Hazardous Materials Transportation Incident Data for Root Cause Analysis

DETAILS

111 pages | | PAPERBACK
ISBN 978-0-309-11810-1 | DOI 10.17226/14336

AUTHORS

Transportation Research Board

HAZARDOUS MATERIALS COOPERATIVE RESEARCH PROGRAM

HMCRP REPORT 1

Hazardous Materials Transportation Incident Data for Root Cause Analysis

Battelle Memorial Institute
Columbus, OH

Subject Areas
Planning and Administration and Environment • Operations and Safety • Freight Transportation

Research sponsored by the Pipeline and Hazardous Materials Safety Administration

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2009
www.TRB.org

HAZARDOUS MATERIALS COOPERATIVE RESEARCH PROGRAM

The safety, security, and environmental concerns associated with transportation of hazardous materials are growing in number and complexity. Hazardous materials are substances that are flammable, explosive, or toxic or that, if released, produce effects that would threaten human safety, health, the environment, or property. Hazardous materials are moved throughout the country by all modes of freight transportation, including ships, trucks, trains, airplanes, and pipelines.

The private sector and a diverse mix of government agencies at all levels are responsible for controlling the transport of hazardous materials and for ensuring that hazardous cargoes move without incident. This shared goal has spurred the creation of several venues for organizations with related interests to work together in preventing and responding to hazardous materials incidents. The freight transportation and chemical industries; government regulatory and enforcement agencies at the federal and state levels; and local emergency planners and responders routinely share information, resources, and expertise. Nevertheless, there has been a long-standing gap in the system for conducting hazardous materials safety and security research. Industry organizations and government agencies have their own research programs to support their mission needs. Collaborative research to address shared problems takes place occasionally, but mostly occurs on an ad hoc basis.

Acknowledging this gap in 2004, the U.S. DOT Office of Hazardous Materials Safety, the Federal Motor Carrier Safety Administration, the Federal Railroad Administration, and the U.S. Coast Guard pooled their resources for a study. Under the auspices of the Transportation Research Board (TRB), the National Research Council of the National Academies appointed a committee to examine the feasibility of creating a cooperative research program for hazardous materials transportation, similar in concept to the National Cooperative Highway Research Program (NCHRP) and the Transit Cooperative Research Program (TCRP). The committee concluded, in *TRB Special Report 283: Cooperative Research for Hazardous Materials Transportation: Defining the Need, Converging on Solutions,* that the need for cooperative research in this field is significant and growing, and the committee recommended establishing an ongoing program of cooperative research. In 2005, based in part on the findings of that report, the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) authorized the Pipeline and Hazardous Materials Safety Administration (PHMSA) to contract with the National Academy of Sciences to conduct the Hazardous Materials Cooperative Research Program (HMCRP). The HMCRP is intended to complement other U.S. DOT research programs as a stakeholder-driven, problem-solving program, researching real-world, day-to-day operational issues with near- to mid-term time frames.

HMCRP REPORT 1

Project HM-02
ISSN 2150-4849
ISBN: 978-0-309-11810-1
Library of Congress Control Number 2009940430

Published reports of the

HAZARDOUS MATERIALS COOPERATIVE RESEARCH PROGRAM

are available from:

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

and can be ordered through the Internet at:

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

Printed in the United States of America

C O O P E R A T I V E   R E S E A R C H   P R O G R A M S

CRP STAFF FOR HMCRP REPORT 1

HMCRP PROJECT 02 PANEL

FOREWORD

By William C. Rogers
Staff Officer
Transportation Research Board

While there has been considerable progress in the development of hazardous materials accident databases, the coverage of incidents reported to the U.S. Department of Transportation is insufficiently comprehensive for identifying root causes or contributors to incidents. This research focused on potential technical improvements to hazardous materials accident databases that are collected and managed by various agencies. The research identified gaps and redundancies in reporting requirements and estimated the extent of the under-reporting of serious incidents. The scope included all transportation modes covered by 49 CFR Parts 100–180. The report can be used by public agencies and industry to identify and prioritize measures that could further enhance the usefulness of hazardous materials transportation incident data. The suggested technical improvements are those of the research team and not the Transportation Research Board or the National Research Council.

Publicly reported incidents can be used for understanding the root causes of, or major contributors to, events involving a spill or release of hazardous materials during, or incidental to, transportation. This understanding can be used by regulators and industry to prioritize areas for attention and to develop or improve safety recommendations, regulations, and programs focused on preventing or reducing the likelihood of future incidents. However, complete, detailed, and accurate data are needed for meaningful analyses that reflect actual issues, and there is concern that the coverage of incidents reported to the U.S. Department of Transportation (under 49 CFR 171.16) is not sufficiently comprehensive.

Under HMCRP Project 02, Battelle Memorial Institute, along with the University of Michigan Transportation Research Institute, Calspan, RA-LUX, Mark Abkowitz, and Christopher Barkan, examined the recent literature on hazardous materials transportation incidents; interviewed carriers, shippers, and federal database managers; conducted detailed database analyses; and provided suggested technical changes to the databases that could improve the availability and quality of hazardous materials transportation incident data. The report also describes a pilot program to link the Hazardous Materials Incident Reporting System (HMIRS) and the Motor Carrier Management Information System (MCMIS) that could show how such an enhancement might more effectively identify root causes.

CONTENTS

SUMMARY

Hazardous Materials Transportation Incident Data for Root Cause Analysis

The objective of this project is to develop a set of potential measures that would enable officials to more effectively use incident reporting databases to identify major contributors (root causes) to hazardous materials (hazmat) transport accidents for all modes of transport. The focus of this study is accidents, namely those incidents in which the vehicle or vessel was involved in a crash event (as opposed to a stationary release from a loose fitting). In the discussion to follow, the terms "accident" and "crash" are used interchangeably. Although multimodal in emphasis, because hazmat truck crashes are dominant, the emphasis in this report is on hazmat truck accidents. For this project, the study team used the following definition of root cause: *One or more contributing factors that lead to the occurrence of a transportation accident and/or affect the severity of its consequences.*

The research team recognized that in order to effectively determine the root cause of a hazmat crash or a series of crashes, data on diverse parameters needed to be collected and analyzed. *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005) developed a matrix listing the parameters believed to provide a more comprehensive understanding of the accident environment. The five categories into which parameters fall are (1) vehicle, (2) driver, (3) packaging, (4) infrastructure, and (5) situational. In some cases, an individual parameter could shed light on root cause but, in many cases, analyses of two or more parameters are needed. In effect, a systems analysis is required. The matrix, although designed specifically for the highway (truck) mode, is applicable to other modes. The five major parameters and key variables under each are shown in Table S-1.

Note that, in addition to these major parameters, institutional characteristics, such as company financial condition, organizational structure, and safety culture, can play an important role in contributing to accident potential. For the purposes of this study, it is assumed that these considerations are embedded in the likelihood that the major parameter variables emerge as causal factors. For example, an organization with a poor safety culture is more likely to utilize a young driver with little experience and an invalid license.

Chapter 2, Literature Review, describes the results of a literature review conducted for the project. The review of relevant transportation accident data collection and analysis literature over the past three decades reveals some important findings and implications regarding the current state of the art of root cause analysis. These can be separated into: (1) recognition of problems and (2) proposed solutions. As early as 1981, there was acknowledgement that analyzing transportation safety using empirical accident data was problematic. Since then, numerous studies have cited five basic problems:

1. Inconsistent reporting practices within and across regions,
2. Non-reporting of reportable accidents,
3. Missing information in accident report records,

Table S-1. Accident parameters.

Vehicle	Driver	Packaging	Infrastructure	Situational
Configuration	Age	Package Type	Road Surface	Pre-Crash Condition
Cargo Body	Experience	Quantity Shipped	Road Condition	Dangerous Event
GVW	Condition	Quantity Lost	Road Type	Vehicle Speed
Vehicle Defect	Valid License	Age (Cargo Tank)	Traffic Way	Impact Location
Vehicle Response	Citation Issued	Rollover Protection	Access Control	Primary Reason
	Driver Response	Inspection History	Speed Limit	Accident Type
	Training	Design Specification	Number of Lanes	Weather Condition
			Location	Light Condition
				Time of Day
				Health Consequences

4. Inaccurate information included in accident report records, and
5. Data elements needed for root cause analysis do not appear on the report form.

The literature also contains suggestions for addressing data quality problems. Among the strategies being implemented or under consideration are the following:

- Posting available accident data on the Internet for review and feedback regarding its accuracy,
- Designing standardized accident reports toward a goal of more uniform data collection,
- Making extensive use of electronic data entry,
- Using sampling techniques to target certain types of accidents,
- Including common identifiers in complementary accident databases so as to integrate key causal information while avoiding duplication of effort, and
- Providing better training for law enforcement officials and other data entry personnel to enable them to collect and process information in a consistent, complete, accurate, and more timely manner.

Many of these strategies offer considerable potential, and are among those that were given careful consideration in the hazmat root cause analysis study.

Chapter 3, Summary of Interviews with Carriers, Shippers, and Database Managers, gives the results of a survey (conducted by mail, phone, and in person) of carriers, shippers, and accident database managers. The results show that when carriers and shippers experience a hazmat accident, several maintain accident databases that contain information that is much more extensive than the information that is required to be reported to federal databases. The investigators record environmental factors and long-term qualitative data that would be helpful in understanding how the hazmat accident occurred and, therefore, is useful for determining how the accident could have been prevented (if prevention was possible). In some instances, factors such as driver criminal history, crash history, and cell phone usage would have helped determine whether the accident was due to the driver, which, if true, could result in an action taken to discipline or suspend the driver. In one case, corrective action was taken by a company to make the driver more aware of these external factors, enabling the driver to prevent future accidents of a similar nature. On the other hand, if factors such as existing traffic/weather conditions and functionality of trucking equipment indicate that the fault of the accident was external to the driver, a change in driving procedures could be considered.

Chapter 4, Database Analysis, focuses on an analysis of the most important available databases for determining root cause. These databases include the Motor Carrier Management Information System (MCMIS); Hazardous Materials Incident Reporting System (HMIRS); Fatality Analysis Reporting System (FARS); Trucks Involved in Fatal Accidents (TIFA); Large Truck

Crash Causation Study (LTCCS); NTSB analyses; Railroad Accident/Incident Reporting System (RAIRS); and Marine Information for Safety and Law Enforcement (MISLE). The MCMIS, TIFA, LTCCS, and FARS databases focus on truck crashes, NTSB investigates all commercial aircraft crashes and certain rail and truck accidents, HMIRS includes all modes, RAIRS focuses on rail, and MISLE is water based. The following provides a brief summary of the characteristics and function of each database and, where applicable, selected potential measures for improving the capability of the database for identifying root causes.

MCMIS includes four major files: Registration, Crash, Inspection, and Company Safety Profile. The Crash file was developed in 1992 to record information on serious accidents involving a truck, bus, or light vehicle transporting hazmat. The process of reporting serious accidents begins with the state agency responsible for filing the MCMIS crash report to screen the Police Accident Reports (PARs) to identify serious heavy truck and bus accidents. Once the state reporting agency finds an accident that meets the requirements for reporting to FMCSA, the information from the PAR for the vehicle is coded into the MCMIS Crash file format and transferred to FMCSA for MCMIS Crash file inclusion.

Although the managers of the MCMIS Crash file have made great strides in improving the quality of the data, additional enhancements are suggested for this database to be more useful in identifying contributing and root causes of accidents. The single biggest improvement in MCMIS crash reporting would be if the existing parameter fields were completely populated. Some fields should be required to be filled, particularly those related to the vehicle, carrier, driver, route characteristics, and point-of-contact information. Since 2000, the *DRIVER_CONDITION_CODE* field has not been required. This field should be filled out again. This is the only field that captures driver performance in MCMIS. For trucks carrying hazardous materials, it is suggested that all five hazmat fields be completely and accurately filled. Presently, in records where one or more of these fields indicates a vehicle as carrying hazardous materials, all five fields are completely filled out less than 15% of the time. The DOT number should be entered for all serious crashes involving hazardous materials. Currently, a DOT number is entered for only 80% of the vehicles carrying hazardous materials. Finally, the *LOCATION* field should be specified in a manner that enables the accident location to be found on a map. Presently this is the case for roughly 30% of the time.

HMIRS is maintained by the Pipeline and Hazardous Materials Safety Administration (PHMSA) and can be considered to be a relational database. In accordance with 49 CFR 171.16, all carriers of hazardous materials by road, rail, water, or air must fill out DOT Form F 5800.1 and submit it to PHMSA within 30 days for a reportable hazmat incident. The reportable incident could occur during loading/unloading, while in transit, or while in temporary storage when traveling between the hazmat shipment origin and its final destination. The great majority of incidents involve a hazmat spill.

For a complete description of the package, vehicle, driver, and roadway characteristics associated with an accident, HMIRS must be joined with MCMIS for trucks and RAIRS for rail. Until the recent restructuring of HMIRS, the biggest impediment to joining the two databases was a lack of common fields. HMIRS now has a field to enter the DOT number and this field is being populated almost 90% of the time. The DOT number also is entered for about 90% of the MCMIS records designated as hazmat placarded. Assuming the non-reporting is random, the DOT number can be used to join about 80% of the accidents that meet both the HMIRS and MCMIS reporting criteria. Since all carriers of placarded quantities of hazardous material must register with both FMCSA and PHMSA, they must have a DOT number and there should be no blank entries in either database.

The following potential measures could enhance the ability of HMIRS to identify root causes of hazmat accidents.

1. Require that the DOT number be an input for all reports filed with PHMSA for in-transit incidents.
2. Perform an additional quality assurance (Q/A) check on carrier names to verify that the name being entered corresponds to the name provided on the annual PHMSA registration form.
3. Require PKGFAIL entries to be filled out for all reports submitted to PHMSA.
4. Continue to emphasize that carriers must file a Form DOT F 5800.1 if there was damage to lading systems on cargo tanks of 1,000 gallons or greater, even without a spill.
5. Capture driver condition information without compromising the confidentiality of the driver.

FARS was considered in concert with the TIFA file. The FARS file is the primary national crash database for fatal traffic accidents. It is a census of all fatal motor vehicle traffic crashes. The TIFA file covers all medium and heavy trucks involved in a fatal crash, and includes virtually all FARS variables for the crash, vehicle, and driver. TIFA survey data supplement FARS data for trucks. The TIFA data include a more accurate identification and description of trucks in fatal crashes, along with details about the cargo, configuration, motor carrier operating the vehicle, and crash type.

Both TIFA and FARS collect information about hazardous materials in the cargo. In the discussion of the variables that identify hazmat cargo in FARS, there are reasons to believe that the TIFA file identifies hazmat cargo more accurately. Therefore, the TIFA program could be modified to add the following additional information about hazmat cargo:

- MC number of the cargo tank. This information has been collected in the past as part of a special data collection effort, so the feasibility of collecting this information has been demonstrated.
- Quantity of hazardous material transported. Cargo weight could be included. The program should consider whether to capture the quantity in terms of liquid measure, where appropriate.

LTCCS was designed as a one-time study to compile a comprehensive set of accident data for approximately 1,000 large truck accidents. The data compilation began in 2001 and was completed in 2003, although analysis of the data is still ongoing.

Comprehensive studies, such as LTCCS, are needed to obtain contributing and root causes of accidents. These initiatives can be focused on a sample of all the accidents occurring in the United States, provided that the weighting of the sampling is known. However, there are significant advantages to performing a selected number of accident investigations annually rather than performing a larger intensive study over a one-to-two-year period as was done for LTCCS.

RAIRS is managed by the Federal Railroad Administration (FRA) as a tool to prevent railroad accidents. This comprehensive accident reporting system was implemented in its present form in 1975. FRA regulations require that all accidents in which damage to track and equipment exceeds a specified monetary threshold (adjusted periodically for inflation) must be reported using Form FRA F 6180.54, the Rail Equipment Accident/Incident Report, which records 52 different variables regarding the circumstances and cause of the accident. Beyond this, major railroads maintain their own internal databases. These typically contain all of the information necessary to comply with FRA reporting requirements, and often have additional data that individual railroads believe is useful for their own safety analysis purposes.

These efforts are significant to root cause analysis in several respects. The FRA reporting requirements ensure that all accidents of consequence are subjected to an analysis of the circumstances of the accident, and that both primary and, if applicable, secondary causes of the accident be determined and reported to FRA. In some cases, these may require fairly intensive analysis of the accident scene if there is some uncertainty about the cause. The major railroads employ specially trained individuals responsible for performing this function. Understanding all of these aspects is pertinent to root cause analysis of hazmat releases caused by railroad accidents.

The MISLE database supports the Marine Safety and Operations Programs. MISLE contains vast amounts of data ranging from detailed vessel characteristics, cargo carriage authorities, and involved party identities, to data on bridges, facilities, and waterways, to records of U.S. Coast Guard activities involving all of these and more. MISLE activities include law enforcement boardings and sightings, marine inspections and investigations, pollution and response incidents, and search and rescue operations. In addition, MISLE manages the information flow involving the administration of all of these activities from the initial triggering event to incident management and response, and the resulting follow-on actions. MISLE development was initiated in 1992 and became fully operational in 2002.

Much of the MISLE database is accessible only to Coast Guard staff. Furthermore, the MISLE data become available to the general public only for closed cases, and it can take several years to close many of the MISLE-reported incidents. This might be one of the reasons why, as part of this study, it was not possible to find common events reported to both HMIRS and MISLE. In its present form, lack of timeliness, access, and interconnectivity are considered insurmountable barriers for MISLE use.

The NTSB accident investigations and reports are investigations of individual accidents. While all commercial aircraft crashes are included, there are certain rail and truck accidents that are also selected for investigation by the NTSB. In order to move toward the identification of root and contributing causes, interested parties need to utilize all available data related to either a single hazmat crash or an entire population. The NTSB has developed a methodology that, through intensive detailed investigation, often leads to the identification of a root cause or causes. NTSB's detailed approach to accident investigation should provide insights to officials and researchers desiring to collect data on a particular hazmat incident or set of hazmat crashes.

Potential Measures for Improving the Identification of Contributing and Root Causes of Hazmat Accidents

The following measures could enhance the ability of the major databases to support more effective identification of the root causes of hazmat accidents.

- *Development of an information system capable of capturing the data for thousands of hazmat accidents that occur each year.*

  The information system would not reside in a single database and would be characterized by the following:
  - The use of a number of relatable databases, analysis tools, and reports contained in an information system that, in their totality, include the information in sufficient detail and quality to identify root and contributing causes of accidents.
  - The capability to build on the databases that currently are used to collect information on hazmat crashes.

  To identify the root and contributing causes of various classes of accidents, the analyst must be able to relate inventory information to accident tables. Inventory information, including the following factors, characterizes the hazmat information system:
  1. Driver characteristics (e.g., age and experience);
  2. Hazmat package characteristics (e.g., tank type and age);
  3. Vehicle characteristics, carrier characteristics, and mileage traveled;
  4. Infrastructure information in sufficient detail to identify causal factors relating to the location of the accident; and
  5. Situational information relating to the driver's decision making.

With this information, a regulatory analyst could mine a dataset and search for factors that are over-represented in one or more classes of accidents. Agency personnel could then use accidents where those factors are present and conduct detailed follow-up investigations to collect the additional information required to identify contributing causes for the selected classes of hazmat accidents.

Currently, PHMSA is completing Phase I of the Multi Modal Hazmat Intelligence Portal (HIP) system. The system is being designed to acquire hazmat information to be stored at a single location. Under HIP, data from a number of agencies will be available, including financial information on the carrier, shipper, manufacturer, and packaging company for Phase I. Only part of the information contained in the system will be made available to the public. Although FMCSA is participating, sharing hazmat crashes in the MCMIS database is not part of the system. The incorporation of this crash file in the future could enhance the usefulness of the system.

- *Add inventory data in databases (truck focus).*

  Inventory data could be added as fields in the major accident databases or supplemental databases (such as PHMSA's Registration Database or the MCMIS Census File) can be coupled to the accident records through the U.S.DOT number and provide insights into the carrier's fleet characteristics and overall safety performance.

- *Link data from HMIRS, MCMIS, RAIRS, and other information sources.*

  Existing fields could be used to link individual hazmat crashes in different databases. Fields that describe the event location, such as latitude/longitude (lat/long) coordinates, street addresses, river and rail mile points, and Federal Information Processing (FIPS) codes, could be added and/or better quality controlled (using GIS technology) to facilitate linking of databases. Another opportunity to accommodate database linking is through the use of commodity codes, provided that a more uniform referencing system can be employed across modes. Common accident identifiers would encourage data integration, validation, and sharing.

  For all hazmat truck crashes, the DOT number must be correctly reported and entered into the crash file. The use of a police accident report (PAR) number would be another possibility for linking traffic accidents. To ensure that this takes place, a copy of the PAR could be submitted with the crash. MCMIS also would need to enforce its rule that the report number submitted to the crash file include the PAR number. From the PAR, the geographic location (using FIPS codes) and the time of the incident could be entered into MCMIS.

