Safety Data: Costs, Quality, and Strategies for Improvement Executive Summary

PUBLICATION NO. FHWA-RD-96-027

SEPTEMBER 1997

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.

Two related reports, completed under the same contract, will be available by the end of 1997. One, entitled *Safety Data: Costs, Quality, and Strategies for Improvement (Final Report)*, provides more details on data collection methodologies and how to package them effectively to meet the needs of individual agencies. The other, *Safety Data: Costs, Quality, and Strategies for Improvement (Research Report)*, identifies, in detail, the issues and costs related to collecting and managing highway safety data and proposes ways to resolve them.

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.

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

NOTICE

Safety Data: Costs, Quality, and Strategies for Improvement

Executive Summary

Introduction

The Intermodal Safety and Transportation Efficiency Act (ISTEA), established in 1992, was designed to improve the performance of statewide and metropolitan transportation systems. To optimize system performance, the ISTEA required development and implementation of six management systems, including the highway Safety Management System (SMS). Although the management systems themselves have been made optional, the philosophy behind the SMS -- to ensure that all "...opportunities to improve highway safety are identified, considered and implemented as appropriate..." -- is still relevant and important. Quality data are needed at a reasonable cost to support effective safety decisions. The objective of this project was to examine the costs and quality of safety data and to identify improvement strategies.

Project Objectives

- Determine the costs of collecting and managing safety data
- Determine the quality of safety data
- Recommend strategies for reducing cost and/or improving quality

Study Approach

Safety data are collected by a variety of organizations and agencies that comprise a variety of highway safety community members. To meet the objectives of this project and to make it feasible given the project resources, the project's scope was narrowed to three key sources: the crash report, roadway inventories, and medical records. The primary focus was on crash reporting. Within the medical source, the focus was upon data regarding driver condition prior to crash, and injury severity, since these are of high interest to the highway community.

Dimensions of Quality

- Coverage
- Accuracy
- Precision
- Consistency
- Timeliness

In examining quality of the data available from these sources, clearly there were several dimensions of quality and the degree to which the sources are maintained will have varying effects, depending upon the user. For example, most geometric-design applications require no more precision about injury severity than the traditional four or five classifications (*killed, severe injury, minor injury, possible injury, and no injury*). However, designers of vehicles and roadside appurtenances wish to know the nature of injuries in great detail and are greatly limited by the standard five-level scale.

Data Sources Studies

- Crash report
- Roadway inventories
- Emergency medical care

The above-mentioned example makes the point that the same piece of data will have a different quality for different users and uses. Therefore, it was necessary to identify the variety of users and uses of highway safety data so that alternative strategies for improving data collection and management could be assessed. The set of users and uses, and the functions that were identified for them, is shown on page 2. For this study, a set of user/use scenarios was developed to represent the range of users and uses. This was a reference base against which candidate improvement strategies were evaluated.

Primary Sources Used for This Study

- Detailed literature review
- Field data-collection visits to 10 States
- Interviews with collectors, managers, and users of safety data at the local, State, and Federal levels

The information used in this study was derived from several sources, including field visits to State and local agencies.

Users and Uses of Highway Safety Data

Data User	Functions of this Entity
Law Enforcement Agencies	Enforcing 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
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

The cost to collect and manage safety data. To provide a framework for examining the costs of collecting and managing crash data, a cost-estimation model has been developed. This model has been used to estimate national crash data collection costs and assist in predicting the cost impact of improvement strategies. The estimates were based upon data obtained from States visited and a limited amount of field data collection. The key variables that affect the costs can have widely ranging values from State to State, making it difficult to arrive at a value that is generally applicable. The national and unit values shown here should be taken as order-of-magnitude estimates.

Key Variables for Determining Cost

- Labor (e.g., time in field, data entry time)
- Employee benefits
- Number of units collected
- Systems hardware & software development
- Overhead

Issues that affect cost and quality. An understanding was needed of current State practices to help identify the strategies needed for controlling costs and improving the quality of safety data. Information gathered from site visits and literature reviews provided the necessary insights. Key issues were identified and are summarized in the sections below. In addition, exemplary practices employed by some state and local agencies were also identified. These practices helped develop the improvement strategies presented in the sections below. It was also encouraging that most agencies visited were:

- Doing a credible job of providing useful data; and
- Transitioning to modern systems based upon current technology, which has only recently become available at affordable costs.

