of this second series of savings has been documented formally.

Impacts.

a. lower costs
b. persons devoted to traffic crash investigation who have the time to complete their work
c. expertise arising from repetition; higher quality investigations
d. freeing of police officers to concentrate on enforcement and patrol

Miscellaneous Practices

17. State-Level Interpretation of Crash Data Reporting
Agency. Pennsylvania Department of Transportation (PennDOT)

Overview. Most of the description of the crash and contributing circumstances in Pennsylvania derives from the crash diagram and narrative. Analysts at PennDOT interpret these two elements of the crash report. That interpretation leads to a coded description of events and manner of collision for all involved units, as well as contributing circumstances other than light, weather, and road surface at the time of the crash.

Cost considerations. No data are available regarding the time required by staff to interpret the narrative and diagram in order to develop contributing circumstances and events as opposed to coding the data submitted by police officers. However, PennDOT has assigned six persons solely to the analytical work. Even though these analysts also determine location where the crash occurred, more than 50 percent of their work involves analysis of events and related crash data. Without such coding, as many as three fewer persons might be required, particularly if the automated location coding increases from 30 percent to 80 percent.

Impacts. Given how the analytical work is performed, the coding of events, location relative to road, and manner of collision probably is more accurate and precise than if the officers in the field had performed the same work. The safety database is far more likely to contain multiple events than what otherwise might be expected. What can not be measured are two elements.

First, is the added value of the data to safety users when given multiple events instead of one event. Second is the difference in overall costs between having dedicated personnel perform the task and using the officer in the field for the same work. That officer may be gathering the data and coding the form as part of other work and the incremental time required, because it might be filling waiting time, may be minimal or none. Without a carefully conducted study of time, comparisons can not be drawn.

18. Study to Establish Proactive Problem Identification
Agency. Office of Traffic Safety, State of California

Overview. The staff at the Office of Traffic Safety (OTS) is responsible for developing, funding and administering the state's highway safety plan. They decided that the safety plan and individual projects should be based upon a more comprehensive picture of a community's safety problems than was currently the case, and that OTS would take on the task of helping localities identify and profile the problems. To do this they sought outside expertise, which they found within state government at the Department of Finance. The Department of Finance maintains a "consulting" group whose role it is to conduct special operational studies for state agencies. OTS contracted with the Department of Finance to develop a proactive problem identification that could replace current practice. The study team developed a microcomputer-based database which combines data from several agencies to allow users to develop profiles on traffic safety conditions in each community within the state.

Cost considerations. No costs were made available to the project team. However,savings, in terms of dollar loss, may accrue as [problems are identified and resolved thereby reducing the number of crashes.

Impacts. Because the report was just being completed at the time this summary was being prepared, it is too early to determine the actual impact. However, early reactions of OTS to the use of the database were reported as enthusiastic. The intended impact is a more optimal Highway Safety Plan and a reduction in crashes.

Roadway Inventory

Automation and Upgrade

19. California Highway Inventory and Performance System (CHIPS)
Agency. California Department of Transportation (CalTrans)

Overview. CalTrans had developed a plan for a corporate database in the mid 1980's. Slowly they have been building it. Most recently they received approval to proceed with the highway inventory and performance system. The project is under the control of a management committee consisting of the CalTrans division chiefs. In addition, a user group has been formed. The department has created a small staff to act as the core resource for the project, but it depends upon significant volunteer work from other staff. A target of December 1993 had been set for the presentation of a concept plan.

Cost considerations. The planning effort is being funded with Highway Planning and Research program funds, No costs were available.

Impacts. Because the system is not yet in existence, there are no impacts to report. The anticipated benefits will be improved decision making, due to improved accessibility to needed information.

20. Automated Updating of Roadway Inventory
Agency. Pennsylvania Department of Transportation (PennDOT)

Overview. The PennDOT uses special software and on-site verification to update its roadway management system (RMS). The system contains data covering more than 100,000 miles of state-numbered highways. Of this amount, approximately 41,000 miles are state maintained; an additional 67,000 are municipally maintained. PennDOT uses a segment and offset method for their system. The segments for each highway start at the county line and are numbered sequentially along the specific highway. They end at physical features and represent lengths of approximately 0.5 miles.[2] Because of the methods used to update the system, and the availability of all data during reconstruction (allowing changes to be incorporated at the time they are done), RMS management believes the system to be more than 95 percent accurate.

When a contract for reconstruction or maintenance is let, the contractor receives all data for the segments included. This eliminates the need for the contractor to record such data before starting work. Further, the system is updated based on the plans as let, then with the "as built" specifications.

Cost considerations. Total costs annually exceed one million dollars. Not included is an amortization of costs for the vans, and for computer operations. The latter include the micro computers and associated hardware for the field, computer time at central headquarters, and the costs of upgrading the central RMS system.

According to the manager of the system, a large proportion of the personnel costs goes to wages and per diem for college students who collect data on roadway condition during the summer. If PennDOT reduces the frequency of gathering roadway condition-data to once every third year instead of bi-annually, the annual cost for the RMS would decrease to approximately 750,000 dollars.

Impacts. The use of dedicated equipment, vans, microcomputers, and software, and scheduled reclassification of the highways has helped PennDOT achieve a 95 percent accuracy for their roadway management system. This accuracy is reflected in their ability to provide data and information for multiple needs. Most importantly, they can save time and money when letting contracts for reconstruction or maintenance. The contractors receive much of the data about the portion of the roadway under contract directly from the system.

---

[2] (Segments for municipally-maintained roadway data start at the municipal boundary. The segments represent points between every intersecting state-number highway within the municipality.)

21. Automated Roadway Inventory Data Collection and Straight-Line Diagram
Agency. Commonwealth of Virginia, Department of Transportation (DOT)

Overview. A 1983, consultant study identified the lack of a coordinated highway inventory and traffic-crash system as a high-priority deficiency to be corrected in DOT. In 1985, development began on a new Highway and Traffic Records Information System (HTRIS) which included provision for a computer-generated straight-line diagram for accessing and interacting with the roadway inventory. Late in development, DOT decided to develop a field data-collection procedure that would use a laptop computer in a moving vehicle with a two-person crew to acquire highway inventory data. The system is seen as an efficient means for collecting data on over 9,000 miles of primary state highways, plus Highway Performance Monitoring System (HPMS) samples and other secondary sections. It is integrated with the automated straight-line diagram which is also linked to the DOT crash-report imaging and location-coding systems.

Cost considerations. The detailed design and implementation of the HTRIS system was performed by a consultant working with DOT. The system development cost was about 4.5 million dollars. There was also significant DOT staff time involved. After the consultant was completed, DOT expended significant internal staff time to revise the system,.

Impacts.

a. System integration and automation greatly facilitates current DOT work.
b. Elimination of redundant storage of administrative inventory data.
c. Reduced duplication of storage of crash data Department of Motor Vehicles (DMV) and DOT. Previously, DOT had directly accessed the DMV crash database. However, the location-coding systems are now incompatible which requires separate storage of the data. Because DOT uses the initial crash file created by DMV to generate its crash file, there is no redundancy of data entry.
d. Traffic-crash analyses were temporarily disturbed when access to the crash-data system was lost, and before all the problems of the new system were resolved. The two systems were not operated in parallel during the transition.

Quality Assurance Programs

22. Quality Improvement Program - Roadway Inventory
Agency. Florida Department of Transportation (DOT)

Overview. A comprehensive transportation data-quality review was instituted by the Florida DOT. As part of this process, the roadway inventory system was evaluated and modified. Procedures were changed and quality control practices instituted which reduced the error rate from approximately one 50 percent to about 10 percent of the items in the file.

Cost considerations. Improvement in the accuracy and completeness of the data will result in less effort on the part of users in correcting erroneous files. The two significant additional continuing cost activities are the annual reviews and the annual meetings. The use of a six-person quality-control team visiting each district in the state for the roadway inventory is a good case in point.

Impacts. Error rates have been reduced from a point where about half the elements in the record contained errors to an error rate of about ten percent. Further savings may result as data collected are subject to justification, and as redundant data sources are merged.

23. Quality Assurance Programs - Crash Location Coding
Agency. Florida Department of Transportation (DOT)

Overview. About three years ago, a data-quality study was performed on the location referencing system, and a 25 percent error rate was discovered when comparing hard copy of crash reports with locations entered into the database. When this became known to the states's Secretary of Transportation, a database quality-review team was established to look at all source data and their processing. An analysis of the Roadway Characteristics Inventory (RCI) revealed that half of the data elements in the file were in error. Standards and procedures were established to improve quality, and a program was implemented.

A multi-level quality-control process has been established for the location-coding process. As a result, the latest estimate of error rate is less than two percent. The program is similar to that instituted for the roadway inventory system. A standard of 95 percent accuracy has been established along with timeliness standards. While the inaccuracies resulting from the coding process have been greatly reduced, the inadequacy of the data on the police crash report (PCR) continues to be a problem. Staff estimates that as many as 40 percent of the crash reports that DOT will record not codable due to lack of information and errors.

Cost considerations. Changes in costs resulting from the new procedures cannot be readily documented. The implementation costs also were not available. However, the systems in place now would seem to be more labor intensive since there is manual intervention in the process to identify and correct errors. Otherwise, the data management systems in place would not seem to represent any unusual expenditure, except that experienced for their development and initial implementation. The current expenditures for salaries and materials associated with maintenance of the location-referencing system, and the location coding of crash reports, is estimated by the DOT to be 207,500 dollars annually.

Impacts. The accuracy of the location-coding system has been increased to greater than 98 percent.

Issues Addressed by Noteworthy Practices

Relationship of Noteworthy to Issues

Agencies have introduced practices described above to address issues that have arisen with the operation of their highway-safety information-systems. A summary of issues affecting the collection and management of data appeared in Chapter 3 in tables 10 through 13 relating to collecting and reporting data for the three sources: highway, roadway inventory, and medical. The remaining tables 14 through 16 describe issues affecting management of those data. This final section relates the noteworthy practices to the issues. It also sets the stage for the strategies which are discussed and evaluated in the next chapter.

Even though the discussion of noteworthy practices was organized by source of the safety data, many of the practices are more universal. For example, state-wide user groups were formed to improve the quality of crash-data collection and management. However, such a group also could give its attention to roadway inventory and medical data.

The assessment of issues affected by the noteworthy practices also is limited to current applications. However, the practice could be extended to cover many of the other issues. For example, the use of state-wide user groups provides an excellent vehicle to expand coverage to many of the issues not presently considered. Whereas a users' group now is working with the implementation of local data entry, it could extend its reach to cover the use of automated resources as part of that process of local data entry.

Practices Affecting Handling of Traffic Crash Data

Eighteen of the noteworthy practices discussed affect issues related to the handling of traffic-crash data. Issues include collection, reporting, and management of those data. As shown in table 38, not all issues are covered. A majority of the efforts appear directed toward issues related to reporting the crash data, with lesser coverage of data management, and even less to gathering the data themselves. For ease of display, several of the practices have been grouped: state-level user groups, communication with local collectors, and non-sworn investigators. This grouping, however, may be somewhat deceiving. Combining them suggests that each of the individual practices affects more issues than might be the case. On the other hand, differences are not substantial.

Of all the practices, the use of non-sworn investigators to handle traffic crashes appears to have the widest potential effect on issues. Most importantly, these persons can act to improve the quality of gathering and reporting the data. The use of state-level user groups, although generally formed to work with the data-management component of the system, also has been able to address other issues, particularly those related to reporting crash data.

The quality-assurance programs appear more oriented toward managing crash data. However, the outreach program in Wisconsin does attempt to work directly with the reporting officer and address some issues related to collecting the data. The program of sponsorship is most appropriate in regard to its potential to resolve issues relating to data contained on multiple databases. The use of linked databases also can help resolve many of the same set of issues.

However as table 38 also indicates, a number of gaps appear. With the exception of using non-sworn investigators, none of the practices have addressed some key problems with data collection including the conflict in roles of the data collector, need to establish fault, inconsistent coverage, and incorrect use of resources. Also ignored are improvements in gathering data away from the scene along with handling and recording data from multiple sources and on multiple documents.

Table 38
Noteworthy Practices Related to Issues - Crash Data

			QUALITY ASSURANCE					AUTOMATION				MISCELLANEOUS	
		1 - 3. State-Level User Groups	4. Data Element Sponsorship	5. Quality Assurance Program	6. Outreach Program	7/8. Commun. w/local collectors	9/10. Scannable forms	11. Vision 2000	12. Local Entry	13. Linked Databases	14-16. Non-Sworn Investigators	17. State-level Interpret.	18. Proactive Problem Identification
DATA COLLECTION	A. COLLECTION												
	1. Conflict in roles										✓		
	2. Need to establish fault												
	3. Inconsistent and biased coverage										✓		
	4. Incorrect use of resources										✓	✓	
	5. Inaccuracy and precision in locating crashes				✓	✓		✓			✓		
	6. Perceived limited usefulness of data				✓						✓		
	7. Danger to data collector												
	8. Inadequate initial and refresher training	✓		✓							✓		
	9. Lack of tools for collecting some data										✓		
	10. Needs for gathering data away from the scene							✓			✓		
	11. Sources of data not readily accessible												
	B. REPORTING												
	1. Making judgements without adequate training	✓			✓	✓					✓		
	2. Severity of crash affects degree of quality in reporting				✓	✓					✓		

Table 38
Noteworthy Practices Related to Issues - Crash Data

		QUALITY ASSURANCE					AUTOMATION				MISCELLANEOUS	
	1 - 3. State-Level User Groups	4. Data Element Sponsorship	5. Quality Assurance Program	6. Outreach Program	7/8. Commun. w/local collectors	9/10. Scannable forms	11. Vision 2000	12. Local Entry	13. Linked Databases	14-16. Non-Sworn Investigators	17. State-level Interpret.	18. Proactive Problem Identification
3. Inadequate initial and refresher training	✓		✓	✓						✓		
4. Poorly defined data elements and codes	✓	✓	✓	✓	✓			✓		✓	✓	✓
5. Too many codes	✓	✓	✓	✓	✓			✓		✓	✓	✓
6. Lack of specific codes and elements	✓	✓	✓	✓	✓			✓		✓	✓	✓
7. Size and complexity of report	✓	✓	✓							✓	✓	
8. Multiple sources of data		✓										
9. Multiple recording of data												
10. Completing report away from the scene												
11. Inconsistency in reviewing reports			✓					✓				
12. Limitations in what is reviewed			✓									
C. ENTRY AND STORAGE												
1. Use of local databases	✓							✓				✓
2. Maintaining paper files												
3. Numbers of steps in handling reports	✓					✓						

Table 38
Noteworthy Practices Related to Issues - Crash Data

		QUALITY ASSURANCE				AUTOMATION				MISCELLANEOUS		
	1 - 3. State-Level User Groups	4. Data Element Sponsorship	5. Quality Assurance Program	6. Outreach Program	7/8. Commun. w/local collectors	9/10. Scannable forms	11. Vision 2000	12. Local Entry	13. Linked Databases	14-16. Non-Sworn Investigators	17. State-level Interpret.	18. Proactive Problem Identification
DATA MANAGEMENT												
4. Separation of duties	✓					✓						
5. Multiple databases	✓	✓							✓			✓
6. Lack of flexibility in changing databases	✓								✓			
7. Editing without access to original facts		✓										
8. Lack of automated assistance						✓	✓	✓	✓			
9. Long periods from time of crash to data entry						✓		✓				
10. Paper files primary source of narrative and diagram												
11. Confidentiality laws												
12. State and local levels not integrated												
13. Quality assurance programs not adequate	✓							✓				✓
14. Changes in forms	✓	✓	✓									
15. Lack of direct access by users	✓		✓					✓				✓
16. Lack of long-term database storage												

Finally, within the phase of data management several key issues remain. One revolves around the maintenance of paper files particularly for later access by users. The other important issue is that of confidentiality, precluding direct access to the data. Without resolution of the latter, data may not be available to the users who need them.

Practices Affecting Handling of Medical Data

None of the sites visited operate a comprehensive system to handle EMS and trauma data. Florida and Pennsylvania are in the process of developing systems which will provide data from a centralized source. Florida is employing OCR and OMR to their process of recording data. Further study of the Florida system should provide useful insight into how a state can address the need to collect and manage medical data related to crashes.

One practice is now being researched based on federal initiative. A number of states are evaluating the use of statistical methods for linking medical and related data covering injuries to traffic crash reports. The project titled CODES attempts to enhance the quality of injury data available for persons who were involved in traffic crashes. An interim report from NHTSA related to CODES was not available at the time of this report. These noteworthy practices along with the need to resolve the issues serve as the basis for the next chapter in which as more comprehensive set of strategies are described and evaluated.

Practices Affecting Handling of Roadway-Inventory Data

Five practices were identified as relating to the handling of roadway inventory data. Table 39 shows how each practice affects issues related to the inventories. With the exception of CHIPS in California and the PennDOT systems, other noteworthy practices do not give attention to one of the more important needs, the linking of roadway inventories to crash databases. Also, none of the programs in place have approached the lack of coverage of the roadway inventory as it relates to coverage of various classes of highways.

Finally, roadway inventories have limited data regarding longitudinal geometric. Some experimentation is underway to integrate pattern recognition from video and GPS to help identify and accurately measure both horizontal and vertical elements. Although no state currently is practicing this task, some individual efforts have been demonstrated.

Summary

In response to the nearly 40 issues identified affecting gathering, recording, and management of traffic-crash data, sites throughout the nation have developed practices to alleviate or resolve the problems. Although many of the practices were developed to address one specific problem, they could be used to address others. The previous discussions has indicated where a specific practice or groups of practices.

Table 39
Noteworthy Practices Affecting Issues Related to Roadway-Inventory Data

		NOTEWORTHY PRACTICES				
		AUTOMATION AND UPGRADE			QUALITY ASSURANCE	
DATA COLLECTION	A. COLLECTION	19. CHIPS	20. Automated Updating	21. Automated Data Collection Program	22. Quality Improvement Program	23. Quality Assurance - Location Coding
	1. Lack of automated assistance	✓	✓	✓		
	2. Longitudinal geometric difficult to measure					
	3. Roadway features not located accurately		✓	✓	✓	✓
	B. REPORTING					
	1. Limited feedback regarding quality of recording data					✓
	2. Construction and maintenance not recorded		✓			
DATA MANAGEMENT	C. ENTRY AND STORAGE					
	1. Not linked to state crash-databases					
	2. Does not reflect construction and maintenance	✓				
	3. Crashes not included with inventory		✓			
	4. Longitudinal geometric not available					
	5. Accuracy and precision affects use of crash data				✓	
	6. Lack of inventory for classes of roadway					
	7. Identification of roadway features imprecise		✓	✓		
	8. Long intervals between updates		✓	✓		

An examination of the issues and the 23 practices shows that many affect the quality of the data, particularly its accuracy, coverage, and timeliness. Proper training of the collectors, recorders, and data management personnel, review of their work, and feedback often are lacking. In many instances, these practices also could reduce the duplication of efforts in collecting and managing data. Where the same element are duplicated on multiple databases, opportunities for error increase. In addition to the potential for errors, duplication generally increases the costs of the process.

7. Evaluating Strategies for a Highway-Safety Information System

Candidate Strategies for Improving the System

As work progressed on the project, the needs for improving the collection and management of safety data became more apparent. Means for resolving most of the issues also became clearer. The literature and site visits conducted as part of the project provided examples of many noteworthy practices which have been adopted in various agencies across the U.S. and elsewhere. These were addressed briefly in Chapter 6.

This chapter identifies and evaluates strategies for improving highway-safety data-collection and management. They apply to the three sources of data studied: crash, medical, and roadway inventory. Most of the chapter is devoted to evaluating the strategies including examining relative cost-effectiveness. The goals, objectives, and measures previously described serve as a basis for assessing the candidate strategies.

The project team compiled an initial list of more than 40 strategies for improvement. Most of the strategies grew from considerations of how to improve the reporting and managing of traffic crash data. However, many of the same strategies also would help resolve issues related to roadway inventory and medical data. The strategies are organized according to the processes of collection and management with further subdivisions in each. Data collection strategies are further grouped into those which alter methods used for collecting or reporting data, and those which assist the collector. Management is divided into strategies which improve administration of the system, data entry, and how data are stored and made available. Table 40, at the end of this section, describing the strategies, lists the all the candidates. Their applicability to each of the three data sources also is indicated.

Data Collection - Change Methods (CC)

Currently data are collected using the police crash report (PCR) or a wide variety of emergency medical services forms. In most states, the PCR is completed at the scene by the responding police officer. EMS personnel complete "run" forms. Strategies described below are directed toward changing how data are collected and reported.

1. **Telephonic reporting.** Reports of minor crashes and defects in the roadways can be telephoned to a central location. It requires that the report be taken by personnel who are trained to ask questions correctly, and it requires some form of answering on a 24-hour basis. Roadway faults also could be reported by drivers and police, then forwarded to the appropriate highway engineering personnel.

2. **Use of sampling.** Use of sampling provides a mechanism by which a large amount of data can be obtained without requiring a substantial increase in time. Through sampling, a selected set of crashes may be investigated in more detail than others. The detail captured in some of the

samples may be even more than is done currently. Lack of certainty about the true universe of crashes occurring makes this strategy problematical.

[crashes] Police use two versions of the crash report to obtain data about all crashes. The shorter version obtains basic data about what, where, and who was involved. Crashes selected using a sampling method would require completion of a longer, more comprehensive document. Stratified sampling techniques also can be employed.