- *Develop a system for each database that will target about 5% of hazmat crashes for more detailed investigation.*

  This approach supports the ability to perform special investigations of a particular class of mode-specific accidents where a statistically significant sample is necessary. The following examples represent potential selections for this targeting:

  – Hazmat crashes involving rollovers of tank trucks in order to more effectively identify root causes so preventive measures could be implemented;
  – Crashes involving a particular hazmat commodity, such as propane, to determine if these accidents are over-represented as a class of accidents and, if so, what measures could be developed to decrease both their number and severity;
  – Hazmat crashes occurring on a particular type of roadway, track class, or involving a certain category of driver based on age and experience; and
  – A random sample of all accidents for further characterization of the entire dataset.

  Although not exclusively for crashes involving hazmat, targeted accident data collection has already been partially implemented. FRA currently supplements selected rail accidents in the RAIRS data with about 100 detailed accident investigation reports published annually. The Coast Guard does a few detailed accident investigations and the NTSB investigates almost all air crashes as well selected serious crashes for rail and truck modes. LTCCS obtained a large number of parameters for approximately 1,000 heavy truck accidents. The TIFA approach that provides detailed analyses for all fatal large truck crashes is a feasible data collection model for this purpose because it is performed annually, relies on telephone calls to check information, and is less expensive than the LTCCS approach. The *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005) effectively applied a similar approach as that used in TIFA to a sample of hazmat crashes found in MCMIS for a particular year. In addition, the project used information for the same crash found in HMIRS, wherever possible.

- *Ensure data completeness and accuracy.*

  Data completeness includes accidents that are not reported as well as accident reports with incomplete data. Incomplete reports can have a negative effect on accident analysis since a complete record of cause and effect is not captured. When the negative consequences of under-reported accidents and incomplete reporting are combined, the ability of an analyst to draw conclusions from the data is significantly compromised. To avoid this from happening, the following measures could be taken—some of which apply to improvements that can be made to data entry while others are focused on the conduct of more thorough accident investigations:
  - Completion of the values for all parameters; no credible information system can operate using records in which many fields are blank.
  - Use of computerized data collection to enforce and validate data coding.
  - Flagging of required fields in the databases such that the system will not accept the record until those fields are complete.
  - Development of incentives to reward those who provide complete and accurate data for a database.
  - Electronic submission of all crash reports to the major databases, such as MCMIS and HMIRS, using the Web to facilitate accuracy.
  - Addition of "error trapping" to all databases in order to eliminate errors by applying a program that would test whether certain logical connections have been made within the same accident record.
  - An effort to ensure that all applicable hazmat accidents are included in a database through such measures as instituting training for all who collect and enter data, and searching media reports for applicable accidents.
- *Add additional information to enhance the ability of the databases to identify root causes of hazmat accidents.*
  - Add latitude and longitude to all databases to provide the exact location of a hazmat crash and assist in identifying the same accident when it appears in more than one database.
  - Add information on the quantity shipped, available from the shipping papers, and an estimate of the quantity spilled if there is a spill, in addition to completing the four HM fields in MCMIS (Placard Y/N, UN Number, Hazardous Material Name, and HM Class).
  - Add the PAR number for serious hazmat crashes so they can be linked to those police reports.
  - Include non-spill hazmat crashes in HMIRS for all crashes involving placarded shipments that meet the criteria for "MCMIS serious crashes."
  - Include digital photos of the accident scene in all of the major databases.
  - For HMIRS and MCMIS, include violations for drivers that occurred during the crash.

Conclusion

The research conducted under this project has demonstrated that there has been considerable progress during the past 20 years in the development and refinement of databases that include hazmat accidents. The project focused on identifying strategies for improving the identification of the root causes of hazmat accidents using these databases. The project findings provide researchers and officials with an overview and analysis of the individual databases and resulted in many potential measures for improving specific databases. Implementation of the major measures—including establishing an information system, linking databases to take advantage of accident descriptions in more than one database, performing detailed sampling of a specific set of crashes to assemble more detailed information, and the adoption of techniques to improve data quality and completeness—could yield the greatest improvements in the ability of interested parties to conduct root cause analysis. This, in turn, will enable officials to identify problems whose solutions or mitigation will result in an improvement in hazmat shipment safety.

Follow-On Project

If implemented, the findings of this report, *Hazardous Materials Transportation Incident Data for Root Cause Analysis,* could lead to the enhanced identification of root and contributing causes of hazmat crashes. Implementation of the potential measures identified will present both technical and institutional challenges. Consequently, the project team suggests that to evaluate the feasibility of implementation, a pilot program be implemented to demonstrate that the system will work effectively in identifying root causes. The team suggests that the pilot test involve linking the HMIRS database, which provides excellent data on the hazmat material and package, with the MCMIS database, which provides superior data on the driver and accident environment. These data could be supplemented by other sources such as telephone calls to carriers. Finally, the pilot test would demonstrate the system's ability to identify root causes and use these results to suggest improvements in hazmat safety.

CHAPTER 1

Introduction

1.1 Project Purpose

The objective of this project is to develop a set of potential measures that would enable officials to use incident reporting databases more effectively to identify major contributors to hazardous materials (hazmat or HM) transport accidents for all modes of transport. The focus of this study is accidents, namely those incidents in which the vehicle was involved in a crash event (as opposed to a stationary release from a loose fitting). In the discussion to follow, the terms "accident" and "crash" are used interchangeably. The goal is to provide the proper data elements, accurately reported, such that root causes of accidents can be determined.

For this project, the research team used the following definition of root cause:

> One or more contributing factors that lead to the occurrence of a transportation accident and/or affect the severity of its consequences.

Inherent in this definition is the assumption that if a contributing factor were not present, then the accident would not have occurred and/or the consequences would not have been realized. However, in reality there is seldom one factor, but often there is a series of causal factors or a causal chain that leads to the accident or the impacts. Furthermore, by identifying several contributing factors to an accident, much can be learned when analyses show that one contributing factor is present in a large fraction of a particular type or class of hazmat accident. Table 1-1 provides an example of a sequence of questions associated with investigating the root cause.

The research team recognized that in order to effectively determine the root cause of a hazmat crash or a series of crashes, data on diverse parameters needed to be collected and analyzed. *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005), which is described in more detail in Section 2.2.16, developed a matrix listing the parameters believed to provide a more comprehensive understanding of the accident environment. These parameters are as follows:

- Vehicle,
- Driver,
- Packaging,
- Infrastructure, and
- Situational.

In some cases, an individual parameter could shed light on root cause but, in many cases, analyses of two or more parameters are needed. In effect, a systems analysis is required. For example, the root cause of a rollover hazmat tank truck crash resulting in a spill could be related to a vehicle problem manifested by faulty brakes, an inexperienced driver with inadequate training, a full load in an obsolete cargo tank with an inadequate inspection history, slick road conditions, and a precipitous lane change by another truck. A matrix showing the

Table 1-1. Root cause questioning.

Question	Response
Why did the truck run off the road?	The driver fell asleep.
Why did the driver fall asleep?	Driver has sleep apnea.
Why were you not aware of this disease?	We did not have an up-to-date medical record.
Why did no one check that he had an outdated medical record?	Our written procedure did not require us to check if employees had not updated their medical records in the last year.

five major parameters with key variables under each is presented in Table 1-2, where the conditions described in this example are shaded.

Note that, in addition to these major parameters, institutional characteristics, such as company financial condition, organizational structure, and safety culture, can play an important role in contributing to accident potential. For the purposes of this study, it is assumed that these considerations are embedded in the likelihood that the major parameter variables emerge as causal factors. For example, an organization with a poor safety culture is more likely to utilize a young driver with little experience and an invalid license. Unfortunately, the science on the relationship of "company organization" or "safety culture" to safety is still sufficiently new that there is no well-defined set of variables capturing the salient characteristics of "company organization" and "safety culture" that could be implemented feasibly.

1.2 Research Approach

For the project, the following research approach was used.

1.2.1 Literature Review

The research team examined the literature related to hazmat crash databases, including those that may be dominated by non-hazmat crashes, to determine how the problem of identifying root causes had been addressed in the past. Part of the literature review was aimed at gaining insight into how root cause analysis should be conducted and lessons learned in other research

Table 1-2. Accident parameters.

Vehicle	Driver	Packaging	Infrastructure	Situational
Configuration	Age	Package Type	Road Surface	Pre-Crash Condition
Cargo Body	Experience	Quantity Shipped	Road Condition	Dangerous Event
GVW	Condition	Quantity Lost	Road Type	Vehicle Speed
Vehicle Defect	Valid License	Age (Cargo Tank)	Traffic Way	Impact Location
Vehicle Response	Citation Issued	Rollover Protection	Access Control	Primary Reason
	Driver Response	Inspection History	Speed Limit	Accident Type
	Training	Design Specification	No. of Lanes	Weather Condition
			Location	Light Condition
				Time of Day
				Health Consequences

Note: Shading reflects contributing factors to root cause of the hypothetical hazmat tank truck crash described in Section 1.1.

efforts. This included National Transportation Safety Board reports that have identified the major contributors and, in some cases, the root causes of severe accidents. For each report, the evaluation looked at the "why" questions that were asked and how the information needed to answer the "why" questions was obtained. A summary of the literature review is presented in Chapter 2 of this report.

1.2.2 Survey of Agencies, Shippers, and Carriers

To learn what quality control measures are being utilized, the project team surveyed agencies that maintain accident databases. Shippers and carriers also were surveyed to gain an understanding of their accident investigation and reporting activities. Agencies were interviewed and questioned concerning the checks that are made to ensure data accuracy and completeness. These interviews are discussed in Chapter 3. In parallel with the discussion with federal agencies, the researchers asked shippers and carriers to address their ability to identify information that would answer "why" questions as well as their willingness to report that information. The results of these interviews also are summarized in Chapter 3.

1.2.3 Analysis of Databases

The team next examined the major crash databases and identified fields that might provide answers to any of the "why" questions associated with identifying root cause. The analysis also included an assessment of data quality—an aspect deemed critical to an understanding of the root causes of hazmat crashes. The consideration of data quality includes both accuracy and completeness. Unless the data is of high quality, any root causes, even if they were reported, could be difficult to uncover. High-quality data enables the analyst to more easily identify trends and relationships; for example, a group of similar accidents, perhaps very severe accidents, can be analyzed for the most common root causes. Even with high-quality data, the results may not be adequate if the pertinent fields are not included in the database.

The following major databases were included in this assessment:

- Motor Carrier Management Information System (MCMIS) managed by FMCSA. The data are compiled by the states from police accident reports (PARs) from serious crashes involving large trucks.
- Hazardous Materials Incident Reporting System (HMIRS) managed by PHMSA. The database only covers shipments of hazardous materials and is self-reported by carriers for the various modes.
- Trucks Involved in Fatal Accidents (TIFA) managed by the University of Michigan Transportation Research Institute (UMTRI). The crashes are culled from the Fatality Analysis Reporting System (FARS) and supplemental data on the crashes are collected by a survey. Only fatal, large truck crashes are included, but data quality is very high.
- FARS managed by NHTSA. The database is designed to include fatal crashes involving any vehicle and is not restricted to trucks.
- Railroad Accident/Information Reporting System (RAIRS) managed by FRA. The data are reported by the carrier and the focus is just rail, although intermodal hazmat shipments are also covered.
- Marine Information for Safety and Law Enforcement (MISLE) managed by the Coast Guard. The dataset is limited to accidents involving an actual or potential violation of the law. Data are closely controlled by the Coast Guard.
- Large Truck Crash Causation Study (LTCCS), a one-time, specific analysis managed by FMCSA and NHTSA. The database involved about 1,000 crashes and included Level I on-the-scene inspections.

In reviewing data quality, agencies were interviewed and questioned concerning the checks that are made to ensure data accuracy and completeness. These interviews are discussed in Chapter 3. In addition to using answers to questions administered to representatives of these agencies, the study team analyzed the databases to gather information on data quality. The most effective method employed was using one database to confirm the data quality of another. That is, the team compared two databases in order to identify missing or incomplete data or, in some cases, crashes that should have been included but were not part of the database.

The project focused on addressing the following issues:

- Availability of accident data for the transportation of hazardous materials,
- Criteria for reporting the accidents,
- Database format and availability (e.g., online, paper, other),
- How data files can be coordinated or integrated,
- How the data are used by each agency to analyze trends,
- How the data are compiled for analysis,
- Methods used to improve data quality,
- Usefulness of the database for identifying root causes,
- Techniques used to link the database to another database describing the same crash, and
- Suggestions for enhancing the usefulness of the database for identifying root causes.

1.3 Effective Methods to Ensure High-Quality Data

The research team relied heavily on the TIFA database for examining truck crashes. This database is recognized in the truck safety community as being both comprehensive and of high quality. The accident records contained in TIFA are the result of careful data checking using police accident reports and an intensive program of additional data collection on a fatal truck crash. This process involves direct contact with key parties such as police, carriers, and tow truck drivers. The research team used TIFA as a benchmark to compare fatal truck accidents occurring in other databases with those in TIFA.

A similar technique of data confirmation and augmentation was used for the Hazmat Serious Truck Crash Project conducted for FMCSA from 2002 to 2005. Police accident reports were checked against a crash description in MCMIS and carriers were interviewed by telephone to supplement the information in MCMIS. This process was followed for approximately 1,000 hazmat accidents identified in the MCMIS database and supplemented where possible with data from the HMIRS database. Supplementing the data provided key information that shed light on the root cause and major contributors to hazmat accidents. For example, one of the carriers interviewed reported that tank trucks, because of the higher center of gravity when full, are more difficult to handle; therefore, they are more prone to rollover. This suggests that driver experience might be important. Consequently, one of the questions asked during carrier interview was the number of years of driving experience the driver had at the time of the accident. The results from this question were significant—drivers with less than five years of experience had a larger fraction of cargo tank accidents resulting in rollovers than did more experienced drivers. The same analysis showed that rollovers were the most common precursor to a hazmat release in an accident.

Based in part on lessons learned from the process of compiling TIFA and the Hazmat Serious Truck Crash Project databases, the research team looked for answers to the following questions:

1. Is it possible to have a high-quality database without follow-up correspondence (contacts) with the reporting entity?
2. How much correspondence is required and what type is most effective?

3. Could the techniques found effective for TIFA and the Hazmat Serious Truck Crash Project be applied to databases that include or focus on hazmat transport?

These questions are considered throughout the analyses of the databases in Chapter 4 and in the potential measures and conclusions presented in Chapter 5.

1.4 Potential Measures to Enhance the Ability of Databases to Identify the Root Causes of Hazmat Crashes

After the team performed the literature review, surveys, and database analyses, it identified potential measures for improving the ability of selected databases to identify the root causes of hazmat crashes. Included are specific measures for improving the quality of data, reducing underreporting, adding fields that will improve the identification of root causes, and linking descriptions of the same accident in different databases. Finally, a method is outlined to supplement the databases through additional data checks and collection of additional data from key sources. The measures are organized in a hierarchy based on whether the focus is on data improvement, institutional challenges, or the cost of implementing particular solutions. The potential measures are provided in Chapter 5.

CHAPTER 2

Literature Review

2.1 Introduction

Concern for transportation accident data collection and the performance of effective root cause analysis is not new. Over the last several decades, policy analysts and researchers have attempted to use crash data to understand what causes accidents and how best to prevent future occurrences. As a result, a body of literature exists with the potential to provide beneficial information to this hazmat root cause analysis study.

The study team conducted extensive online searches for relevant literature, focusing on studies of transportation accidents and, more particularly, on the quality of information utilized and the types of analyses that have been performed. As a result, a variety of sources were identified and subsequently reviewed. The remainder of this chapter describes the results of that process.

2.2 Synopses of Relevant Studies

The discussion below contains synopses of relevant literature that was obtained and reviewed. In each case, background is provided on the study objectives, followed by a description of findings, conclusions, and recommendations. The synopses appear in no particular order. Section 2.3 contains a summary discussion of key lessons learned from the literature review and how this information relates to the objectives of the hazmat root cause analysis study.

2.2.1 *Rail Equipment—Train Accident Data*

Rail Equipment—Train Accident Data (Bureau of Transportation Statistics) is a document that describes reporting requirements for rail equipment, train accidents, and issues associated with data collection. Railroads are required by regulation (49 CFR 225) to report monthly to FRA all such accidents that meet a certain dollar threshold. This damage amount does not include loss of lading, cleanup costs, societal costs, loss of main line, personal injury, or death. Data must be updated when the costs associated with the accident are 10% higher than initially reported.

The Bureau of Transportation Statistics (BTS) acknowledges that non-sampling errors exist in this reporting system due to

1. Non-entry error,
2. Duplicate entry error (when more than one railroad is involved),
3. Missing data error,
4. Response/measurement error (e.g., accuracy of repair records),
5. Coding/recording error, and
6. Non-coverage error (railroad systems that are excluded from reporting requirements).

It is further noted that such errors are less likely to be associated with the more severe accidents since they receive greater investigative scrutiny.

The following recommendations are made for conducting verification and validation at various levels of the reporting process:

1. Improvements in the railroad's internal control plan to ensure that missing and corrected data are provided to the railroad safety officer,
2. Review of reports by railroad safety officer prior to submission to FRA,
3. Use of edit checks within FRA's data entry system, and
4. Performance of cross-field and cross-record checks.

The information is posted on the FRA Internet site, providing others with an opportunity to review entries and comment on their authenticity.

2.2.2 *Project 5 Overview—Developing Common Data on Accident Circumstances*

Project 5 Overview—Developing Common Data on Accident Circumstances (Bureau of Transportation Statistics) describes a project undertaken to evaluate data currently available from which to identify the factors and circumstances that are present in transportation crashes and incidents. A comparison also is made to what is needed by investigators and researchers to improve analysis effectiveness, leading to recommendations that are made for how to enhance data quality. The overall objective of the activity was to identify those data elements needed for adopting a common framework of factors across a wide variety of events and modes.

Included within the scope of the study were crashes or mishaps meeting all of the following conditions:

1. Involving the movement or operation of a vehicle, vessel, aircraft, pipeline, or other conveyance in the course of moving people or goods,
2. Occurring within a U.S. jurisdiction or involving a U.S. commercial carrier,
3. Being either intentional or unintentional in nature, and
4. Resulting in substantial property damage or injury, or the death of a passenger, crewmember, pedestrian, other worker, or bystander within 30 days of the event.

Data reviewed as part of the project included reports filed with U.S.DOT agencies, other federal agencies, and some non-federal agencies (e.g., state medical examiner offices). The basis for performing an evaluation of these data was the Haddon Matrix, a conceptual framework used to analyze risk factors or prevention measures for mishaps and injuries. The Haddon Matrix divides an event into three chronological phases (1) pre-event (contributing to event likelihood), (2) event (influencing likelihood and severity of an injury), and (3) post-event (affecting likelihood of survival/recovery). Each of these phases is further divided into four groups of risk factors (1) operator, (2) vehicle, (3) physical environment, and (4) social/cultural/organizational circumstances.

Among the data gaps and limitations discovered from applying this methodology were the following:

- Some important data elements are rarely collected, such as data on the injury mechanism, operator fatigue, distractions, and alcohol use;
- Lack of information on injury type and severity;
- Lack of a narrative description in reports, or information contained in narratives is not used;
- Lack of detail on human factors;

- Absence of guidelines for law enforcement officers and others who are expected to file incident/accident reports; and
- A linkage between a crash investigation report and death certificates and autopsy data is typically missing.

Among the recommendations for addressing these inadequacies are

- Make greater use of sampling to obtain more detailed information on events of interest, including performing supplemental studies in conjunction with sampling.
- Perform special studies using other databases (e.g., Consumer Product Safety Commission's National Electronic Injury Surveillance System) to address transportation-related injuries for which data are not routinely collected by DOT agencies.
- Improve data collection/reporting details about crash severity and mechanisms of injury.
- Add photographic evidence to crash files.
- Make greater use of geographic information systems (GIS) to identify more precisely where the event occurred and to relate the location to surrounding features.
- Incorporate data from non-DOT sources (e.g., information on a death certificate) into DOT data records.

Other recommendations were associated with how to make greater use of technology to improve data collection and included

- Provide crash investigators with handheld devices containing drop-down menus for on-scene data entry.
- Incorporate the use of event data recorders into the police accident reporting process.
- Encourage the installation of automatic crash notification in road vehicles and have this data included in the investigation.

2.2.3 "National Crash Data Bases Underestimate Underride Statistics"

"National Crash Data Bases Underestimate Underride Statistics" (*Road Management & Engineering Journal* 1999), summarizes the results of a study that selected 275 fatal truck-car crashes reported in both the Fatality Analysis Reporting System (FARS) and National Accident Sampling System—Crashworthiness Data System (NASS/CDS) for the purpose of evaluating the frequency of crashes that are characterized as underrides. Data from NASS/CDS showed that the percentage of fatal underrides of large trucks by passenger vehicles was much higher in NASS/CDS (27%) than in FARS (7%). The NASS/CDS statistics were considered to be more reliable because a larger amount of resources and personnel are devoted to investigating a crash in NASS/CDS than in FARS.