National Costs for Collecting and Managing Crash Data

- Total - $130 million
- Data collection - $50 million
- Data management - $80 million
- Average unit cost - $21 per crash

Crash Data Collection

The general patrol officer has other, higher priorities than collecting safety information both at the scene of a crash and in the general context of the job. State patrol officers usually have greater motivation to perform well on crash reporting, because a substantial part of their job is crash investigation, and it is directly reflected in job performance review. The degree of coverage of crashes not known, neither is the manner in which it may vary over time and under different conditions. This has a substantial affect on the ability to evaluate accurately the effectiveness of any safety improvement action. Officers are asked to perform a task for which they receive little preparation or support, both in terms of training and motivation. The reporting officer is often asked to make judgments and provide information (e.g., roadway and injury data) which exist in other data sources at a higher level of quality than can be attained by the officer. Officers currently do not generally have portable computing devices available to guide and facilitate data collection, thus limiting quality and productivity. Location data are often inaccurate, incomplete, or missing. This is a problem of major importance, since the location is the principal means for linkage to other spatial data (e.g., roadway inventory).

Issues of Crash Data Collection

- Patrol Officer has higher priorities
- Degree of coverage of crashes is not known
- Patrol Officer not adequately trained
- Data are collected that are available elsewhere
- Lack of automated assistance
- Lack of support from law-enforcement management
- Inaccurate & imprecise location of crash

Issues of Crash Data Management

- Lengthy delays for processing
- Significant redundancy of data management
- Difficult to attain linkages
- Inflexible databases
- Editing remotely from source
- No comprehensive quality assurance
- Entire process labor-intensive and inefficient
- Lack of general systems orientation

Noteworthy Practice

The Wisconsin DOT developed a quality assurance program that established a quality-control group focusing on data management. The group identified and prioritized quality assurance strategies for implementation.

Issues of Roadway Data Collection

- Lack of automated assistance
- Longitudinal geometrics not collected or well maintained
- Imprecise and/or incompatible location data
- Data collectors have other higher priority tasks
- Inconsistent use of project data
- Limited coverage of the system

Crash Data Management

Delays of several months can occur between the time of the crash and the time data become available from the data store. These delays lead to significant redundancy of data management since each agency that needs the data has created its own separate system. In some cases, two local agencies and four State agencies are entering and maintaining the same data. The independent and outmoded nature of many database management systems severely limits the ability to link complementary data stores. This is a barrier to having additional data and higher quality data available for users. Furthermore, it restricts productivity during the data entry and edit process. In general, data edits are performed at central locations remote from the originating officer. This discourages effective follow-up and adds to inefficiency when corrections are pursued. While most agencies perform editing and correction processes, rarely is a comprehensive program for quality assurance pursued. The existence of a quality assurance group was rare. Due to several of the factors above and outmoded processing practices, the data-management process generally involves too many people and too much handling of paper. The latter issue results from an historical lack of systems orientation.

Roadway Data Collection

Several key issues listed above are also applicable to roadway data collection and management. Lack of automated assistance is one of those, but is repeated here because it is even more important to the collection of field data about the road than it is to the crash. Linkages are generally not made between photologs, traffic data, and the inventory. Longitudinal geometrics (e.g., curvature and grade) are an example of data elements that are difficult to measure cost effectively. They are often missing from roadway inventories, and where present, are inaccurate or out-of-date. Accurate location data are also a significant issue for roadway data. This is sometimes due to varying location reference systems being used by different governmental entities. It is also due to insufficient care taken by data collectors. Data collection is often done by district personnel who have other higher priority tasks. Job performance review for these individuals does not usually include roadway data collection. District offices are usually the source for updates from complete projects. This aspect of job completion is also not high priority and is, therefore, not consistently performed. Finally, but not least, funding limitations have significantly restricted the coverage of the system to only the highest classification, leaving little or no data on many other roads along which safety analyses are being performed.

Roadway Data Management

Data entry is not usually a formalized process. It is either done at the district office or central headquarters. Data are keyed in to a computer from field forms. Edit routines are minimal, often limited to checking field type (i.e., numeric, alpha). Resources allocated to this activity are restricted, making frequent updates infeasible. The reliance upon project data updates is not sufficient, as mentioned above. There is almost no quality control performed on the roadway inventory file.