[medical] For the medical database, sampling can be used to provide greater coverage on selected injuries, including outcomes of injuries which normally are described only in terms of treatment required.

3. **Increase reporting thresholds.** Increasing the reporting threshold reduces the number of crashes which the police must report to the state safety agency. For example, police may only have to report those crashes resulting in an injury or for which a tow resulted because of damage sustained. To be effective, such increases in threshold also require dispatcher intervention to reduce the number of responses made by police.

4. **Non-sworn crash investigator/Data collection specialists.** The specialist is one specifically trained in gathering the required data. That task is his primary function.

   [crashes] Instead of using patrol or beat officers, trained non-sworn crash investigators would handle crashes. Such a position is being employed successfully in communities throughout the nation.

   [roadway inventory] For roadway inventory, technicians can be hired and trained specifically for the purpose.

5. **Use of non-automated/automated technologies** Emphasis, particularly for crash reporting will be on the use of non-automated technologies. The limitation has been employed because the Federal Highway Administration (FHWA) currently is pursuing the use of technologies through other research projects. However, for roadway inventory and medical data, integration of portable computers and relational databases show significant promise.

   [crashes] Investigators can record data about a crash on tape recorders, notebooks, portable computers, and video cameras. These media then can be integrated to prepare the crash report.

   [roadway inventory] Employ portable computers in specially designed vans with fifth-wheel equipment to verify roadway features and allow entry and correction of those features. Recording can be enhanced through use of global positioning systems (GPS) and video recognition software. Portable computers allow recording of data from construction plans during and after completion.

   [medical] Portable computers, particularly linked to trauma centers, can provide basis for comprehensive, centralized medical databases.

6. **Common, state-wide EMS and trauma reporting form.** To help ensure capture of key EMS and medical data, all EMS and trauma personnel need to use forms which allow recording of data which are needed for the highway safety information system.

7. **EMS collects all injury-related data.** All data related to type of injuries and their potential causes will be completed by the EMS personnel who are on the scene and treating the injured. These data will not be collected by the police but key elements from the EMS data will become part of the crash safety database. Linkages between EMS databases and the traffic crash data are required.

8. **Must-move legislation.** Requires the removal of vehicles from the scene as soon as possible. This legislation is designed to reduce congestion and the potential for secondary crashes. Crash data collection takes place off-road eliminating the potential for further harm to the reporting officers and involved persons.

9. **More extensive supervisor review.** Supervisors give all crash reports more intensive review. Blank spaces are not accepted; accuracy is verified. Errors, when detected, then are corrected.

   [crash] The strategy emphasizes local-level quality review. It needs to be coupled with efforts to reward effective investigation and reporting.

   [roadway inventory] The reviewer is responsible for ensuring that all attributes of quality are met.

10. **Completion by specially trained personnel.** Reports are completed at the local level by persons specially trained in crash analyses or injuries and their treatment. The data collector and reporter provides notes for a specialist who completes the crash or medical data report.

11. **Use of short-form and long-form reports.** Both the police crash report (PCR) and EMS reports come in a short and long form. Elements contained on the short form are collected for every crash reported by investigators or EMS personnel. The long form contains additional elements which describe the crash, drivers, vehicles, environment, and injuries in a more comprehensive manner. This longer form is completed only when the crash severity exceeds a specified threshold, e.g. an injury or one of the vehicles required a tow as a result of damage sustained.

12. **Sworn crash investigators.** Instead of using patrol or beat officers, specially-trained police crash investigators would handle most crashes. The police officer receives advanced training in crash investigation and carries all tools needed to perform his job. Where multiple crashes occur, beat patrol would be dispatched to and handle the less severe crash.

Data Collection - Assist the Collector (CA)

In addition to changing the methods used to collect and report data, tools can be provided which assist the collector in the performance of the work. The purpose of these strategies is to help reduce time required while encouraging higher quality collection and reporting of the data.

1. **Improved instruction manuals.** Prepare simple, pictorial manuals which show how to investigate a crash and report each element. The manual must contain a good index which makes it particularly useful for the officer who has just begun to report crashes or who handles a crashes infrequently. It is helpful for training and re-training. Properly prepared instruction manuals can be applied to reporting data for all three sources: crash, medical, and roadway inventory.

2. **Pre-printed diagrams.** Most crashes are similar in that they involve either a single vehicle striking a fixed object, or two vehicles in an angle, rear-end, or sideswipe collision. The officer selects the appropriate diagram and adds features and measurements. The diagrams can be part of the crash report or contained on a separate page.

3. **Simplified form design.** Simplifying the form can be accomplished by reducing the number of data elements collected as well as the number of codes within elements. In simplifying the report, fewer codes are available along with fewer elements to report. The form should be limited to one side, where possible.

   Simplified forms are not appropriate for either roadway inventory or the medical community. Roadway inventory data require a significant number of technical details which do not lend themselves to a simple report. The medical providers need comprehensive data to provide adequate treatment.

4. **Revised/improved design and format.** The crash report is completely revised and redesigned to help guide the collection and recording of data. The goal is a form which is easier to read and reduces the need for reference manuals. During revision, extensive field input, testing, and evaluation are performed. During redesign, care must be taken to ensure continuity of elements and their coding from one design to the next. Failure can result in discontinuities in the data and an inability to perform long-term studies.

5. **Improved definitions of elements/codes.** Errors often appear to stem from lack of clarity in descriptions. A task force reviews each element and code based on an analysis of how the codes have been reported and how many errors made, along with evaluation of planned coding by impartial third parties. Once the task force has completed its review, the PCR is revised to incorporate the findings.

6. **Priorities given to reporting.** Agency management increases the priority of investigating and reporting crashes. Investigators are sent to the scene of all crashes and provided time and tools to complete reporting of the crashes effectively. Reviewers have time to check for quality. They validate the data reported, particularly in terms of accuracy and consistency, as well as insuring completeness. On-going training for investigating crashes, and collecting and recording data are important tasks.

7. **Enhanced training.** At the state level, employ audit and training teams. Their responsibilities include first-time and recurrent training, and review of issues in crash reporting with individual agencies and officers. Larger agencies can develop the training for themselves. Training for smaller agencies occurs on a regional or state level. Ongoing training, where properly per-

formed for all three of the sources, crash, medical, and roadway inventory data, can provide appropriate support for improved quality.

8. **Collector incentives.** Quality reporting can be rewarded in a number of ways. A most common example for law enforcement is preference given to the assignment of overtime. Other rewards such as additional training, choice of shift or assignments, additional time off, and even material awards, can be valuable. Rewards can be forthcoming from both the agency and state level.

9. **Interact with counterparts.** Personnel who collect and report the data can be provided opportunities to interact with others who perform similar functions. These sessions, particularly when linked with training can provide significant opportunities to improve the collection and recording of data. Additionally, involved persons have opportunities to recommend changes or improvements to the process.

   [crashes] For crash reporting, the state can oversee local task forces which attempt to identify important issues related to investigating crashes and using crash data.

   [medical] Reporters of medical data have an opportunity to meet with counterparts who perform the same tasks. Specialized programs can be initiated at these sessions, particularly where such programs may require short-term collection of additional data.

10. **Conferences and training programs.** The state can provide funds for selected officers to attend conferences or special training sessions. Persons performing quality work also have opportunities to participate in specialized training for handling crashes. Conferences and specialized training also provide interaction among persons with similar interests. They also give the collectors a wide range of contacts and may involve users of the data who can provide feedback regarding the quality of reporting.

11. **Field participation in forms design.** Officers are selected to assist the state in ongoing redesign and updating of the PCR. They also can be responsible for testing as versions are developed. Whenever a state agency reviews the PCR, it needs to include field personnel in that review. Similar steps can be done with collectors of roadway inventory and medical data.

12. **Minimum wage personnel for data collection.** Roadway inventory is updated on regular basis through the use of persons on minimum wage or students to walk the roads and record specific data. This step makes the assumption that the features are located accurately along the roadway and that personnel are verifying those features including road surface and painting. The strategy can supply substantial data about roadway features and enhance the usefulness of the roadway inventory.

13. **Common coding nationally.** Use the standards that have been developed for crash reporting on a national basis (e.g., ANSI, NGA, and CADRE). States have made the attempt to employ these standards when updating their PCR's.

[crash] The greatest difficulty in standardization and increase in codes lies with those who collect crash data. Crash reporting is a task that must be performed, but has a low priority for completion. Adding data or complexity to a report does not encourage improved quality.

[roadway inventory] Although no national reporting standards exist for reporting data on roadway inventory, the FHWA Highway Performance Monitoring System can provide some guidance.

[medical] The abbreviated injury scale has been proposed as a means of achieving national consistency in reporting injury severity. However, the coding system is oriented more toward supporting appropriate treatment than for safety analysis.

14. **Computer directed interview (AI/expert systems).** Police use a computer program to help guide their interview of persons involved in the crash, as well as officers who have investigated it. This strategy employs expert systems which identify elements needed dependent upon the type of crash. It helps ensure completeness in collection of appropriate data.

    For EMS and medical personnel, the computer directed interview could enhance the comprehensiveness of data. The interview would be able to probe for responses, most of which the personnel would be trained in providing.

Data Management - Administrative (MA)

Data management refers to recording data to databases, then maintaining and making the data available. Guiding the process is data administration. Strategies are suggested which can reduce the effort required to manage the data while also improving the availability and quality of the data captured.

1. **Security for privacy.** Improved security must be implemented to prevent inappropriate access so that data can be made more available. Particularly important is the security of medical data to prevent inappropriate or inadvertent invasion of privacy, yet allowing access to data research.

   Security for privacy is often confused with ownership of the data. In many cases, security measures are applied because the owner does not believe that the data should be shared, or worse that "someone may misinterpret them." Additionally, laws designed to protect individual privacy often are misinterpreted preventing release of data in which the individual can not be identified.

2. **Quality-assurance programs.** Programs for quality assurance place emphasis on ensuring that data received are of high quality. Elements of this strategy include appropriate training and auditing of field work, editing of the data, and feedback regarding issues as well as high quality work. Each of these aspects provides tools to help people who work with the data enhance usefulness as well as quality. An overall program of quality assurance requires supporting staff to coordinate efforts and operate programs, and is given high priority.

3. **Management support.** For data administration, management support includes a priority for ongoing review of how data are collected and processed. It includes interaction with users to determine who are using data and how effective is that use. Management, through this interaction, can identify issues with data collection and storage that need to be resolved. Additionally, management needs to oversee regular review and updating of the database along with vehicles used to collect and process the data.

4. **System task-force.** A task force is responsible for overseeing operation of the system. In this role, it reviews the process of collecting highway safety-data, the tools used in collection, and the data themselves. It receives input from all persons associated with the system, including those collecting, recording, and coding the safety data, as well as the users. A system task-force is concerned with accuracy, timeliness, and simplicity of operations. It helps insure that data from various sources can be used efficiently and accurately.

   Where linked databases are coming into operation, a system task force becomes a necessity. The group must oversee the relationships of the bases so that the data may be shared properly. If a medical database is one of those linked, an added burden is to insure that those data are available while respecting privacy guidelines.

5. **Reorganization of management.** Managerial structure, particularly in regard to database design and development, has evolved out of the single-organization, traditional, vertical structure. This structure does not encourage fast response to issues, particularly where issues cross departmental or agency lines. More importantly, a single organization can not have in-depth knowledge of needs from others outside that group; it attempts to satisfy what it perceives to be needs of others. The management team must have authority to manage and maintain the system, not just one database. Along with the team is a data administrator who has a broader view and understands the role that data play in the overall picture.

Data Management - Data Entry (ME)

Data entry represents the largest recurring costs associated with managing the three sources. Strategies are directed toward reduction in cost and improvement in the efficiency in record quality data.

1. **Scannable forms.** Use of optical mark recognition (OMR) and optical character recognition (OCR) assist with capturing data from reports. The introduction of this strategy requires revision of the crash reports. OMR and OCR technology has applicability for all three sources of data; crash reports, EMS/trauma reporting, and roadway inventory.

   Scannable forms using OMR, however, may be an interim step. Automated methods and distributed database will reduce or eliminate the need for coded forms to collect data.

2. **Online and batch editing.** Online and batch editing are part of data entry. Online editing helps catch errors either in coding or in data entry while the form is immediately available to the data-entry staff. They also may be able to make immediate contact with the reporting agency to clarify errors. Batch editing allows the use of more sophisticated editing rules. This

can help reduce some of the inconsistency found within forms, e.g., weather and road conditions which do not match. Automated editing often can replace missing values.

On the other hand, online and batch editing can not correct errors of accuracy where an erroneous code value has been used and is consistent with other values. The codes still must be reviewed and corrected.

3. **Linked sources of data.** By linking sources of data which apply to crashes, including driver license, vehicle registration, roadway inventory, and medical data, fewer elements need to be gathered and coded by a collector of crash data. More importantly, the linked databases can provide more data about an incident than had a single individual attempted to collect them. Finally, the databases are maintained by those who have the largest stake in quality, e.g. driver license data are maintained by the licensing agency. This process helps improve the quality of all data.

4. **Local data entry.** Instead of forwarding all reports to a single source for data entry, the agencies which collect data also enter them. When local processing is done, crash reports are part of the safety information system shortly after they have been completed. This process requires communications links to the central system along with software and database design to allow interaction between the local and central systems. To be effective, however, the local agency also must have access to the data.

5. **Imaging systems.** The translation of visual images to digital format of different types facilitates the management, and access, of data.

   [crash and medical] Instead of microfilming or filing paper copies of the data reports, the report itself becomes a digital image. The image can be called and scanned from multiple sources. Although imaging systems make available copies of the reports for use, they do not capture individual data elements. Use of these systems for statistical analyses is limited. However, use of imaging can help reduce data entry because not all of the data need to be captured in a database, only those needed for statistical work.

   [roadway inventory] Introduce technical equipment which can be used to interpret photographs and video material particularly for capturing distances and dimensions. GPS can play a key role.

6. **Motivation and feedback.** The personnel performing data entry handle repetitive tasks, and often are faced with incomplete, or poorly written or prepared reports. Use of various forms of motivation, feedback, and training can help these persons perform their jobs with greater satisfaction.

7. **Cross-training personnel.** By cross-training personnel, each person on the safety-database staff has an opportunity to perform tasks other than the one to which he is assigned. This provides a break from repetition and available help to reduce backlogs in portions of the process.

8. **Sponsored elements.** Collectors and users of the data assume responsibility for various elements in the databases. This can include elements which may only be collected and used by a small subset of the users. Responsibility for sponsored elements includes:

   - developing the elements, their codes, and description of how to collect them,
   - auditing their use and maintaining quality control, and
   - ongoing review of its need for inclusion.

   Use of sponsored elements means that those who have the greatest need for quality in specific elements play an important role in maintaining that quality. Sponsored elements may extend across multiple databases.

9. **Comprehensive requirements for development and redesign.** Comprehensive design of a system includes the use of joint requirements planning (JRP) and joint application design (JAD), both of which are structured processes used during development. The process uses persons representing data gathers, recorders, and users. The result is design of a system which provides the requisite information with the fewest possible steps required to process the data. A significant emphasis is placed on quality.

10. **EMS/trauma database.** EMS and trauma databases are needed for analyzing the response to, and handling of, injuries resulting from a crash. Such a base requires the use of common reporting along with training of collectors and reporters in proper reporting similar to what is done for officers who handle the crash. The trauma database will maintain information about the nature and severity of injuries, length of hospitalization, and outcomes from treating those injuries.

11. **Use of data analysts to interpret descriptive material.** The collector and reporter record most of the data about a crash using narrative and a diagram. Personnel at a central location then interpret these items and complete the report. This is particularly useful for providing data about the manner of collision, events, and contributing causes. However, specially trained personnel also can receive EMS or other medical reports and interpret them according to the needs of the database. In either case, the personnel are able to interpret the data, often more accurately and precisely.

Data Management - Storage (MS)

Storage of data generally represents investment in computer hardware and software which is expensive, but can have a long-term payoff in terms of better availability of the data. Appropriate design of data storage can result in improved quality.

1. **Relational database systems.** For purposes of improved information management, particularly for local entry and open access to the data, the use of a relational database is important. In a relational database, fields of data which belong together are maintained as a unit and entered only once. Linkages between the groups of fields allow rapid access to any one element. However, relational databases require a database administrator to help control how the database is built and maintained. It also requires substantial investment in time and funds when the system currently does not have such database software. Use of a relational database is critical

for integrated use of micro-computers for data management at the local level, and with data storage, including distributed data handling, at a central level. Finally, relational databases can enhance the ability of users to generate information about crashes.

2. **Distributed databases.** Distributed databases provide unique sources of data which can be collected for purposes of providing information about specific subjects, e.g., drivers licenses. Each source is designed and maintained by the organization which has responsibility for the subject. Multiple locations can exist for the same class of data, e.g. crash data, each representing what was collected by the location. However, a centralized administrative responsibility helps ensure integrity of the many separate sources. Data are made available through communications links to other databases, such as a crash safety database. The system is expensive to build and is most productive when the entire system is has a single manager or management group with authority to make decisions. Such databases require appropriate security tools to ensure adequate control over availability of the data. They also are seen as a threat to "ownership" or loss of power over when and to whom data are released.

3. **Object-oriented database.** These databases are created from "objects" which are sets of related descriptions and processes. For example, an automobile is an object which has attributes and properties some of which are fixed, e.g. "is self-propelled," and some of which vary, e.g. color. The object itself will call for needed attributes. Objects are not fixed to a specific item of software; rather they are called and used for a specific purpose. Changes can be made to the objects with requiring changes to the database itself. Further, various types of objects can be combined. Text and images can be accessed together and treated as a single item.

   Object oriented programming, however, remains in its infancy. A number of companies are beginning to provide software capable of providing and using objects. Building applications from objects has yet to be commonplace.

4. **Quality control.** Quality control is used to help ensure that the data stored in the file are accurate, complete, and consistent. Design of standards should be done by a group of persons who represent the collectors, recorders, and users. The standards must be directed to high quality, but not so difficult as to prevent collecting and storing data in a cost-effective and timely manner. The application of quality-control standards also necessitates ongoing review which can be performed by a quality-assurance team.

   Effective quality control also involves several other strategies noted earlier including enhanced training, feedback to the collectors, and enhanced local review. The process requires a group of persons dedicated to it. Users and collectors should be part of the team.

The summary of strategies discussed above appear as table 40 on the following pages. The table indicates the source of the data to which the strategy applies.

Table 40
Strategies for Improving Collection and Management of Data

Strategy	Crash Reporting	Roadway Inventory	Medical Services
Data Collection - Change Methods			
CC1. Telephonic reporting	x	x	
CC2. Sampling techniques	x		x
CC3. Increase reporting thresholds	x		
CC4. Non-sworn carsh investigators/Data collection specialists	x	x	
CC5. Use of non-automated technology	x	x	x
CC6. Common, state-wide EMS/trauma reporting form			x
CC7. EMS collects all injury-related data	x		x
CC8. Must-move legislation	x		
CC9. More extensive supervisor review	x	x	
CC10. Collection by specially trained personnel (local level)	x		
CC11. Use of short-form and long-form reports	x		
CC12. Sworn crash investigators	x		
- Assist the Collector			
CA1. Improved instruction manuals	x	x	x
CA2. Pre-printed diagrams	x		
CA3. Simplified form design	x		
CA4. Revised/improved design and format	x		
CA5. Improved definitions of elements/codes	x	x	x
CA6. Priorities given to reporting	x		
CA7. Enhanced training	x	x	x
CA8. Collector incentives	x	x	x
CA9. Interact with counterparts	x		*
CA10. Conferences and training programs	x	x	x
CA11. Field participation in forms design	x	x	x
CA12. Minimum wage persons for data collection		x	

Table 40
Strategies for Improving Collection and Management of Data

Strategy	Crash Reporting	Roadway Inventory	Medical Services
CA13. Common coding nationally	x	x	x
CA14. Computer directed interview (AI/expert systems)	x		*
Data Management - Administrtaive			
MA1. Security for privacy	x	x	x
MA2. Quality-assurance programs	x	x	x
MA3. Management support	x	x	x
MA4. System task-force	x	*	*
MA5. Reorganization of management	x	x	x
- Data Entry			
ME1. Scannable forms	x		x
ME2. Online and batch editing	x	x	x
ME3. Linked sources of data	x	x	x
ME4. Local data entry	x	x	x
ME5. Imaging systems	x	x	x
ME6. Motivation and feedback	x	x	x
ME7. Cross-training personnel	x		
ME8. Sponsored elements	x	x	
ME9. Comprehensive requirements for development and redesign	x	x	
ME10. EMS/trauma database			x
ME11. Use of data analysts to interpret descriptive material	x		x
- Data Storage			
MS1. Relational database systems	x	x	x
MS2. Distributed databases	x	x	x
MS3. Object-oriented database	x	x	x
MS4. Centralized quality control	x	x	x

Evaluation of Strategies

Classification of Strategies for Evaluation

A total of 46 different candidate strategies were identified, applicable to one or more of the three data sources being studied. Before a final conclusion could be reached regarding strategies to recommend, they needed to be assessed. This process was done in two steps: classification as to immediacy of applicability and then evaluation of appropriate candidates. Three classifications were used:

1. **Proven:** This classification comprises strategies that the project team believed to be sufficiently **proven** that they did not require further evaluation, especially because they did not require major investments or otherwise create significant concomitant impacts in the agency. These strategies primarily derived from site visits and include some of the documented noteworthy practices. The extensive use of these strategies by agencies was generally considered enough demonstration of their cost-effectiveness. Feedback from agency experience was considered a better criteria for selection than the rating system used for many of the other candidate strategies.