This discrepancy in underrides as a crash characteristic was attributed in part to a lack of available information in the FARS police reports to determine whether the crash involved an underride. To help alleviate this problem, it was suggested that the interview skills of FARS analysts be enhanced to help guide them in identifying and coding underrides.

2.2.4 *Transportation Research Circular 231: Truck Accident Data Systems: State-of-the-Art Report*

Transportation Research Circular 231 (TRB 1981) summarizes the proceedings of a workshop that addressed

1. Issues that should guide the collection of truck safety data,
2. Data available to address these issues,
3. Quality and completeness of available data, and
4. Potential sources of additional data.

It was concluded that meaningful data to support heavy truck accident causation studies existed in a variety of sources. However, being able to locate, verify, and collate such data was considered challenging. Two important areas of data deficiency were noted as (1) the role of economic factors in truck operations and driving practices and (2) coarse categorization of truck accident data.

2.2.5 *Accident Models for Two-Lane Rural Roads: Segments and Intersections*

Accident Models for Two-Lane Rural Roads: Segments and Intersections (Vogt and Bared 1998) describes the collection, analysis, and modeling of accident and roadway data pertaining to segments and intersections on rural roads in the states of Minnesota and Washington. A comprehensive review of data quality was performed as part of this effort. This included comparisons of values of multiple variables for consistency and flagging unusually large values of variables. Of particular interest were findings that there are inconsistencies in how attributes are defined in different accident databases as well as variations in reporting thresholds, making it difficult to conduct direct comparisons. The authors also note that the reliability of reported accident characteristics depends on the acumen of the report officer/official and witnesses.

2.2.6 *The Human Factors Analysis and Classification System—HFACS*

With human error cited as the cause of the vast majority of civil and military aviation accidents, an argument is made that a more comprehensive accident analysis and classification framework for collecting data investigating human error is needed. HFACS was developed with this objective in mind. HFACS describes the following four levels of human failure:

1. Unsafe acts,
2. Preconditions for unsafe acts,
3. Unsafe supervision, and
4. Organizational influences.

Unsafe acts are comprised of *errors* (mental or physical activities of individuals that fail to achieve their intended outcome) and *violations* (willful disregard for the rules and regulations that govern safe operations). Errors are further subdivided into those that are *skill based, decision oriented,* and *perceptual,* while violations are segmented into *routine* and *exceptional.*

Preconditions for unsafe acts are based on the premise that unsafe acts are often symptoms of a deeper problem. Preconditions are divided into *substandard conditions of operators* and the *practices they commit,* respectively. Substandard conditions of operators are subdivided into *adverse mental states, adverse physiological states,* and *physical/mental limitations.* Substandard practices of operators are categorized as *crew resource management* and *personal readiness.*

Unsafe supervision traces causation of events back to the supervisory chain of command. Four subcategories of unsafe supervision are defined as

1. Inadequate supervision,
2. Planned inappropriate operations,
3. Failure to correct a known problem, and
4. Supervisory violation.

The final category, organizational influences, addresses the institutional culture and how the organization is structured to perform. Organizational influences are subdivided into the following three groups:

1. Resource/acquisition management,
2. Organizational climate, and
3. Organizational process.

It is argued that accident databases can be reliably analyzed using HFACS and, in doing so, objective, data-driven intervention strategies can be identified. The authors (Shappell and Wiegmann 2000) state that application of HFACS has been proven effective and the approach is now being utilized by multiple military and civilian organizations.

2.2.7 "Human Factors Root Cause Analysis of Accidents/Incidents Involving Remote Control Locomotive Operations"

"Human Factors Root Cause Analysis of Accidents/Incidents Involving Remote Control Locomotive Operations" (FRA 2005) documents a human factors root cause analysis (RCA) of six train accidents/incidents involving remote-control locomotive (RCL) operations in U.S. railroad switching yards that occurred in 2004. RCA used a modified version of the Human Factors Analysis and Classification System (HFACS), in which operator impacts, preconditions for operator acts, supervisory factors, organizational factors, and outside factors were defined as concentric category influences. Data collection and analysis tools included information gathered from participating railroads, interviews and surveys, travel to the accident/incident site, and the development of decision trees designed around the HFACS taxonomy. A total of 36 probable contributing factors were identified among the 6 case studies, from which several key safety issues emerged.

2.2.8 *Large Truck Crash Causation Study (LTCCS) Analysis Series: Using LTCCS Data for Statistical Analyses of Crash Risk*

The Large Truck Crash Causation Study (LTCCS) was undertaken jointly by FMCSA and NHTSA, utilizing a representative sample of nearly 1,000 injury and fatal crashes involving large trucks that occurred between April 2001 and December 2003. This report (Hedlund and Blower 2006) focuses on how statistical analyses of the LTCCS database can be used to investigate crash causes and contributing factors.

Within this context, data limitations are discussed. These include issues involving data accuracy and completeness. The authors conclude that variables that are directly observable by investigators are likely to be more accurate and complete, such as most vehicle and non-transitory environmental data. By contrast, variables that depend on interviews are more suspect in terms of accuracy and completeness (even if investigators have checked other sources to confirm the interview reports). An example of this latter consideration is whether the truck driver was in violation of the federal hours-of-service rules at the time of the crash.

2.2.9 *Highway Safety: Further Opportunities Exist to Improve Data on Crashes Involving Commercial Motor Vehicles*

The process for collecting, entering, and processing commercial motor vehicle crash data to meet federal reporting requirements involves several steps. Crash data initially are collected by local law enforcement then sent to the state for processing before being uploaded by the state into FMCSA's data system. The objective of this study (GAO 2005) was two-fold: to examine what is known about the quality of commercial motor vehicle crash data and what states are doing to improve it, and to evaluate the results of FMCSA's efforts to facilitate the improvement of the quality of commercial motor vehicle crash data submitted to the agency.

Sources of information utilized in the study included data reported by FMCSA; previous studies on the quality of commercial motor vehicle crash data; interviews with FMCSA officials, developers of FMCSA crash data tools, commercial vehicle industry researchers, and public interest organizations; grant documentation for 34 states that participated in FMCSA's safety data improvement program in fiscal year 2004; case studies of six states that participated in that program; and interviews with states that had not participated or were no longer participating in the safety data improvement program.

Overall, GAO concluded that commercial motor vehicle crash data do not yet meet general data quality standards of completeness, timeliness, accuracy, and consistency. More specifically, for fiscal year 2004, nearly one-third of commercial motor vehicle crashes that states are required to report to FMCSA were not reported and, of those that were reported, there were problems with accuracy, timeliness, and consistency (e.g., 15% of crash records reported to FMCSA could not be matched to the carrier's DOT number). Data quality problems most often stemmed from errors or omissions either by law enforcement officers at the scene of a crash or in the processing of crash reports to a state-level database. Among the specific problems cited were the following:

1. Infrequent opportunities for officers to receive training on filling out crash reports,
2. Unfamiliarity with what and how to report that result from infrequent occurrences of commercial motor vehicle crashes in an officer's jurisdiction,
3. Competing priorities at the officer level (where safety is a higher priority than data collection at the crash scene),
4. Use of manual crash reporting forms (compounded when the commercial vehicle crash report is a supplemental form),
5. Complex processes some states use to transform a report into the FMCSA format, and
6. An overall lack of quality control during data entry.

To combat this problem, individual states are engaged in the following activities, utilizing federal funds allocated by FMCSA to support state efforts to collect and report commercial motor vehicle crash data:

- Analyzing existing data to identify problems and develop plans for addressing them,
- Reducing report backlogs that have not been entered into state-level databases,
- Developing and implementing electronic data systems for collecting and processing crash information (e.g., on-scene reporting using handheld computers), and
- Providing training on the definitions and criteria for commercial motor vehicle crashes and emphasizing the importance of data quality.

To date, improvements in both the timeliness and number of reportable crashes have been observed, as measured by FMCSA's data quality rating system. However, GAO found that this system contains some flaws that can mask the true effectiveness of crash reporting and made specific recommendations for how to address these shortcomings.

2.2.10 *In-Depth Accident Causation Data Study Methodology Development Report (SafetyNet)*

This report (Paulsson 2005) was prepared for the European Commission in order to develop a system for taking an independent, in-depth accident causation database and creating an investigation process that identifies the main risk factors leading to a crash. The main difference between the proposed and existing systems is that this system would be constructed from the ground up with the sole purpose of determining the causes of accidents, unlike the multitude of existing databases that have to be cross-referenced, when even possible, to accomplish this objective.

One major concern that this report recognizes is the need for accurate and consistent data. To address this concern, the report recommends conducting interviews and issuing questionnaires to confirm all aspects of an incident as well as implementing systems to review the procedures that data collectors are using at crashes.

2.2.11 *Comprehensive Safety Analysis 2010: 2006 Listening Session*

This listening session (Coray Gurnitz Consulting and Abacus Technology 2007) enabled participants to supply ideas on how FMCSA could improve its commercial motor vehicle safety compliance and enforcement programs. Among the suggestions made were the need for higher quality data (including crash causation determination and type of accident), consistent data submission and enforcement across states, and making the data visible immediately after it is submitted.

2.2.12 *Safety Report: Transportation Safety Databases*

This report (NTSB 2002) evaluated the data quality issues of the many external databases used to perform accident investigations, safety studies, and special investigations. The main purpose of this report was to identify information gaps and establish data quality standards to ensure compatibility between databases and increase the usability of these databases.

Aside from developing a new database that would contain all of the necessary information for the various analyses, NTSB felt that it was most important to modify existing databases to be more compatible with each other (namely the NASS, FARS, and state databases), and improve the accuracy or completeness of submitted information (many databases have fields for information that are not recorded by the data collector).

2.2.13 *Illinois Department of Transportation Crash Data Process Audit*

This report (Scopatz 2006) was compiled after a study team collected information about the processes the Illinois Department of Transportation (IDOT) uses to collect motor vehicle crash data. It was concluded that the current accident reporting system was not working well.

This audit was not conducted because of incorrect or incomplete information, but rather due to untimely information. Because of inefficient recording processes, IDOT was experiencing a backlog of nearly six months for reporting crashes to the necessary databases (FARS and MCMIS). Recommendations included reducing the number of unnecessary reports that are filed (for crashes that do not meet the FARS or MCMIS reporting requirements) and implementing electronic file transfer instead of printing out reports and hand keying them into the necessary database.

2.2.14 *User's Guide to Federal Accidental Release Databases*

This report (EPA 1995) focuses on the incompatibilities of the various federal hazmat databases hosted by agencies such as NRC, EPA, and DOT. It was concluded that it is difficult to evaluate the overall effect of an accident without gathering information from more than one database, which can be time consuming. It was recommended that, in the future, the databases be linked by key identifiers to give users access to all of the available information for a given accident.

2.2.15 *Comparative Risks of Hazardous Materials and Non-Hazardous Materials Truck Shipment Accidents/Incidents*

Although this report (Battelle 2001) is a risk assessment for hazmat and non-hazmat accidents, it includes discussion of the federal databases being used in hazmat root cause analysis. From reviewing these databases, the following recommendations were made:

- Standardize the definitions of what constitutes an accident, what accidents must be reported, and what information must be reported.
- Include common fields in various databases so that pertinent information can be shared and not duplicated.
- Implement electronic filing for the major databases to reduce any backlog time.

2.2.16 *Hazardous Materials Serious Crash Analysis: Phase 2*

This report (Battelle 2005) details the process that the study team implemented in order to develop a hazmat accident database by combining data from MCMIS, HMIRS, state police accident reports (PARs), and interviews of carriers involved in the accidents. By joining these data, a higher level of understanding about the details and possible cause of the crash were obtained. This would not have been possible by using only the MCMIS or HMIRS database because of often-missing or inaccurate data. For instance, the MCMIS database classified 569 crashes as accidents involving hazmat Class 3 cargo. Once this was combined with the other sources, only 465 of these actually contained hazmat Class 3 cargo, and 69 of the 569 crashes did not even represent hazmat shipments.

2.2.17 *Unified Reporting of Commercial and Non-Commercial Traffic Accidents*

The objectives of this study (Shupe Consulting 2001) were to document the current business processes, forms, and data used for accident reporting in South Dakota and on national databases, and to develop a design specification for implementing a single system that could record, manage, and track accident information. It was concluded that the existing system was not well integrated with national databases (conflicts with state and FARS reporting), needed greater analysis capabilities, was time consuming to support (too much dependence on manual entry), contained inaccurate data, and lacked user accessibility. In proposing an improved system based on electronic data entry, it was recognized that challenges remain with redesigning crash report forms, establishing uniform reporting policies and procedures across the state, and providing adequate accident data collection training for law enforcement officers.

2.2.18 "Crashes Involving Long Combination Vehicles: Data Quality Problems and Recommendations for Improvement"

The author (Scopatz 2001) performed a study for the AAA Safety Foundation to identify barriers to analysis of longer combination vehicles (LCVs)—doubles and triples operating on our nation's highways. The states of Florida, Idaho, Nevada, Oregon, and Utah participated in a review and evaluation of their data collection and analysis practices. Oregon and Utah also participated in an audit of completed crash reports for crashes involving LCVs. It was concluded that none of the five states had a crash reporting system that adequately supports an analysis of LCV safety. Of particular note was a lack of reliable data on the specific configuration of vehicles involved in crashes. The report also contains recommendations for improving the quality of data for crashes involving large trucks and a state's ability to analyze LCV crashes.

2.2.19 "Use of Emerging Technologies for Marine Accident Data Analysis Visualization and Quality Control"

This paper (Dobbins and Abkowitz 2009) focused on performing analyses of allisions, collisions, and groundings on the inland waterway system. (Note: a vessel collides with another moving vessel but allides with a fixed object such as a bridge.) The source of accident information was U.S. Coast Guard marine casualty data from 1980 through mid-2007. During that time, the Coast Guard transitioned between three major system designs. The authors found significant quality issues with the U.S. Coast Guard accident data, specifically reporting inconsistencies among regions, missing data elements, and inaccurately reported information (including geographic location). Visualization using satellite imagery (in programs such as Google Earth) proved valuable in validating accident locations and understanding how the characteristics of each location may have contributed to accident causation and consequence. Recommendations are made as to how emerging technologies can be meaningfully applied to marine casualty data validation and analysis.

2.3 Summary of Findings and Implications

A review of relevant transportation accident data collection and analysis literature over the past three decades reveals some important findings and implications regarding the current state of the art of root cause analysis. These can be conveniently separated into recognition of problems and proposed solutions. These considerations are addressed, in turn, in the following discussion. It should be noted that a disproportionate number of prior studies has focused on the truck mode. However, where other modes have been considered, the findings and implications are remarkably consistent.

2.3.1 Data and Analysis Problems

As early as 1981, there was acknowledgement that analyzing transportation safety using empirical accident data was problematic. Beginning then, and continuing to the present time, numerous studies have cited the following five basic problems:

1. Inconsistent reporting practices within and across regions,
2. Non-reporting of reportable accidents,
3. Missing information in accident report records,
4. Inaccurate information included in accident report records, and
5. Data elements needed for root cause analysis not appearing on the report form.

A variety of reasons have been provided for why these problems exist, most notably

- Low law enforcement priority of data collection at the accident scene when compared with protecting public safety;
- Lack of understanding of how to complete an accident report involving vehicles hauling hazardous materials due to the low frequency of filling out these reports for police and many carriers;
- Reliance on manual data entry;
- Different reporting forms used by entities to serve different interests; and
- Disagreement or misunderstandings regarding the definition of terms.

Whatever the case, until these problems are adequately resolved, the ability to perform highly effective root cause accident analysis will be compromised.

2.3.2 Solutions Being Implemented or Under Consideration

Fortunately, the literature also contains suggestions and indications that some progress is being made in addressing data quality problems. Among the strategies being implemented or under consideration are the following:

- Posting available accident data on the Internet for review and feedback regarding its authenticity;
- Designing standardized accident reports toward a goal of more uniform data collection;
- Making extensive use of electronic data entry;
- Using visualization technologies to more precisely locate where an event occurred;
- Having data collection requirements influenced by available root cause analysis methodologies (e.g., HFACS, Haddon Matrix);
- Using sampling techniques to target certain types of accidents;
- Including common identifiers in complementary accident databases so as to integrate key causal information while avoiding duplication of effort; and
- Providing better training for law enforcement officials and other data entry personnel to enable them to collect and process information in a consistent, complete, accurate, and more timely manner.

Many of these strategies offer considerable potential, and are among those that were given careful consideration in the hazmat root cause analysis study.

CHAPTER 3

Summary of Interviews with Carriers, Shippers, and Database Managers

Carriers, shippers, and accident database managers were interviewed to ascertain their knowledge of reporting requirements, learn how they investigate accidents, and gather their suggestions for improving accident databases to support root cause analysis. Database managers were asked to discuss process changes or new initiatives for the databases they are managing.

3.1 Introduction

To learn about their internal process for investigating accidents and determining their root causes, carriers and shippers were questioned about their knowledge of accident databases. Representative examples of the questionnaires are included in Appendix A (available on the TRB website at www.TRB.org by searching for HMCRP Report 1). Appendix B (also available on the TRB website at www.TRB.org by searching for HMCRP Report 1) contains the results of this effort. In several instances, based on the wishes of the interviewee, the questioning became more "free form" and did not precisely follow the questioning order. Nevertheless, the interviews provided the project team with valuable insights into the carrier or shipper's process for identifying the root causes of hazmat accidents.

Table B-1 in Appendix B (available on the TRB website at www.TRB.org by searching for HMCRP Report 1) displays shipment and operator information for the 13 carriers who responded to a request to complete an interview from the National Tank Truck Carriers (NTTC) and the three extremely large carriers that were interviewed via telephone. Table B-2 in Appendix B displays the freight carriers' responses to questions involving a hypothetical hazmat accident in which the vehicle drove off the road. The names of the carriers have not been included in the table in order to protect the confidentiality of the respondents. Each carrier response is grouped in somewhat arbitrary categories based on the number of power units operated by the company. The following are the categories used for the table:

- Small, less than 100 power units;
- Medium, 100 to 299 power units;
- Large, 300 to 499 power units;
- Very large, 500 to 999 power units; and
- Extremely large, 1,000 or more power units.

The carriers are listed in the order that the questionnaires were received by the researchers. Appendix B also includes the text of an interview with a major water carrier.

3.2 Summary of Responses from Carriers

When carriers experience a hazmat accident, all of the details and events of the accident are recorded thoroughly. Several of the responding companies maintain accident databases that contain information that is much more extensive than the information that is required by the federally maintained databases. The carriers that attempt to get to root causes of the accident utilize these more extensive data sets. The investigators record environmental factors and long-term qualitative data that would be helpful in understanding how the hazmat accident occurred, and in determining how this type of accident may be prevented (if prevention is possible). In some instances, factors such as driver criminal history, crash history, and cell phone usage would have helped determine whether the accident was due to the driver, which, if true, could result in an action taken to discipline or suspend the driver. In one case, corrective action was taken by a company to make the driver more aware of these external factors, enabling the driver to correct for them and thereby prevent future accidents. On the other hand, if factors such as existing traffic/weather conditions and functionality of trucking equipment indicate that the fault of the accident was external to the driver, a change in driving procedures might be made.

Carriers would also like to see PHMSA play a more active role in communicating with the companies that reported the accident in order to get complete and accurate information. Most of the companies said they would use training offered by PHMSA to better fill out the crash form. More accurate forms will add to a more thorough database. This will help greatly for those that wish to use the database for research in prevention of serious hazmat accidents and mitigating crash impacts.

The questionnaire results provided some insights into how carriers think accident causation analysis should be performed. Although carriers conduct their own investigations of major hazmat accidents and search for the "root causes" of their crashes, they also believe that the authorities have a responsibility to do the same. Carriers suggested that the following steps should be taken during an investigation:

1. Obtain vehicle operator statements of evidence for the hazmat accident.
   - Carriers think that both they and the authorities should collect as much information as possible in order to identify accident causation. This includes collecting witness statements, consulting police reports, and reviewing the driver's log history, license records, and records of violations. Determining when the driver last rested would be especially valuable.
   - The driver's cell phone use and satellite tracking records should be accessed to collect additional information.
   - Driver's actions should be compiled in order to look for causal events.
   - Several carriers believe that assigning fault for a crash to a particular driver would be beneficial in that often the truck driver is not responsible for causing a crash.
   - A large water-based carrier contacted by the research team investigates accidents and requires the pilot (or tanker man) to complete a special investigations form they have devised to obtain information that could lead to the identification of the root cause for the accident under investigation. Employees receive training in the proper procedure for completing this form.
2. Search for defects in the vehicle.
   - As part of this effort, vehicle equipment (such as brakes) should be examined during a post-crash inspection.
   - Pertinent maintenance records should be reviewed for any insights into the cause of the crash.
3. Examine vehicle operator history in order to identify health problems that could have contributed to the accident.

- One extremely large motor carrier looks at a driver's log book (to try to identify fatigue factors) and checks the driver's performance history as part of their internal investigation.
- Some carriers examine a vehicle operator's health history for indications of such conditions as sleep apnea, diabetes, or alcoholism.