It is often the users who raise questions about accuracy before a correction is made. The database management systems employed are usually outdated technology. The inflexibility that results will restrict the data to current conditions. Thus, a history of changes, needed for long-term safety studies, is not available.

Issues of Roadway Data Management

- Inefficient data entry
- Inadequate edit routines
- Long intervals between updates
- Lack of comprehensive quality assurance programs
- Limited, or no, history file maintained

Emergency Medical Data Collection

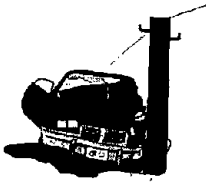
The medical data of interest to this project were "driver condition prior to the crash" and "injury severity." Law enforcement personnel can not be expected to provide accurate data on this. They are not adequately trained. Furthermore, the definitions used on crash reports, while useful for many highway engineering purposes, do not provide sufficient precision for users interested in injury outcomes (e.g., vehicle interior designers). Emergency care personnel keep records that could be accessed to provide the desired data at a better level of quality than usually available in police crash reports. However, definitions used by these agencies are designed to facilitate selection of treatment, and are not equally useful for safety analyses. Making EMD useful is further exacerbated by the lack of centralized control on reporting exhibited most clearly by lack of standardization of EMD reports and records within a State.

Issues of Medical Data Collection

- Law enforcement personnel are not trained to gather medical data.
- Definitions for injury severity on most crash reports have limited value.
- Definitions for injury severity in most EMD records are not oriented to safety analyses.
- EMS forms are not standardized.

Emergency Medical Data Management

The lack of centralized data sources, or a distributed data base for the various EMD sources, is generally due to the absence of EMD central control. Part of the problem comes because EMD falls squarely within the realm of privacy and confidentiality laws. Linkage to EMD is greatly complicated because of this. Coverage is also an issue with this data source. Not all injured parties are handled by the emergency care system. Injured parties will seek direct help of a personal doctor or other medical practitioners outside the emergency care system. Reports for those are not required from the attending physician or other practitioner.



Issues of Medical-Data Management

- Centralized emergency medical data sources do not exist.
- Confidentiality issues create barriers to linkage of EMD and PCR data.
- Not all injured parties are handled by EMS or EMD sources.

How do we resolve these issues, providing users with better quality data and helping data managers control costs?

More than 40 strategies were identified for improving safety data collection and management. Some of these involved relatively low cost and required little change, while others required significant resources and time (i.e., more than a year) to implement. Several technologies were also identified as being applicable.

Set out below is a discussion of many of the strategies in groups called "packages." Packages are generally formed because of some common attributes of the strategies in the group. However, any combination of strategies can be used in a given situation. The strategies are generally applicable to the three data sources studied, except where obviously meant for only one of them.

Noteworthy Practice

Fort Collins, CO; Fort Lauderdale, FL; and Rapid City, SD are examples of communities that are using non-sworn crash investigation specialists to complete crash report forms at crash sites.

Improved Data Collection Instruments

- Revised report format
- Revised set of elements
- Revised definition of elements
- Revised codes for elements
- Improved user materials
- Development of pre-printed diagrams

Improved Data Management Process

- Minimization of number of "stations" through which the report must go
- Cross training of personnel to allow single individuals to handle all aspects of the process
- Enhanced training
- Outreach meetings
- Newsletters
- Recognition and rewards to collectors who perform well
- Focus upon accurate location coding

Readily Implemented Improvement Strategies

Improved Data Collection Instruments. Most data collection is done using paper forms. Until automated means replace forms completely, changes in their format and content can make the collection process more efficient and produce higher quality data. Even if automated devices were used, many of these changes would still be applicable. Some data that are collected are no longer needed. Many forms are too crowded. Elimination of unneeded data elements from the form allows a cleaner design that field personnel can complete faster and with a greater degree of completion. Standardization of forms, at least throughout a State, is important to ensure consistency of data. Accuracy and precision can be helped by clarifying codes, assuring that alternative codes are mutually exclusive, and providing manuals and training materials that effectively communicate how to handle the more complex cases. Preparing diagrams of the crash as part of the crash report is particularly distasteful to many general patrol officers. Providing pre-printed diagrams for both basic cases and the complex locations in a jurisdiction would encourage officers to properly complete them.