2. **Further evaluation:** Classed are strategies which require **further evaluation** before they can be recommended to operating agencies. These strategies may have been used by several agencies and may have sufficient impact on agency operations, especially in terms of cost, that the project team believed further evaluation was required. These strategies were reviewed by members of the project team and rated according to their appropriate contributions to the goals and objectives discussed previously in Chapter 6.

3. **Further development:** In this classification are strategies which require **further development** before they adequately can be assessed. These include strategies which were not sufficiently defined or definable, and for which there was limited experience. The concepts will need further definition and refinement or more field experience before an adequate judgement can be made by the project team. Some of these could become subjects for future FHWA research efforts. Some are already subjects of research and demonstration programs.

Table 41 provides a summary listing of the strategies according to the categories just listed. The columns on the left side of the list indicate into which category the strategy was placed. Unless otherwise noted, however, those strategies listed as "evaluate further" are assessed against all data sources as listed on the right side.

Methodology Employed for Strategies Requiring Further Evaluation

Strategies selected for further evaluation were subjected to a rating by the project team. Ratings were then compared to potential changes in costs to establish a "cost-effectiveness

Table 41
Classifying Strategies for Evaluation

Proven	Evaluate Further	Develop Further	Strategy	Crash Reporting	Roadway Inventory	Medical Services
			DATA COLLECTION - CHANGE METHODS (CC)			
	x (rdwy)	x (crash)	CC1. Telephonic reporting	x	x	
	x		CC2. Sampling techniques	x		x
x			CC3. Increase reporting thresholds	x		
	x		CC4. Non-sworn crsh invetig./Data collection specialists	x	x	
	x		CC5. Use of non-automated technology	x	x	x
	x		CC6. Common, state-wide EMS/trauma reporting form			x
	x (EMS)	x (crash)	CC7. EMS collects all injury-related data	x		x
		x	CC8. Must-move legislation	x		
x			CC9. More extensive supervisor review	x	x	
	x		CC10. Completion by specially trained personnel (local level)	x		
x			CC11. Use of short-form and long-form reports	x		
	x		CC12. Sworn crash investigators	x		
			DATA COLLECTION - ASSIST COLLECTOR (CA)			
x			CA1. Improved instruction manuals	x	x	x
x			CA2. Pre-printed diagrams	x		
x			CA3. Simplified form design	x		
	x		CA4. Revised/improved design and format	x		
x			CA5. Improved definitions of elements/codes	x	x	x
		x	CA6. Priorities given to reporting	x		
	x		CA7. Enhanced training	x	x	x
x			CA8. Collector incentives	x	x	x
	x		CA9. Interact with counterparts	x		*
x			CA10. Conferences and training programs	x	x	x
x			CA11. Field participation in forms design	x	x	x
	x		CA12. Minimum wage persons for data collection		x	

Table 41
Classifying Strategies for Evaluation

Proven	Evaluate Further	Develop Further	Strategy	Crash Reporting	Roadway Inventory	Medical Services
	x (medical)	x	CA13. Common coding nationally	x	x	x
	x		CA14. Computer directed interview (AI/expert systems)	x		*
			DATA MANAGEMENT - ADMINISTRATION (MA)			
		x	MA1. Security for privacy	x	x	x
	x		MA2. Quality-assurance programs	x	x	x
x			MA3. Management support	x	x	x
	x		MA4. System task-force	x	*	*
		x	MA5. Reorganization of management	x	x	x
			DATA MANAGEMENT - DATA ENTRY (ME)			
	x		ME1. Scannable forms	x		x
	x		ME2. Online and batch editing	x	x	x
	x		ME3. Linked sources of data	x	x	x
	x		ME4. Local data entry	x	x	x
	x		ME5. Imaging systems	x	x	x
x			ME6. Motivation and feedback	x	x	x
x			ME7. Cross-training personnel	x		
	x		ME8. Sponsored elements	x	x	
	x		ME9. Comprehensive requirements for development and redesign	x	x	
	x		ME10. EMS/trauma database			x
	x		ME11. Use of data analysts to interpret descriptive material	x		x
			DATA MANAGEMENT - STORAGE (MS)			
	x		MS1. Relational database systems	x	x	x
	x		MS2. Distributed databases	x	x	x
		x	MS3. Object-oriented database	x	x	x
x	x(rdwy. &med.)		MS4. Centralized quality control	x	x	x

assessment." Because the effectiveness of a given strategy will vary dependent upon its application, the team developed a set of "typical" scenarios that require the use of traffic safety data. These scenarios are shown in table 42 and described in detail in Appendix B. Given the three components of highway safety, driver, vehicle, and roadway, and site-specific and systematic applications, a total of six general conditions were possible. Two other scenarios were added to account for budgeting and liability.

Table 42
Framework for Scenarios

	Highway-Safety Component		
Application	Driver	Vehicle	Roadway
Site-Specific	#1 Police Selective Traffic Law Enforcement	#2 Commercial Vehicle Inspection for Overweight and Defective Equipment	#3 High-Hazard Location Identification
Systematic	#4 Use of Crash Data to Evaluate a Driver Improvement Program	#5 Effectiveness of Anti-Lock Brakes	#6 Changes in Crashes Related to an Increase in the Speed Limit

Note: Two additional scenarios were added which are more general in nature
- Supporting a Budget Allocation
- Assisting with Defending a Tort Liability Action

Each strategy was then rated against the eight scenarios shown in table 42. A nominal rating system was employed. Project team members assessed how well each strategy would achieve the goals and objectives described in chapter 6 within the context of each scenario.

The discussion following describes how the ratings were formulated. Then the results of the ratings are summarized.

Assessment of the costs associated with each strategy also was performed using a nominal rating system. However, the cost model (developed in the early stages of this project and described in chapter 5) also was used to estimate the impacts of costs of implementing and operating strategies applicable to crash data collection and management.

Table 43
Values Used for Qualitative Ratings

-2	substantially detrimental to the goal or objective
-1	somewhat detrimental
0	has no or negligible effect
+1	somewhat contributes
+2	substantially contributes

A five-point rating scale was employed based upon values ranging from -2 to +2 shown in table 43. Even though the team treated the values as nominal, for computational purposes they were assumed to be continuous. Members of the project team examined how a specific strategy would affect the objectives for each scenario and rated the effect accordingly. These ratings were then averaged across all raters and across all scenarios for each objective to obtain an average value. Finally, an average value for each goal was computed based on the individual objectives contributing to the goal. Weighting the results with respect to each goal was considered, but the team found no clear basis to establish the weightings. The combined, averaged value of effectiveness provided one side of the cost-effectiveness picture. A highly-rated strategy would have a high positive-effectiveness rating.

Costs were divided into implementation and operation. Implementation requires both expenses and time. Both can range from very low to very high. Even if a strategy appears to provide a significant contribution to the highway-safety information-system, costs associated with its implementation may prove to be a significant barrier. Implementation costs of strategies were represented through a nomative scale. A strategy which has negligible costs-of-implementation and could be implemented almost immediately would be rated a "1" for both the cost and time dimensions. On the other hand, very high implementation costs, or very long implementation periods, were assigned a value of "5." Table 44 provides a summary of the values employed.

Table 44
Values Used to Assess Implementation Costs and Time

1	negligible cost or immediately implemented
2	low (up to $0.25 per report) or less than 6 weeks
3	moderate ($0.25 to $1 per report) or up to 6 months
4	high ($1 to $5 per report) or up to 1 year
5	very high (>$5 per crash report) or long (> 1

The degree of change in cost of operation that would result from implementing a given strategy was assessed based upon a qualitative rating, also ranging from -2 (substantial increase in costs) to +2 (substantial decrease). Table 45 displays the descriptions of the values used

Table 45
Changes in Costs of Operating the System

-2	substantially increases costs of operation
-1	somewhat increases costs
0	has no or negligible effect
+1	somewhat decreases costs
+2	substantially decreases costs

Assessment of the Strategies

Strategies for Improved Crash Report Data Collection and Management

This section presents and discusses the results of a qualitative cost-effectiveness analysis conducted of the candidate strategies for improving the collection and management of safety data. Only those proven strategies or those selected for further evaluation are included. In the sections below, those selected for further evaluation are discussed first after which they are combined with the proven strategies to arrive at a comprehensive list.

Results of ratings. Ratings from all members of the team were averaged across the team then combined to form an overall rating. A review of the ratings form "effect on the organization" showed a wide variation among the raters stemming from differences in how they interpreted the goal. As a result, this goal was dropped leaving the three remaining goals. Ratings for each goal and the summary across the goals appears as table 46.[5]

To achieve a better differential among the rated strategies, the overall rating was a summation rather than average; Therefore, the maximum range would have lain between -6.0 and +6.0. Only one strategy, linked data sources with a rating of +1.8, achieved an overall ratings approaching ±2.0. This seems to reflect both the conservative nature of the raters and the reality that any one of the strategies will not be the "silver bullet" that resolves all issues. Use of scannable forms, on the other hand, was the only strategy with an overall negative rating.

Results of cost analysis. Cost ratings were established for all three data sources being studied. The ratings for crash data appear in table 46. However, because of the availability of a cost model for estimating collection and management costs for crash reports, the model was preferable to using the qualitative ratings. The subjective ratings, however, are useful for comparisons among strategies or for a specific strategy across multiple data sources.

In applying the model for estimating cost impacts of crash reporting strategies, the team made assumptions about the effect on the strategies on each of the variables. In most cases, strategies affected only a few of the variables in the model. The use of assumptions also allowed the team to apply the model to the "proven" strategies along with those requiring further evaluation. The resulting cost impacts, in terms of total and unit cost, are summarized in table 47. Changes in unit-costs resulting from increasing costs per unit by approximately one dollar to a savings approaching 2.50 dollars. The largest increase would result from using specialists at the local level to complete a significant portion of the crash reports. On the other hand, substantial savings would accrue from reducing the amount of data collected through the use of long and short forms. Generally, those strategies scoring highest in meeting the objectives also tended to be the most costly.

Benefit-cost results. In order to develop a benefit-cost structure, some assumptions were made regarding both the effectiveness and cost ratings. Any strategy which had an effectiveness rating lower than 0.0 was consider to have a "low" effectiveness. At the other end, high effectiveness arbitrarily was set at a minimum rating of 1.0.

Strategies rated as "low" for costs were those which increased costs per unit. A "high" rating was given to savings of 1.00 dollar or more. By combining both sets of classifications, each

---

[5] Later in the analysis, the impact of strategies on organizational effectiveness were incorporated when examining the overall perspective.

Table 46
Traffic Crash Data
Summary Scores for Goals Across all Scenaros
and Across All Strategies (those selected for further evaluation)

STRATEGY	GOALS						COST ANALYSIS		
							Implementation*		Change in Cost of Operation**
	QUALITY	SUITABILITY	SAFETY	SUM OF SCORES	EFFECT ON ORGANIZ.	SUM OF SCORES	Costs	Time Required	
cc2 Sampling	-0.2	-1.0	1.5	**0.3**	[redacted]	[redacted]	2.5	1.8	-1.0
cc4 Non-sworn crash investigators	0.8	0.2	0.2	**1.2**	[redacted]	[redacted]	1.8	3.6	-0.8
cc5 Use of non-automated technologies	0.4	0.1	0.5	**1.0**	[redacted]	[redacted]	2.4	3.4	-0.3
cc10 Review by specialists	0.5	0.0	0.0	**0.5**	[redacted]	[redacted]	2.6	3.4	0.9
cc12 Sworn crash investigators	0.8	0.2	0.0	**1.0**	[redacted]	[redacted]	2.4	3.3	0.0
ca4 Revised PCR design	0.3	0.2	0.2	**0.7**	[redacted]	[redacted]	2.9	3.3	0.4
ca7 Enhanced training	0.5	0.1	0.1	**0.7**	[redacted]	[redacted]	2.0	3.2	0.3
ca9 Interact with counterparts	0.4	0.1	0.0	**0.5**	[redacted]	[redacted]	1.6	2.2	0.1
ca14 Computer directed interviews	0.6	0.1	0.0	**0.7**	[redacted]	[redacted]	2.4	4.1	0.7
ma2 Quality-assurance programs	0.6	0.1	0.0	**0.7**	[redacted]	[redacted]	2.3	3.1	0.8
ma4 System task force	0.5	0.3	0.0	**0.8**	[redacted]	[redacted]	2.4	3.0	0.0
me1 Scannable forms	-0.1	0.0	-0.2	**-0.3**	[redacted]	[redacted]	2.8	4.5	-1.1
me2 Online/batch editing	0.3	0.1	0.0	**0.4**	[redacted]	[redacted]	3.0	3.8	0.5
me3 Linked data sources	0.8	0.4	0.6	**1.8**	[redacted]	[redacted]	3.7	4.5	-0.1
me4 Local data entry	0.4	0.7	0.0	**1.1**	[redacted]	[redacted]	3.2	4.2	0.2
me8 Sponsored elements	0.1	0.0	-0.1	**0.0**	[redacted]	[redacted]	2.6	3.4	-0.1
me9 Comprehensive system requirements	0.3	0.3	0.1	**0.7**	[redacted]	[redacted]	3.3	4.2	-0.2
me11 Use of analysts for interpretation	0.3	-0.1	0.2	**0.4**	[redacted]	[redacted]	2.9	3.5	1.0
me5 Imaging & OCR/OMR	0.3	0.9	0.3	**1.5**	[redacted]	[redacted]	4.5	5.0	-1.5
ms1 Relational database	0.2	0.4	0.0	**0.6**	[redacted]	[redacted]	3.8	4.2	-0.2
ms2 Distributed databases	0.4	0.5	0.2	**1.1**	[redacted]	[redacted]	4.0	4.8	-0.6

Nominal scoring scale:
- -2 substantially detracts from the goal or objective
- -1 somewhat detracts from the goal or objective
- 0 has minimal or no effect on the goal or objective
- +1 somewhat contributes to the goal or objective
- +2 substantially contributes to the goal or objective

* range: 1 lowest to 5 highest
** range: -2 to +2

Table 47
Collection and Management of Traffic Crash Data
Analysis of Cost Impacts of Candidate Strategies

Strategy	VALUES					CHANGE (%)					CHANGE (Value)				
	Safety Data					Safety Data					Safety Data				
	Collection Cost	Management Cost	Total	Frequency of PCR's (000000's)	Unit Cost	Collection Cost	Mngmt Cost	Total	Frequency of PCR's	Unit Cost	Collection Cost	Management Cost	Total	Frequency of PCR's (000000's)	**Unit Cost**
Base	$72.9	$49.1	$122.0	6.361	$19.20										
CC1															
CC2	$62.9	$40.8	$103.7	5.282	$19.60	-14%	-17%	-15%	-17%	2%	($8.6)	($6.9)	($15.5)	(1.078)	**$0.40**
CC3	$54.9	$34.1	$89.0	4.420	$20.10	-25%	-31%	-27%	-31%	5%	($13.5)	($10.4)	($23.9)	(1.078)	**$0.90**
CC4	$67.2	$49.1	$116.3	6.361	$18.30	-8%	0%	-5%	0%	-5%	($5.2)	$0.0	($5.2)	(1.078)	**($0.90)**
CC5	$69.4	$49.1	$118.4	6.361	$18.60	-5%	0%	-3%	0%	-3%	($3.3)	$0.0	($3.3)	(1.078)	**($0.60)**
CC6															
CC7															
CC8															
CC9	$74.3	$49.1	$123.4	6.361	$19.40	2%	0%	1%	0%	1%	$1.5	$0.0	$1.5	(1.078)	**$0.20**
CC10	$66.2	$62.6	$128.8	6.361	$20.20	-9%	28%	6%	0%	5%	($6.1)	$17.3	$11.2	(1.078)	**$1.10**
CC11	$62.9	$45.7	$108.6	6.361	$17.10	-14%	-7%	-11%	0%	-11%	($8.6)	($3.2)	($11.8)	(1.078)	**($1.90)**
CC12	$72.9	$49.1	$122.0	6.361	$19.20	0%	0%	0%	0%	0%	$0.0	$0.0	$0.0	(1.078)	**$0.00**
CA1	$72.9	$49.1	$122.0	6.361	$19.20	0%	0%	0%	0%	0%	$0.0	$0.0	$0.0	(1.078)	**$0.00**
CA2	$65.5	$49.1	$114.5	6.361	$18.00	-10%	0%	-6%	0%	-6%	($6.7)	$0.0	($6.7)	(1.078)	**($1.10)**
CA3	$65.6	$49.1	$114.7	6.361	$18.00	-10%	0%	-6%	0%	-6%	($6.6)	$0.0	($6.6)	(1.078)	**($1.10)**
CA4	$65.6	$49.1	$114.7	6.361	$18.00	-10%	0%	-6%	0%	-6%	($6.6)	$0.0	($6.6)	(1.078)	**($1.10)**
CA5	$72.9	$49.1	$122.0	6.361	$19.20	0%	0%	0%	0%	0%	$0.0	$0.0	$0.0	(1.078)	**$0.00**
CA6															
CA7	$72.9	$51.8	$124.7	6.361	$19.60	0%	6%	2%	0%	2%	$0.0	$2.9	$2.9	(1.078)	**$0.40**
CA8	$72.9	$49.9	$122.8	6.361	$19.30	0%	2%	1%	0%	1%	$0.0	$0.8	$0.8	(1.078)	**$0.10**
CA9	$72.9	$50.8	$123.7	6.361	$19.50	0%	4%	1%	0%	2%	$0.0	$1.8	$1.8	(1.078)	**$0.30**
CA10	$72.9	$52.6	$125.5	6.361	$19.70	0%	7%	3%	0%	3%	$0.0	$3.8	$3.8	(1.078)	**$0.50**
CA11	$72.9	$49.1	$122.0	6.361	$19.20	0%	0%	0%	0%	0%	$0.0	$0.0	$0.0	(1.078)	**$0.00**
CA12															
CA13															
CA14	$72.9	$42.4	$115.2	6.361	$18.10	0%	-14%	-6%	0%	-6%	$0.0	($5.8)	($5.8)	(1.078)	**($1.00)**
MA1															
MA2	$76.2	$51.8	$128.0	6.361	$20.10	5%	6%	5%	0%	5%	$3.5	$2.9	$6.4	(1.078)	**$0.90**
MA3	$72.9	$49.1	$122.0	6.361	$19.20	0%	0%	0%	0%	0%	$0.0	$0.0	$0.0	(1.078)	**$0.00**
MA4	$72.9	$50.4	$123.3	6.361	$19.40	0%	3%	1%	0%	1%	$0.0	$1.4	$1.4	(1.078)	**$0.20**
MA5															
ME1	$72.9	$44.4	$117.2	6.361	$18.40	0%	-10%	-4%	0%	-4%	$0.0	($4.3)	($4.3)	(1.078)	**($0.80)**
ME2	$72.9	$51.8	$124.7	6.361	$19.60	0%	6%	2%	0%	2%	$0.0	$2.9	$2.9	(1.078)	**$0.40**
ME3	$65.6	$49.1	$114.7	6.361	$18.00	-10%	0%	-6%	0%	-6%	($6.6)	$0.0	($6.5)	(1.078)	**($1.10)**
ME4	$72.9	$49.1	$122.0	6.361	$19.20	0%	0%	0%	0%	0%	$0.0	$0.0	$0.0	(1.078)	**$0.00**
ME5	$72.9	$42.4	$115.2	6.361	$18.10	0%	-14%	-6%	0%	-6%	$0.0	($5.8)	($5.8)	(1.078)	**($1.00)**
ME6	$72.9	$49.6	$122.5	6.361	$19.30	0%	1%	0%	0%	1%	$0.0	$0.5	$0.5	(1.078)	**$0.10**
ME7	$72.9	$49.6	$122.5	6.361	$19.30	0%	1%	0%	0%	1%	$0.0	$0.5	$0.5	(1.078)	**$0.10**
ME8	$72.9	$49.6	$122.5	6.361	$19.30	0%	1%	0%	0%	1%	$0.0	$0.5	$0.5	(1.078)	**$0.10**
ME9	$72.9	$52.5	$125.4	6.361	$19.70	0%	7%	3%	0%	3%	$0.0	$3.6	$3.6	(1.078)	**$0.50**
ME10															
ME11	$66.2	$59.9	$126.1	6.361	$19.80	-9%	22%	3%	0%	3%	($6.1)	$13.2	$7.1	(1.078)	**$0.60**
MS1	$72.9	$51.8	$124.7	6.361	$19.60	0%	6%	2%	0%	2%	- $0.0	$2.9	$2.9	(1.078)	**$0.40**
MS2	$66.2	$50.5	$116.7	6.361	$18.30	-9%	3%	-4%	0%	-5%	($6.1)	$1.4	($4.6)	(1.078)	**($0.90)**
MS3															
MS4	$76.5	$51.8	$128.3	6.361	$20.20	5%	6%	5%	0%	5%	$3.8	$2.9	$6.7	(1.078)	**$1.10**

1 Based upon estimated national ferquency of crashes

2 I strategies CC2 and C3 were unadjusted for change in frequency, unit costs would $16.30 and $14.00 respectively

strategy was placed into a category of "relatively low," "moderate," or "high." None of the benefit-cost ratings for crash reporting fell into the "relatively low" category.