4. Examine roadway/runway geometry for accident causes.
   - Road conditions including weather, obstructions, and traffic flow should be part of the investigation.
   - The infrastructure configuration such as lane width, curves, and slopes can also provide indications of crash causation.
5. Take pictures of the accident scene.
   - Carriers recommend photographing the crash scene, including images of the roadway, vehicle positioning, crash damage, spill location, and any environmental damage. These pictures can be used to assist in assessing crash causation.

3.2.1 Carrier Satisfaction with HMIRS

The great majority of the carriers were satisfied with the process of filling out the PHMSA's HMIRS, although two of these companies actually used a separate spill center to fill out the form. Most companies think that PHMSA's criteria for filing the hazmat report is clear. One suggestion called for the form to be updated in order to increase clarity. Nevertheless, seven of the carriers currently provide training for completing the 5800.1 report. The great majority of these carriers believe that PHMSA-provided training would be useful for their staff completing the 5800.1 report. For those that would take advantage of such training, if offered, suggestions included online training/web conference, by CD-ROM, seminars/classroom setting, and training similar to that given to NTSB inspectors. One carrier believes that PHMSA should consider providing training in identifying root causes of accidents. In this regard, some of the techniques used by the NTSB investigators would be valuable for the carriers as they seek answers to the causes of their own crashes.

Most of the companies do not feel that additional data should be added to HMIRS. One exceptionally large company pointed out that for less-than-truckload cargo, there is often more than one type of hazmat being transported. Consequently, they suggested that there should be provisions in the form for listing more than one type of hazmat. Furthermore, the responses indicated that sometimes there are undeclared hazmat shipments found on vehicles and there should be a provision in the 5800.1 report for describing these as hazmat. (Note: provisions for handling both of these situations were added after 2004, and several other carriers take advantage of these provisions.) If an undeclared shipment is detected, there is an "Undeclared HM Shipment" box to check when filling out the form. If there are multiple, less-than-truckload shipments, many carriers check the "Additional Pages" box and complete the form for each of the partial shipments. The responses to include provisions that already exist indicate that there is need for additional training or clarification in the instructions for filling out the form.

A medium-sized company said they wanted more detail in Section 6 of the 5800.1 form, including whether other parties and other environmental factors (weather, road conditions, obstructions, fatigue, maintenance history, hazmat training, hazmat experience, age of equipment, other human factors) were involved. Those carriers that recommended changes to HMIRS also suggested that it be reworded in terms of carrier industry terminology.

Most carriers think that PHMSA has an obligation to contact carriers who do not complete their 5800.1 report properly. They believe that PHMSA should contact the carriers by telephone, letter, or e-mail. Carriers would also like to see PHMSA play a more active role in communicating with the companies involved in the accident in order to get more complete and accurate information. They believe that improved data in the database will enable them to more effectively use the database to detect contributing factors being experienced by other carriers and thereby address those factors in their own safety training programs.

3.2.2 Carrier Satisfaction with MCMIS

Although most carriers do not use the MCMIS crash data, of those that had, the following suggestions were made:

- Add specific information as to which party was responsible for a crash. Distinguish between an accident that was preventable and one that was not preventable.
- Determine the specific cause, as well as contributing circumstances, as determined by investigators.
- Designate repeat offenders in the crash reports.

3.3 Shipper Responses

Two interviews were conducted with two major shippers of chemicals in the United States. The following summarizes some of the major points and recommendations made by the officials.

3.3.1 Shipper 1

Shipper 1 is an extremely large corporation that conducts formal investigations of accidents involving its hazmat shipments. It is a standardized process that includes an auditing component. There are several triggers for conducting a formal investigation, including severity and potential consequence (e.g., how much leaked, type of hazard involved, injuries, media attention, and traffic shutdown).

When an incident occurs while the shipment is in the custody of a carrier, CHEMTREC is immediately notified. The distribution leader at the shipper's plant site where the shipment originated creates an incident report. Based on the contents of the report, the incident is classified as one of the following:

- Category A, warranting CEO attention,
- Category B, investigated within 24 hours with investigation led at the vice president or director level, or
- Category C, investigated within 72 hours.

Also note that the size of the investigative team increases at each higher level of review.

For Category A and B events, the incident investigation may include on-site data collection, but only if deemed necessary for the shipper to have confidence in knowledge of the situation. The results of an investigation are recorded in a database. Among the data elements contained in an incident record are fields for both Surface Cause (e.g., transportation accident due to human error) and Root Cause fields. The interviewee estimated that they identify the correct root cause in 70% to 80% of the cases. When problems arise, it is usually associated with the carrier's involvement in the process. It was mentioned that bulk carriers tend to provide better reporting than less-than-truckload (LTL) carriers.

The outcome of an incident investigation or an analysis performed on multiple incidents in the database is a list of recommended action items to implement in mitigating future risks. Examples of such actions include more targeted training and auditing.

It also was suggested that because of recognized issues related to HMIRS data quality, a credible root cause analysis should be performed by PHMSA. This could be achieved by implementing a verification protocol whereby PHMSA/DOT conducts follow-up audits of those accidents meeting a certain severity threshold. From the shipper's perspective, a desirable criterion would be the material hazard, with a focus on TIH (toxic inhalation hazard), flammable, and reactive materials.

Although interested in doing so, this shipper has not devoted the time or resources to investigate the merits of using mode-specific hazmat accident data (e.g., MCMIS, RAIRS, MISLE) for performing risk analysis. This interviewee was not willing to provide an opinion regarding the potential value of these databases when investigating root causes of accidents.

3.3.2 Shipper 2

Shipper 2 also is an extremely large corporation. They make approximately 40 million placarded shipments annually, of which about 50% are bulk.

Several triggers have been defined that warrant the launching of an accident investigation. Examples include whether a spill occurs, a personal injury is involved, an evacuation is ordered, or if the potential existed for a major impact. In such instances, CHEMTREC is to be immediately notified and an internal accident investigation file is opened.

When a trigger is met and the shipment is in the custody of the carrier or logistics service provider (LSP), the carrier or LSP is responsible for leading the accident investigation. Shipper 2 may be a part of the carrier's (LSP's) accident investigation team. Regardless, the company expects to be kept apprised of the investigation and updates its internal investigation file accordingly. The updates are entered into the database and tracked through an event-in-action tool (ENAT). Shipper 2 personnel are highly trained in this aspect of data input and analysis. Moreover, the shipper has established modal experts (warehouse, road, rail, bulk marine, terminal) to assist in collecting and evaluating relevant information. This data collection process and repository have proven to be very important and, in many instances, demonstrates that what is reported to CHEMTREC does not align with what really happened.

It is important to note that the shipper's philosophy on the need for establishing root cause has evolved over time. Although in the past there may have been an emphasis on establishing an "ultimate" root cause, the company now recognizes that the true root cause may be a combination of factors that collectively lead to accident occurrence and impact severity. This approach also helps in being able to identify a control point (or points) where improvements can be made.

Regarding the use of outside databases, Shipper 2 echoed other stakeholder sentiments that HMIRS cannot be relied upon to provide credible information. The company believes that the problem of inaccurate reports that appear in HMIRS is more significant than accidents that go totally unreported. Some fields are notoriously unreliable, but of great analysis interest, such as the type of emergency response. Shipper 2 believes that the inaccuracies and missing elements that appear in HMIRS could be corrected as accident investigations proceed, but that HMIRS records are rarely updated once originally filed.

Given these circumstances, the extent to which Shipper 2 utilizes HMIRS is to identify accidents that should have been reported to them and were not or, vice versa, to identify accidents that were reported to them but do not appear in HMIRS. The American Chemistry Council aids in this process by providing company-specific HMIRS reporting records to its member companies on an annual basis. Although Shipper 2 does not routinely use modal-specific accident databases (e.g., MCMIS, RAIRS, MISLE), its general impression is that inconsistencies exist between accidents reported in these databases and what appears in HMIRS. The company feels that for all rail incidents, RAIRS data is pretty reliable, whereas it prefers to go directly to truck carriers for accident data rather than rely on MCMIS.

Regarding ways to improve HMIRS, Shipper 2 believes that the highest priority should be to make the data contained therein more accurate and consistent, only adding more reporting elements if they have a direct connection to establishing root cause. Improved accuracy could perhaps be accomplished by having PHMSA request that the reporting entity update the HMIRS record after a certain amount of time has elapsed since the date of the accident. Techniques to improve data consistency might be to utilize tools that facilitate more automated data entry (e.g., Web-based data entry) and to employ checks and balances in judging whether reported information makes logical sense. Part of the problem with reporting, the company believes, is a lack of education on the part of shippers/carriers/LSPs in terms of the importance of filing accurately with HMIRS, how best to accomplish this task, and pitfalls to avoid. To that end, PHMSA could provide better training to these stakeholders.

Although perhaps a bit removed from the task at hand, the shipper's vision for HMIRS is that motivation for reporting should be based on industry desire to put improvements in place to mitigate future risk rather than feeling obligated to report in order to achieve regulatory compliance. To that end, the company would recommend that the federal government produce data that not only accentuates failures, but successes as well.

3.4 Interviews with Database Managers

The research team also conducted interviews with selected agencies that are responsible for managing key databases. The agencies include PHMSA, FMCSA, and FRA. The complete interviews are found in Appendix B (available on the TRB website at www.TRB.org by searching for HMCRP Report 1).

The discussion below presents a summary of findings deemed most significant.

3.4.1 Interviews with Agencies Maintaining Databases (PHMSA)

The PHMSA officials indicated that reports go right into the database, which includes high-level quality control processing. They employ character-to-character checks to ensure that their process translated the paper form properly.

During this process, they examine the form for personally identifiable information (PII), business rule inconsistencies, invalid dates, and invalid commodities (by cross-checking with the commodities in the database). To determine if they were caused by the hazardous materials, fatalities and injuries are validated by PHMSA using their own subprocess. Additional checks include cases such as when the report shows that 5.5 gallons were spilled from a 5-gallon container. PHMSA will go back and ask the filer whether there were multiple packages that failed and request that they file a supplemental report or they will sometimes use an e-mail reply as confirmation to correct the data themselves. They look for city/county inconsistencies. If a shipper is an individual, they will not put that name in the database.

The PHMSA official believes there would be a benefit from more verification, not just form-based validation. Sometimes important information such as costs, injuries, or other important information is left off. Sometimes this is because many big companies hire spill centers to complete the forms for them. Rather than checking all submittals, the officials believe that checking the significant or serious incidents would provide the most benefit.

With respect to underreporting, PHMSA staff use a Web crawler and look for incidents, which they put in the HMIRS. In addition, they match things up with the telephonics [National Response Center (NRC) record]; if there is no match, it is flagged. They wait 60 days and then marry up the unreported incidents (URIs) to telephonics.

Both highway and rail show a greater incidence of URIs than expected by the modal distribution of incident reports. For example, 91% of all URIs are highway incidents whereas 86% of all reported incidents are highway incidents. For rail, it is 9% and 4%, respectively. Air has nearly 0% of the URIs (only 1 incident), but 10% of the reported incidents.

The official was not sure as to which mode had the most complete incident reports, but indicated that many highway incident reports were not completed.

The PHMSA official had the following suggestions for improving data collection effectiveness and quality.

1. More companies should report online to reduce errors and
2. More business rules should be used in online tools so a filer could not submit an inaccurate report.

Another official indicated that there are two aspects to the reporting requirements, the regulations and the report itself. The rulemaking aspect is an impediment, primarily because it is the rule itself that specifies who has to report. This official said it took 10 years to change the form the last time, making sure all stakeholders were heard, etc. To simply change the form itself, all that is required is to go through Office of Management and Budget's (OMB) information collection procedures, which include 30-day and 60-day notices and the justification required by the Paperwork Reduction Act. During the conversation, this interviewee determined that if the specification of *who has to report* was added to the form itself, PHMSA would no longer have to go through the rulemaking process to make changes in the form and that they could do so more often and more quickly.

3.4.2 Interviews with Agencies Maintaining Databases (FMCSA)

The interview below was conducted with a key administrator responsible for the management of the MCMIS database at FMCSA.

When accident reports are received, states upload crash reports through SafetyNet. States extract the data, either through an automated system or manually. That is, the data can be extracted using a computer program, or the cases can be keyed in directly. Certain fields are mandatory, such as carrier name and address. All fields are required, although blanks in non-mandatory fields do not result in rejecting the case.

FMCSA evaluates the accuracy of the submitted records through the following:

1. Use of a data quality module,
2. UMTRI evaluations of the completeness and accuracy of the MCMIS Crash data,
3. On-site data reviews,
4. NISR (contractor) evaluations of state crash report forms for compliance and accuracy,
5. NISR evaluations of state extraction logic and methods, and
6. Crash data collection training for enforcement personnel.

In addition, FMCSA utilizes the State Data Quality Improvement Program to help insure completeness and accuracy. The program includes the following:

1. Independent evaluation of the completeness and accuracy of the MCMIS Crash data,
2. On-site (at the state) reviews of state processes,
3. Evaluation of the accuracy and sufficiency of state crash forms to collect the MCMIS data,
4. On-site (at the state) evaluation of the data extraction logic and methods,
5. On-site (at the state) training for enforcement and other personnel, and
6. Three-day Data Quality and Training Conference in San Antonio for representatives of all the states.

FMCSA has the following suggestions for improving the effectiveness and quality of data collection, and in all cases the administrator believes that suggestions must be tailored to individual states:

- Provide continuing funding and technical assistance to the states.
- Improve data collection by police officers by providing in-depth training by qualified individuals.
- Improve data handling and processing at the state level by developing programs to meet individual state needs.
- Improve data handling and processing at FMCSA by recognizing and correcting system bugs immediately.
- Maximize electronic data collection and processing, as well as integrating other databases such as driver history, CDLIS, etc.

When asked if additional training was needed for MCMIS, the administrator answered *yes* but stated that FMCSA already has an extensive training program. They attempt to train personnel at each step of the process of data acquisition, from the officer who collects the data (through direct training, visor cards for police vehicles, and train the trainer) to the state personnel who extract and upload the data, to FMCSA personnel who prepare the file.

When asked if improvements should be made to MCMIS to improve data collection for hazardous materials crashes, the FMCSA database administrator responded that it would be unrealistic to collect any additional hazmat data such as quantity and package type than is currently collected.

3.4.3 Interviews with Agencies Maintaining Databases (FRA)

The interview was conducted with an FRA official responsible for administering the RAIRS database. When questioned about whether the RAIRS database includes the consideration of root cause analysis and/or root cause releases, the official responded that the RAIRS database and normal accident reports have had limited use for root cause analyses (RCAs) because they are "event" reports rather than detailed investigations of specific incidents. Their principal purpose is to enable accumulation of a statistically valid database on accidents for analysis of historical trends. The official added that about 150 accidents per year are subject to more intensive investigation out of a total of about 3,000 annually. Nevertheless, even these more detailed analyses are not fully developed, in-depth RCAs. In 2008, these reports were made available online, in addition to databases which have been available for many years.

There are also some concerns about the consistency between railroads in how they interpret primary and secondary causes. Even within a railroad, there can sometimes be problems. Such inconsistencies can interfere with, confound, or complicate analyses.

The official pointed out that root cause analysis now falls within the new Risk Analysis Division. Although the risk analysis initiative is new, it may lead to suggestions for changes.

When asked about barriers, either institutional or other, to implementing these changes, the FRA official replied that one barrier relates to regulations. There is a statutory limit to what railroads must report so significant changes are not easily implemented. The two paths for significant changes would be via rulemaking process or the Rail Safety Advisory Committee (RSAC).

When questioned, the official pointed out that there have been suggestions external to FRA to change or improve the database with respect to root cause analysis. They have received inquiries from labor about more detail regarding operations data such as RCL (remote-controlled locomotives). NTSB has suggested expanding the cause codes. There is a notice of proposed rulemaking (NPRM) (Docket No. FRA–2006–26173) calling for technical clarifications, expanding the scope of the instructions and improving certain definitions, proposing some new accident cause codes and collection of some additional data.

In order to provide information to inspectors in a more timely manner, FRA is constantly monitoring railroad compliance with reporting requirements.

When questioned about suggestions for improvement or changes that could be made to the database, the FRA official responded that FRA is especially looking for ways to improve turn-around time. Currently, they use a monthly "batch" process. The vision is for continuous flow of information into the database, enabling rapid detection of trends or evidence of potential problems. Presently, there is about a two-month delay. Another area that FRA would like to see better utilized is reporting of incident location using latitude/longitude (lat/long) coordinates. The record layout permits that, but compliance is voluntary so it is inconsistently reported. The FRA Geographic Information System (GIS) group is developing a linkage between lat/long coordinates and linear locations along rail lines. The official also cited a problem relating to yard-switching miles that are not recorded directly, but estimated based on person-hours worked by the crew. This has a potential impact on the reliability of this parameter for normalization of accidents.

3.5 Summary of Findings from Interviews

Many carriers and shippers, particularly the larger ones, have a formal process that triggers a graded response to accident investigation when an employee reports that he or she has been involved in an accident. For the more serious accidents, supplemental information is obtained in an effort to identify the root and contributing causes of accidents. One company collects information from witnesses, reviews the driver's log, the driver's cell phone usage, the driver's actions during the course of the accident, inspects the vehicle for defects, examines the vehicle operating history, examines the roadway geometry, and takes pictures of the accident scene. Some believe they identify the root causes of accidents for between 70% to 80% of the accidents. In many cases, corrective actions are recommended. Carriers and shippers have a vested interest in preventing accidents, and many of the accident reports recommend corrective actions that will reduce the frequency—and perhaps the severity—of future accidents.

Based on interviews with the organizations maintaining federal databases, although there is a commitment to improve both the quality and completeness of the data, there has not been a significant long-term commitment to capture information that is capable of identifying root and contributing causes of accidents. The most relevant hazmat database, HMIRS, focuses on the adequacy of packaging standards. MCMIS and RAIRS have a broader focus than hazmat accidents and would require a major refocusing if they were to begin collecting the information required to identify the root and contributing causes of hazmat accidents. Potential measures for achieving this objective are included in both the discussions of the individual databases in Chapter 4 and the potential measures presented in Chapter 5.

CHAPTER 4

Database Analysis

This section analyzes the major crash databases and identifies fields that can provide answers to "why" questions in determining root cause. Each of the following major databases is included in this discussion:

- Motor Carrier Management Information System (MCMIS)
- Hazardous Materials Incident Reporting System (HMIRS)
- Fatality Analysis Reporting System (FARS)
- Trucks Involved in Fatal Accidents (TIFA)
- Large Truck Crash Causation Study (LTCCS)
- Railroad Accident/Incident Reporting System (RAIRS)
- Marine Information for Safety and Law Enforcement (MISLE)

In addition, the NTSB and *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005) approaches to accident analysis are described.

4.1 Motor Carrier Management Information System (MCMIS)

4.1.1 MCMIS Database Description

MCMIS includes four major files named Registration, Crash, Inspection, and Company Safety Profile. For this project, the Registration and Crash files are the most relevant. Although the focus of this analysis will be on the Crash file, when trying to identify the contributing and root causes of accidents, information in the Registration file can provide useful supplemental information. Specifically, the Registration file has carrier information on the number and configuration of vehicles, the number of drivers, annual miles driven, and accident rates. Such information is needed to determine the extent to which a class of accidents with similar contributing or root causes might be occurring annually. Vehicle miles traveled (VMT) helps estimate the risk of occurrence by providing a measure of exposure. Identifying the risk enables officials to target resources for reducing the number of accidents.

The design of the Crash file was developed in 1992 to record information on serious accidents involving a truck, bus, or light vehicle transporting hazmat. The types of information collected on each heavy truck or bus involved in these serious accidents has changed little since it was developed. The process of reporting serious accidents begins with a law enforcement official filling out a police accident report (PAR). The state agency responsible for filing the MCMIS crash report screens the PARs to identify serious heavy truck and bus accidents. Once the state reporting agency finds an accident that meets the requirements for reporting the incident to FMCSA, the information for the vehicle from the PAR is coded into the MCMIS Crash file format and electronically transferred to FMCSA or manually entered by the state agency for MCMIS Crash file inclusion.

The MCMIS dataset records are at the vehicle level, as opposed to the accident level. Therefore, if multiple trucks or buses are involved in the same crash, there will be one MCMIS crash record for each truck and bus. The vehicle focus means that there will always be more records than accidents. Information coded into the MCMIS Crash file and obtainable from the PAR includes the date, time, and location of the accident; a description of the vehicle; the name and address of the carrier; personal information, including driver licensure; parameters that describe the roadway at the accident location; and the accident impacts. Because the database contains driver information—name, address, sex, and birth date—the full MCMIS Crash file is publicly available but only after data that can identify individuals have been stripped out. However, since driver information was used in this project for various purposes, the project team signed a confidentiality agreement and obtained a version of the file that included the driver fields.

4.1.2 Location and Ownership of Data

The MCMIS Registration and Crash files are maintained by FMCSA and available only to authorized governmental users on a routine basis.