Improved Data Management Process. Several aspects of data management can be immediately addressed. Data handling can be simplified in many agencies. Improved training of the data edit and entry personnel can reduce costs and improve quality. It is important to minimize the number of times a form is handled in the process. An outreach program can also be effective for improving the quality of the collector's product. Simple feedback and rewards programs have been found effective. Finally, if data management were to focus only upon requiring improved location data and coding, data quality would improve significantly because of its increased usefulness.

Longer-Term Improvement Strategies

Decreased Reliance on the General Patrol Officer as Collector of Crash Data. If high quality crash data are desired, the general patrol officer, who is overloaded with other tasks and not adequately trained for the crash reporting task, should be replaced with a person whose job is primarily oriented to this task. Since State patrol officers perform crash investigations as a primary job responsibility, they tend to be better motivated and trained for this than officers in urban and suburban units. Replacement of the latter with non-sworn or sworn crash investigation specialists has resulted in improved reporting. Use of non-sworn personnel is also a cost-saving action. It is possible to reduce the reliance on general patrol by just raising the reporting threshold or employing sampling methods. The reduction of sample size has negative impacts that must be considered in light of offsetting cost savings and the possibility of better attention to the reporting of remaining crashes. Another approach retains officers as a reporter, but relies upon specially trained personnel for interpretation that officers are not equipped to provide. Officers would focus upon a full narrative and diagram that would be coded by the specialists. Driver condition and injury severity is a particularly difficult area for officers to report accurately. It is possible to use emergency medical data sources for this information, but only where such favorable conditions as interagency cooperation, trauma registries, and computer linkages are available.

Comprehensive Quality Assurance Program. Quality assurance involves monitoring the stream of data entering the database, identifying problem areas, providing feedback and programs to improve quality, and constantly improving means for error trapping, as well as cost-effective means for correction. To do this job in a comprehensive way will require the creation of a small group, probably at the State level, dedicated to this effort. The quality assurance team should also address maintenance of, and access to, files and documents. Capabilities are needed in data and statistical analysis, public relations, and training.

Apply Re-engineering Methods. Systems design and operation principles need to be applied, both within each entity of the highway safety data system and between them. Creating a task force of collectors, managers, and users of highway safety data is the first step. Systems specialists can be used to define a re-engineering process to be carried out under the leadership and monitoring of the task force. The result could include major re-organization to bring important entities of the system under a single umbrella for more cost-effective operation. A significant part of the process should be focused upon defining users and data needs. This includes establishment of a data element sponsorship program to assure that those who depend upon specific data for their functions are represented and providing resources as appropriate.

Decrease Reliance on the Officer as Crash-Data Collector

- Use non-sworn investigators
- Use sworn officers as specialists
- Increase reporting threshold
- Sampling
- Use trained interpreters to code from narrative and diagram
- Use specialists to review crash reports where they originate
- Use EMD sources to supply data on severity

Comprehensive Quality Assurance Program

- Create a quality assurance team
- Conduct regular quality review studies of data files
- Communicate identified problems to collectors and managers (e.g., with a newsletter and workshops)
- Communicate to collectors about users and uses of data
- Recognition and rewards for performance
- Conduct regular state-of-the-art refresher training
- Improve form design
- Improve user materials
- Establish linkages to alternative data sources
- Upgrade error trapping and editing processes
- Enhance document control and access
- Maintain history files covering 10 years