One issue that had to be resolved was that effectiveness ratings were obtained formally for those strategies which were classified as requiring further evaluation. The team believed that sufficient information was available from interviews and reports by those employing "proven" strategies, that more formal evaluation was not an effective use of time. To establish benefit-cost ratings for "proven" strategies required a different approach. The team reviewed material obtained during visits and other contact with agencies who already had implemented the strategies. This review of experience established a basis for dividing the effectiveness into low, moderate, and high. The effectiveness ratings obtained in this manner then were joined to the cost ratings established earlier allowing the proven strategies also to be placed into cost-effectiveness categories.

Table 48 displays the results of the cost-effectiveness assessment for all strategies that can be employed to improve the collection and management of crash data. Of the 35 strategies identified, 13 were considered to have relatively high cost-effectiveness. Of the 13 only three came from the "proven" strategies. The remainder had been selected for further evaluation. A possible reason for proven strategies falling into the "moderately cost-effective" classification is that these actions generally affect only limited aspects of the task of collecting and managing crash data. Further, the nature of the strategies suggests that each will make some difference with a relatively low cost of implementation. An example would be cross training personnel to help improve data entry. Obviously, use of cross trained personnel provide additional persons to help reduce backlog. It also can help reduce costs by eliminating the need for additional personnel. On the other hand, such cross training probably has a limited effect on quality, and it may result in paying higher wages because of the increased skills. For the most part, the collectors and managers of crash data appear unwilling to incorporate changes which may appear complex, costly to implement (without a known substantial savings), or which may not appear to have an immediate benefit.

Strategies for Improved Roadway Inventory Data Collection and Management

Results of ratings. Those strategies for roadway inventory as a data source requiring further evaluation were rated in the same manner as those for crash reporting. The resulting ratings are summarized in table 49. The unweighted ratings (as discussed in the previous section) for each goal are shown along with a sum. Also, the scores for "effect on organization" were dropped from initial use because, as was the case with crash data, team members varied in the manner in which they rated the goal. As was the case for crash data as a source, linking data sources showed the greatest possible contribution to the goals and objectives.

Ratings of proven strategies were developed in the same manner as those had been for crash data. The team reviewed the information gathered as part of visits and contacts with

Table 48
Relative Cost-Effectiveness of Strategies for Improved Collection and Management of Crash Report Data

Relatively Higher Cost-Effectiveness		Relatively Moderate Cost-Effectiveness	
CC4 -	*Non-sworn investigators*	*CC2* -	*Sampling*
CC5 -	*Use of non-automated technology*	cc3* -	Increase reporting thresholds
CC12 -	Sworn crash investigators	cc9 -	More extensive supervisor review
ca2 -	Pre-printed diagrams	*CC10* -	*Completion by specially trained personnel (local level)*
ca3 -	*Simplified form design*	cc11 -	Short and long forms
CA4 -	*Revised PCR design*	ca1 -	Improved instruction manuals
CA14 -	*Computer directed interview*	ca5 -	Improved definitions
MA4 -	*System task force*	*CA7* -	*Enhanced training*
ME3 -	*Linked data sources*	ca8 -	Collector incentives
ME4 -	*Local data-entry*	*CA9* -	*Interact with counterparts*
ME5 -	*Imaging systems & OCR/OMR*	ca10 -	Conferences and training programs
MS2 -	*Distributed databases*	ca11 -	Field participation in form design
me6 -	Motivation and feedback	MA2 -	Quality-assurance programs
		ma3 -	Management support
		ME1 -	*Scannable forms*
		ME2 -	*On-line & batch editing*
		me7 -	Cross training
		ME8 -	*Sponsored elements*
		ME9 -	*Comprehensive requirements for development*
		ME11 -	*Use data analysts to interpret*
		MS1 -	*Relational database*
		ms4 -	Centralized quality control

* Lower-case codes apply to proven strategies. Upper-case codes and italicized strategies represent those evaluated further.

personnel responsible for collecting and managing roadway inventory data. From these reviews, the team was able to establish low, moderate, and high effectiveness for these proven elements.

Because no cost models were designed for roadway, the potential changes shown in the cost analysis portion of table 49 were used. For each scenario, the team estimated the relative level of costs for implementing the strategy, the time required, and the change in operating costs which could be expected to occur. The costs and time for implementation were rated on a five-point scale from 1 which meant practically no cost or immediate implementation, to 5 for very high cost or several-year implementation. Generally, those strategies which required use of technologies such as linked databases resulted in the highest costs and time for implementation. On the other hand, the team generally believed that such strategies would have a significant effect on reducing operating costs.

Table 49
Roadway Inventory Data
Summary Scores for Strategies and Goals
Across All Scenarios

	GOALS						COST ANALYSIS		
							Implementation*		Change in Cost of Operation
STRATEGY	QUALITY	SUITABILITY	SAFETY	SUM OF SUM OF SCORES	EFFECT ON ORGANIZ.	SUM OF SCORES	Costs	Time Required	
cc1 Telephonic reporting	0.4	0.2	0.0	0.0	[redacted]	[redacted]	1.0	3.0	0.3
cc4 Data collection specialists	0.6	0.0	-0.2	0.0	[redacted]	[redacted]	1.3	3.0	0.3
cc5 Use of Technologies	0.8	0.5	1.1	0.0	[redacted]	[redacted]	2.0	3.7	-0.3
ca7 Enhanced training	0.6	0.0	0.2	0.0	[redacted]	[redacted]	0.7	3.0	0.1
ca12 Minimum wage persons	0.0	0.2	-0.6	0.0	[redacted]	[redacted]	1.0	2.7	-1.3
ma2 Quality-assurance programs	0.6	0.0	-0.3	0.0	[redacted]	[redacted]	0.7	3.0	0.9
me2 Online/batch editing	0.5	0.1	-0.3	0.0	[redacted]	[redacted]	1.3	3.3	0.9
me3 Linked data sources	0.7	0.6	0.0	0.0	[redacted]	[redacted]	2.0	4.0	-0.6
me4 Local data entry	0.3	0.3	0.0	0.0	[redacted]	[redacted]	2.3	4.7	-0.4
me11 Imaging & OCR/OMR	0.3	0.5	0.0	0.0	[redacted]	[redacted]	3.0	4.5	-1.0
me8 Sponsored elements	0.4	0.0	-0.3	0.0	[redacted]	[redacted]	0.7	3.3	0.0
me9 Comprehensive system reqmnts.	0.6	0.3	0.0	0.0	[redacted]	[redacted]	1.7	4.0	-0.4
ms1 Relational database	0.2	0.4	0.0	0.0	[redacted]	[redacted]	1.7	4.3	-0.6
ms2 Distributed databases	0.3	0.4	0.0	0.0	[redacted]	[redacted]	2.3	4.3	-0.6
ms4 Central-level quality control	0.5	0.3	-0.3	0.0	[redacted]	[redacted]	1.3	3.7	0.8

* range - 1, lowest to 5, highest

Nominal scoring scale:
- -2 substantially detracts from the goal or objective
- -1 somewhat detracts from the goal or objective
- 0 has minimal or no effect on the goal or objective
- +1 somewhat contributes to the goal or objective
- +2 substantially contributes to the goal or objective

Benefit-cost results. Table 50 displays the union of costs and effectiveness. The combination was done in the same manner as that used for strategies applied to collecting and managing crash data (discussed in the previous part). Because of a greater range in costs of implementation and operation, and because of lower ratings of effectiveness, some strategies fell into the category of "relatively low cost-effectiveness." No proven strategies fell into the category of relatively higher cost-effectiveness.

Table 50
Relative Cost-Effectiveness of Strategies for Improved Collection and Management of Roadway Inventory Data

Relatively Higher Cost-Effectiveness		Relatively Moderate Cost-Effectiveness		Relatively Lower Cost-Effectiveness	
				CC1 -	*Telephone reporting*
CC5 -	*Use of Technologies*	*CC4 -*	*Data collection specialists*		
		cc9 -	More extensive supervisor review	ca8 -	Collector incentives
ME3 -	*Linked data sources*			ca10 -	Conferences and training programs
ME9 -	*Comprehensive system requirements*	ca1 -	Improved instruction manuals		
		ca5 -	Improved definition of codes	*CA12 -*	*Use of minimum- wage field-personnel*
		CA7 -	*Enhanced training*		
		ca11 -	Field participation in forms design		
				MA2 -	*Quality assurance programs*
		ma3 -	Management support		
				ME2 -	*Online/batch editing*
		ME4 -	Local data entry	me6 -	Motiviation and feedback
		ME8 -	Sponsored elements		
		ME11	*Imaging & OCR/OMR*	*MS4 -*	*Centralized quality control*
		MS1 -	*Relational databases*		
		MS2 -	*Distributed databases*		

* Lower-case codes apply to proven strategies. Upper case codes are those strategies evaluated further.

Strategies for Improved Emergency Medical Data Collection and Management

Results of ratings. The candidate strategies for emergency medical data collection and management were rated in the same manner as those for crash reporting. The resulting ratings are summarized in table 51. The ratings for each goal are shown along with a sum. Ratings of effectiveness were fairly conservative here, as well, and "effect on organizations" was not included for reasons previously noted. In the case of medical data as a source, sampling was considered to have the greatest potential contribution, particularly because of the ability of collectors to provide very detailed data for a select number of crashes.

Even more difficult than for roadway inventory would have been determining the costs for collecting and managing medical data. Not only do two separate sources, EMS and hospitals, handle the data, but the kinds of databases on which data reside are disparate. Further, data

Table 51
Emergency Medical Services/Trauma Data
Summary Scores for Strategies and Goals
Across all Scenarios

STRATEGY		GOALS						COST ANALYSIS		
								Implementation*		Change in Cost of Operation
		QUALITY	SUITABILITY	SAFETY	SUM OF SCORES	EFFECT ON ORGANIZ.	SUM OF SCORES	Costs	Time Required	
cc2	Sampling techniques	0.6	-0.7	2.0	0.0	[redacted]	[redacted]	1.3	3.7	-0.3
cc5	Use of technologies	0.2	0.3	0.7	0.0	[redacted]	[redacted]	3.3	3.7	-0.3
cc6	Common state-wide form	0.9	0.4	0.0	0.0	[redacted]	[redacted]	2.7	3.7	0.3
cc7	EMS collect all injury data	1.1	0.1	-0.3	0.0	[redacted]	[redacted]	2.3	4.0	-0.3
ca7	Enhanced training	0.5	0.0	0.3	0.0	[redacted]	[redacted]	1.7	3.0	-0.3
ca13	Common coding nationally	1.0	0.2	-0.3	0.0	[redacted]	[redacted]	1.7	4.3	1.0
ma2	Quality-assurance programs	0.5	0.0	0.0	0.0	[redacted]	[redacted]	2.3	3.0	1.0
me1	Scannable forms	0.0	0.0	0.0	0.0	[redacted]	[redacted]	3.3	4.3	-1.4
me2	Online/batch editing	0.2	0.0	0.0	0.0	[redacted]	[redacted]	2.3	3.3	0.7
me3	Linked data sources	0.7	0.3	0.3	0.0	[redacted]	[redacted]	4.0	4.7	-0.3
me4	Local data entry	0.3	0.3	0.0	0.0	[redacted]	[redacted]	3.7	4.3	-0.3
me10	EMS/Trauma database	0.6	0.4	0.0	0.0	[redacted]	[redacted]	4.0	4.3	-0.7
me11	Use of analysts to interpret	0.4	-0.1	0.3	0.0	[redacted]	[redacted]	2.0	3.3	0.7
me12	Imaging & OCR/OMR	0.3	0.5	0.0	0.0	[redacted]	[redacted]	4.5	4.5	-1.0
ms1	Relational database	0.1	0.4	0.0	0.0	[redacted]	[redacted]	2.7	4.3	-0.3
ms2	Distributed databases	0.3	0.3	0.0	0.0	[redacted]	[redacted]	3.7	5.0	-0.3
ms4	Central-level quality control	0.5	-0.1	0.0	0.0	[redacted]	[redacted]	2.7	3.0	0.0

Nominal scoring scale:
- -2 substantially detracts from the goal or objective
- -1 somewhat detracts from the goal or objective
- 0 has minimal or no effect on the goal or objective
- +1 somewhat contributes to the goal or objective
- +2 substantially contributes to the goal or objective

* range - 1, lowest to 5, highest

are collected on a wide variety of forms, not all of which are transferred to automated sources. Finally, so many of the data are for medical purposes and have only limited relationship to traffic safety. The final cost assessment was performed in the same manner as that for the roadway inventory, by using qualitative ratings of effect on costs. The results of the ratings also appeared in table 51.

Benefit-Cost Results. The results from the ratings were translated into an overall cost-effectiveness framework, in the same way as explained above for the roadway inventory data source. The resulting table of cost effectiveness for strategies appears as table 52. The "proven" strategies have been added to the table, as well. All of the most highly-rated strategies were from those originally selected for further evaluation. None of the proven strategies fell into the higher class. More importantly, with one exception, all of the highly rated actions also are related to the use of advanced technologies.

Conclusions from the Ratings of Strategies

An overall summary of the resulting cost-effectiveness rankings is summarized in Table 53. Each of the three sources of data is shown along with two of the three classes for the strategies: proven and requiring further evaluation. A strategy is shown in terms of its relative cost-effectiveness, lower, moderate, and higher. Those strategies which have cost-effective ratings shown in upper-case and italicized letters represent ones selected for further evaluation and which have been discussed at the beginning of this chapter. None of the candidate strategies was found to be detrimental to achieving the objectives that have been posited for the system.

With the exception of using pre-printed diagrams, simplified form design, and motivation and feedback, all proven strategies fell into the classification of either moderate or relatively low cost-effectiveness. Some possible reasons given for the lack of higher effectiveness from already implemented actions is that they have been relatively limited in their scope, e.g. only assisting the data managers, limited in their extent, and limited in the degree of technical difficulty. Most managers of data sources usually are unwilling to implement major programs, not only because funds are limited, but because such programs also are unproven.

Discussion of the Strategies

Crash Report Data

Strategies which have been ranked as relatively high cost-effectiveness generally are designed to reduce or eliminate the role of the general patrol officer in the collection of crash data. This is done either through non-technological mechanisms which replace the general patrol officer with specialists or by supporting the officer with improved tools. Database strategies are designed to both reduce the amount of data that must be collected by officers, and to

Table 52
Relative Cost Effectiveness of Strategies for Improved Collection and Management of Emergency Medical Data

Relatively Higher Cost-Effectiveness	Relatively Moderate Cost-Effectiveness	Relatively Lower Cost-Effectiveness
CC2 - Sampling techniques *CC5 - Use of technologies* *CC7 - EMS collect all injury data* *CA7 - Enhanced training* *ME1 - Scannable forms* *ME3 - Linked data sources* *ME10- EMS/trauma database* *ME5 - Imaging and OCR/OMR*	*CC6 - Common statewide form* ca9 - Interaction with counterparts ca10 - Conferences and training programs ca11 - Field participation in forms design *ME4 - Local entry* *MS1 - Relational database* *MS2 - Distributed databases* *MS4 - Central quality control*	ca1 - Improved instruction manuals ca5 - Improved definitions of elements & codes ca8 - Collector incentives *CA13 - Common coding nationally* *MA2 - Quality assurance programs* ma3 - Management support *ME2 - Online & batch editing* me6 - Motivation & feedback *ME11 Analysts to interpret and code reports*

* Lower-case items numbers indicate classification is based upon interpretation of input received from operating agencies visited. Upper-case and italicized titles are based upon ratings of effectiveness and results of cost projections using the project cost-model.

minimize the handling of data at different levels of government. Imaging systems that employ OCR and OMR technology are considered effective means for cost reduction in the interim (5 to 10 years) until the majority of reports are completed on automated devices.

The formation of a system task force, a highly rated strategy for data management, can be accomplished quickly, and can be a foundation which facilitates and supports many others. Most other highly-rated strategies involve use of technological advances to assist with managing and storing data.

A final classification for strategies identified those which could be implemented readily (in less than 6 months as an example), and those which would take a longer period or were technologically advanced (generally implying a longer period to complete). Table 54 provides a display of the strategies using both the time-classification and their relative cost-effectiveness.

The strategies identified as having rapid implementation include some important means for improving quality and reducing costs associated with data collection and management. Many

Table 53
Effectiveness of Strategies for Improving Collection and Management of Highway Safety Data

Proven	Evaluate Further	Strategy	Crash Reporting	Roadway Inventory	Medical Services
		DATA COLLECTION - CHANGE METHODS (CC)			
	x (rdwy)	CC1. Telephonic reporting		L	
	x	*CC2. Sampling techniques*	*M*		*H*
x		CC3. Increase reporting thresholds	m		
	x	*CC4. Non-sworn crsh invetig./Data collection specialists*	*H*	*M*	
	x	*CC5. Use of non-automated technology*	*H*	*H*	*H*
	x	*CC6. Common, state-wide EMS/trauma reporting form*			*M*
	x (EMS)	*CC7. EMS collects all injury-related data*			*H*
		CC8. Must-move legislation			
x		CC9. More extensive supervisor review	m	m	
	x	*CC10. Completion by specially trained personnel (local level)*	*M*		
x		CC11. Use of short-form and long-form reports	m		
	x	*CC12. Sworn crash investigators*	*H*		
		DATA COLLECTION - ASSIST COLLECTOR (CA)			
x		CA1. Improved instruction manuals	m	m	l
x		CA2. Pre-printed diagrams	h		
x		CA3. Simplified form design	h		
	x	*CA4. Revised/improved design and format*	*H*		
x		CA5. Improved definitions of elements/codes	m	m	m
		CA6. Priorities given to reporting			
	x	*CA7. Enhanced training*	*M*	*M*	*M*
x		CA8. Collector incentives	m	l	l
	x	*CA9. Interact with counterparts*	*M*		m
x		CA10. Conferences and training programs	m	l	m
x		CA11. Field participation in forms design	m	m	m
	x	*CA12. Minimum wage persons for data collection*		*L*	
	x (medical)	*CA13. Common coding nationally*			*L*
	x	*CA14. Computer directed interview (AI/expert systems)*	*H*		h

Table 53
Effectiveness of Strategies for Improving Collection and Management of Highway Safety Data

Proven	Evaluate Further	Strategy	Crash Reporting	Roadway Inventory	Medical Services
		DATA MANAGEMENT - ADMINISTRATION (MA)			
		MA1. Security for privacy			
	x	*MA2. Quality-assurance programs*	*M*	*L*	*L*
x		MA3. Management support	m	m	l
	x	*MA4. System task-force*	*H*	h	h
		MA5. Reorganization of management			
		DATA MANAGEMENT - DATA ENTRY (ME)			
	x	*ME1. Scannable forms*	*M*		*H*
	x	*ME2. Online and batch editing*	*M*	*L*	*L*
	x	*ME3. Linked sources of data*	*H*	*H*	*H*
	x	*ME4. Local data entry*	*H*	*M*	*M*
	x	*ME5. Imaging systems*	*H*	*M*	*H*
x		ME6. Motivation and feedback	h	l	l
x		ME7. Cross-training personnel	m		
	x	*ME8. Sponsored elements*	*M*	*M*	
	x	*ME9. Comprehensive requirements for development and redesign*	*M*	*H*	
	x	*ME10. EMS/trauma database*			*H*
	x	*ME11. Use of data analysts to interpret descriptive material*	*M*		*L*
		DATA MANAGEMENT - STORAGE (MS)			
	x	*MS1. Relational database systems*	*M*	*M*	*M*
	x	*MS2. Distributed databases*	*H*	*M*	*M*
		MS3. Object-oriented database			
x	x(rdwy. &med.)	*MS4. Centralized quality control*	h	L	*M*

Table 54
Classification of Strategies for Collection and Management of Crash Report Data

	Relatively Higher Cost-Effectiveness	Relatively Moderate Cost-Effectiveness
Readily Implemented	ca2 - pre-printed diagrams ca3 - simplified form design CA4 - Revised PCR design MA4 - System task force me6 - Motivation and feedback	cc3 - Increase reporting thresholds cc9 - More extensive supervisor review ca1 - Improved instruction manuals ca5 - Improved definitions CA7 - Enhanced training ca8 - Collector incentives CA9 - Interact with Counterparts ca10 - Conferences and training programs ma3 - Management support me7 - Cross training ME8 - Sponsored elements
Longer-Term &/or Technologically Advanced	CC4 - Non-sworn investigators CC5 - Use of non-automated technology CC12 - Sworn Crash Investigators **CA14 - Computer directed interview** **ME3 - Linked data sources** **ME4 - Local data-entry** **ME5 - Imaging systems & OCR/OMR** **MS2 - Distributed databases**	CC2 - Sampling CC10 - Completion by specially trained personnel (local level) cc11 - Short and long forms ca11 - Field participation in form design MA2 - Quality-assurance programs **ME1- Scannable forms** **ME2 - On-line & batch editing** ME9 - Comprehensive requirements for Development ME11 - Use Data Analysts to Interpret **MS1 - Relational Database** ms4 - Centralized quality control

- Lower-case items numbers indicate classification is based upon interpretation of input received from operating agencies visited. Upper-case values are based upon ratings of effectiveness and results of cost projections using the project cost-model.
- No strategies were classified as having relatively low cost-effectiveness
- **Bolded** items are considered technologically advanced

of the strategies in this group involve common-sense actions, which would be implemented in any comprehensive improvement program. Finally, most are related to training and education, in different forms.