4.1.3 Database Format

Initially, the MCMIS Crash file contained only one table, but recently it has been divided into the four tables (*CRASH_MASTER, CRASH_CARRIER, CRASH_DRIVER,* and *CRASH_EVENT*) shown in Figure 4-1.

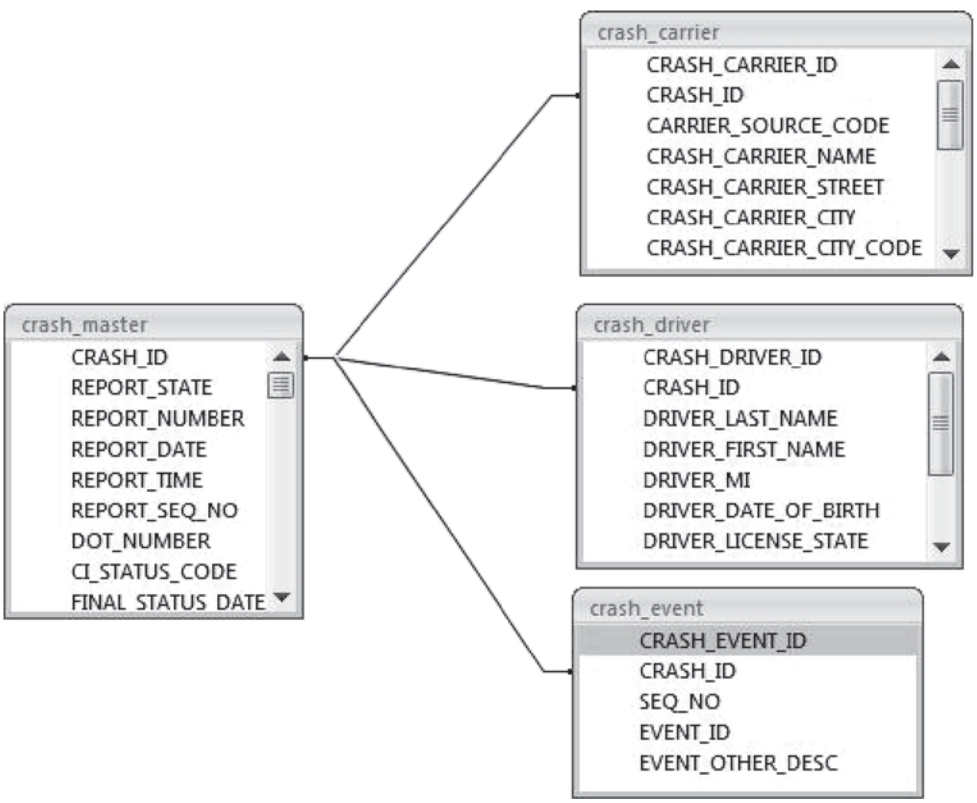

Figure 4-1. Tabular relationship in MCMIS Crash file.

The *CRASH_MASTER* Table contains information on the date and location of the crash, roadway type, road surface condition, weather conditions, and vehicle configuration. It also contains information indicating if a hazmat release occurred, the number of fatalities, and serious injuries (those requiring transport to a medical facility for treatment). Since the state is required to provide a report for trucks and buses involved in either intrastate or interstate commerce, this table has two *yes* and *no* fields to flag *FEDERAL_REPORTABLE* and *STATE_REPORTABLE* crashes. The *CRASH_MASTER* Table has five fields that are used when hazardous materials are being carried by the vehicle involved in the crash.

The *CRASH_CARRIER* Table lists the carrier name and address information. The *CRASH_DRIVER* Table provides the name, address, sex, and birth date of the driver. The table also contains licensure information. The *CRASH_EVENT* Table contains information on the accident sequence, beginning with the first dangerous event (e.g., crossing the median). These events do not consider pre-crash conditions such as speeding, slippery road conditions, poor visibility, etc., which are captured in other fields. There is a one-to-one relationship between the records in the *CRASH_CARRIER, CRASH_DRIVER,* and *CRASH_MASTER* tables. Although any number of events can be entered under a single vehicle incident listed in the *CRASH_MASTER* Table, for the majority of records, only one event is listed. In 2005, the maximum number of events associated with a single vehicle crash was four. (In the original structure of the data, up to four events could be coded.)

4.1.4 Threshold for Exclusion or Inclusion

The MCMIS Crash file was developed to capture vehicle data on all serious crashes of trucks and buses involved in commerce. A crash is considered serious if there is a fatality, an injury requiring prompt medical attention at a facility away from the accident location, or if one of the vehicles involved in the crash had to be towed from the scene due to disabling damage. Since the database reports vehicle involvements, if two qualifying vehicles are involved in an accident, there will be two records, one for each qualifying vehicle.

4.1.5 Years of Coverage

DOT started to provide funding to states to report serious heavy truck and bus accidents in 1992. In the early years, the underreporting of accidents was high. Although it has improved in recent years, the non-reporting rate is currently estimated to be about 20%. Some states have a much higher reporting rate than others.

4.1.6 Types of Fields Covered

The *CRASH_MASTER* Table contains information about the location of the crash, the vehicle configuration, the type of roadway, and the weather at the time of the crash. It also contains the DOT number of the carrier and crash consequences (number of fatalities, injuries, and if it was a tow-away accident). Most importantly to this project, this table contains the following fields:

- *VEHICLE_HAZMAT_PLACARD,*
- *VEHICLE_HAZMAT_NUMBER,*
- *VEHICLE_HAZMAT_MATERIAL,*
- *VEHICLE_HAZMAT_CLASS_ID,* and
- *HAZMAT_RELEASED.*

The first and last fields are yes/no flags. The intent of the middle three is to provide the UN number, the name of the hazardous material, and the one-digit hazard class number, respectively. The intent of these three fields is to provide information consistent with the

172.101 Hazardous Materials Table in 49 CFR Part 172. There is no requirement that the name of the material be taken from the Part 172 Table and, as a result, the trade name of the hazardous material is often entered.

The *CRASH_CARRIER* Table contains information on the carrier and the *CRASH_DRIVER* Table contains information on the driver of the vehicle described in the *CRASH_MASTER* Table. The *CRASH_EVENT* Table can have any number of events for each entry in the *CRASH_MASTER* Table. This is where the event sequence for the accident is specified. Prior to the restructuring that occurred in about 2003, four separate fields named EVENT_1 EVENT_2, EVENT_3, and EVENT were provided for the reporter to describe the accident sequence following the initiating event. The event codes are shown in Table 4-1. For example, if the accident sequence was a collision with a motor vehicle in transit, followed by the vehicle running off the road, followed by the vehicle rolling over, there would be three EVENT_ID records entered and the entries for SEQ_NO 1, 2, and 3 would be 13, 1, and 3, respectively.

Note that these are all events and not conditions that would be useful for identifying contributing causes of the accident. Conditions would have to be inferred from the *CRASH_MASTER* Table in fields such as *WEATHER, ROAD_CONDITION,* and *LIGHT_CONDITION.* Although there is an accident description narrative in all PARs, this narrative must be captured in code numbers in the *CRASH_MASTER* and *EVENT* tables.

4.1.7 Database Purpose and Function

FMCSA is responsible for ensuring the safety of commercial interstate motor vehicle truck and bus safety. In order to ensure safety, the MCMIS Crash file has been developed to provide data on the number of serious truck and bus crashes that are occurring each year. From such information, FMCSA can monitor trends, evaluate the effectiveness of current regulations, and

Table 4-1. Event codes and descriptions.

Event Codes	Event Description
1	Non collision ran off road
2	Non collision jackknife
3	Non collision overturn (rollover)
4	Non collision downhill runaway
5	Non collision cargo loss or shift
6	Non collision explosion or fire
7	Non collision separation of units
8	Non collision cross median/centerline
9	Non collision equipment failure (brake failure, blown tire, etc.)
10	Non collision other
11	Non collision unknown
12	Collision involving pedestrian
13	Collision involving motor vehicle in transit
14	Collision involving a parked motor vehicle
15	Collision involving a train
16	Collision involving pedalcycle
17	Collision involving animal
18	Collision involving fixed object
19	Collision with work zone maintenance equipment
20	Collision with other movable object
21	Collision with unknown movable object
98	Other

provide information that can be used to develop more effective regulations. Note that the database was never designed to identify contributing and root causes of accidents.

4.1.8 Data Collection

Data collection for MCMIS is a complex process that begins with the police officer filling out and filing a PAR. These reports are compiled and sent to the appropriate state agency, where a determination is made of whether an accident involving a truck, light vehicle, or bus also met the crash severity definition and should be reported to FMCSA. For it to be classified as serious—and therefore reportable—the crash must either have resulted in a fatality, required that someone be transported to a remote facility for emergency medical treatment, or required that one of the vehicles involved in the crash had to be towed from the scene. Once it has been determined that the accident should be reported to FMCSA, the information is transcribed from the PAR either into an electronic file that is transmitted to FMCSA or manually entered through a Web interface. FMCSA then performs certain checks and enters the data into the MCMIS Crash database.

Complexity arises not only from the process of going from the PAR to the MCMIS Crash form, but also from the number of agencies and individuals involved in the reporting. A query of the individuals or organizations filling out the PARs in 2005 totaled over 60,000 entries for approximately 145,000 crash records entered (only about 2% of these are hazmat crashes). The exact number of individuals filling out the form could not be determined because many of the entries were organization names, not the names of the individuals filling out the form. If any one of the thousands of police officers filling out a PAR fails to record the value for a parameter or any of the state agency staff fail to transcribe a parameter value or transcribe it incorrectly, the data submitted in the MCMIS Crash file is incomplete or inaccurate.

4.1.9 Data Compilation

Over the years, efforts have been made to develop some standardization in the format used by each state and territory for its PARs. Although there have been some successes, there are still vast differences among the forms. Over time, there also has been an effort to keep the types of information reported in the Crash file consistent. Although the MCMIS file has been segmented into several tables from the one table used initially, the information requested has remained the same. This has enabled the states and territories to develop a standard protocol for translating the data in their PARs into the MCMIS format. Essentially, there remain 56 different translation protocols since each state or territory has a slightly different PAR. Although the MCMIS report file prepared by the state or territory is electronically transmitted to FMCSA, there appears to be little automation on the front end of the process. In the majority of the states and territories, the PARs are prepared by hand by the police officers and the translation of the PAR information into the MCMIS electronic file structure is also a manual operation. If the police officer filling out the PAR uses an abbreviated notation for the name of the carrier, unless the person transcribing the data into the MCMIS Crash file format realizes it is a shortened carrier name, the spelling of the carrier name listed in the PAR becomes the carrier name listed in the MCMIS Crash file.

4.1.10 Accuracy and Completeness of Data

Studies have shown that the complicated process of filling out PARs, identifying those truck and bus accidents that meet the definition of serious accidents, and then entering the data from the PARs into the MCMIS format is neither complete nor accurate. Over the last five years, the University of Michigan Transportation Research Institute (UMTRI) has been under contract to FMCSA to assess the accuracy and completeness of the MCMIS crash reporting system on a state-by-state basis. To date, 27 states have been evaluated. It is known that states are taking the evaluations into consideration in improving their systems and have instituted changes to correct the problems identified by UMTRI. The UMTRI program is just one facet of a comprehensive program at FMCSA to assist the states in improving their accident reporting. In general, significant progress has been made, but the completeness and accuracy of the data in MCMIS remains a serious issue. Moreover, there are many fields in the PAR that are either not filled out or not translated into the MCMIS format, and when fields that must be used to identify serious accidents are left blank or inaccurately filled out, then either serious accidents that should be reported are not reported or accidents that are not serious are filed because the inaccurate information provided makes them appear serious. Additional analyses reflecting on the accuracy and completeness of the data in the MCMIS Crash file are found in Appendix C (available on the TRB website at www.TRB.org by searching for HMCRP Report 1).

4.1.11 Identification of Hazmat Incidents in MCMIS

Table 4-2 lists some of the key parameters recorded in MCMIS and the percentage of the entries for which no information is presented in calendar year (CY) 2005. Overall, there is about a 20% underreporting rate. As shown in Table 4-2, the percentage of blanks in the MCMIS Crash file tables varies from zero to about 30%. For fields like *FATALITIES* and *INJURIES,* there are no blank entries because zero is entered if there were no injuries or fatalities. Similarly, there are no blank entries for Y/N fields such as *TOW AWAY*.

There is one parameter, *DRIVER CONDITION CODE,* which is left blank in all crash records after CY 2001. In 2001, the field was coded as "driver appeared normal" for about 94% of the crashes. It is the other 6% of the crashes where the contributing cause, or even the root cause, might have been "driver condition." Since there is a location on the PAR to enter this information and since a police officer is trained to observe a person's behavior, some weight can be assigned to the officer's opinion regarding the driver's suitability to be operating a motor vehicle. Since this is the

Table 4-2. Percentage of entries blank by parameter name.

Parameter Name	Percentage Blank
Carrier Name Provided for Each Incident	0%
Fatalities	0%
Injuries	0%
Tow Away	0%
County Code	1%
Driver Name Provided for Each Vehicle	2%
Vehicle Configuration	4%
Weather	7%
Light Condition	8%
Road Surface Condition	8%
Cargo Body	9%
No Event Sequence Provided for Each Incident	16%
Vehicle Identification Number	17%
Traffic Way	18%
Access Control	23%
Accident Location Adequate	23%
Vehicle Hazmat (Y/N)	24%
GVWR	26%
DOT Number	31%

only driver condition parameter included in the MCMIS Crash file, leaving this parameter blank is considered a significant loss.

For the remaining parameters that have less than complete coverage, either the PAR did not provide the value or the state staff person translating the information into the MCMIS crash format did not enter a value for the parameter. In *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005), PARs were obtained for all vehicles that were believed to be transporting hazardous materials in 2001, over 1,800 PARs. For almost all the parameter fields that were left blank in the MCMIS crashes in 2001, it was possible to fill in missing parameter values from the PARs. There were some notable exceptions. One state did not provide the commercial vehicle truck supplement to the PAR and, without that, it was impossible to fill in the vehicle configuration and gross vehicle weight rating (GVWR). That supplement also contained the hazmat data. If the staff person has access to all the PAR data, all the blank percentages would probably be less than 10%. The reasons for the information being lost are unknown.

The last four entries in Table 4-2 provide some statistics regarding how well the information is coded into MCMIS. The MCMIS Catalog, published on the FMCSA website, provides no standard regarding how to fill out the location field. The information provided is sufficient to locate the accident on a map only 23% of the time. These entries typically occurred at intersections or the location was specified as a route name or number along with the number of the nearest milepost. There are a variety of reasons for this low percentage. In many cases, only the route name is given and if the state and county are given, which is normally the case, then the best that can be done is to locate the accident on a route somewhere in a county through which the route passes. Evidently, at some point in the translation between the PAR and the MCMIS Crash file, a 30-character limit was imposed on the location field. There were numerous cases where the entry stopped mid-word and, as a result, truncated the milepost information needed to locate the accident on the route.

A comparison of the entries in the *CRASH_EVENT* and *CRASH_MASTER* Tables reveals that no crash event is provided for 16% of the crashes. Of the remaining 84% of the crashes with event sequences, 57% list one event, 13% list two events, 5% list three events, and 9% list four events. Based on these percentages, for 16% of the crashes, it is impossible to even identify the type of crash. For slightly more than one-half of the crashes, one event sequence is provided. Based on the statistics shown in Table 4-3, the use of one event to define the accident often can be justified. The dominant single-event sequence accident in which a truck is involved is coded as *EVENT_ID*=13, "collision involving motor vehicle in transit." This is the event code for 82% of the vehicle incidents with a single event listed in the *CRASH_EVENT* Table. The next most likely entry is "other," occurring 4% of the time. Table 4-3 lists, in order of decreasing percentages, the name of the single event followed by the percentage of crash records and the number of crash records listing this single event to describe the accident sequence.

In going down the list shown in Table 4-3, there are several significant features. If the number of single events that fall into the equipment failure category were totaled, although individually they each show a zero percentage, the combined number would be above 1%. There also are a fair number that could be better described using more than one sequence. For example, it is highly unlikely that striking an animal would cause sufficient damage to the vehicle to result in an injury to one of the truck occupants or result in the truck being towed from the scene, making it a serious accident that would be reported to MCMIS. Did striking the animal result in a jackknife or the truck running off the road or overturning? A two-element event sequence probably should have been used for that class of accidents. Overall, the biggest concern remains underreporting, closely followed by failure to enter parameter values for a significant fraction of the records.

Table 4-3. Crashes described by a single event.

Event Description	Percentage of Single Events	Number of Crash Records
Collision involving motor vehicle in transit	82%	73,591
Other	4%	3,358
Collision involving a parked motor vehicle	3%	2,462
Collision involving fixed object	3%	2,424
Non collision overturn (rollover)	2%	2,147
Collision with other movable object*	1%	1,074
Non collision other	1%	901
Non collision ran off road	1%	792
Collision involving pedestrian	1%	767
Collision involving animal	1%	606
Non collision explosion or fire	0%	352
Non collision jackknife	0%	308
Non collision cargo loss or shift	0%	299
Non collision equipment failure (brake failure, blown tire, etc.)	0%	254
Collision involving pedalcycle	0%	237
Collision involving a train	0%	134
Non collision separation of units	0%	96
Collision with unknown movable object	0%	70
Non collision downhill runaway	0%	44
Collision with work zone maintenance equipment	0%	35
Non collision unknown	0%	33
Non collision cross median/centerline	0%	31

*Previously "collision involving other object."

4.1.11.1 Overall Reporting Rating for the States

Table 4-4 provides a summary of state reporting rates for different factors for those states that UMTRI has evaluated. The rates are recorded as the percent of reportable cases that were actually reported, by crash severity (in the MCMIS scale). Table 4-4 shows that reporting rates have tended to improve in recent years and that the lowest rates are associated with earlier years.

For completeness, UMTRI has compiled the percentage of missing data for the primary variables. They also compare variables as reported in MCMIS to how they appear in the state file. Unfortunately, comparing the values doesn't reveal if the MCMIS information is accurate, just whether it is the same as what was reported to the state. To get a further handle on accuracy, it would be necessary to compare the information with an independent source. This is accomplished in the next section for fatal crashes using the TIFA database. In the TIFA survey, individuals are called and asked about the crash. These contacts include the driver, owner, safety director, and reporting police officer.

4.1.11.2 Comparing MCMIS with TIFA to Evaluate Crash Data Accuracy

The TIFA file maintained by UMTRI provides a unique opportunity to evaluate the accuracy of the data reported to the MCMIS Crash file. TIFA data are collected independently from the MCMIS Crash file data, using a different methodology. MCMIS Crash file data are extracted by the states from their crash data and uploaded to the MCMIS Crash file using the SafetyNet System. Some states extract the data from a supplementary crash reporting form, while others incorporate the required data on their primary crash report. Many states use a computer algorithm to identify reportable crashes, while others manually identify reportable crashes and extract the data. Whatever the method, the data originate with the officer responsible for filing the crash report.

Table 4-4. Reporting rates of states to MCMIS Crash file compiled from UMTRI reports.

State	Data Year	Overall Reporting Rate	Fatal	Injury	Tow Away
Alabama	2005	76.0	91.4	76.4	75.0
Arizona	2005	78.2	93.8	83.4	75.6
Connecticut	2005	Likely <30%			
California	2003	72.0	84.2	73.9	70.9
Florida	2003	24.0	55.6	26.5	20.0
Georgia	2006	68.1	78.8	68.4	67.4
Idaho	2006	72.9	92.3	90.5	60.7
Illinois	2003	43.0	71.0	42.3	42.6
Indiana	2005	80.5	90.3	81.9	79.6
Iowa	2005	71.6	94.1	86.4	61.4
Louisiana	2005	56.6	79.6	57.0	54.7
Maryland	2005	31.1	84.6	56.0	15.6
Michigan	2003	73.7	92.4	73.1	73.4
Missouri	2000	60.9	76.8	63.7	58.8
Missouri	2005	83.3	94.6	84.9	81.8
Nebraska	2005	86.8	100.0	82.0	82.7
New Jersey	2003	82.5	67.4	81.5	83.2
New Mexico	2003	9.0	27.5	11.0	6.8
North Carolina	2003	48.2	63.3	49.4	47.1
Ohio	2000	38.8	50.7	58.2	28.6
Ohio	2005	42.5	85.4	52.7	32.3
Pennsylvania	2006	77.0	91.7	74.5	77.6
South Dakota	2005	66.4	78.9	64.9	66.4
Tennessee	2004	51.3	93.5	54.8	47.4
Washington	2003	37.6 to 53.7	67.2	About 40	

In contrast, the TIFA protocol uses state crash reports, but primarily to identify persons with knowledge of the crash for a telephone interview. Interviewers typically contact the driver, owner, or safety director of the carrier for a detailed interview about the truck, driver, and carrier that operated the truck. If the driver or carrier can not be contacted or refuses to cooperate, the interviewer will contact the reporting officer, tow operator, or other witness. But the great majority of interview information comes from sources that actually operated the truck. In addition, all survey information is reviewed by experienced editors who decode the vehicle identification number (VIN), look up the manufacturer's original specifications for the vehicle to compare with interview information, and also consult a library of information on typical cargoes, trailers, and carrier operations. Information that is ambiguous or unusual is clarified by return calls to the original respondent. Rarely, if no other information is available, some limited descriptive information may be coded from the police report. But the overwhelming majority of information in the TIFA file is collected directly from the vehicle operators.