Technologically Advanced Improvement Strategies

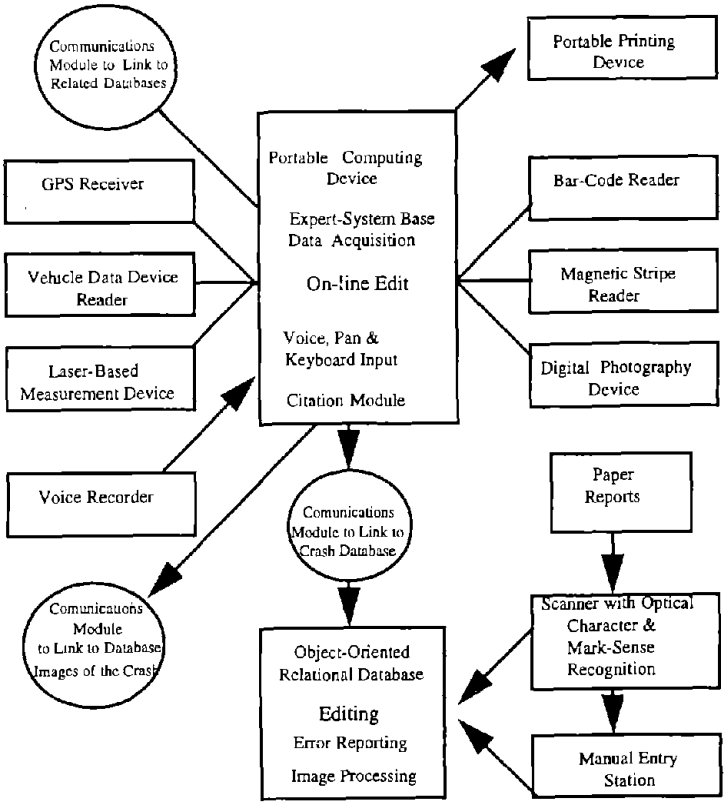
Apply Re-engineering Methods

- Establish task force of data collectors, managers, and users
- Establish comprehensive systems requirements
- Conduct process and methods review
- Conduct systems technology review
- Develop requirements for sponsored elements
- Re-organize for more cost-effective operation

Maximize Automated Collection of Data

- Use portable computers to guide data collection
- Develop expert-system based interview protocols for guiding field data collection
- Minimize reliance on keyboard input (i.e., use pen and voice)
- Integrate digital instrumentation with the portable device (e.g., GPS receiver, digital photography, and laser measurement devices)
- Provide communications links with remote databases for uploading data
- Maximize point-of-entry error trapping and correction
- Establish links with alternative data stores for direct reference
- Employ satellite imaging for roadway data-collection
- Use instrumented vehicles to collect roadway data

Maximize Automated Collection of Data. Technology is driving the advancement of business operations today. It is only recently that the technology applicable to safety data collection and management has developed to a point where it may be implemented. The timing is good, however, since many State and local data processing systems have become antiquated to the point that changes are inevitable. The most appealing aspect of new technology for field data collection is the major advances in portable devices. It is now feasible to use a portable computing device to guide data collection and record it. Sophisticated portable measurement devices have been developed to complement portable computers. Applicable devices for safety data collection include global positioning systems, bar-code readers, magnetic stripe readers, digital photography (still and motion), and laser-based distance-measuring instruments. Use of a portable computing device allows point-of-origin control on data quality through automated guidance of the collector and on-board error-trapping and correction modules. Cellular communication allows the collector to connect to other data sources for reference purposes (e.g., check existing roadway data and check current driver records) and to up-load to a central processing location.



Prototype Portable Crash Data-Collection System

Futuristic views of safety data collection, using advanced technology, include use of:

- Expert systems to guide the data collector, using an interview-type method.
- Satellite imaging to capture data about the crash and roadway.
- Instruments, on-board vehicles equipped for intelligent transport systems purposes, to capture vehicle-dynamics through the crash sequence.
- Instrumented vehicles to collect longitudinal geometrics and other data along.
- Image processing to analyze and measure digital photographs.

While the methods in this list are not necessary for successful safety data collection and management today, they should be considered in establishing future directions toward which the system might evolve.

Maximize Automated Management of Data. Several basic areas of data management effectively apply to new technology. Database design and communications are two key areas. The desire for interconnection and flexibility among databases suggests that a relational database design should be employed. Use of object-oriented databases allows the storage of data in a variety of formats, including images. By establishing a comprehensive communications infrastructure, data stores can be readily linked, leading to the feasibility of distributed databases. One concomitant of a distributed database and/or an advanced communications system is that entry, editing, and correction can be maintained at the local level where it is best done. Furthermore, all qualifying agencies can access the data at the earliest possible time. Where paper documents continue to be used, document imaging, with both character and mark recognition, can substantially reduce processing delays and costs. However, these automated data entry approaches must be accompanied by appropriate safeguards to assure that there are replacements for the human intervention for edit and correction that have been replaced.