The longer-term and technologically-oriented strategies cover the gamut of impacts desired. including lower cost from scanning of reports, better quality through comprehensive editing systems, and greater flexibility through use of relational database designs. Quality-assurance and quality-control activities are included in this grouping. The effects of these are probably quality by a dedicated quality-assurance team. Where current management systems include effective pre-processing systems and rudimentary batch-edit routines, the quality assurance effort will be less cost-effective than where these elements are not present.

Because the project team was able to estimate costs associated with the collection and management of crash data, they were able to compute changes in unit costs for each of the strategies which applied to the crash reporting. These unit costs are ordered from the greatest reduction to the greatest increase and are shown in table 55. Along with the changes in costs are the relative effectiveness for each strategy. Of the 35 strategies applicable to crash reporting, 11 were estimated to have a positive effect on costs. Their implementation would help reduce costs associated with collection and management (in some cases either, in some both) of the data. The greatest savings appears to result from introducing long and short forms. Savings results from the police performing less data collection at the scene rather than less costs for data management. However, the cost effectiveness of this strategy is only moderate because the reduction in data collection also reduces the amount of data available for safety information.

Roadway Inventory Data

The strategies ranked relatively high in cost-effectiveness focus on use of technology, both for measurement and acquiring data, in order to limit what must be collected in the field. Comprehensive system requirements, once established, will provide an important framework for developing a roadway-inventory system that can be fully integrated with other systems.

Strategies for improving the collection and management of roadway-inventory data also were classified by ease of implementation in the same manner as was performed for crash data collection and management. Either the strategies were easily implemented, or they were longer term or required advanced technologies. Table 56 displays the results of the classification coupled with the relative cost-effectiveness. Unlike strategies for crash data, those for roadway-inventory data fell into three categories, lower, moderate, and higher cost-effectiveness.

The more moderately cost-effective strategies involve a combination of non-technological assistance to the data collector and the use of specialists instead of district engineers who have many other responsibilities. Strategies directed toward data management tend to focus upon technological approaches, either to reduce labor costs, or to institute improved database designs. While the technological strategies are considered to be relatively effective, they also have relatively high costs associated with their implementation.

Table 55
Effect of Strategies on Unit-Cost of Collecting and Managing Crash Data

Index No.	Description	Change in Unit-Cost %	Change in Unit-Cost $/PCR		Rating of Relative Cost-Effectiveness
CC11	Use of short-form and long-form reports	-11%	($2.10)	Reduced Costs	□
CA2	Pre-printed diagrams	-6%	($1.20)		●
CA3	Simplified form design	-6%	($1.20)		●
CA4	Revised/improved design and format	-6%	($1.20)		●
ME3	Linked sources of data	-6%	($1.20)		●
CA14	Computer directed interview (AI/expert systems)	-6%	($1.10)		●
ME5	Imaging systems	-6%	($1.10)		●
CC4	Non-sworn crash investig./Data collection specialists	-5%	($0.90)		●
MS2	Distributed databases	-5%	($0.90)		●
ME1	Scannable forms	-4%	($0.80)		□
CC5	Use of non-automated technology	-3%	($0.60)		●
CA1	Improved instruction manuals	0%	$0.00	No Chg	□
CA5	Improved definitions of elements/codes	0%	$0.00		□
ME4	Local data entry	0%	$0.00		●
CC12	Sworn crash investigators	0%	$0.00		●
MA3	Management support	0%	$0.00		□
CA11	Field participation in forms design	0%	$0.00		□
ME7	Cross-training personnel	1%	$0.10	Incr.	□
CA8	Collector incentives	1%	$0.10		□
ME6	Motivation and feedback	1%	$0.10		●
ME8	Sponsored elements	1%	$0.10		□
MA4	System task-force	1%	$0.20		●
CC9	More extensive supervisor review	1%	$0.20		□
CA9	Interact with counterparts	2%	$0.30		□

Table 55
Effect of Strategies on Unit-Cost of Collecting and Managing Crash Data

Index No.	Description	Change in Unit-Cost			Rating of Relative Cost-Effectiveness
		%	$/PCR		
ME2	Online and batch editing	2%	$0.40	Inc. Costs	□
CA7	Enhanced training	2%	$0.40		□
CC2	Sampling techniques	2%	$0.40		□
MS1	Relational database systems	2%	$0.40		□
CA10	Conferences and training programs	3%	$0.50		□
ME9.	Comprehensive requirements for development and redesign	3%	$0.50		□
ME11	Use of data analysts to interpret descriptive material	3%	$0.60		□
MA2	Quality-assurance programs	5%	$0.90		□
CC3	Increase reporting thresholds	5%	$0.90		□
CC10	Completion by specially trained personnel (local level)	5%	$1.00		□
MS4	Centralized quality control	5%	$1.00		●

● = Relatively high cost-effectiveness □ = relatively moderate cost-effectiveness

Several of the strategies ranked with relatively low cost-effectiveness focus upon activities designed to encourage or educate the data collector. Because the collector of roadway data is directly accountable to the people who use the data, this is considered enough incentive to do a good job. Furthermore, if the person collecting the data is doing this as a low-priority task, not part of their basic job description, enhanced training is not likely to affect that person's incentives may be meaningless. Quality-assurance programs and centralized quality control were rated relatively low; however, the rating is highly dependent upon how the agency treats and uses the roadway data.

Emergency Medical Data

The relatively higher cost-effectiveness rankings for strategies to improve emergency medical data fall mostly with longer-term technological applications. These are primarily designed to improve data management, both through linkage and reducing labor-intensive activities. These are nicely combined with complementary activities on the collection side, which designate the better-trained EMS personnel to collect injury data needed for safety analysis.

Table 56
Classification of Strategies for Collection and Management of Roadway-Inventory Data

	Relatively Higher Cost-Effectiveness	Relatively Moderate Cost Effectiveness	Relatively Lower Cost-Effectiveness
Readily Implemented		cc9 - More extensive supervisor review ca1 - Improved instruction manuals ca5 - Improved definition of codes CA7 - Enhanced training ma3 - Management support	CC1 - Telephone reporting ca8 - Collector incentives ca10 - Conferences and training programs me6 - Motiviation and feedback MS4 - Centralized quality control
Longer Term &/or Technologically Advanced	**CC5 - Use of Technologies** **ME3 - Linked data sources** ME9 - Comprehensive system requirements	CC4 - Data collection specialists ca11 - Field participation in forms design ME4 - Local data entry ME8 - Sponsored elements ME11 Imaging & OCR/OMR **MS1 - Relational databases** **MS2 - Distributed databases**	CA12- Use of minimum wage field-personnel MA2 - Quality assurance programs **ME2 - Online/batch editing**

- Lower-case items numbers indicate classification is based upon interpretation of input received from operating agencies visited. Upper-case values are based upon ratings of effectiveness and results of cost projections using the project cost-model.
- **Bolded** items are considered technologically advanced

The other selected strategies provide those personnel with better training and tools to simplify their job and make them effective providers of data for those outside the medical system, as well.

Table 57 displays the classification of strategies for collecting and managing medical data. These too are divided into those easily implemented and those which require a longer period or use of technologies to assist with their implementation. The strategies ranked relatively moderate in cost effectiveness are less technologically oriented than the previous group. They focus on form design, quality assurance, and improved database technology.

The strategies having a relatively lower cost-effectiveness ranking involve, in the short term, basic support for the collector and those managing the data. For the longer run, national uniformity was not considered relatively cost-effective, nor was the use of special interpreters. The latter was because the personnel in the emergency medical area are generally highly

Table 57
Relative Cost Effectiveness of Strategies for Improved Collection and Management of Emergency Medical Data

	Relatively Higher Cost-Effectiveness	Relatively Moderate Cost-Effectiveness	Relatively Lower Cost-Effectiveness
Readily Implemented	CA7 - Enhanced training	ca9 - Interaction with counterparts ca10 - Conferences and training programs ME4 - Local entry	ca1 - Improved instruction manuals ca5 - Improved definitions of elements & codes ca8 - Collector incentives ma3 - Management support me6 - Motivation & feedback
Longer-Term &/or Technologically Advanced	CC2 - Sampling techniques CC5 - **Use of technologies** CC7 - EMS collect all injury data ME1 - **Scannable forms** ME3 - **Linked data sources** ME10 - **EMS/trauma database** ME5 - **Imaging and OCR/OMR**	CC6 - Common statewide form ca11 - Field participation in forms design MA2 - Quality assurance programs MS1 - **Relational database** MS2 - **Distributed databases** MS4 - Central quality control	CA13 - Common coding nationally ME2 - Online & batch editing ME11 - Analysts to interpret and code reports

- Lower-case items numbers indicate classification is based upon interpretation of input received from operating agencies visited. Upper-case values are based upon ratings of effectiveness and results of cost projections using the project cost-model.
- **Bolded** items are considered technologically advanced

trained. On-line editing appears with relatively low rank because fairly effective manual editing is assumed extant.

Grouping of Strategies

Because the strategies represent a broad range of activities, at different levels of investment and for different parts of the system, an important consideration is how they can be combined for overall better results than if implemented individually, or in an uncoordinated manner. In fact, some of the strategies are interdependent, some are inclusive of others, and some are mutually exclusive of others. The relationship between the various strategies is demonstrated in table 58. The stronger relationships are indicated through symbols, however, the blank cells are also indicative of a positive, yet not strong, relationship (a symbol was not used to avoid cluttering the table).

Table 58
Relationships Between Strategies

	CC1 Teleph. Reporting	CC2 Sampling	CC3 Increase Thresh-olds	CC4 Non-sworn Investig	CC5 Non-automated Technol.	CC6 Common EMS Form	CC7 EMS Collect Injury	CC8 Must Move Legisl.	CC9 Extens. Superv Review	CC10 Compltn by Trained Persnl.	CC11 Short & Long Forms	CC12 Sworn Investigs.
CC1. Telephonic reporting	---									○	○	
CC2. Sampling techniques		---									●	○
CC3. Increase reporting thresholds	○		---									
CC4. Non-sworn crsh invetig./Data collection specialists				---	○		○					
CC5. Use of non-automated technology	☮			○	---				○	○		○
CC6. Common, state-wide EMS/trauma reporting form					○	---			○	○		
CC7. EMS collects all injury-related data					○	●	---		○	○		
CC8. Must-move legislation								---				
CC9. More extensive supervisor review						○	○		---	☮		
CC10. Completion by specially trained personnel (local level)				○		○	○		☮	---		○
CC11. Use of short-form and long- form reports		○							○	○	---	
CC12. Sworn crash investigators					○		○		○			---
CA1. Improved instruction manuals						○	○					
CA2. Pre-printed diagrams	☮					☮	☮					
CA3. Simplified form design	○				○	○	○		○	○	○	
CA4. Revised/improved design and format	○				○	●	○		○	○	○	
CA5. Improved definitions of elements/codes						○	○		○	○		
CA6. Priorities given to reporting	☮	☮	☮	○	○	☮	○		○	○	○	○
CA7. Enhanced training				○		○			○	○		○
CA8. Collector incentives	☮			○		☮	☮		○	○		○
CA9. Interact with counterparts				○								○
CA10 Conferences and training programs				○			○					○
CA11 Field participation in forms design				○		○						
CA12 Minimum wage persons for data collection	○	☮	☮	●	○	☮	☮	☮	●	☮	☮	☮
CA13 Common coding nationally				○	○	○	○				○	○
CA14 Computer directed interview (AI/expert systems)							○					
MA1. Security for privacy												
MA2. Quality-assurance programs				○	○				✔	✔		○
MA3. Management support									○	○		
MA4. System task-force						○	○					
MA5. Reorganization of management												
ME1. Scannable forms				○		○			○	○		○
ME2. Online and batch editing									○	○		
ME3. Linked sources of data						○	○					
ME4. Local data entry									○	○		
ME5. Imaging systems												
ME6. Motivation and feedback				○								○
ME7. Cross-training personnel												
ME8. Sponsored elements												
ME9. Comprehensive rqrmnts for development and redesign												
ME10 EMS/trauma database						○	○					
ME11 Use of data analysts to interpret descriptive material				○		○	○			✔		○
MS1. Relational database systems						○						
MS2. Distributed databases						○						
MS3. Object-oriented database												
MS4. Centralized quality control				○	○	○			○	○		○

Table is to be read horizontally. That is, each row is the referenced strategy and each cell in the row indicates the relationship of the other strategy to it

- ● The presence of the strategy above is necessary, or extremly important, to the success of the referenced strategy
- ○ The presence of the strategy above is important, or very useful, to the success of the referenced strategy
- ✔ The referenced strategy is a subset of the strategy above
- ☮ The two strategies are usually inappropriate to be practiced together, or it is not applicable

Table 58
Relationships Between Strategies

	CA1 Imprvd. Instrucn. Manuals	CA2 Pre-Printed Diags.	CA3 Simplified Form Design	CA4 Revised Form Design	CA5 Improved Defs.	CA6 Priority to Reportg	CA7 Enhanced Training	CA8 Collector Incentives	CA9 Interact with Others	CA10 Confs. & Trng Progs.	CA11 Assist in PCR Design	CA12 Minim Wage Collcrs.	CA13 Common National Coding	CA14 Computer Directed Intervw.
CC1. Telephonic reporting	○	○	○	○		⊛	○	○			○	✔		●
CC2. Sampling techniques	●			●		⊛	●	○	○	○		⊛		
CC3. Increase reporting thresholds						⊛						⊛		
CC4. Non-sworn crsh invetig./Data collection specialists	○					✔	○		○	○				⊛
CC5. Use of non-automated technology	○						○							
CC6. Common, state-wide EMS/trauma reporting form	○	⊛	⊛	⊛	○	⊛	○		○	○	○	⊛		○
CC7. EMS collects all injury-related data		⊛	⊛	⊛	○	⊛	○	○	○	○		⊛		
CC8. Must-move legislation												⊛		
CC9. More extensive supervisor review			○			✔	○					⊛		
CC10. Completion by specially trained personnel (local level)			○			✔	○					⊛		
CC11. Use of short-form and long- form reports	○		✔	✔	○	⊛	○				○	⊛		○
CC12. Sworn crash investigators	○					✔	○		○	○		⊛		
CA1. Improved instruction manuals	---					✔								
CA2. Pre-printed diagrams		---		○										
CA3. Simplified form design	○		---	✔	✔		○				○	⊛		○
CA4. Revised/improved design and format	○		✔	---	✔		○				○	⊛		○
CA5. Improved definitions of elements/codes	●				---		●				○			○
CA6. Priorities given to reporting	○		○	○	○	---	○	✔			○	⊛		
CA7. Enhanced training			○		○	○	---							
CA8. Collector incentives						○		---	○	○	○	○		
CA9. Interact with counterparts							○	○	---	○	⊛	⊛		
CA10. Conferences and training programs	○					○		○	✔	---	⊛			
CA11. Field participation in forms design			✔	✔	✔	○		○	○	○	---			
CA12. Minimum wage persons for data collection	●	⊛	⊛	⊛	●	⊛	●	○	⊛	○		---		⊛
CA13. Common coding nationally	●		○	●	✔		●	●	○	○			---	○
CA14. Computer directed interview (AI/expert systems)	●			●	●					○	○	⊛		---
MA1. Security for privacy														
MA2. Quality-assurance programs	○	○			○	○	○	○		○		⊛		
MA3. Management support								○				⊛		
MA4. System task-force											○	⊛		
MA5. Reorganization of management														
ME1. Scannable forms	○		○	●			○				○	⊛		⊛
ME2. Online and batch editing	○			○	○									
ME3. Linked sources of data														
ME4. Local data entry	○		○	○	○		○							
ME5. Imaging systems		○												
ME6 Motivation and feedback						✔		●	○	○	○	⊛		
ME7. Cross-training personnel	⊛	⊛					●					⊛		
ME8. Sponsored elements				●	○									
ME9.														
ME10 EMS/trauma database		⊛	⊛	⊛	○	⊛				○		⊛		
ME11 Use of data analysts to interpret descriptive material	○	○	○	○	○		○			○		⊛		⊛
MS1. Relational database systems					○									
MS2. Distributed databases					○									
MS3. Object-oriented database		○		○										
MS4. Centralized quality control					○									

Table is to be read horizontally. That is, each row is the referenced strategy and each cell in the row indicates the relationship of the other strategy to it.

● The presence of the strategy above is necessary, or extremly important, to the success of the referenced strategy

○ The presence of the strategy above is important, or very useful, to the success of the referenced strategy

✔ The referenced strategy is a subset of the strategy above

⊛ The two strategies are usually inappropriate to be practiced together, or it is not applicable

Table 58
Relationships Between Strategies

	MA1 Security for Privacy	MA2 Quality Assurance Progs.	MA3 Mgmt. Support	MA4 System Task Force	MA5 Reorg. of Managmt.	ME1 Scannable Forms	ME2 Online & Batch Editing	ME3 Linked Data Sources	ME4 Local Data Entry	ME5 Imaging Systems	ME6 Motivn & Feedbk.	ME7 Cross Trng. Persnl.	ME8 Sponsrd Elements	ME9 Compre. Reqmts.	ME10 EMS/ trauma Database	ME11 Data Analysts Interp.	MS1 Relanl. DB's	MS2 Distributd. DB's	MS3 Object Oriented DB's	MS4 Centralzd Quality Control
CC1. Telephonic reporting		○						○	○			○								
CC2. Sampling techniques																				
CC3. Increase reporting thresholds																				
CC4. Non-sworn crsh invetig./Data collection specialists		○	○		○															
CC5. Use of non-automated technology				○																
CC6. Common, state-wide EMS/trauma reporting form		○		○		○		○		○										
CC7. EMS collects all injury-related data		○	○	○			○	●		○		⊗	⊗	⊗						
CC8. Must-move legislation																				
CC9. More extensive supervisor review		✔	○												⊗					
CC10. Completion by specially trained personnel (local level)		○													⊗	⊗				⊗
CC11. Use of short-form and long- form reports			⊗	○											⊗					
CC12. Sworn crash investigators		○	○												⊗					
CA1. Improved instruction manuals				○																
CA2. Pre-printed diagrams																				
CA3. Simplified form design				○				○							⊗	○				
CA4. Revised/improved design and format				●		○		○							⊗					
CA5. Improved definitions of elements/codes		○		○			○	⊗					○	○						
CA6. Priorities given to reporting			●	○										○						
CA7. Enhanced training			○																	
CA8. Collector incentives		○	○								✔									○
CA9. Interact with counterparts		○	○																	
CA10. Conferences and training programs		○	○	○							○									
CA11. Field participation in forms design			○	○																
CA12. Minimum wage persons for data collection						⊗						⊗			⊗	⊗				
CA13. Common coding nationally		○		○			○													○
CA14. Computer directed interview (AI/expert systems)				○			⊗	●		○				●	○		○	○	○	
MA1. Security for privacy	---			○				○						○	○		○	●	○	
MA2. Quality-assurance programs		---	●	○			✔	●			✔									✔
MA3. Management support			---	○	○															
MA4. System task-force			●	---	○															
MA5. Reorganization of management			●	○	---															
ME1. Scannable forms				○		---				○					⊗					○
ME2. Online and batch editing							---	○	○	○		○		○			○	○	○	○
ME3. Linked sources of data			●	○				---		○				●	○		●	○		
ME4. Local data entry			○	○		○	○	○	---	○		○		○			○	○		
ME5. Imaging systems				○		●				---				○						
ME6. Motivation and feedback		○	○								---									○
ME7. Cross-training personnel			○									---			⊗					
ME8. Sponsored elements			○	●	○								---		⊗					
ME9. Comprehensive rqrmnts for development and redesign			○	●	○									---	⊗					
ME10 EMS/trauma database			●	○				○		○		⊗	⊗	⊗	---		○	○		○
ME11 Use of data analysts to interpret descriptive material								○								---	○			
MS1 Relational database systems			○	●				✔						○			---			
MS2 Distributed databases			○	●				●						○			●	---		
MS3. Object-oriented database			○	●		○		○		○				○			○		---	
MS4 Centralized quality control			○	○	○		○	○				○					○			---

Table is to be read horizontally. That is, each row is the referenced strategy and each cell in the row indicates the relationship of the other strategy to it

● The presence of the strategy above is necessary, or extremly important, to the success of the referenced strategy.
○ The presence of the strategy above is important, or very useful, to the success of the referenced strategy
✔ The referenced strategy is a subset of the strategy above.
⊗ The two strategies are usually inappropriate to be practiced together, or it is not applicable

Symbols also show mutual exclusivity, i.e., where the two strategies are inappropriate to be practiced together. An example of such inappropriate selection is simplified report design and EMS/trauma database. Because of the use of these data for many specialized purposes beyond the needs of the safety community, and because of the training of the collectors, a simplified form would not enhance the process of collecting often complex medical data.

On the other hand, some strategies require others to be in place. An example is the use of computer-directed interviews. Linked data systems need to be active before practical use can be made of computers to assist with interviews. Likewise, redesign of a crash report should take place under the direction of a system task force.

The table is designed for the user who, although focussing upon a specific issue, e.g., use of non-sworn crash investigators, can quickly scan across to identify highly complimentary strategies. Examples include: relieving the investigator of reporting injury severity by assigning that responsibility to EMS personnel, and re-organization of management to insure that leadership is directed toward the new group, rather than just being an added responsibility of an already overloaded administrator.

Further Research and Demonstration Needed

Throughout this project a number of aspects of the topic have been identified as items requiring further study. The level of study can be divided simply into demonstration and research. Many of the strategies recommended here have demonstrated their effectiveness because they are currently in use in one or more operating agencies. However, not much has been done to quantify the effectiveness of the strategy. This project has used qualitative ratings to make a first estimate, but field demonstrations under monitored conditions will be necessary to provide enough documentation to satisfy others seeking guidance. Therefore, a set of field evaluations is needed.