Because TIFA data are assembled by a completely separate process and entity, the TIFA file can serve as a relatively independent view of trucks involved in traffic crashes that are reported to the MCMIS Crash file. We say *relatively* because there are occasional instances where certain information may be coded from a police report. But there are good reasons for regarding TIFA data as reasonably accurate. Most cases have more than one source for the data (unless all information can be collected from a single source). Any ambiguous or contradictory information is clarified by further calls. TIFA editors have over 20 years of experience in working with large trucks, and they are knowledgeable about the variety of vehicles and operations. Finally, all cases are checked using computer algorithms for consistency, and to identify unusual cases for further review. The TIFA file is not free of errors, but there are multiple layers of checks to keep the error rate low.

Accordingly, the TIFA file can be used to check the accuracy of fatal truck crashes in MCMIS. The TIFA file includes only fatal crash involvements, significantly limiting the number of MCMIS records that can be checked. Nevertheless, since fatal crash involvements often receive the most detailed investigations by police, the records in MCMIS for fatal involvements arguably should be the most accurate. In that sense, the results of the comparison with TIFA represents the best-case scenario.

To perform the comparison, records in four years of TIFA crash data were matched to the corresponding records in the MCMIS Crash file. Crash involvements for 2002 through 2005 were matched, producing 11,914 records for comparison.

Table 4-5 shows the comparison of the identification of hazardous material in the MCMIS Crash file and the matching record in the TIFA file for the period from 2002 to 2004. In the TIFA file, the value recorded was whether the cargo body had amounts of hazardous materials sufficient to require a placard. In the table, cases where that information was unknown for all cargo bodies are counted as no hazardous materials. In MCMIS, the unknowns are shown as "(blank)," to indicate the value was left unknown. Valid entries for the variable in MCMIS are "Y" or "N." The one case with an "M" is considered to be a typo in which "N" was meant.

Table 4-5 shows some disconcerting patterns. Of the cases where the TIFA interview indicated the truck had placarded amounts of hazmat, the MCMIS hazmat placard was coded "Y" for 212, "N" for 172, and left blank for 50. There were also 95 cases where the hazmat placard was coded "Y" in MCMIS, but the TIFA interview showed that the truck did not have hazmat cargo. Thus, of the 434 cases in the TIFA file recorded as carrying hazmat, only 212 or 48.8% were coded as carrying hazardous material in the MCMIS file. Similarly, of the 307 cases coded as carrying hazardous material in the MCMIS file, only 212 or 69.1% actually had hazardous material.

One possible explanation for the discrepancy could be that the MCMIS variable captures trucks showing a hazmat placard, regardless of whether the truck actually had hazmat. In fact, there were 60 cases identified in the MCMIS Crash file with hazmat placard coded "Y" for trucks that were empty at the time of the crash. Another 34 had some cargo, but the TIFA interview showed that it did not include hazmat.

Comparisons also were made for several other variables. For truck cargo body and configuration, a comparison was made of the detailed code levels. The TIFA file includes more detailed types of cargo bodies and truck configurations than allowed in the MCMIS Crash file but, for the purpose here, both were aggregated to the levels permitted in MCMIS. Table 4-6 shows the distribution of

Table 4-5. Identification of hazmat cargoes in TIFA and MCMIS, 2002–2004.

TIFA Code	MCMIS Code	N
Hazmat Cargo	**Hazmat Placard**	**Total**
Yes	N	172
	Y	212
	(blank)	50
No	M	1
	N	8,484
	Y	95
	(blank)	2,900
Total		11,914

Table 4-6. Proportion of cargo body type in MCMIS coded correctly, based on comparison with TIFA data.

Cargo Body Type	Frequency	% Correct
No cargo body (e.g., bobtail)	251	49.0
Van	5,432	71.9
Flatbed	1,713	70.1
Tank	1,011	77.7
Auto carrier	77	68.8
Dump	1,947	65.5
Refuse	288	64.9

cargo body types from the TIFA file, the number of such body types in TIFA, and the percentage identified correctly in MCMIS.

Table 4-7 makes a similar comparison for truck configuration. The primary truck types, straight trucks with no trailer and tractor-semitrailers, are identified accurately 87.5% and 75.5% of the time, respectively. Less recognizable types like straight trucks pulling a trailer or bobtail tractors are less often accurately identified in MCMIS.

Finally, Table 4-8 shows the percentage of selected variables that are coded the same in TIFA and the MCMIS Crash file. Variables shown in Table 4-8 are drawn from the FARS file and not from the TIFA interview. GVWR class in MCMIS aggregates the classes to 1 to 2, 3 to 6, and 7 to 8. The variable is left unknown in 62% of the cases, so the last row of the table shows the accuracy of the variable in MCMIS *excluding* unknowns.

4.1.12 Quality Control Process

The MCMIS reporting methodology presents a difficult quality control process. First, there are a large number of jurisdictions filling out PARs that vary from state to state. Although many reporting agencies do not break down the reporting to the officer's level by providing a badge number, the 151,000 reports filed in 2005 were filled out by more than 61,000 agencies or individual officers. This means that, on average, a police officer from a specific agency might fill out less than three truck PARs in a given year. Assuming there are about 3,000 placarded shipments involved in crashes each year, the probability that a police officer will have to fill out a PAR for a placarded truck is on average, less than once every 20 years (60,000/3,000). This poses a significant training problem if the officer will be filling out the hazmat supplement only a few times in his or her career. Requiring or sponsoring a formal training program in 50 states for an event that occurs a few times in an officer's career is probably not cost effective. Providing a guide to

Table 4-7. Proportion of truck configuration in MCMIS coded correctly, based on comparison with TIFA data.

Truck Configuration	Frequency	% Correct
Straight truck	2,839	87.5
Straight truck plus trailer	373	42.9
Other straight truck	8	75.0
Bobtail tractor	175	61.7
Tractor-semitrailer	7,956	75.5
Tractor doubles	439	76.1

Table 4-8. Proportion of selected variables in MCMIS coded correctly, based on comparison with TIFA data.

Variable	% Correct
Interstate carrier	83.8
Number of vehicles in crash	85.2
Number of fatalities	99.2
Light condition	89.7
Road surface condition	91.0
Weather	84.0
Trafficway flow	41.0
GVWR class	34.8
GVWR class where known	91.5

filling out the state PAR and the MCMIS crash form would be more appropriate and is the avenue recommended for consideration. Although FMCSA has an elaborate training program, it does not focus on hazmat crashes. Because of unique characteristics of hazmat crashes, training should include special attention for this category of cargo.

Tests to evaluate the effectiveness of the current (quality assurance) Q/A process are challenging because of the number of agencies providing data. Each agency likely has its own Q/A process, so queries at the national level contain considerable uncertainty regarding their quality. A few years ago, when one compared carrier names and DOT numbers, the agreement was poor and the number of ways the carrier name was presented ran for pages, particularly for a carrier with a name that can be easily misspelled or mistyped, especially considering that most of the entries are being entered from a handwritten PAR. Fortunately, many of these previously observed problems are no longer present.

The overall Q/A process was checked selecting a single-truck carrier, "Schneider Trucking." This company was picked because in 2001, the number of ways this company's name could be spelled ran on for pages. Within the umbrella of "Schneider Trucking," the company appears to be operating under three DOT numbers, "Schneider Specialty Carriers," "Schneider Bulk Carriers," and "Schneider National Carriers Inc." Although there are slight variations in the names and addresses for the divisions, for the DOT number reported for Schneider Specialty Carriers, all the variations are in the name and address and none of the names or addresses are the same as those reported for the other two Schneider divisions. While the Q/A could be improved so only one name and address was recorded for each of the Schneider divisions, the variations are not considered a major impediment toward carrier-specific analysis since the DOT number appears to always be reported correctly. Since it is known that the source of the variation in the name and address starts when the carrier's name and address are handwritten on the PAR, the only way the variation could be eliminated would be to not faithfully record the information in the PAR but instead refer back to a pick list taken from the MCMIS Carrier Registration file where the DOT number was assigned. Since this is not a serious analysis impediment, improvement in other areas, such as reducing the number of blank entries and ensuring the consistency among current entries, would be more cost effective.

4.1.13 Interconnectivity with Other Databases

The MCMIS Crash file could be connected with HMIRS, TIFA, and—for grade-crossing accidents—RAIRS. The date and state where the incident occurred, along with the carrier's DOT number, provides a way to link accidents reported in the other databases. An attempt was made to couple MCMIS with the LTCCS, but the coder of the accident data in the causation study have obscured any information that can be used to link the data in the LTCCS to the other databases. The data are only given by month and year, no state or day of the month is provided. The VIN number of the vehicle has been truncated, so it is difficult to match the numbers shown in the LTCCS with the numbers in MCMIS. No other common fields could be found.

One useful improvement that FMCSA could make to the MCMIS Crash file that would increase the possibility of linking the file to other databases would be to restore the rule for how the *REPORT_NUMBER* field is constructed. Prior to 2001, the instructions to the state were to use the police report number in the *REPORT_NUMBER* field. That rule is no longer required, although some states clearly embed the police crash report number in the *REPORT_NUMBER* field. The actual police report number of a crash would permit a hard link to a specific crash, not just the probabilistic link obtained by using date, time, and geographic location.

4.1.14 Analyses Using Database

The previous sections have described some of the characteristics of the MCMIS Crash database. Since the focus of this analysis is hazmat accidents, here the focus will be on techniques to identify accidents involving hazmat vehicles and then to show the characteristics of those crashes. The ways to join MCMIS crash records with the records from other accident databases will be discussed in this section. The results from joining two datasets when the same accident is reported in both (for example, MCMIS and HMIRS) will be summarized separately after the individual databases have been described.

Hazardous Materials Serious Crash Analysis: Phase 2 (Battelle 2005), described in Chapter 2 of this report, developed a matrix listing the parameters that were believed to provide a comprehensive understanding of the accident environment. The project reviewed PARs and made telephone calls to the carrier to obtain data on those parameters. The parameters, divided into the five classifications of vehicle, driver, packaging, infrastructure, and situational are shown in Table 4-9.

The unshaded boxes are not recorded in MCMIS and, for those that are recorded, the color codes show the percentage of the hazmat accidents that correspond to entries in MCMIS. When

Table 4-9. Accident parameter coverage in MCMIS based on percentage not null.

Vehicle	Driver	Packaging	Infrastructure	Situational
Configuration	Age	Package Type	Road Surface	Pre-Crash Condition
Cargo Body	Experience	Quantity Shipped	Road Condition	Dangerous Event
GVW	Condition	Quantity Lost	Road Type	Vehicle Speed
Vehicle Defect	Valid License	Age (Cargo Tank)	Traffic Way	Impact Location
Vehicle Response	Citation Issued	Rollover Protection	Access Control	Primary Reason
	Driver Response	Inspection History	Speed Limit	Accident Type
	Training	Design Specification	No. of Lanes	Weather Condition
			Location	Light Condition
				Time of Day
				Health Consequences

Key:
> 95%
50% to 95%
< 50%

attempting to identify root and contributing causes of accidents, any cells in Table 4-9 that are not shaded could not be shown to be contributing or root causes of accidents. The table clearly shows that the MCMIS dataset must be supplemented considerably before it can be a good source of information for identifying root or contributing causes of hazmat truck accidents.

4.1.15 Summary and Potential Measures for Improving Root Cause Analysis

The MCMIS Crash file is formed using a complex set of operations that vary from one PAR originator to another, and Crash file preparation that varies from one state to another. Although the managers of the MCMIS Crash file have made great strides in improving the quality of the data, additional improvements are required for this database to be a useful tool in an information system that is capable of identifying contributing and root causes of accidents.

The single biggest improvement in MCMIS crash reporting would be MCMIS parameter fields that are completely populated. Some fields should be required to be filled out, particularly those related to the vehicle, carrier, driver, route characteristics, and point-of-contact information. Additional improvements include the following:

- Require that the *DRIVER_CONDITION_CODE* field be filled out. In *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005), the code "Appeared Normal" was the common entry for about 94% of the vehicle crash records. The codes for "Asleep" and "Fatigued" totaled about 3%, and the total for "Drugs or Alcohol Impairment" is about 1% and less than 1% for "Being Ill." The *DRIVER_CONDITION_CODE* field is clearly valuable, especially for programs designed to improve driver performance. This is the only field that captures driver performance in MCMIS and provides a valuable indicator of whether, in the opinion of the police officer filling out the PAR, the vehicle driver was truly incapacitated.
- Fill all five hazmat fields completely and accurately for trucks carrying hazardous materials. Presently, in records where one or more of the fields indicates a vehicle as carrying hazardous materials, all five fields are completely filled out less than 15% of the time. When several of the fields are filled out, the entries are often inconsistent, making it difficult to make an accurate determination of when a truck was transporting a hazardous material, and the UN number, class, and/or name of the hazardous material actually being transported. Although it is normally possible to identify the name of the hazardous material from the data reported in the *VEHICLE_HAZMAT_MATERIAL* field, it should be noted that in either the recording of the information or in the electronic transmission of the data, the field is often truncated.
- Enter the DOT number for all serious crashes involving hazardous materials. Currently, a DOT number is entered for only 80% of the vehicles carrying hazardous materials. A carrier transporting hazardous materials, even if not involved in interstate commerce, must register with FMCSA and be assigned a DOT number. For hazmat shipments, it should always be possible to assign a DOT number.
- Fill out the *VEHICLE_CONFIGURATION_ID* and *ROAD_CONFIGURATION_ID* fields.
- Specify the *LOCATION* field in a manner that enables the accident location to be found on a map. Presently, this is the case roughly 30% of the time. Specifying the route number or street name followed by the longitude and latitude is a straight-forward way to present location information. The difficulty in identifying the accident location on a map is aggravated by truncation of the field occurring somewhere in the recording or record transmission process, thereby eliminating key information.

Give state personnel entering the data into the MCMIS crash record system access to databases containing information such as the MCMIS Registration file and the 49 CFR Part 172 Hazardous

Material Table. Having access to this information could enable the state personnel to verify the hazmat entries and even fill in any information missing from the PAR. Linking the data entry process with these and other files so the data entry personnel could choose from "pick lists" that are narrowed down as additional characters are entered could make it easier to accurately populate fields in the MCMIS Crash file.

The MCMIS Crash file data dictionary could be enhanced so it contains not only the definition of a parameter and the format for the field in the database but also the format of the data to be entered. Specifying the format in the database does not necessarily define the data entry format as evidenced by past records. A section answering some commonly asked questions would be valuable as well. One question might be: *If the PAR lists the carrier location as one of the carrier's freight depots, should that address be entered in the MCMIS Crash file or should the address of the carrier's home office, taken from the MCMIS Registration file, be entered?* Another question might be related to the choice of entering a street address or a postal box number. Questions asked about the many situations that occur when filling out the *LOCATION* field also would be worthwhile given the different formats currently being listed in the MCMIS Crash file. For example, an Interstate route could be designated as I-70, IR70, I070, I70, or some other format. If the potential measure is adopted to use longitude and latitude when specifying a location, then the format and accuracy must also be specified. If the coordinates were expressed in decimal degrees, then specifying the longitude and latitude to two decimal places would place the accident on a highway, but if specified to three decimal points the location would be shown as either being on the left- or right-hand side of the right-of-way.

Build data quality consistency checks into the data entry process. For example, if a number is entered into the *DOT_NUMBER* field that is not in the MCMIS Registration file or is inconsistent with the carrier's name and address in the Registration file, then the number should be flagged and held until the correct number can be determined. If the UN number is not listed in the 49 CFR 172 Hazardous Material Table, it should not be possible to enter it into the *VEHICLE_HAZMAT_NUMBER* field.

4.2 Hazardous Materials Incident Reporting System (HMIRS)

HMIRS is maintained by PHMSA. In accordance with 49 CFR 171.16, all carriers of hazardous materials by road, rail, water, or air must fill out DOT Form F 5800.1 and submit it to PHMSA within 30 days of a reportable hazmat incident. The reportable incident could occur during loading, while in transit, during unloading, or while in temporary storage when en route between the origin and the final destination for the hazardous material. An incident is reportable if (1) the National Response Center (NRC) was notified, (2) there is an unintentional release of a hazardous material or the discharge of any quantity of hazardous material, (3) a cargo tank with a capacity of 1,000 gallons or greater containing any hazardous substance suffers structural damage to the lading retention system or damage that requires repair to a system intended to protect the lading system (even if there is no release of hazardous material), or (4) an undeclared hazardous material is discovered. In accordance with 49 CFR 171.15(b), NRC must be notified immediately if there is (1) an injury or fatality directly resulting from exposure to the hazardous material, an evacuation lasting more than one hour, a major artery closed for more than one hour, or alteration of an operational flight pattern or aircraft routine; (2) a fire, breakage, or spillage of a radioactive material; (3) a fire, breakage, or spillage of an infectious substance; (4) the release of a marine pollutant; or (5) a situation exists that poses a continuous danger to life at the scene.

4.2.1 Database Description

Prior to 2005, the HMIRS database consisted of three tables. The *CON* Table provided information on the incident. the *MAT* Table contained the name and address of the shipper, origin and destination address, the hazardous material being shipped, the amount released, and any damage to the packaging. The *RMK* Table was used for remarks. In most cases, there is one *MAT* entry for every *CON* entry. The *RMK* Table limits the text field to 80 characters, a legacy from the 80-character entry on an IBM card. Thus, there are often several *RMK* entries for each *CON* entry.

Beginning in 2005, the HMIRS database was significantly restructured. The new database structure is shown in Figure 4-2.

Figure 4-2 shows that all of the tables are related to the *IREPORT* Table. This table assigns a unique report number to each carrier-reported hazmat incident. *IREPORT* contains information on the carrier, the incident location, and impacts in terms of fatalities, injuries, and economic damage. It also provides contact information so that PHMSA data entry personnel can request additional data when certain fields are left blank. The *IEVENT* Table performs a function similar to the old *RMK* Table. The big difference is that after 2005, each line defines part of the event sequence. Although most *IRECORDS* have only one *IEVENT* record, about 10% have four *IEVENT* records. In 2005 and 2006, there were no *IREPORT* entries that had more than four *IEVENT* records. The *IACTION* Table was new in 2005 and gives the carrier the opportunity to identify changes that have been made to its operations as a result of the incident. An examination of the action statements demonstrates that some carriers have prepared thorough accident investigations and probably know the contributing causes and root causes of the incident. This table provides a way of identifying improvements that have been made without providing evidence of negligence that could be used in any litigation arising as a result of the incident, which is a major concern to the accident reporter.

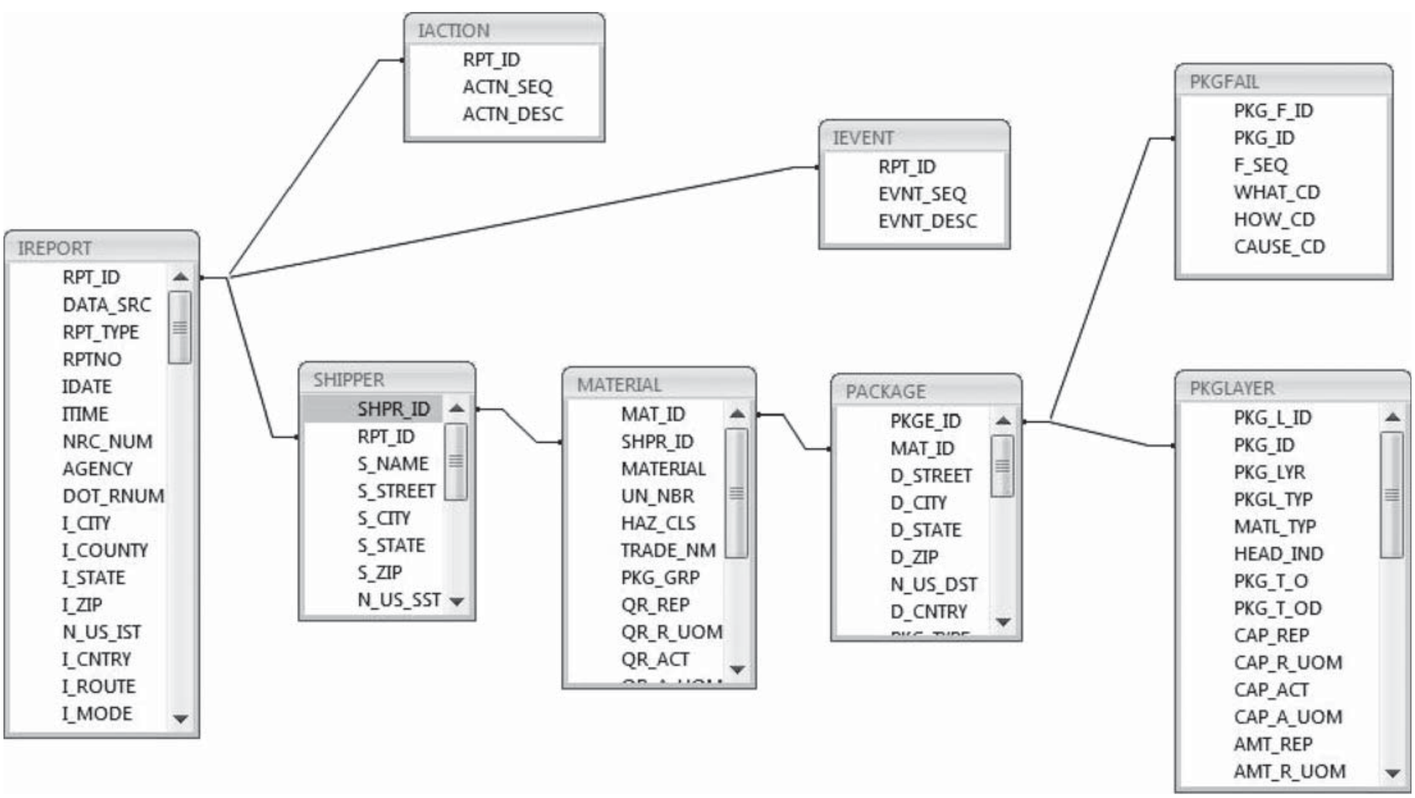

Figure 4-2. Relationship among HMIRS tables from 2005 onward.