Noteworthy Practice

The California Highway Patrol has adopted a plan for automation of all field-data collection, calling it "Vision 2000." It includes mobile digital computers, GPS, and magnetic stripe readers.

Maximize Automated Management of Data

- Develop a comprehensive communications infrastructure
- Establish links with alternative data stores for reference and download
- Implement local data entry
- Install relational databases
- Install object-oriented databases
- Implement distributed databases
- Create linkages to computer-aided design systems to get roadway data
- Provide document imaging with OCR & OMR

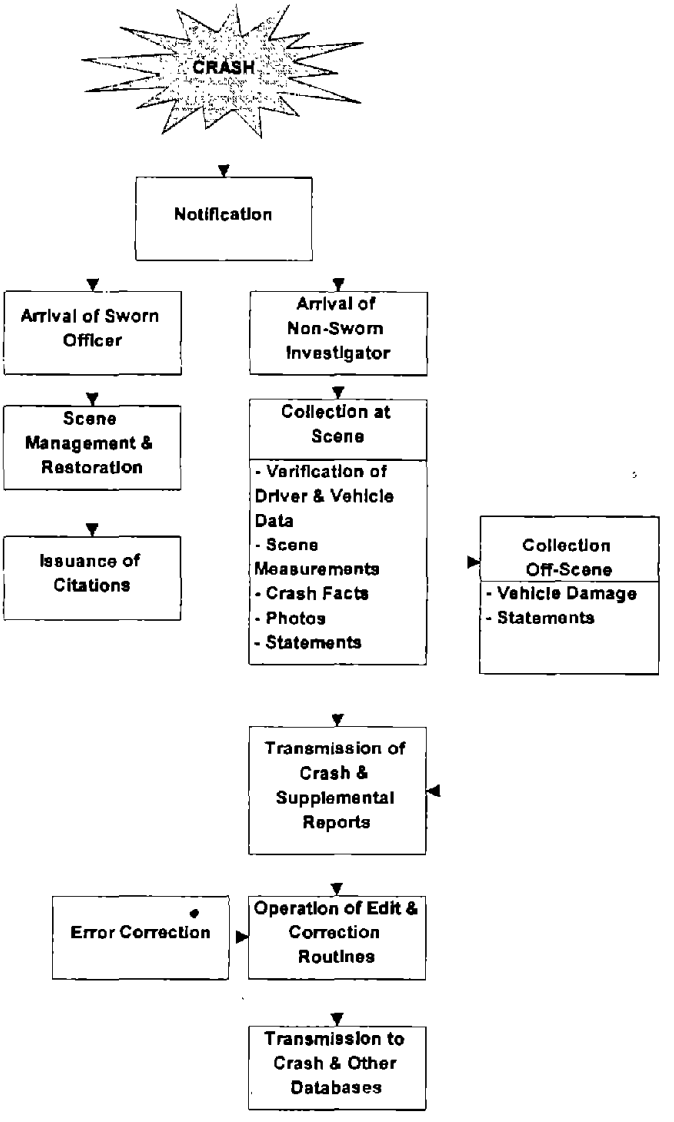
Noteworthy Practice

PennDOT uses special software and on-site verification to update its roadway management system. It allows them to provide details about the roadway for maintenance bidding. The system is also updated based upon plans and as-built drawings.

Conclusion

The process shown schematically in the diagram provides a framework for a system of data collection and management which can embody all of the desired strategies identified as part of this project. It involves the division of labor between sworn officers who can manage the scene and others who can be trained to collect the desired data in a cost-effective manner. It enables the use of highly automated devices for data collection and validation and the latest technologies in communication for its transmission and merging with the existing database. It allows for a broad range of types of data to be included in the database.

Virtually every agency contacted as part of this research effort was committed to improving their collection and management of safety data. This project has identified many ways that an agency can act immediately and comprehensively to improve the process. It is important, however, that the actions be taken with a systems perspective, rather than as a set of unconnected ad-hoc actions. Therefore, it is desirable to develop a detailed form of the prototype displayed here, toward which the agencies involved may evolve. The best result will be a plan that combines technological and non-technological methods, materials and processes. While the application of technologies does not guarantee good results, technologies can help, in the right context, achieve the desired objectives faster and more effectively.



Prototype Crash Data Collection and Management Process