Field Evaluations

Strategies are not implemented in a vacuum, and usually not as a single measure (except where the strategy is really a composite of several sub-strategies). Therefore, field evaluations should seek to implement a "package" of strategies. Each package should have an underlying theme, or core. Examples of themes include:

1. Replacement of the general patrol officer as a crash reporter in urban and suburban areas.

2. A basic re-design of the crash report.

3. A basic re-design of the database management system.

4. A comprehensive quality-assurance program.

5. Complementing a well-designed, operating collection and management system by applying advanced technology throughout the process.

6. Reducing crash data-reporting requirements of the officer to an absolute minimum (without losing needed data).

7. Providing complete longitudinal geometrics in the roadway inventory database.

8. Creating a roadway-inventory history-database.

9. Use of EMS personnel to collect injury-severity and driver-condition data in a form suitable for highway safety analyses, without using "blind" matching i.e., as being done under the CODES project.

10. Localizing data entry and data management.

11. Implementing major re-organization to more effectively direct and control the data collection and management process.

12. Achieving maximum inter-connectivity of, and minimum duplication between, safety databases.

13. Developing and implementing a comprehensive and continuing effort to train safety data collectors, using state-of-the-art training techniques (to demonstrate the real potential of adequate training).

Each of these "themes" can involve several, if not many, of the strategies that have been suggested. The specific combination will be highly dependent upon conditions in the jurisdiction involved. The ideal approach would be to package strategies considered highly complementary, applying them to a site that has an appropriate base condition. The evaluation would involve developing and quantifying measures of effectiveness designed to address issues of cost, quality, and institutional or organizational impacts.

Research

A complementary research effort is needed to investigate some strategies and issues more fully before FHWA can suggest other new or improved strategies for agencies to try. Research needed includes:

1. Identifying ways for a jurisdiction to estimate the degree of coverage of crashes of various types, how they vary over time and with respect to other variables, and how an agency can use that information to adjust its evaluation designs, analytical methods, and decision making.

2. Further study of the strategies identified as requiring more development.

3. Expand current FHWA efforts to study emerging technologies, to include (possibly using the current FHWA program for traffic-records technology-application and demonstration):

   a. Database management, e.g., advanced databases, advanced communications.

   b. Roadway-inventory data collection, e.g., using satellite imaging and instrumented vehicles, especially to obtain longitudinal geometrics, and to maintain current conditions.

   c. Use computer-based interviews for crash and medical-data collection (including possible application of expert systems).

   d. On-going projections and monitoring of technological advances to allow early experimentation, e.g., satellite imaging, data from intelligent-vehicles.

8. Review and Recommendations

Results of the Investigations

Issues Related to Quality and Costs

The evaluation of highway safety data set out to identify issues associated with its collection and management, to determine the quality associated with the efforts, and calculate costs of performing the work. These goals were met and have been addressed in detail in this report. In a final step, the research team used the information related to issues, quality, and costs to develop and evaluate a set of strategies designed to improve the system. This final chapter highlights key findings and describes those strategies which have the greatest potential cost-effectiveness.

Throughout the study, three sources of data have been considered: highway crash, roadway inventory, and crash injury (also called "medical" or "emergency medical services and trauma") data. Most of the attention in this report was devoted to crash data because most of the efforts expended and issues related to quality and costs are associated with handling these data.

As part of the research, the team examined what the literature has said, talked with experts throughout the Nation, and visited 10 sites to view and collect first-hand information about operation of the data sources. Descriptions of the efforts and findings have appeared in various reports produced during the course of the project. Salient points were summarized in this report.

Issues of quality and costs of data collection were found to be the most critical among those studied. Within quality, most important was availability of appropriate, accurate data in a timely fashion. Cost issues reflected the growing gap between a larger need to collect and use highway-safety data and the resources available to do the work.

Quality of collecting and reporting data has been treated thoroughly in the literature. Hauer and Hakkert, and later O'Day both have written significant treatises addressing some critical aspects regarding the quality of crash data. The important finding is that users of the data must ensure that their use reflects those conditions which affect quality. This finding also reflects the fact that quality varies not only among groups of users but also within specific classes. Police use of crash data reflect a different set of conditions than a researcher studying injury-producing crashes. Likewise, within a hospital the billing organization and those measuring the effectiveness of the trauma unit have widely differing requirements.

Field interviews and focus groups clearly identified the differences in needs along with and concerns related to both quality and costs. Those persons who must respond to the public, e.g., a mayor, needed timely data showing how, when, and where crashes occurred within the community. On the other hand, a State agency receiving funds for reducing driving-under-the-influence required data describing the role of alcohol in crashes at the end of an extended period of evaluation. Finally, the highway engineer facing the need to distribute safety-improvement funds required comprehensive, accurate, and precise data relating crashes to highway characteristics. None of the organizations involved in safety-data collection and management were able to satisfy these diverse needs, although some were doing better than others. In fact, a number of noteworthy programs were being operated many of which were attempting to address improvements in such aspects of quality as accuracy and timeliness.

At the same time there is a diverse demand for quality data, there is an ongoing drive to reduce the costs of collecting and managing them. Police, for example, are being called upon to perform more and varied services with reduced staff. Collection of crash data, which is seen by many to be "for someone else" is one early victim of reduced services. At the management level, reductions in staffing are mirroring reductions in budgets available to support the State bureaucracy, while still meeting social mandates, e.g. medical assistance to the poor. As with issues of quality, approaches are being taken nationwide to achieve cost efficiencies while continuing to meet various needs for the safety-data.

Costs of Collecting and Managing Data

Of the three sources of data, crash, roadway inventory, and medical, costs could be estimated only for collecting and managing crash data. Two conditions precluded estimating costs for roadway inventory and medical data. First, few organizations were able to identify efforts and total costs associated with collecting these other data. Second, and perhaps more important, most of the data collected for these latter two sources are unrelated to highway safety.

As the research team discovered, collectors and managers had a vague idea of what costs were involved in handling highway safety data. At a macro-level, police managers could determine how many resources were being used for crashes and make estimates of their costs. Some of the managers of State systems had a good idea of costs either because their department was the sole manager of the statistics (and they knew the total budget), or because they had recently been upgrading their systems and had obtained a better estimate of costs for that purpose.

However, sufficient cost information was available at the local, State, and national level to make some estimations. What was lacking was time required to perform some tasks, particularly computing the costs of collecting data. Past research by others had computed times for handling the crash and did not account for the various tasks performed by the responder at the scene. Police managers assumed, and which subsequently proved to be an incorrect assumption, that costs incurred by their agency were used to satisfy safety needs, particularly those prescribed by the State. However, separating costs for safety data from those for crash (and case) management proved extremely difficult. On-scene measurement was limited to a few cases simply because of the very small likelihood of a crash occurring when an observer was available to measure time accurately. Attempts to use self reporting from officers proved highly unreliable for the needs that had to be made.

Even with the limitations placed on collecting timing data, the research team was able to develop some approximations of time for reporting crash data which ranged from 16 minutes for single-vehicle, property-damage-only (PDO) crash in an urban area to more than 60 minutes for a multiple-vehicle crash involving severe injuries in a rural area. These times reflected between 30 (the more frequent situation) and 70 percent of the time spent at the scene.[1] The remaining percentage of time was devoted to crash and case management.

From these costs, and from estimates regarding the amount of time spent by local and State agencies in managing the data a cost estimate was computed.[2] On a national level, for all types of crashes, the estimated average cost per crash report recorded was $19.20. Using the national estimates of crashes by severity and the estimates of time derived from the study, a range in unit-costs from $17.50 for a PDO to approximately $32.00 for a severe-injury crash could be shown.

The costs estimated for collecting and managing data for PDO crashes substantially exceeded earlier estimates, such as by NHTSA. On the other hand, the costs associated with a severe-injury crash were substantially lower, e.g. the NHTSA estimate of $185 for a fatal crash.

The relatively small differences in costs of collecting and managing crash data, particularly when compared to earlier estimates results from several situations. First, unless collectors used a "short" and "long" form of the crash report for collecting data, the amount of data collected for any crash, regardless of severity, would be the same. Data collection accounts for approximately two-thirds of the total. Although an argument could be raised that the costs associated with the Fatal Accident Reporting System (FARS) should be included for all fatal crashes, the reporting requirements had a negligible effect on the work required to collect and manage data at the State and local level. As a result, the incremental cost for a fatal crash would be small.

Second, assumptions generally held by police officers which could affect how they view the time spent on collection of safety data suggested significant time was spent on the fatal crash. Yet,

---

[1] The percentages are based on the time spent by the initial responding (and usually reporting) officer and did not include time spent by other responding officers, supervisors, and specialist. Subsequent personnel generally provided support for crash and case management. Rarely did they assist in collecting safety data. Excluding their time reduced the overall percentage of time spent for safety data. At a fatal crash, for example, the time spent collecting and reporting crash data might only be 5 percent of the total.

[2] In many cases, the research team found that the same data were being processed by multiple organizations. The costs computed for crash-data management assumed that the data were handled only once.

almost all of the additional time, along with additional resources, collected data used to manage the case and did not appear in the crash-safety data base. Even though the difference in time required to collect data for a PDO and a severe-injury crash was on the magnitude of 2 to 1, the severe crashes only account for 5 percent of the total crashes. On a per-minute basis, the difference in time will not add substantially to the overall time spent collecting and reporting crash data.

A more thorough study of the time required for data collection might provide more accurate for estimating the costs associated with various levels of crash severity. However, the resultant changes to the average costs estimated above would likely be relatively small primarily because, as the research team learned, most of the costs are incurred regardless of the level of severity with a limited increment arising from increasing severity itself.

Selection of Strategies

Evaluating the Strategies

Investigation and review of operations at selected sites nationally and discussion with key personnel in the field of highway safety information led to the development of a number of strategies for improving the quality and reducing the costs of handling the data. These strategies were classed as those which affected the collecting and reporting of data, or their management (or in some cases, both). For collecting and managing traffic-safety data, the research team limited their analysis to those strategies which were non-technological. Technological and non-technological strategies were evaluated as they applied to collecting and reporting roadway inventory and crash injury data. Both classes of strategies were considered for managing data from all three sources.

Each of the strategies was examined from two perspectives: Effect on meeting the goals and objectives set for the highway safety data, and effect on costs. Highly cost-effective strategies would be those which substantially met the goals and objectives, *and* which reduced costs associated with handling the data. A further division among the strategies was between those which were proven and easily implemented, and those which required further evaluation.[3]

The evaluation of each strategy was performed by rating the strategies against 6 scenarios representing site specific and systemic applications for the driver, vehicle and roadway, and an additional two generalized scenarios (budgeting and defending tort liability). These scenarios were important because the likelihood of any one strategies meeting a set of goals and objectives,

[3] A third category represented strategies which required further development because of their complexity or technological aspects. Although they were noted in the report, these strategies were not evaluated. An example of one that fell in this last category was the use of object-oriented data bases.

e.g., the accuracy attribute of quality, is dependent upon how the strategy applies to a specific use. Nominal ratings of how well each strategy met goals and objectives were combined across all scenarios and all raters. The resulting average strategies were then summed to present a rating of relative effectiveness.

This rating then was combined with relative changes in costs to present a cost-effectiveness rating. For crash data, the changes in costs derived from the cost model; for the roadway and crash injury data, the changes were attributed to estimated magnitude of effects on costs of operation and implementation.

Collecting and Managing Crash Data

In terms of reducing costs, the use of short and long forms for recording crash data was substantially superior to all other strategies. However, it suffered from a limited ability to meet goals and objectives, particularly those of quality and coverage. As a result the use of two different crash reports, was rated only moderately cost effective. Several strategies, including simplified form design, revised format and design, and pre-print diagrams showed relatively high cost-effectiveness. These strategies also were non-technological. The most high-rated technological strategy was the use of linked data sources. It also appeared as relatively high cost-effectiveness for collection and management of roadway-inventory and for crash-injury data. Table 59 lists those strategies which were considered relatively highly cost-effective and which either resulted in no change or reduced costs.

Table 59
Crash Data Collection and Management:
Relatively High Cost-Effectiveness

Pre-printed diagrams
Simplified form design
Revised/improved design and format
Linked sources of data
Computer directed interviews (AI)
Imaging systems
Use of non-worn crash investigators
Distributed data bases
Use of non-automated technology
Local data entry
Sworn crash investigators

Collecting and Managing Roadway Inventory Data

Unlike the evaluation of strategies for collecting and managing crash data, a cost model was not available for use in assessing effect on roadway inventory data. Instead, the research team rated the strategy on its potential to reduce costs of operation, along with its ease of implementation both from a perspective of costs

Table 60
Roadway Inventory Data Collection and Management:
Relatively High Cost-Effectiveness

Use of technologies for collecting and reporting data
Linked data sources
Comprehensive system requirements

and time (see table 60). A further division was between those strategies which could be implemented readily, and those which required longer term introduction or were technologically advanced. Because much of the data used for roadway inventory comes from measurements, technology becomes important. As a result, none of the strategies considered to have relatively high cost-effectiveness were also considered as being readily implemented. Only three strategies fell within the highest rating:

Collecting and Managing Emergency Medical Data

A large number of strategies for improving collection and management of emergency medical (crash injury) data fell into the category of "relatively high cost-effectiveness." As was done with the roadway inventory, the research team evaluated each strategy based on its contribution to the goals and objectives, then its relative effect on costs of operation, and on ease and cost of implementation. A major issue surrounding the current collection and management of this last source of safety data is that few central data bases exists, and there are no apparent standards for how and what data will be collected and managed.

One highly rated strategy fell into the category of being readily implemented, enhanced training designed to improve what data are collected and the format in which they are reported. An additional seven highly-rated strategies were considered either longer term for introduction or technologically advanced. All highly-rated strategies for collecting and managing emergency medical services data appear in table 61. Attention is called to the need for developing EMS/trauma databases. These generally are non-existence; their development would provide a substantial improvement to the quality and accessibility of these data, not only for the medical community but for the larger community of users.

Table 61
Emergency Medical Services Data Collection and Management: Relatively High Cost-Effectiveness

- Enhanced Training
- Sampling techniques
- Use of technologies
- EMS collect all injury data
- Scannable forms
- Linked sources of data
- EMS/trauma data base
- Imaging and OCR/OMR

Applying the Strategies

The assessment of strategies concluded with an overview which showed how strategies could be grouped as packages and introduced as a unit. The purpose of this analysis was to indicate where joint implementation would be necessary or valuable. An example of a necessary presence of one strategy to another is the need for a common EMS form in order for EMS personnel to collect all injury data (the latter considered to be a highly rated in terms of cost-effectiveness). Additionally, if EMS personnel do collect all injury data, then the implementation of several other strategies including non-automated technologies, extensive supervisory review, and enhanced training (as examples) would prove very useful to the strategy.

Packages of strategies then need to be evaluated based on field-level trials. For roadway inventory and medical data, a cost model should also be developed in order to allow future researchers to examine both the effectiveness of the strategy (or package) in terms of meeting goals and objectives, and the ability of the strategy to reduce costs associated with the process.

Effort needs to be directed toward technological applications, particularly where these applications can be done using current technologies and at a relatively low cost of implementation. Such a step is important to applying several of the high-rated strategies which can, in particular, help meet quality goals and objectives. This process compliments the work on emerging technologies in that it provides basic technologies which can be improved once proven advanced technologies become available.

In conclusion, the collection and management of highway safety data from its three sources appears to be relatively expensive and does not always meet levels of quality expected. A number of agencies using these safety data have introduced practices which are exemplary. They do attempt to reduce some of the problems either associated with costs or the quality of the data. Additionally, a number of other strategies can be implemented which will assist in the reduction of costs while enhancing quality. Many of the recommendations do not require technology to undertake, and where they do, the technologies are generally available and can be instituted readily. Some of the potential outcomes can result in substantial cost savings while having a positive effect on meeting goals and objectives related to the improvement of highway safety data.

Appendix A
Comments from Users Regarding Quality of Crash Data

Coverage	• Reporting of PDO crashes varies among jurisdictions; local thresholds vary from the State threshold. • Pedestrian and bicycle crashes are not well-reported. • Hit and run crashes are not reported or incorrectly reported. • Not entering data for PDO's, either because they are not reported, or they are reported on a short form which the State does not process for data entry, eliminates important samples for location-specific analyses. • Alcohol involvement for pedestrians often is ignored.
Accuracy	• Officers tend to be accurate on reporting the "when, here, and who" aspects of the crash, but less accurate with the "how and why" aspects. • Locations are mis-coded a significant proportion of the time with distance to or from the intersection often inaccurately measured. • The officer either does not have sufficiently good judgement for determining when a crash may be intersection related, or does not consider it important. • Seat-belt, alcohol, and speed prior to the crash upon is primarily based on occupant statements rather than officer investigation of evidence. • Changes in the codes on the format are not being reflected in the coding process; the use of "other" and "unknown" occur too frequently. • Elements related to the roadway and actions related roadway features, e.g. distinguishing between "cross-median" and "u-turn across median," are inaccurately reported. • The type of crash is often inaccurately coded. • Officers often mark the wrong scannable-report (OMR) "bubbles", resulting in values "ten" times what is correct, e.g., number of occupants.

Completeness	• The seating position of passengers is not provided for on the PCR. • Hit and run crashes are a significant proportion of total but result in many incomplete elements on the PCR. • Occupant attributes are not recorded if uninjured, but are desired for safety analyses. • System reports do not tabulate the frequency of occurrence of "other" and "unknown." • A significant percentage of PCR's have one or more elements not completed. • Alcohol involvement is sometimes implied in the narrative, but the area of the form for reporting alcohol involvement is not completed. • Lack of details on contributing circumstances for roadway elements significantly limits engineering analyses. • Up to 20 percent of the reports do not have any primary contributing circumstances indicated. • Helmet usage was not reported on up to two thirds of motorcycle crashes. • Passenger age is frequently missing or indicated as "unknown."
Precision	• Great precision is not needed for classifying injury severity for highway-engineering use, but it is needed for motor-vehicle administration uses. • Either the design of the PCR, or the reporting practices, do not allow analysts to identify initial engagement points of vehicles and objects. • Narratives and diagrams tend to be oriented to establishing fault, rather than recording facts.

Precision	•	Some collision factors and elements may not be precise enough resulting in a frequent use of "other."
	•	When changes are made of codes on the form, to accomplish greater precision, the new codes can confuse officers; increasing the number of codes, however, does not necessarily increase precision.
	•	Officers often provide insufficient information to locate crashes precisely - e.g., street name is often not recorded with sufficient precision to identify it.
	•	The typology used for pedestrian actions is not adequately related to the real world.
	•	Some codes for contributing circumstances are imprecise, e.g., "careless driving" or "speed too fast for conditions."
	•	Indicating the part of the body injured is not needed for highway-engineering safety analyses.
	•	Recording a complete sequence of events is important, especially when the PCR record that is created is based upon self-coded data on the PCR, rather than interpretation of a narrative and diagram.
	•	Injured parties in a school bus crash are not reported with sufficient precision to know if the injured parties were on the bus, getting on or off, or passing a bus on a bicycle.
	•	Imprecise descriptions of roadway elements or technical aspects, e.g. hydroplaning preclude use of PCR for critical studies.
Timeliness	•	Crash and citation data are available barely in time for preparation of the State Highway Safety Plan each year.
	•	The media wants very current data
	•	One State patrol indicated that a 2-month lag is okay for strategic safety analyses, but want immediate notification of crashes with severe injuries and current data for selective enforcement planning
	•	Safety researchers want no longer than a two-month lag on the availability of data from the system
	•	Users from law departments want current data
	•	Count data are often too late for the annual planning cycle
	•	Selective-enforcement planners in local law-enforcement agencies want current data
	•	Lack of timeliness in having automated files available is what drives many local agencies to create their own crash reporting systems.

Consistency	• Injury severity is not consistently defined across the State • Officers are not consistent when reporting location relative to road • Street classification is not reported consistently across the State • When reports are altered, incomplete retraining results in inconsistent reporting. • Alcohol involvement is not consistently reported between agencies • Officers are inconsistent in indicating direction of travel on their diagrams • Inconsistent reporting thresholds occur across States • Possible contributing circumstances are indicated in a higher percentage of the cases, with increasing severity of the crash • Officers indicating "impaired" as a possible contributing circumstance, often do not complete the "driver condition" area • Inconsistent use of street and highway names on reports results in as many as 20 percent of the crashes on State trunk highways not being location coded - a bias among non location-coded crashes seems to be toward over-representation of urban crashes

Appendix B
Scenarios to Be Used for Evaluating Strategies

1. Police Selective Traffic Enforcement

Overview

Probably the most common use of traffic crash information by law enforcement agencies is the planning for and evaluation of selective traffic law enforcement programs. Law enforcement agency managers have long recognized that it is not possible to apprehend all traffic law violators. Also, most of those managers recognize that the need for enforcement is greater in some geographic areas than in others. Selective traffic law enforcement reflects both of these issues. Defined most simply, selective traffic law enforcement is the focusing of traffic law enforcement efforts at those locations and times where the need is greatest, as determined by the incidence of traffic crashes. In addition, enforcement should also be directed toward traffic law violations that are over-represented in crashes.