Table 4-10. Total number of incidents reported by phase and mode, 2005 and 2006.

T_PHASE	Air	Truck	Rail	Water
En Route	1,049	4,871	1,303	83
Loading	553	4,542	30	6
Unloading	595	19,487	70	24
En Route Storage	1,868	1,709	46	24

The *SHIPPER* Table provides a listing of the shippers with hazardous materials involved in the incident. If the shipment was not "exclusive use," there could be several *SHIPPER* records for each *IREPORT* record. Similarly, a shipper could put more than one *MATERIAL* in the shipment, meaning that if an incident occurs, there could be several *MATERIAL* records for each *SHIPPER* record. The restructured database provides the opportunity to identify individual packages of the same material by having multiple *PACKAGE* records entered for a single *MAT_ID* record in the *MATERIAL* Table. Lastly, it is possible to describe where, how, and why each of the individual packages failed in the *PKGFAIL* Table. If there are multiple layers to the package, those can be described in the *PKGLAYER* Table. Since both the *PKGLAYER* and *PKGFAIL* Tables are related to the *MATERIAL* Table and not to each other, it is not possible to describe how the packaging layers failed in the incident.

The post-2005 structure of the HMIRS database provides the opportunity to report the performance of individual packages in shipments carrying multiple packages of more than one kind of hazardous material from more than one shipper. Although this level of detail is not often needed or used when reporting an incident, the structure of the database permits it. In all other databases, when there are multiple classes of hazardous materials in the shipment, their structure permits only one hazmat entry. The reporter must choose which hazardous material to designate, and this can become a source of disagreement when the records in two different databases report that the vehicle contained different classes/divisions of hazardous materials. In actuality, both types of hazardous materials were present.

The structure of the HMIRS database is ideally suited for examining package behavior in both the normal shipping environment and following an accident. Table 4-10 shows the total number of incidents reported in 2005 and 2006 by mode and phase, and Table 4-11 shows the total number of accidents reported by mode and phase for the same two-year period.

A comparison between the columns in Table 4-10 and Table 4-11 shows that all of the incidents reported for water and air are related to normal transport and are not related to accidents. Even with truck and rail, approximately 10% of the en route incidents are related to accidents. The focus of HMIRS is clearly not the transport accident environment. For the normal transport

Table 4-11. Total number of accidents reported by phase and mode, 2005 and 2006.

T_PHASE	Truck	Rail
En Route	558	91
Loading	1	
Unloading	13	
En Route Storage	1	1

Table 4-12. Further breakdown of en route accidents, 2005 and 2006.

Report Type	Truck	Rail
Incidents	521	87
Undeclared Shipments	4	4
Cargo Tanks, No Release	33	0
Total In-Transit Accidents 2005-2006	558	91

environment, the root cause of spills is probably related to the handling of the packages, information that HMIRS captures very well. The focus of this project, however, is transportation accidents. Specifically, these include the 558 truck and 91 rail transportation accidents that occurred en route over the two-year period shown.

Table 4-12 presents a summary of the HMIRS database accidents for 2005 and 2006. Table 4-12 shows eight incidents of undeclared shipments. Five of these shipments were empty cargo tank shipments and the regulations require them to be placarded even when empty. The other three contained hazardous materials and were not placarded. The new requirement to report damage to cargo tanks having a capacity of 1,000 gallons or greater when they suffer damage to the lading system or its protective system even though there is no release, manifests itself in the recording of 33 additional truck accidents over a two-year period.

4.2.2 Purpose and Function

HMIRS was developed in the early 1980s and has been maintained ever since. Its purpose has always been to provide regulatory agencies with the information they need to monitor the safety of hazmat transport, document the effectiveness of current regulations, and—if shown to be warranted—provide the data required to support new regulatory initiatives. As stated previously, the focus is on package performance. Most of the incidents reported occur during normal transport and are not related to accidents. Since this project focuses on accidents, the majority of the records reported in HMIRS are not referenced in the following analysis.

4.2.3 Data Collection

When an HMIRS reportable incident occurs, the carrier is required to fill out DOT Form F 5800.1 and submit it to PHMSA. As follows, there are four filing methods, but most of the reports are received by the first three of these:

1. XML submissions—five carriers do this on a *nightly* basis and most follow up with a paper copy,
2. Online 5800.1 incident reporting application,
3. PDF—some follow up with an e-mail that includes a PDF attachment, and
4. Faxes from some filers and package carriers from other filers who deliver, on a *monthly* basis, the paper forms for all incidents within the past month.

A carrier has 30 days in which to file a report. PHMSA does Web searches of newspapers and also receives a list of incidents in which NRC was notified. By comparing the list of incidents that have been reported by these sources, carriers that have not filed within the 30-day period are identified. Those carriers are notified by phone or mail regarding their delinquent status.

4.2.4 Data Compilation

Reports filed by XML submission or online 5800.1 go directly into the database, which includes high-level quality control processing. For PDFs and faxes or paper forms delivered by package carriers, PHMSA scans and performs optical character recognition (OCR) for accuracy and then enters the record into the database. PHMSA also employs character-to-character checks to ensure that the OCR process translated the paper form properly. As part of the data entry process, PHMSA examines the form for personally identifiable information (PII), and such information is scrubbed from the report. For example, if the carrier name is the name of a person, the name is blanked out in the report. The HMIRS database is updated monthly.

4.2.5 Accuracy and Completeness of Data

Aside from the inconsistencies that are discussed in Section 4.2.6, data accuracy is mostly controlled by the accuracy of the carrier reporting the information. The carrier has access to first-hand knowledge of the incident through personal interviews with the vehicle operator involved in the accident and repair information that documents the extent of the damage and the costs to complete the repairs. PHMSA staff, in interviews, complained that it is evident that many of the reports are reviewed by a lawyer prior to submission and they believe this review greatly weakens the value of the information reported. No individual or business wants to put information on the record that could be used against them in a civil suit. This is particularly true for the accident reports that make up a small fraction of the HMIRS database.

When questionable or incomplete information is provided, PHMSA staff contact the carrier and request additional information or clarification of the information they have received. In the case of injuries and fatalities, since HMIRS distinguishes among injuries and fatalities that are the direct result of exposure to the hazardous material and those as a result of the traffic accident, all fatalities and injuries are validated by their own process to determine if they were caused by a hazardous material. This includes obtaining coroner reports and death certificates. Data compilers also look for any mention of injuries if none were marked on the form.

In fiscal year 2009, PHMSA will introduce an online Incident Reporting System that will require filers to fix incorrect data before the submission will be accepted. However, since the carriers will also be able to file the reports via the other methods available, the effectiveness of these checks will be limited to electronically submitted reports.

4.2.5.1 Comparing HMIRS with TIFA to Evaluate Crash Data Accuracy

As was the case for MCMIS, the TIFA file provides a unique opportunity to evaluate the accuracy of the crash data reported by carriers to PHMSA and entered in the HMIRS database. Although the TIFA data is collected differently, the data associated with fatal truck crashes are checked rigorously for data quality. Consequently, TIFA data is highly reliable.

In this discussion, and in subsequent comparative analyses of the accuracy of HMIRS data, it is important to recognize that not every attribute, including some cause codes, in the HMIRS database is considered. Rather, those that can be compared because similar information theoretically exists in multiple databases are used to estimate the level of accuracy of the HMIRS data that are being collected.

The TIFA file provides an opportunity to evaluate the completeness of reporting of a subset of hazmat incidents. The TIFA file includes all trucks that were involved in a traffic crash that included at least one fatality. Data collected includes whether the cargo on the truck was placarded as hazardous material and whether the cargo spilled. The TIFA file is based on NHTSA's FARS file, which is a comprehensive database of all fatal traffic crashes. UMTRI receives police reports on all crashes involving a truck. Interviewers contact the truck owner, its operator, the carrier's safety director, the original reporting officer, or any other person knowledgeable about the truck at the time of the crash. The interviewers collect a set of detailed information about the configuration of the vehicle, which is cross-checked with multiple sources, such as manufacturer's specifications. As such, the TIFA file should provide a complete record of hazmat releases in fatal truck crashes, incidents that should be in the HMIRS data.

Given the seriousness of fatal truck crashes, one would expect these events to be among the most likely to be reported to HMIRS. Accordingly, estimates of reporting completeness in this subset may be regarded as close to the upper boundary of the plausible range for the true reporting rate.

TIFA and HMIRS data for 2005, the most recent year when both were available, were used in the comparison. The TIFA file includes all traffic accidents involving a medium or heavy truck and a fatality that occurred as a consequence of the accident, whether from hazmat release or not. Cases required to be reported to HMIRS involve either a release or damage to the lading system. The TIFA file records whether there was a hazmat spill, so HMIRS cases that meet that criterion can be identified in TIFA. TIFA, however, does not capture information that indicates damage to the lading system, although it does capture rollover, which is probably strongly associated with damage to the lading system. Thus, the TIFA file should include all HMIRS-reportable cases that involve a fatality and truck transportation. However, only the cases reportable because there was a hazmat release can be specifically identified in TIFA. Cases that are reportable because of damage to the lading system cannot be specifically identified, although many are very likely among those in which the truck rolled over.

Of the 53 HMIRS highway cases in 2005, 50 were matched to a record in the TIFA file. The match had to be done manually; that is, by reviewing individual records, since there are no case identifiers in common between the two files. In addition, other potential match variables—location and time—all had various problems. The time and even the date recorded in TIFA varied from the self-reported data in HMIRS. Time varied most frequently, often by +/–30 minutes. The only location information useful for the match that was common between the two files was the county name. In the TIFA file, county is captured as a numeric code (the Federal Information Processing Standards [FIPS] code, as found at http://www.census.gov/geo/www/fips/fips.html, which is part of a standard and widely used geographic identification system). In HMIRS, county is an alphabetic string, and therefore subject to the vagaries of spelling. There were variations in spelling the county name and, in one case, an incident was reported as occurring in Albany County, NY, in the HMIRS data, but actually occurred in Albany County, WY. Hand-matching is unavoidable when such materials are all that is available.

Three HMIRS cases did not appear in the TIFA file. This may occur if the vehicles transporting the hazardous material were not medium or heavy trucks. Each case was searched for in the 2005 FARS file. One case from Bronx, NY, was in FARS, but not TIFA. The vehicle was coded as a truck, so the case should have been included in TIFA. The reason it was not included is that it was added to FARS after the complete FARS file was released to the public. A corrected version of FARS was released with the case after the TIFA file was itself completed.

It is likely that the two other records in HMIRS but not TIFA were inaccurately reported. One record was from Douglas County, KS, on 12/22/2005 at 6:30. This record seemed to be matched by three records in the 2005 FARS, which occurred on the same date at 6:38. But all three vehicles were light vehicles, and none coded in FARS as carrying hazmat cargo. However, it is known that FARS underreports hazardous material in cargo. Finally, a record from Lynchburg City, VA, on 1/31/2005, was not found in FARS at all. No fatal crash occurred in Virginia on that day. There may have been an error in the date of the incident in HMIRS, or some other error.

Table 4-13. TIFA hazmat crash involvements matched with HMIRS TIFA 2005/HMIRS 2005.

Cargo Spillage	Matched		Total
	No	Yes	
None	106	9	115
Spill of hazmat	27	40	67
Unknown	3	1	4
Total	136	50	186
	Row Percentages		
None	92.2	7.8	100.0
Spill of hazmat	40.3	59.7	100.0
Unknown	75.0	25.0	100.0
Total	73.1	26.9	100.0

One criterion for reporting to the HMIRS file is a spill of hazardous materials. All fatal truck crash involvements in which hazmat cargo is released should be reported to the HMIRS. The TIFA data show that there were 67 trucks in a fatal crash in which hazardous material spilled (see Table 4-13). Forty of the records were found in HMIRS, for a reporting rate of 59.7%. Cargo spillage clearly is associated with reporting, as the data in the table show that only 9 of the 115 trucks with hazmat cargo that did *not* spill were reported. Overall, 26.9% of trucks carrying hazardous materials that were involved in a fatal crash were reported to HMIRS.

Non-spill incidents, which nonetheless included damage to protection for the lading retention system, are also required to be reported. This could account for the nine non-spill cases that were found in HMIRS and, likewise, the 106 non-spill cases that were not. The TIFA data do not include any information on whether there was damage to the system protecting the cargo retention system of the truck, so there is no way to determine directly if such damage accounts for the observed pattern of reporting. However, the TIFA data include a variable that identifies whether the vehicle rolled over. Rollover should be strongly associated with damage to the truck, including the system protecting the cargo retention system, so trucks with hazardous material that rolled over as part of the crash would be expected to be reported to HMIRS. Again, there is no way to determine directly if they *must* be reported, but certainly a high proportion would be expected to qualify.

In Table 4-14, the results of matching the HMIRS to the TIFA file are shown disaggregated by cargo spillage and whether the truck rolled over. In the top half of the table, match results are shown for trucks that did not have cargo spill in the crash. Of the cases that rolled over, and thus likely damaged the protection for the cargo retention system and qualified for reporting to HMIRS, only 4 of the 15 (26.7%) were actually reported to HMIRS. Only 5% of the 100 hazmat trucks in a fatal crash, with no spill and no rollover, were reported to HMIRS. Note, however, that trucks with a hazmat spill and rollover were reported at only a 61.2% rate. Crashes in which there was a spill, but no rollover, were reported at a 55.6% rate, somewhat lower but not practically different.

The TIFA file includes some limited information about the motor carrier, including whether the carrier was private, operating trucks incidental to another business, or for-hire, and whether the carrier operated in interstate or intrastate commerce.

Table 4-14. TIFA/HMIRS match results, by cargo spillage and rollover, TIFA 2005/HMIRS 2005.

No Cargo Spill			
No spill	Matched		Total
Rollover	No	Yes	
No roll	95	5	100
Rollover	11	4	15
Total	106	9	115
	Row percentages		
No roll	95.0	5.0	100.0
Rollover	73.3	26.7	100.0
Total	92.2	7.8	100.0
Hazmat Spill			
	Matched		Total
Rollover	No	Yes	
No roll	8	10	18
Rollover	19	30	49
Total	27	40	67
	Row percentages		
No roll	44.4	55.6	100.0
Rollover	38.8	61.2	100.0
Total	40.3	59.7	100.0

Considering all records matched, there was little difference in reporting rates between private and for-hire carriers or between interstate or intrastate carriers. Fatal crash involvements of a truck carrying hazmat cargo operated by an intrastate carrier were reported 19.5% of the time, compared to 29.8% of involvements where the carrier was operating interstate. The rate for interstate carriers is somewhat higher, but there are too few cases for the difference to be considered statistically significant. Rates for private and for-hire carriers are virtually identical, 26.5% and 28.3%, respectively.

Although it is not possible to determine definitively whether each case qualified for reporting either because of a spill or damage to the protection to the lading retention system, the results here at least bound the probable HMIRS reporting rate. On the high side is the 59.7% reporting rate observed for cases in which there was a hazmat release. All of these cases definitively qualified for reporting, yet less than three-fifths were actually reported. The lower bound for HMIRS reporting might be the overall reporting rate of 26.9% of fatal truck crash involvements with hazmat cargo.

However, it is quite unlikely that the overall reporting rate is anywhere near as high as 59%. In the case of the MCMIS crash file, the overall reporting rate is about 75% of the reporting rate for fatal crashes. If the same ratio is applied to the HMIRS file, this would mean the overall reporting rate is about 45%. But the MCMIS crash file is reported by state agencies, and the FMCSA has an intensive program to increase reporting, including paying for changes to systems to increase reporting. The obligation to report to HMIRS falls on the thousands of private hazmat carriers, and there is no systematic program to make sure that all appropriate cases are entered into the database. Thus, the overall reporting rate may be even lower than that of HMIRS. However, given current data, it is not possible to provide a more realistic estimate.

4.2.6 Quality Control Process

Reports that are not submitted electronically are checked twice. The first check ensures that the translation from the submitted form to the electronic record has been completed accurately. After the records have been placed in an electronic form, the records are checked for business rule inconsistencies, invalid dates, and invalid commodities (by cross-checking with the commodities in the database).

Additional checks include cases of city/county inconsistencies or when the report shows that 5.5 gallons were spilled from a 5-gallon container. In these instances, the filer is contacted and asked whether there were multiple packages that failed, following which, the information is corrected.

An analysis of the data demonstrates that some obvious Q/A checks are not being performed. For example, several carriers that file thousands of reports each year file under more than a dozen names and several DOT numbers. The multiple DOT numbers are probably valid and are the result of acquisitions and mergers. However, since all hazmat carriers must annually register with PHMSA, the name of the carrier could be required to be selected from the Registration file. In cases where no entry of the DOT number is provided, the submission should be rejected and the carrier required to resubmit the report with the DOT number completed. Unfortunately, several large carriers that are submitting thousands of reports never provide a DOT number on a single report.

4.2.7 Interconnectivity with Other Databases

The interconnectivity of HMIRS with other databases varies by mode. For trucks, the DOT number, date of incident, and county of incident provides a fairly good way to link records in HMIRS and MCMIS databases. Incident date, county, and the occurrence of a fatal accident might be a good technique to link HMIRS and TIFA. Although some effort was made, there seems to be no good way to link HMIRS and MISLE for ship or barge hazmat accidents. For certain accidents, FRA and/or NTSB might select an accident for more study. In such cases, the additional information that is collected may result in the identification of some contributing causes or even a root cause for the accident.

4.2.8 Analyses Using Database

The HMIRS database underwent a major structural change that took effect at the start of 2005. The first part of this evaluation will look at the effect of the HMIRS structural change. This will be followed by an evaluation of some of the parameters that could identify contributing causes of hazmat accidents. Since HMIRS really focuses on package behavior, it is likely that any capability for identifying contributing causes will be limited to package behavior. The final section will look at the gains that might be realized when HMIRS is coupled with other databases such as MCMIS, TIFA, RAIRS, and MISLE.

4.2.8.1 Effect of HMIRS Structural Changes Taking Effect in 2005

The major structural change made to HMIRS at the beginning of 2005 was to break up the three main tables, commonly labeled *MAT, CON,* and *RMK* into a series of smaller linked tables titled: *IREPORT, IEVENT, IACTION, SHIPPER, MATERIAL, PACKAGE, PKGLAYER,* and *PKGFAIL.* From an overall perspective, the material previously found in *MAT* is now found in *IREPORT* and *SHIPPER.* The material previously found in *CON* is now found in *PACKAGE, PKGLAYER,* and *PKGFAIL. IEVENT* appears to capture the descriptive text previously found in *RMK,* and *IACTION* is a new table that asks the carrier what actions have been taken to reduce the likelihood or the consequences of this accident should the conditions that initiated the accident be present in the future.

4.2.8.2 Main Features of the Restructured HMIRS Database

In the restructured database, the basic accident information is contained in the *IREPORT* Table. This table contains information regarding the date, time, and location of the incident, and the resulting consequences expressed in terms of hazmat- and non-hazmat-related fatalities and injuries, road closure durations, evacuations, and damage costs. The changes to *IREPORT* following the restructuring are described in Table 4-15.

Similar tables could be prepared for the *SHIPPER* and *PACKAGE* Tables. In these cases, the data are restructured but there does not appear to be a significant expansion of the data fields. Overall, the restructured HMIRS provides a much-improved description of accident consequences. However, there is limited information on the driver, description of the route, and conditions at the accident scene. To obtain more accurate information on the accident scene, such as whether the accident occurred on a curve or while turning at an intersection, one would have to be able to identify the location of the accident from the route and location fields and then refer to map software to determine the road geometry. The other alternatives would be to find the accident in MCMIS or access the specific PAR. Privacy concerns may limit the availability of some personal driver information. To make it publicly available, personal driver information could reside in a separate file that would be kept confidential.