Some law enforcement agencies will assign personnel to be at selected locations at specific times; others will make potential selective enforcement locations known to officers responsible for those areas and leave the amount of enforcement time to the officers' discretion. To evaluate the effectiveness of their programs, law enforcement agencies will look at crash incidence after the implementation of a selective enforcement program to learn if:

- Crashes have decreased in number at the selected locations.
- Crashes have decreased in number at the selected times.
- Specified violations are no longer over-represented in crashes that do occur.

Law enforcement agencies with greater analytical capabilities will also look at factors such as displacement of crashes both geographically and temporarily, and develop weighting schemes for measuring effectiveness by factoring crash severity into their analyses. These agencies will also be able to determine if reductions in numbers of crashes are statistically significant and to assure that reductions are not simple regression to the mean artifacts.

Detailed Description

Objectives of the analysis. The objectives of these analyses are to identify potential selective traffic law enforcement locations and to evaluate selective traffic law enforcement programs. In addition to identifying locations that are over-represented with crashes, the times when those crashes are occurring, and traffic law violations associated with the crashes should also be determined. Agencies identify locations with the greatest frequency of crashes, or adjust the "raw" numbers by traffic volume or crash-severity factors.

Intended end-user of the information. The intended end-users of this information are the law enforcement agency managers who have responsibility for assignment of personnel and for the evaluation of agency deterrence projects. Ultimately, results of selective traffic law enforcement programs are usually presented in reports to agency (or city, county or State) administrators, or presented in agency annual reports. The person who identifies the selective traffic law enforcement locations and evaluates such programs is usually a manager of an agency's traffic section, or a civilian employee of the agency's research or planning units.

Data sources to be used for this analysis. The primary data source will be the agency's traffic crash report file. If it is maintained as it should be, reports will be either filed directly by location, or cross-referenced by location through either card or computer files. If the agency wants to adjust their locations based on exposure, traffic volume information will need to be obtained from appropriate city, county, or State traffic engineering offices. Depending on the sophistication of the agency's data system, information on crash-related traffic law violations may be available in the electronic file, or the source documents (crash reports) will need to be manually reviewed for the information.

Analysis to be performed. This program calls for two analyses: the identification of selective traffic law enforcement locations (times and violations) and the assessment of the effectiveness of the selective traffic law enforcement efforts at each location in terms of reducing the number of crashes, and impact on crashes community or area wide. Site identification can be as simple as a crash count, or can be adjusted based on crash severity or exposure in terms of traffic volume. Analyses of program effectiveness will include tests for significant reductions in crashes at the specific locations, or cumulatively across all locations.

Information to be produced. The initial information to be produced is the listing of selective traffic law enforcement locations times and violations for dissemination to enforcement officers and their officers. Later information will be the basis of reports on the effectiveness of the program in terms of crash reduction.

2. Commercial Vehicle Inspection for Overweight and Defective Equipment

Overview

Almost any police traffic enforcement program based on analyses of crash data can be called "selective traffic law enforcement." Scenario #1, described above, describes how crash data are used for the implementation and monitoring of a "classic" selective enforcement program. This scenario describes a variation of a comprehensive selective traffic law enforcement program where the analysis of crash data and subsequent enforcement program focus on specific vehicle types rather than "moving" traffic law violations.

This scenario assumes that someone in a law enforcement agency (or even the community at large) has observed that large commercial vehicles appear to be involved in a disproportionate number of crashes in an area. Based on this observation, crash records would be analyzed and, presumably, it would be determined that, in the area in question, crash involvement of large commercial vehicles would be something on the order of more than three times their crash involvement rate citywide. As the analysis continues, it would also be noted that in most of the crashes, the commercial vehicle driver was indicated as being at fault, with the most common violations noted as either "speed too fast for conditions" or "disobeying a traffic control device."

Examination of the roadway inventory file or a visit to the area of the crashes would show that the direction of travel of most of the crash-involved commercial vehicles was a 4 percent downgrade. A site visit would also confirm that there were a number of freight terminals and/or gravel pits in the area where the crashes were occurring. The person doing the crash analysis should then take the findings of the large commercial vehicle crash over involvement to someone who was experienced with commercial vehicle operations and enforcement.

That person's review of the crash situations might indicate a supposition that many of the crash-involved vehicles were either overweight or had defective braking equipment. Either situation could result in the large commercial vehicles gaining too much speed on the downgrade to be able to stop or slow in time to avoid a collision either at an intersection or with vehicles operating a lower speeds (either as through traffic or entering/exiting the roadway).

The result of all these analyses would be the implementation of a commercial vehicle enforcement program upstream of the crash area focusing on both weighing vehicles and inspecting braking equipment. The results of the program should include:

- Decrease in the number of crashes involving large commercial vehicles.
- Overall decrease in the number of crashes in the area due to increased police presence.
- Specified commercial vehicle violations are no longer over-represented in crashes that do occur.

Law enforcement agencies with greater analytical capabilities will also look at factors such as displacement of crashes to learn if the commercial vehicle drivers are altering their routes to avoid the increased overweight/defective equipment enforcement efforts.

Detailed Description

Objectives of the analysis. The objectives of this analysis are to confirm that a large commercial vehicle crash problem does exist, to determine reasons for that large commercial vehicle crash over involvement, and to develop an appropriate countermeasures program. This program will be both location and time specific in most instances.

Intended end-user of the information. The eventual end-users of this information are the law enforcement agency managers who have responsibility for assignment of personnel and for the operational law enforcement personnel who will have the responsibility for carrying out the countermeasures program(s). Information from the countermeasures program will subsequently be used in an evaluation of projects. Ultimately, results of this type of selective traffic law enforcement programs would be presented in reports to agency (or city, county or state) administrators, or contained in agency annual reports. In a special situation such as the one described here, data access could be made by differing individuals. The person who usually identifies high crash incidence locations and evaluates such programs is usually a manager of an agency's traffic section, or a civilian employee of the agency's research or planning units.

Data sources to be used for this analysis. As with a standard selective traffic law enforcement program, the primary data source will be the agency's traffic crash report file. If it is maintained as it should be, reports will be either filed directly by location, or cross-referenced by location through either card or computer files. The type of vehicles involved in a crash (passenger car, large truck, motorcycle) is usually coded on the crash report, but may not be entered into a city's crash information data base. If the information is not is the data base, a manual review of all crash reports for the involved area will need to be made. Depending on the sophistication of the agency's data system, information on crash-related traffic law violations may be available in the electronic file, otherwise the source documents (crash reports) will again need to be manually reviewed for the information. In addition, physical information on the crash location will be needed from either the roadway inventory of from a visit to the site to determine if characteristics such as grade are relevant factors.

Analysis to be performed. This program calls for several analyses: confirmation that large commercial vehicles are over involved in crashes, determination of potential reasons for that over involvement (e.g. traffic law violations such as "disobeying traffic control devices" or "speed too fast for conditions"), examination of potential environmental factors (e.g., road grade) , and a conclusion as to other factors that are contributing to the crashes (e.g., overweight vehicles or defective brakes). Analyses of program effectiveness will include tests for significant reductions in crashes at the specific locations, or cumulatively across all locations.

Information to be produced. The initial information to be produced is the listing of commercial vehicle-involved crashes are the suspect site. A detailed examination of those crash reports will provide additional information as to probable crash causative factors and candidate enforcement countermeasures. Later information will result in reports on the effectiveness of the program in terms of reducing the over involvement of large commercial vehicles in crashes at the specified location.

3. High-Hazard Location Identification

Overview

Engineering (to a greater extent) and law enforcement agencies seek to identify the most hazardous locations on the highway system. This is done to make the most cost-effective use of the limited resources available for highway safety improvements. Hazard is defined in different ways by different agencies. It ranges from frequency of crashes to a sophisticated measure of excessively high crash experience relative to similar types of roads.

The analysis relies heavily upon data from police crash report (PCR) files and roadway inventory files. Important variables include location, severity, and descriptors of the event(s). A "site" may defined with respect to an intersection, segment, corridor, or area. Furthermore, the hazard sought may become focused, such as on wet weather, speed-involved, or night crashes. The information is often initially presented as a prioritized listing of sites as a "high-hazard location" (HHL) report. This may be followed by development of collision diagrams and drawings of site physical characteristics. Countermeasures are selected based upon analysis of information on crash and site attributes. Cost-benefit analyses are sometimes performed using crash information to project reductions in costs as a result of improvements.

The users of this information usually are mid-level engineers who will place highly summarized results into materials for high-level decision-makers who will program funds for specific improvements. Law enforcement personnel using this information are usually planning officers who provide the results to mid-level command personnel for final selection and assignment of personnel for enforcement.

Detailed Description

Objectives of the analysis. This is a basic traffic engineering use of traffic safety-data and its objective is, stated simply, to identify and prioritize locations for safety improvement actions. The analysis will specifically focus on those locations where engineering improvements should help to ameliorate the crash problem.

Intended end-user of the information. Likely users are high-level engineering management who will make funding decisions, and mid-level law enforcement command personnel who will make staff allocation decisions.

Persons who will access the data. Data processing is usually done by personnel in a management information services (MIS) department. Rarely will end users have direct access to the data. The initial HHL list is usually a pre-defined report which is regularly produced by the MIS agency, either at a designated time, or upon request. When access is available to individuals outside the MIS agency, the user is often either a mid-level engineer/analyst or a law-enforcement planning staff member who does the analysis and provides summary information to the end user.

Data sources to be used for this analysis. Sources to be used for these analyses are the central crash information repository for the entities in question (e.g., State files for State routes, county files for county routes, etc. depending on how the various records repositories record location information. Roadway inventory files are an additional data source frequently used for this type of analysis. Often the engineers refer to individual crash reports using the narrative and diagram to help in identifying counter measures to be employed.

Analysis to be performed. Analyses used to identify HHL's can vary considerably depending upon the definition of "hazard" and the statistical basis desired. Frequency of crashes is often used. The use of frequency can be further enhanced by using the cost of crashes or some other device for weighting according to severity. The definition can be extended to account for severity and exposure (using a rate). Statistical tests can be added to determine if a site exceeds the average by a sufficient amount to be likely due to more than chance. The group upon which the average is based can be defined to create a universe of peers, having common characteristics. The definition of a site will vary depending upon the focus of the analysis. There can be an interest in specific points in the system such as intersections, segments between intersections, lengths of routes (corridors), and whole areas (counties). The analyses involving a length of highway can involve application of specific criteria for determining the termini of the section. There is a trend among State and local agencies to use the more sophisticated approaches for HHL identification.

Information to be produced. Information for the HHL sites may be based upon more than one year of data. HHL listings may include the following information for each site:

1. Location and identification.
2. Averages for the measure(s) of hazard.
3. Peer group assignment.
4. Critical rate, for the peer group, of the measure(s) of hazard.
5. Ranking for the site.
6. Summary descriptive statistics for the site.
7. Listings of key variables for individual crashes at the site (usually sufficient to allow manual preparation of a collision diagram).
8. Listing of key roadway attributes for the site (including traffic volumes).

4. Use of Crash Data by a Motor Vehicle Administrator.

Overview

The agency responsible for driver licensing in a State typically develops a data acquisition and processing activity to identify deviant drivers. This is usually followed up by a sanction program intended to improve driver performance. This particular scenario involves the use of such information to evaluate the effectiveness of alternative sanctions.

The computer record is centered on the driver. Information attached to each driver record includes the incidence of traffic violations, crashes (usually only when this driver was at fault), and a listing of any sanctions imposed. Crash data are often processed by another agency in the State, and entries are made in the driver record using the driver license number as a linking variable.

An important potential use of such driver-centered data is in the evaluation of counter measures. In this scenario, several experimental countermeasures are postulated, and the data used to determine which countermeasure(s) should be selected for more general application. The ultimate users of the data and the analyses are the administrators of the driver licensing authority. They depend heavily on complete and accurate reporting from the field.

This scenario will follow through the activities associated with such a countermeasure evaluation. There are many possible driver improvement areas that could be addressed; as an example we will consider a program to reduce future crashes and violations among drivers who have already had two crashes over the past year. The measure of effectiveness in this program will be a comparison of crashes in the year after the sanction to those in the year before, and a relative assessment among groups for the after period.

The processes described in this scenario are adapted from many programs conducted by Ray Peck and his predecessor, Ron Coppin, at the California Department of Motor Vehicles. California often had a numerical advantage in conducting such experiments. But similar experiments have been and can be done in smaller states with equally useful results.

Detailed Description

Objectives of this analysis. The objectives of this analysis are to compare the driving behavior of four groups of motorists who been involved and at fault in two crashes over a 1-year period. Each group (other than the control) will be subjected to a particular sanction:

- Group 1 will receive a friendly letter stating that the motor vehicle administrator hopes that driving performance will improve.

- Group 2 will be receive a more threatening warning letter suggesting possible future sanctions if performance does not improve.

- Group 3 will be required to attend a 2-hour seminar at a motor vehicle department facility where they will given pointers on how to avoid crashes along with discussing their driving records, the dangers of speeding, and the consequences if they are arrested again.

- Group 4 which is a control group where no actions are taken.

Driver records will be compared for a specified period after the countermeasure is imposed, and the cost or benefit of any crash reduction determined will be compared with the cost of executing the countermeasure the program. Ultimately one or more countermeasures may be chosen for widespread application. A properly conducted experiment will lead to the greatest value for the dollar.

Intended end-user of the information. The intended end-user of this information is the State's motor vehicle administrator, who is responsible for driver improvement activities. This administrator wishes to spend his budget in the most effective way, and is looking for the most cost effective choice. The persons who will access the data could be on the research staff of the motor vehicle administrator or could be a research contractor (e.g., in the psychology department at a State university). If the person is a motor vehicle department staff member, she should have some statistical training and should be knowledgeable about the accuracy and completeness of the data being analyzed.

Data sources to be used for this analysis. The major data source for this effort is the driver record file. This is a driver-centered file, with personal information as well as entries for each pertinent action that occurs. This would include the driver's name, address, age, and sex. Sometimes the file will include a notation if the driver has had a formal driver education program. The license category (commercial or private, motorcycle endorsement, etc.) will be recorded. Finally there will be entries for convictions for traffic violations and some crash involvements. In many States crash involvement will be entered in the driver record only when that driver is at fault.

The driver record file normally contains only a brief abstract of the data available in the crash file. The analyst may be interested in more detail regarding the type of crash, the injuries, or the kinds of vehicles involved. The driver license number may be used as a linking variable to the crash file. Some States also have a citation file that might be useful in this analysis, since it could contain citations that never proceeded to convictions.

Analysis to be performed. The analyses produce sets of crash rates, involvements divided by time, for each of the four subject groups. Before any analysis begins, however, there is the important step of experimental design.

In order to have a viable experiment for this project, it is necessary to identify the subjects in the driver record, and to divide them into three randomly selected experimental groups and one control group. Crash information has been collected by police agencies and forwarded to the motor vehicle department where it is entered into the individual driver records. A first step in this experiment is to search the driver records to identify those with two and only two crashes within the past 12 months.

The sample size required depends on the crash frequency in the after period, on the difference in the crash rate among the three groups, and on the confidence interval desired or acceptable to the researcher. There will surely be a regression to the mean effect, so that the total number of crashes for these groups will be smaller in the post period. A 2-year driver file might be analyzed at the outset to get an estimate of the number of crashes in a second year for drivers who had two in a first year. If the proportion of drivers with crashes in the post period were 20 percent, the designer might anticipate that this would approximate the value for the new experiment. A sample size of the order of 10,000 for each treatment group would lead to a standard error of about 1 percent. Differences among the groups performance would have to be larger than this 1 percent to permit a conclusion that one was better than the other.

Information to be produced. The data in the driver file includes a continuing record of violations and (usually) at-fault crashes, each entry accompanied by a date, and perhaps by details of the crash (injury or no injury, drinking or no drinking, etc.) The basic measure of performance is determined by counting crashes and violations for each group over a period of 1 year following the sanction, and comparing the frequency among the three groups.

5. Use of Crash Data by a Vehicle Manufacturer to Evaluate the Effectiveness of Anti-Lock Brakes

Overview

All vehicle manufacturers are concerned about the safety impact of design changes and devices that affect the risk to occupants riding in their cars and trucks. Safety is becoming an increasingly important factor in the marketplace today, and manufacturers recognize that it is desirable for changes in design or features to lead to fewer crashes and injuries.

Manufacturers consider safety in the design stage, simulating crashes by computer, and by conducting sled and instrumented crash tests. In addition, many manufacturers sponsor indepth crash investigation programs designed to look in detail at the effectiveness of components such as anti-lock brakes to reduce the incident of loss of control and collisions resulting from loss of control because of improper braking.

Ultimately it would be desirable to confirm the results of their work by analyzing the mass crash data representing the real world. Police crash data, containing information about injuries and crash types, represent a large data source for such an investigation. Most manufacturers would presumably be interested in determining whether predicted injury reductions are occurring.

In this scenario, the manufacturer is considering adding anti-lock braking systems (ABS) to a vehicle class and including them as standard equipment. They are concerned that if the brakes are not effective, the added cost may not be accepted by the public. Furthermore, by providing anti-lock brakes as standard equipment, and implying their safety, the manufacturer may be liable for driver errors induced by unwarranted assumptions about the safety of using such brakes.

The question at hand is whether the predicted changes can be detected in the police crash data, and how large a change has occurred.

Detailed Description

Specific objectives of this analysis. The objectives of this analysis are to acquire a set of crash data from one or more States, to identify those crashes in which this new vehicle is involved, and to compare the types of crashes between vehicles equipped and not equipped with anti-lock brakes. The manufacturers have estimated that anti-lock brakes will reduce crashes evolving from loss of vehicle control by 30 percent.

Insurance companies also have an interest in the analysis because they are offering lower collision insurance rates for vehicles with anti-lock brakes. If the brakes are not effective, insurers will have to revise their rates or risk losses because of inadequate premiums.

Intended end-user of the information. The intended end users of this information are the vehicle manufacturing company's management and the safety design personnel. The data analysis will be done by analysts in the manufacturer's safety office. The introduction of ABS has been widely heralded. However, the manufacturers want to change their marketing strategy promptly if the changes in crashes do not occur. The manufacturers will also want to identify how they will continue to use the brakes on other models and vehicle lines.

Data sources to be used for this analysis. The desired data files are State crash data collections from several large States. A large number (100,000) of these vehicles have been sold. With 225 million persons in the U.S. the sum of California (30 million), New York (18 million), Texas (17 million) and Florida (13 million) account for about one-third of the people in the country. If these vehicles were distributed in the country in proportion to population, about one-third of the 100,000 vehicles produced with ABS would be in one of these four States. Of these 33,000 vehicles about 12 percent (3,960) would be expected to be involved in a crash which resulted from loss of control.

If all of the data from each State were used, most statistical tests are inappropriate for publication, although they may be useful to the analyst. More important however, is the ability to identify those crashes which resulted from loss of control. This means that crash investigators must include elements on he crash reports which allow the researchers to identify crashes in which loss of control was a contributing factor.

Analysis to be performed. The analysts must first identify the vehicle populations of interest. There are two groups: (1) the new cars with the anti-lock braking, and (2) vehicles of the same make and model without the brakes. The most effective way to identify these with the required precision is by using the Vehicle Identification Numbers (VIN) assigned by the manufacturer. In addition, the data must be further subset to provide a file that fits the criteria of "loss of control."

Since each State uses different elements for identifying harmful events, manners of collision, and contributing factors, the analysis files will be developed State by State. For each State the data will be tabulated in the following categories:

- Two vehicle groups: new cars, last year's cars.
- First harmful event is loss of control.

Other elements of importance may be weather, road surface, injury severity, and other contributing factors. Amount of damage to the vehicle would also be of value to insurance companies, but few States gather data which allow an assessment based on severity of damage.

Information to be produced. The first assessment of the data will be a two-way table comparing number of crashes in which loss of control was the first harmful event for vehicles with and without anti-lock brakes. If sufficient data are available, the analysts may examine subsets of the same table based on severity of injury and on numbers of vehicles involved. It is entirely possible that no difference may exist in loss of control collisions as a whole, but those vehicles with anti-lock brakes were involved in less serious crashes.

If no differences are found, the analyst may have to search for reasons. There may be confounding factors on the assumption that the brakes worked as manufactured. For example, drivers with anti-lock brakes may drive more aggressively trusting the anti-lock brakes to help them avoid a crash (although this "risk compensation" theory was largely discredited in studies involving seat belt use). Drivers may not know that they can steer their vehicle, even with the brake pedal depressed. The manufacturers and insurers may find that they need to conduct education and training.

6. Measuring Changes in Crashes Related to an Increase in the Speed Limit

Overview

Although the law permitting States to increase their maximum speed limit on certain controlled access highways to 104.6 km/h has been in place for several years, there are States that have not chosen to do so. In addition, other States have not raised the limit on all eligible roadways, and changing demographics can affect the eligibility of roadways to have a 104.6-km/h maximum speed limit.

While there have been studies on the safety impact of increasing speed limits from 88.5 km/h to 104.6 km/h, the results of these studies have not been consistent. While many roadways showed increased fatality and/or other crash-related rates after their maximum speed limit was raised, this has not always been the case. It is therefore possible that a State without full implementation of the 104.6-km/h limit on eligible roadways would want to do their own study before raising the limit on any (or all) roadways.

For this scenario, it will be assumed that a State increased the speed limit on some, but not all eligible roadways and now wants to do a study on the effects of the increased limit on the affected roadways. Since not all eligible roadways had their speed limit raised, it will be assumed that sections of roadway that have retained their 88.5-km/h speed limit can be found that are otherwise comparable to those that were raised to 104.6 km/h. This will enable researchers to crease a simple four cell assessment grid describing a pre- post-implementation condition with a control.

The treatment can be a before/after comparison with a control area. The measure used can be overall crash rate, fatality rate, change in crash severity or other, similar measure. If the data analyst has suitable capabilities and sufficient historical data are available, more sophisticated time series analyses can be conducted.

Detailed Description

Objectives of this analysis. Objectives of this analysis are to conduct a before and after or time series study of the crash or crash severity rate on a roadway (the experimental section) on which the speed limit was increased from 88.5 to 104.6 km/h while also evaluating those rates on a similar roadway (the control section) where the speed limit remained at 88.5 km/h.

Crash records will be compared for two specified time periods before and after the speed limit was increased on the experimental roadway. If possible, the data collection periods would be at least the year prior to the speed limit change and a 1-year period beginning at least 6 months after the speed limit changed. (The gap between the data collection periods will not affect data validity if a full 12 months are used for the post-change data collection, this will eliminate any seasonal biases. The interim period will allow motorists to fully adjust to the new speed limit prior to the post change data collection interval. It will be desirable to look as several years' pre-speed limit change information so that trends in the crash or crash-severity rates can be noted.

Intended end-user of the information. The intended end-users of this information are the state's highway department and legislature. The latter may want to use the data to consider legislation to permit 104.6-km/h speeds anywhere in the State. The highway department will use the information to help determine the roadways on which roadways the limit should be increased, if increases are authorized by the legislature. The persons who will access the data could be on the research staff of the highway department, legislative analysts, or it could be a research contractor (e.g., in the psychology department at a State university). If the persons are highway department staff members or legislative analysts, they should have some statistical training and should be knowledgeable about the accuracy and completeness of the data being analyzed.

Data sources to be used for this analysis. The major data source for this effort is the State crash record for the roadway sections being studied. This is a crash-centered file, with information on crash severity, types of vehicles involved, maneuvers, harmful events, and information about the roadway and environment. There are usually few personal identifiers in this file, but demographic information on vehicle occupants and injury severity is usually listed. Finally there will sometimes be entries for violations of traffic laws or other notations of possible driver error. Some States even include estimates of the pre-crash speeds of vehicles, but this is generally not considered to be reliable information.

Assuming that crash data collection and recording has been consistent in both the experimental and control areas for the entire study period, there should be little difficulty with evaluating the data needed. Other than specifying the road segments needed, all data should be in the state file in a form that is usable.

Analysis to be performed. The before and after analyses produce sets of crash rates, involvements divided by time, for each of the four subject groups. A time series analysis will provide information as to the significance of any change in crash or crash-severity rates in the experimental area, determine if the changes are different in the control area, determine if the change is different from what was predicted by historical trends.

The crashes subjected to analysis will depend upon data available in the State's crash file. The simplest analysis is to look at the total crash rates for the both areas in both time periods. However, since it can be hypothesized that the higher speed limit will result in more severe crashes (as opposed to simply more crashes), assigning severity scores to all crashes and evaluating the severity rate may be a more fruitful approach. Since it can be argued that higher speeds can make any crash more severe, limiting the study to only certain crash types is probably not appropriate.

Information to be produced. The user information produced from this study will be a report showing the change in crash or crash severity rates for the experimental, the significance of that change, if that change was different from the change noticed on the control roadway, and if the change differed from that forecast from historical trends. The specific nature of what is produced will depend upon the sophistication of the analytical test(s) used.

7. Use of Crash Data by a State Legislative Committee to Support a Budget Allocation

Overview

A State legislative budget committee is considering a request from an association of county sheriffs for financial support of their crash investigation activities on the State and Federal highway system. The sheriffs argue that, since they are supported solely by county taxes, the state should provide some form of reimbursement for their investigative work on State routes. The committee has agreed in principle, but would like to seek data to support a particular funding distribution.

The State's crash data file contains information regarding crash severity (injury, non-injury, fatal), crash types (multi-vehicle or single vehicle), road class, and the type of investigating officer (city, county, State).

The committee has requested that a State police analyst investigate and write a report detailing the costs (for the past year) assignable to county officers investigating crashes on State and Federal routes.

Detailed Description

Objectives of this analysis. The objectives of this analysis are to develop a cost model for the crash investigation portion of police activities, estimating the amounts in a two-way tabulation of investigator type and road class. Crash data will be used to estimate the numbers of crashes investigated; external data will be used to estimate the per crash investigation costs.

Intended end-user of the information. The intended end-user of this information is the legislative budget committee that wishes to develop a bill for fair restitution of the counties for their efforts. The bill will ultimately be voted on by the legislature. The investigator using the information is assumed to be a staff analyst with the State Police. This person should have easy access to the States crash data files, and also to records of police salary and other operating costs.

Data sources to be used for this analysis. The primary data source is the State crash file. If it is up to date it should contain records of all the police-reported crashes over the time period of interest. In addition to the computer file, data on police salaries, overhead, and overtime pay rates are needed to complete this analysis. There may be no central location for these data, but they should be available from public records.

Analysis to be performed. One variable in the crash data file reports the type of agency that conducted the investigation. Although there may be more detail (such as police post number, county or city name), the data will be recoded into three levels: State police, county sheriffs, and city police. A second variable of interest it the road class. This will be recoded into just three levels: State and Federal highways combined, county roads, and city streets.

To develop a cost figure, it is necessary to estimate the number of staff hours necessary for each investigation, multiply this by the appropriate hourly investigator cost, add the overhead, and sum the tabular information for each cell.

Information to be produced. The information necessary to answer the sheriff's question is the estimated cost to the counties of investigating crashes that occurred on States routes. The other cells are of interest, too, since the State may sometimes investigate crashes that occur in roads under county jurisdiction.

8. Use of Crash Data by the Courts for Determining Law Violations and Civil Liability

Overview

Information about traffic crashes is frequently used in the courts in traffic, criminal and civil cases. While the adjudication of traffic law violation cases resulting from crashes is the most common of these, the criminal and civil cases can be much more serious. While the court's role of special, case-specific reports prepared as part of a crash reconstruction has been examined, no study is known which has examined the role of data found on both the individual crash report and in crash data files.

Many law enforcement agencies have a policy that generally requires officers to issue traffic law violation citations at almost all crashes. These citations are usually issued to the driver that the investigating officer considers to be "at fault" in the collision, and the citation is usually for a traffic violation such as disobeying a traffic control device, driving too fast, or failing to yield.

When these cases come to court, if the cited driver chooses to contest the citation, the citing officer will often use the crash report as a basis for his/her testimony regarding the violation. The admissibility of the use of the crash report by the officer depends on court practice and how the report is prepared. For example, a report prepared at the crash scene by the officer can probably be used by the officer as "notes prepared contemporaneously" by the officer. On the other hand, if the officer made only field notes at the scene and later prepared a final report, the court may prohibit use of the crash report on the witness stand, and the officer would have use only of the field notes.

In criminal cases (e.g., a crash situation where a driver is charged with a traffic-related felony such as aggravated DUI where an impaired driver seriously injured another person), the use of the original crash report can be even more critical. Again, if it was prepared contemporaneously at the scene, a testifying officer can use the report on the stand as a tool for refreshing his/her memory. The accuracy of the report, and statements copied at the scene can make a critical difference as to the successful prosecution of the case. While additional testimony from crash reconstructionist will probably also be made, that testimony will be based on the results of the reconstruction. In those cases, the expert will probably not be permitted to refer to any reports. However, the expert may well be presented with a copy of the original crash report while on the witness stand and asked to explain various aspects of that report.

The scenario proposed here, however, is based on potential civil litigation. As for criminal cases, initial investigating officers will sometimes use their original crash report in testimony. Experts many also be asked to comment on the original crash report. This scenario is not concerned with the specific report of the crash that led to the litigation. Rather, the court will be asked to look at aggregate crash data to determine if a highway department permitted an unsafe condition to exist (and persist) which was a proximal cause to the injuries suffered by the plaintiff in the case.

The aggregate data to be examined could come from an already existing data source such as the State central repository, the local police department, or State or local highway departments. Rather than relying on already existing data bases, however, it is likely that the attorneys for the plaintiff would subpoena all original crash reports for the location in question from either the State or the community in question, analyze those reports, and then construct their own data base or other aggregation of the information sought. In most cases, this information would be other crashes with characteristics similar to the case in question.

For this scenario, we will assume a serious injury crash occurred where a motorist failed to negotiate a curve, left the roadway and struck a large tree. The issues in question could be warning signs (including advisory speeds) ahead of the curve, lighting in the area, and the geometry and surface of the roadway itself.

Detailed Description

Objectives of this analysis. The objective of the analysis will be to learn the incidence of other crashes involving drivers who failed to negotiate the curve in question and other particulars of those crashes. The goal of the attorneys for the plaintiff will be to show that there had been a sufficient number of crashes prior to the crash in question to alert the appropriate authorities that an unsafe condition existed and that they did nothing the eliminate the condition. The attorneys for the plaintiff will also want to look at roadway inventories to lean the specifics of the location, and also review various engineering warrants to learn what is recommended in terms of signing, lighting and geometrics for such a location.

The key evidence, however, will be crash data showing that crashes were occurring at that location and that most were similar in nature and that the curve in the road was either inherently unsafe or there was insufficient warning prior to it.

Intended end-user of the information. The initial end user of the data will be the attorneys for the plaintiff in the case. They will have commissioned the analysis to bolster their case. The ultimate end-user of the information will be the court. Whether the case is being decided by a judge or jury, it will be the trier of fact that makes the decision as to the adequacy of the data and whether it shows what it is intended to show. The person who identifies the selective traffic law enforcement locations and evaluates such programs is usually a manager of an agency's traffic section, or a civilian employee of the agency's research or planning units.

Data sources to be used for this analysis. The primary data source will be the agency's traffic crash report file. If it is maintained as it should be, reports will be either filed directly by location, or cross-referenced by location through either card or computer files. If the agency wants to adjust their locations based on exposure, traffic volume information will need to be obtained from appropriate city, county, or State traffic engineering offices. Depending on the sophistication of the agency's data system, information on crash-related traffic law violations may be available in the electronic file, or the source documents (crash reports) will need to be manually reviewed for the information.

Analysis to be performed. This program calls for two analyses: the identification of selective traffic law enforcement locations (times and violations) and the assessment of the effectiveness of the selective traffic law enforcement efforts at each location in terms of reducing the number of crashes, and impact on crashes community or area wide. Site identification can be as simple as a crash count, or can be adjusted based on crash severity or exposure in terms of traffic volume. Analyses of program effectiveness will include tests for significant reductions in crashes at the specific locations, or cumulatively across all locations.

Information to be produced. The initial information to be produced is the listing of selective traffic law enforcement locations times and violations for dissemination to enforcement officers and their officers. Later information will be the basis of reports on the effectiveness of the program in terms of crash reduction.

References

"Accident-Analysis Software Needs Identified." The Urban Transportation Monitor. May 25, 1990.

American Association of Motor Vehicle Administrators. Draft, ANSI D20.1 Data Elements Dictionary. Arlington, VA: 1989.

An Evaluation of Traffic Accident Records Systems in Texas and Other States. University of Texas at Austin, Policy Research Project Report no. 65. Austin, TX: 1984.

Baker, Susan P. "Medical Data and Injuries." Status Report, Insurance Institute for Highway Safety 18:12, 5-6 Aug. 16, 1983 .

Blincoe, Lawrence J. and Faigin, Barbara. The Economics of Motor Vehicle Crashes. 1990. National Highway Traffic Safety Administration. Washington, DC: September 1992.

Bozak, David J. Accident Data Collection-Duplication, One State Experience. Portland, OR: July 16, 1991

Brown, David, *et al.* State of New Jersey Traffic Records Assessment. National Highway Traffic Safety Administration Technical Assessment Team. November 16-20, 1992.

Chipman, M.L. Motor Vehicle Accident Fatality Statistics: An Investigation of Reliability 74:6 381-385 November - December 1983.

Council, Forrest M. and Paniati, Jeffrey F. "The Highway Safety Information System." Public Roads 54:3, 234-240 December 1990.

Federal Highway Administration. "Recording and Reporting of Accidents." Motor Carrier Safety Regulations. Federal Register 37:207 22868-22870 October 26, 1972.

Federal Highway Administration. Fatal and Injury Accident Rates on Public Roads in the United States. Office of Highway Safety and Highway Information Management. Washington, DC: 1990.

Fife, Daniel and Barancik, Jerome F. "Discrepancies in Vehicular Crash Injury Reporting: Northeastern Ohio Trauma Study IV." Accident Analysis and Prevention 17:2 147-154 April 1985.

Gerrard, J. Peter, Mosher, and Walter W. Jr. Analysis of Reliability of Accident Information Obtained From Off-The-Scene Sources, pt.1. School of Engineering and Applied Sciences Reports Group 68-23. Los Angeles, CA: December 1960.

Graves, Richard A. Development of A Traffic Accident Analysis System. U.S. Dept. of Commerce. Springfield, VA: NTIS, February 1972.

Hall, J.W. "Deficiencies in Accident Record Systems." Transportation Planning and Technology 9 199-208, 1984.

Hargroves, B.T. and Hargroves, J.M. Accuracy of Virginia Accident Data. Virginia Highway and Transportation Research Council VHTRC-82-R13. Charlottesville, VA: September 1981.

Harris, S. "The Real Number of Road Traffic Accident Casualties in The Netherlands: A Year-Long Survey." Accident Analysis and Prevention 22:4, 371-178, August 1990.

Hauer, E. and Hakkert, A.S. "Extent and Some Implications of Incomplete Accident Reporting." Transportation Research Record 1185 1-10, 1988.

Hoetner, Jerald J. Police, Fire, and Refuse Collection Expenditures. International Association of City Management. 1989.

Howard, B. V. ,Young, M. F., and Ellis, J. P. Appraisal Of The Existing Traffic Accident Data Collection and Recording System. Department of Transport Office of Road Safety. Melbourne, South Australia: 1979.

Hughes, Warren E. *et al.* New and Emerging Technologies for Improving Accident Data Collection. Federal Highway Administration, Report No. RD-92-097. March 1993.

Hutchinson, T. P. Road Accident Statistics. Rumsby Scientific Publishing. Adelaide, Australia: 1987.

Hutchinson, J.W. and Kennedy, T.W. "Use Of Accident Records In Highway Research." Highway Research News 13, 1-8, June 1964.

Ibrahim, K. and Silcock, D.T. "The Accuracy of Accident Data." Traffic Engineering and Control 492-497, September 1992.

Kelsh, William E., Heitzler, Carter P., and Rauth, Susan G. Cost Analysis of Virginia System for Processing Accident Data. Virginia Highway & Transportation Research Council. Charlottesville, VA:1984.

Law Enforcement Management and Administrative Statistics. US Department of Justice, 1990.

Lucke, Roy E. and Stenzel, William W. Final Evaluation of Phase 1 of the Traffic Safety Law Enforcement and Disposition System of New York State. Evanston, IL: September, 1981.

Lund, Adrian K., Thum, Denise, and Preusser, Carol W. Completeness of Driver Records. Insurance Institute for Highway Safety. Arlington, VA: February 1991.

Massie, Dawn L. and Kenneth L. Campbell. CARDfile Evaluation. University of Michigan. Transportation Research Institute. Ann Arbor, MI: March 1990.

Miller, Ted., *et al.* The Costs of Highway Crashes. The Urban Institute. Federal Highway Administration. Springfield, VA: NTIS, October 1991.

Mounce, Nancy Hatfield and DeLucia, Barbara Hilger. Analysis of Accident Data Quality. Texas Transportation Institute. College Station, TX: August 1991.

National Highway Traffic Safety Administration. Report from the Conference on the Collection and Analysis of State Highway Safety Data. Conference on the Collection and Analysis of State Highway Safety Data. Washington, DC:1990.

National Highway Traffic Safety Administration. The Economic Cost to Society of Motor Vehicle Accidents. Washington, DC:1983.

National Highway Traffic Safety Administration. Crashworthiness Data System, Data Collection, Coding, and Editing Manual. National Accident Sampling System. US Department Center for Transportation Information. Cambridge, MA: January 1992.

National Highway Traffic Safety Administration. A Decade of Progress: Fatal Accident Reporting System. Washington, DC:1991.

National Highway Traffic Safety Administration. A Design Manual for State Traffic Records System, 2 vols. Washington, DC:1973.

National Highway Traffic Safety Administration. Problem Identification Manual for Traffic Safety Programs. National Traffic Safety Programs, Office of State Program Assistance. Washington, DC:1976.

National Highway Traffic Safety Administration. National Summary Report on the Findings and Recommendations of the Accident Data Improvement Plan. Office of State Program Assistance. Washington, DC:1981.

National Highway Traffic Safety Administration. Fatal Accident Reporting System, 1990. Washington, DC:1992.

National Safety Council. Accident Facts. Chicago, IL:1991.

National Highway Traffic Safety Administration. Coding and Validation Manual. Fatal Accident Reporting System. Washington, DC: 1992.

National Highway Traffic Safety Administration. Study Report of Methods to Improve the Application of State Traffic Records Systems. Transportation Research Board, National Research Council. Washington, DC:1987.

National Highway Traffic Safety Administration. Accident Investigation and Reporting: Highway Safety Program Manual. Washington, DC: September 1972.

National Highway Traffic Safety Administration. A Review of Information on Police-Reported Traffic Crashes in the United States. General Estimates System (National Accident Sampling System). Washington, DC: March 1990.

National Safety Council. Critical Automated Data Reporting Elements For Highway Safety Analysis - CADRE. NSC CADRE Task Force. May 31, 1991.

National Highway Traffic Safety Administration. A Review of Information: Fatal Traffic Crashes in the United States. General Estimates System (National Accident Sampling System). Washington, DC: March 1990.

National Governors' Association. NGA Motor Carrier Accident Reporting Project: Data Dictionary. 1989.

O'Day, James. Accident Data Quality, National Cooperative Highway Research Program Synthesis of Highway Practice.. Transportation Research Board, National Research Council. Washington, DC:1993.

Orsay, Elizabeth M., Lucke, *et al.* "The Impaired Driver: Hospital and Police Detection of Alcohol and Other Drugs of Abuse in Motor Vehicle Crashes." Presentation at the Society for Academic Emergency Medicine Annual Meeting. San Francisco, May 1993.

Paniati, Jeffrey F. and Council, Forrest M. "The Highway Safety Information System: A New Tool for Safety Analysis." Compendium of Technical Papers. Institute of Transportation Engineers. Washington, DC:1991.

Paniati, Jeffrey F., and Council Forrest M. "The Highway Safety Information System: Applications and Future Directions." Public Roads 54:4, 271-278 March 1991.

Pfefer, Ronald C. Information Needs and Data Relationships for an Illinois CCSRS: Technical Memorandum Number 4. "Feasibility Study of a Comprehensive Computerized Safety Recordkeeping System (CCSRS)." Illinois Department of Transportation and Deloitte, Haskins, and Sells. February 1988.

Raub, Richard A. and Ferguson, Alex. Time Spent by State Police at the Scene of a Traffic Accident. Illinois Department of State Police. Springfield, IL:1983.

Raub, Richard A., Pfefer, Ronald C., and Lucke, Roy E. Task Report: Issues and Noteworthy Practices Related to Highway Safety-Data Collection and Management. Northwestern University Traffic Institute. Federal Highway Administration contract DTFH61-91-C-00051. unpublished: February, 1994.

Roy Jorgensen Associates, Inc. A Comprehensive Review of the Traffic Records Keeping System. State of Tennessee Department of Transportation: August 28, 1986.

Shinar, D., Treat, J.R. and McDonald, S.T. "The Validity of Police Reported Accident Data." Accident Analysis and Prevention 15:3, 175-191, June 1983.

Site Data Book: A Compilation of Summaries and Materials Resulting from Site Visits. Northwestern University Traffic Institute. Federal Highway Administration contract DTFH61-91-C-00051. Unpublished interim report: 1993.

Stein, Howard S. "Comparison of OMC Accident Data with Independent Truck Accident Data from Washington State." Paper Presented at the 72nd Annual Meeting of the Transportation Research Board, Washington, DC: January 1993.

Stenzel, William W. and Lucke, Roy E. Police Personnel Allocation Manual: Sheriffs' Departments. U.S. Department of Transportation, DTNH22-88-C-05026. Evanston, IL: September 1991.

Turner, Ed and Peers, J.B. Road Accident Statistics Working Party: Interim Report. Greater London Council Department of Highways and Transportation. London, England: 1968.

Turner, Daniel S. and Mansfield, Edward R. "Variability in Rural Accident Reporting." Transportation Research Record 910, 8-14 1983.

Turner, Daniel S. "The Police Chief's Role in Deficient Accident Reporting." Paper presented at the 63rd Transportation Research Board Annual Meeting. Washington, DC: 1984.

Turner, Daniel S. and Colson, Cecil, W. "Accident Data as a Tool for Highway Risk Management." Transportation Research Record 1172, 11-22 1988.

Turner, Daniel S. "Factors Contributing to Abnormal Accident Reporting." Transportation Research Record 974, 1984.

"Uniformity in Motor Carrier Accident Reporting." Capital Ideas. National Governors Association. Lexington, KY: June 15, 1988.

Wu, Chong Qing. Uniformity of Highway Safety Data in the United States. Thesis in Partial Fulfillment of the Requirements for the Degree of Master of Science, Northwestern University, Evanston, Illinois, June 1993.