The packaging information provided before and after the restructuring of the database was extensive and has only improved. Capacity and quantity shipped is now requested as part of the report.

4.2.8.3 Relevant en Route Accident Statistics

Table 4-16 shows the statistics for only those records with *T_PHASE* = 261, signifying the phase of operations is "en route" when *ACCIDENT* = T, signifying that the record is submitted because a reportable accident occurred. Tables 4-10 and 4-11 both summarized the records for the years 2005 and 2006, the first two years of reporting after the database reconstruction.

Table 4-15. Main features incorporated into the restructured database.

General Topic Area	Variables in Tables after Restructuring	Status Prior to 2005
Report Referencing	IREPORT contains several fields that could be used to relate incident reports to reports in other databases.	There were no fields that would enable linking to records in other databases.
Contact Information	IREPORT contains numerous fields listing names, addresses, and phone numbers of the person filling out the report and provides entries to list the police accident report number.	No information regarding person filling out the form recorded with the accident record.
Deaths and Injuries	IREPORT breaks down the deaths and injuries into public, employee, and emergency responders for both HM and for non-HM.	Previously, only deaths and injuries from HM were listed and no breakdown into classes of individuals was possible.
Evacuation	Both the number and type of individuals evacuated is given, as is the duration of the evacuation. If a major road was closed, the duration of the closure is also given.	Previously, just the number of individuals evaluated was listed.
Conditions	Weather condition is now listed.	Previously no listing of weather conditions was recorded.
Phase	In addition to the transport phase, for air, additional information is given as to the step in the multimodal operation where the spill occurred.	Only the phase is given, en route, loading, unloading, or en route storage.

Table 4-16. Summary of IREPORT records for en route accidents.

RPT_TYPE	Year	Truck	Rail	Totals by Year	Totals by Report Type
A	2005	274	48	322	
A	2006	247	39	286	608
B	2005	0	0	0	
B	2006	4	4	8	8
C	2005	20	0	20	
C	2006	13	0	13	33
	Totals	**558**	**91**		**649**

Table 4-12 provided an expansion of part of Table 4-11, focusing on en route accidents and the reason for reporting as determined by the *RPT_TYPE* parameter. Any accident that results in a hazmat spill is classified as "A," an undeclared hazmat spill is classified as a "B," and new reporting category "C" indicates that a cargo tank with a capacity of 1,000 gallons or greater was involved in an accident in which (1) there was damage to the lading or the safety system protecting the lading such that repairs had to be made and (2) there was no spill of hazardous material.

The totals in Table 4-16 are mostly from spills, as can be seen by comparing the rows for RPT_TYPE = A with the totals row in Table 4-16. Note, that all the air and vessel records are considered spills and not due to an accident. Of the 7,306 reports in Table 4-10, Table 4-11 shows that only 649 are accidents. Note that the absence of air and vessel columns in Table 4-16 indicates that all the air and vessel records shown in Table 4-10 are spills not associated with any en route accident.

The new requirement to report non-spills associated with cargo tanks having capacities of greater than 1,000 gallons if the ladling or the system protecting the ladling is damaged is shown under Report Type C. There are a total of 33 reported accidents that were coded under this new classification. The remainder of this analysis will look at these 649 reports.

The first evaluation will focus on whether the 649 accidents associated with the *IREPORT* records have corresponding records in the *SHIPPER, MATERIAL, PACKAGE,* and *PKGFAIL* Tables. By successively linking the tables to the *IREPORT* records it is possible to determine if there are accidents in *IREPORT* that do not have corresponding records in the linked tables. When *SHIPPER* is linked, the total number of records increases from 649 to 730. However, none of the *IREPORT* records are dropped. This emphasizes that one of the features of the restructured database is its ability to separately provide an accurate description of the spills of hazardous materials and their behavior for those offered for transport by several shippers. The 558 truck accidents represent material from 596 shippers while the 91 rail accidents were associated with 134 shippers. The only limitation found in the *SHIPPER* Table is that the destination field has been left blank for 105 *SHIPPER* records. From the standpoint of identifying contributing and root causes, the lack of destination information for a significant fraction of the records is probably not a limitation.

When the *MATERIAL* Table is added, the number of records increases to 750 records, indicating that several accident records in *IREPORT* involve multiple types of hazardous material. Once again, all 649 accidents described in *IREPORT* are represented. Adding the *PACKAGE* Table increases the number of number of records to 762, indicating that some materials in multiple packages have been described accurately. The number of accidents associated with these 762 packages remains 649, indicating that package information is available for all 649 accidents listed in *IREPORT*. When the *PKGFAIL* Table is linked to the other tables, the situation changes significantly. There are 130 accidents described in *IREPORT* that have no *PKGFAIL* records and there are another 6 that have no failure cause for the package failure. Thus, the cause of package failure cannot be described for 78% of the accidents. A check was made to see if the package failure might be missing because these were Report Class C accidents (and thus no spill) but although 24 of the 33 Class C accidents are among the 136 with no *PKGFAIL* records, there are many where the information is missing. A query to look at the number of accidents with missing *PKGFAIL* records that were recorded as spills showed that all were flagged as spills. Clearly, the carrier should have provided a *PKGFAIL* record for all 130 of the accidents.

4.2.8.4 Significant Parameters

HMIRS records request that the carrier reporting an accident fill out the *PKGFAIL* Table and from the *WHY FAILED* field, it is possible to identify why the package failed. Note that this is not why the accident occurred. A shipper could place some corrective actions in the *IACTION* Table that may be indicative of the contributing causes for the accident but the causes do not have to be listed in the database. HMIRS does not contain other driver information. Therefore, the only driver information in the database is that incidentally found in the *IACTION* Table.

HMIRS also has very limited information on the location where the accident occurred. There are four relevant parameters that help identify the location of the accident: *I_STATE, I_COUNTY, I_CITY,* and *I_ROUTE.* The user guide requests that the *I_ROUTE* parameter specify the "street location on which incident occurred." Based on the 558 truck accidents that occurred in 2005 and 2006, the carriers filing interpret *I_ROUTE* to be a street address, a route designation, and frequently a mile marker or mile-post designation, or an intersection of two named or numbered highways. The *I_ROUTE* listings that were not considered adequate to identify the location of the accident provided were a blank entry, just the route designation, or an incomplete phrase that designates the route location as some distance from an undefined point. Of the 558 truck accidents, the actual location of the accident could only be identified for 347 cases, just over 60% of the accidents. Thus, if route characteristics were a contributing cause of the accident, it would be impossible to identify those causes for 40% of the accidents reported in HMIRS.

HMIRS also contains limited information on the vehicle characteristics. If the package type is a C, a cargo tank, and the volume shipped is larger than 5,000 gallons, one can infer that the vehicle configuration is a semitrailer hauling a cargo tank. If the vehicle configuration is a contributing cause for the accident, the HMIRS record must be coupled with data in MCMIS. In MCMIS, if the data fields are fully populated, the vehicle, driver, and road characteristics documented in MCMIS can be coupled with the package information in HMIRS to get a comprehensive picture of the driver, vehicle, package, and route characteristics present at the time of the accident.

The restructured HMIRS database provides fields to enter the carrier's DOT number, a parameter not requested prior to the 2005 restructuring. If an accident meets the reporting criteria in both HMIRS and MCMIS, and the data fields are all filled out, then an analyst has a good description of the accident expressed in terms of driver, vehicle, package, and route characteristics. Coupling HMIRS with RAIRS and TIFA would provide a comprehensive picture of the accident, just as is the case with coupling HMIRS with MCMIS. For a two-year period, there were no accidents involving hazardous materials that were reported on the water in both MISLE and HMIRS. The reporting criteria seem to be too different so for the few water spills of DOT-defined hazardous materials that occur as a result of ship or boat accidents, it is not possible to supplement the package data shown in HMIRS.

For the DOT number field to be useful, it must be comparable with the corresponding field in MCMIS. HMIRS records the number as a number field and it is recorded in MCMIS as a text field. In MCMIS, sometimes there are letters like DOT and US preceding the number. If the letters are ignored, it is easy to convert the text field into numbers and look for matches based on the date and state. The county and time can then be used to verify the matches. To begin with, the DOT number is not supplied for 60 of the 558 accidents, reducing the total number of possible matches to 498. When the DOT number, date, and state of accident occurrence were used to match up the 498 accidents identified in HMIRS for years 2005 and 2006 with the much more extensive MCMIS file, a total of 110 matches were found. Four of the records were eliminated because the accident times were different by several hours and the counties were also different. Thus, for slightly more than 20% of the HMIRS records, the accident reporting criteria overlap sufficiently to require the accident to be reported in both databases. Of the 106 matched records, there were 8 records where the times were within 15 minutes but the counties were different. Using mapping software, and the location field as a final determiner, it was concluded that in six of these eight differences, HMIRS recorded the wrong county and in the other two, the county reported in MCMIS was incorrect. In all of the cases, the accident location was very close to the county border, normally less than three miles. In the six cases where it is suspected that the county recorded in HMIRS is wrong, one possible cause is the extra step used in recording the county. Rather than entering the county name, the county FIPS code was used. Entering the actual name introduces the wide variety of misspellings possible. This impedes computer matches. The use of the FIPS code is much better. If the name was entered as a string, it could be used to cross-check with FIPS, which would be useful to guard against errors in entering the code. But the best approach may be to avoid having to use probabilistic matching and use a case identifier, such as police report number. Every time there is an extra step, an additional source of errors can arise.

A single query was used to identify 110 accidents reported to both MCMIS and HMIRS that occurred for the same carrier on the same date, in the same state. These accidents were then screened to see if there were any cases where the MCMIS record should not have been joined with the HMIRS record. There were 12 records that were suspect and, in the end, 4 were believed to represent different accidents. One way to eliminate this 10% error rate would be to join additional fields. Requiring the county to be the same would have eliminated all 12 records, including the 8 where either HMIRS or MCMIS reported the wrong county. Time was more difficult to use because even for those believed to be reported in both databases, the times were often different by 15 minutes or more. Location was quickly rejected because of the vast difference in the types of information recorded for that field. Of the 106 MCMIS records that were joined to HMIRS, one-half were not flagged as being a hazmat shipment in any of the four hazmat descriptive fields in MCMIS. Thus, if the join between HMIRS and MCMIS had just looked at the MCMIS records that were flagged as being hazmat, one-half of the incidents would have been missed.

Information technology is rapidly advancing to the point that it is feasible to require the longitude and latitude of the accident to be included as part of the accident record. Many states already have longitude/latitude as an entry on their police accident reports but few police officers populate the field when filling out the form. It is not believed that the reason is their inability to know the coordinates of their position. Most police cars can be located geographically by their dispatcher and, for many companies, this capability also exists for trucks and trains. Thus, provided that a format for the longitude and latitude are specified—decimal degrees or degrees, minutes, and seconds—it should be possible to use these two fields to join records and thereby eliminate the screening step. Many handheld GPS devices will provide this information and, for those without a handheld device, the accident can be located on numerous free and commercial software packages. For these free packages, all that is needed is Internet access.

4.2.8.5 Class C Record Types

In the restructured database, the 5800.1 form imposes a new requirement. If a cargo tank having a capacity of greater than 1,000 gallons is involved in an accident and the lading or the system protecting the lading is damaged, the carrier is to report the accident even if there is no spill. The way these accidents are coded is to enter a "C" under the RPT_TYPE variable in the *IREPORT* Table. The other designations are "A" for a spill and "B" for an undeclared hazmat shipment. Of the 649 en route hazmat accidents, 33 shipments, all truck mode, were coded as RPT_TYPE = C for the years 2005 and 2006.

The rate of carrier compliance with this new requirement to report some non-spill hazmat accidents for cargo tanks is difficult to determine. Would one expect more than 33 incidents in two years? Focusing just on the truck accidents, in 2005 and 2006, there are 558 records. A query for the cargo tank configuration shows that 442 of those accidents involved cargo tanks. One approach for investigating underreporting is to examine the rollovers reported as spills and determine if they are being over-represented. In the *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005) study, spills occurred in 66% of the rollovers. Of the 442 cargo tank accidents reported in HMIRS, 357 rolled over and of those rollovers, 291 were coded as spills. Thus spills occurred in 82% of all rollovers. Because rollovers are likely to damage the rollover protection system, spills from rollovers seem to be over represented in the HMIRS database. The number of cargo tank rollovers reported as Class C records was 23. The non-spill cargo tank rollovers would have to be increased from 23 in two years to 110, an increase of about 85 non-spill accidents to lower the spill rate to 66%. Is it reasonable that 110 cargo tanks would rollover in a two-year period and only 23 would experience damage to their rollover protection system that was serious enough to require repair? Given that cargo tank rollovers frequently result in a release, it might be anticipated that damage would occur in more than 21% (23/110) of the non-spill rollover accidents. Perhaps a more fundamental result is that if all cargo tank accidents that met the MCMIS definition of *serious* were required to be reported, the number of additional records being reported would increase by at most a few hundred. Given that there are more than 30,000 records added to HMIRS every two years, requiring all serious cargo tank accidents to be reported would increase the record load on PHMSA by less than 1%. By making the HMIRS reporting requirements have some of the same requirements as MCMIS, the added benefit is that more accidents would be reported in both databases. Since MCMIS has more information on the road configuration, environmental parameters, and driver characteristics, it should be possible to perform more analyses that move toward identifying contributing and root causes of hazmat accidents.

4.2.8.6 Inclusion of IACTION Table

One of the major changes in the restructuring of the HMIRS database was the addition of an action statement table that can be used by carriers to state the changes they propose to make to prevent or reduce the likelihood that such accidents would occur in the future. There are 649 accident records in *IREPORT* for 2005 and 2006 in which *T_PHASE* = 261 and *ACCIDENT* = T. These consisted of 558 and 91 truck and rail records, respectively. There are 453 records with action entries, 411 for truck and 42 for rail. Thus, the percentage of carriers providing action statements is 74% for truck and 46% for rail. In looking at the action statements, while the contributing cause is not given, it is clear that the contributing cause that resulted in the action statement is known by the carrier, but it just is not stated. Like the causes of failure that are used in the *PKG_FAIL* Table, a table of contributing causes could be developed and added to the *IACTION* Table that would provide the basis for the actions taken; they would not have to be assumed. The tables developed in the RAIRS might be a useful place to start when developing this table of contributing causes. Even though no driver information is present in HMIRS, many of the actions focus on increased driver training, so it is clear that one of the contributing causes for many actions is inadequate driver training. Another goal would be to increase the number of carriers providing action statements, particularly for rail. A goal could be to have a compliance rate in excess of 90%.

4.2.8.7 Differences before and after Changes in HMIRS Structure

Prior to the restructuring of HMIRS, hazmat records had to be duplicated to account for the presence of a carrier hauling hazardous material for several shippers in the same vehicle and for designating more than one destination for some of the packages. A query of the database for *PHASE* = *261* and *ACCDR* = *True,* shows that there are a total of 603 primary incident records for 2003 and 2004 and an additional 146 records to account for multiple shippers and multiple destinations. The total number of records is therefore 749 records, remarkably similar to the total of 747 records reported for 2005 and 2006 combined. Although not important when considering root causes, it is easier to understand the shipment logistics after the restructuring. The total number of incidents is quite close as well, considering that in the 649 accidents reported for 2005 and 2006, 33 were associated with the new requirement to report non-spills if a cargo tank having a capacity of 1,000 gallons or greater was involved in an accident and there was damage to the cargo or the equipment protecting the cargo. If those were not considered, the number of crashes in 2005 and 2006 would total 616, again quite similar to the 603 reported in 2003 and 2004. A more detailed evaluation of the records is found in Appendix D (available on the TRB website at www.TRB.org by searching for HMCRP Report 1).

4.2.9 Summary and Potential Measures for Improving Root Cause Analysis

4.2.9.1 Summary of Database Evaluations

The restructured HMIRS database can be considered to be a relational database and, except for the PKGFAIL Table, the record set for an en route accident is complete. Even in the case of the PKGFAIL Table, the data are available for about 80% of the HMIRS en route accidents. From the point of view of identifying route and contributing causes, this is not believed to be a significant limitation.

For a complete description of the package, vehicle, driver, and roadway characteristics associated with an accident, HMIRS would have to be joined with MCMIS for trucks and RAIRS for rail. Until the restructuring of HMIRS, the biggest detriment to joining the two databases was the lack of common fields. HMIRS now has a field to enter the DOT number, and this field is now being populated almost 90% of the time. The DOT number is also entered for about 90% of the MCMIS records designated as showing a hazmat placard. Assuming the non-reporting is random, the DOT number can be used to join about 80% of the accidents that meet both the HMIRS and MCMIS reporting criteria. Since all carriers of placarded quantities of hazardous material must register with both FMCSA and PHMSA, they must have a DOT number and there should be no blank entries in either database.

Information technology has advanced to the point where both the carrier reporting in HMIRS and the police officer reporting to MCMIS have the capability to report the longitude and latitude of the accident. Providing that a common format is used in both databases, it is believed that one query could be used to identify accidents reported to both HMIRS and MCMIS and the data-scrubbing step could be eliminated or significantly reduced.

The main reason why a fraction of the HMIRS and MCMIS records cannot be linked is the difference in reporting criteria. Some of the difference between the number of records in MCMIS that can be joined with HMIRS records can also be attributed to underreporting.

It is suspected that carriers are not reporting all of their Report Type C accidents but that statement can not be made with certainty. Lastly, the carriers are providing action statements for 74% of the truck accidents and 46% of the rail accidents. An increase of the carrier reporting rate to at least 90% would be highly desirable. In addition, the action statements given are quite positive and indicate that the carriers have done enough accident investigation to identify some changes that would decrease the frequency of similar accidents in the future. The usefulness of this information would be greatly improved if a cause table, similar to the *WHAT FAILED* Table, was developed so the carrier could list some contributory causes from a pick list. Although there might be some resistance to adding that field because of liability issues, moving toward being able to routinely list contributing causes would be helpful.

4.2.9.2 Potential Measures for Improving Root Cause Analysis

The following potential measures would enhance the ability of HMIRS to identify the root causes of hazmat accidents.

1. Require that the DOT number be a mandatory input for all reports filed with PHMSA for en route incidents.
2. Perform an additional Q/A check on carrier names to verify that the name being entered corresponds to the name provided on the annual PHMSA Registration form.
3. Require PKGFAIL entries to be filled out for all reports submitted to PHMSA.
4. Continue to emphasize the new requirement that carriers must file a 5800.1 form following an accident if there was damage to lading and lading protection systems on cargo tanks of 1,000 gallons or greater, even though there is no loss of hazardous material. This is the new requirement to report Class C accidents. Such a notice might be given to carriers when PHMSA notifies them that it has received and approved their annual hazmat registration application.
5. Capture driver condition information without compromising the confidentiality of the driver. The following design option from MCMIS can be enhanced for use in HMIRS. Based on analysis of the data, the list of options can be enhanced by using the following driver condition categories:
   1 = Appeared Normal,
   2 = Had Been Drinking,
   3 = Illegal Drug Use,
   4 = Sick,
   5 = Fatigue,
   6 = Asleep,
   7 = Medication, and
   8 = Unknown.

The project team believes that adopting the potential measures above would decrease errors in data entry and make it easier to query the database for potential causes of accidents.

4.3 Fatality Analysis Reporting System (FARS)

This section briefly describes the FARS file. Since the TIFA database incorporates the FARS records for trucks involved in fatal accidents, to avoid a fragmented analysis, much of the detailed evaluation is covered in Section 4.4, which describes TIFA.

The FARS file is the primary national crash data file for fatal traffic accidents. It is a census of all fatal motor vehicle traffic crashes. The TIFA file covers all medium and heavy trucks involved in a fatal crash, and includes virtually all FARS variables for the crash, vehicle, and driver. TIFA survey data supplements FARS data for trucks (hereafter the word "trucks" will be used to refer to medium and heavy trucks, i.e., trucks with a gross vehicle weight rating [GVWR] over 10,000 lbs). The TIFA data include a more accurate identification and description of trucks in fatal crashes, along with details about the cargo, configuration, motor carrier operating the vehicle, and crash type.

Both TIFA and FARS collect information about hazardous materials in the cargo. In the discussion of the variables that identify hazmat cargo in FARS, it will be shown that there are reasons to believe that the TIFA file identifies hazmat cargo more accurately. Since TIFA incorporates virtually all FARS variables, discussion of those variables and their usefulness will be discussed in Section 4.4, which focuses on the TIFA file.

4.3.1 Agencies/Organizations Responsible for Data Collection and Entry

FARS is compiled by the National Center for Statistics and Analysis at NHTSA.

4.3.2 Database Years of Coverage

The FARS file was initiated in 1975 and has been in continuous operation to the present time.

4.3.3 Criteria for Reporting and Inclusion of Data

The FARS file includes all traffic crashes involving

- A fatality that occurs as a result of a crash, or
- A fatality that occurs within 30 days of a crash, and
- At least one motor vehicle in transport on a public road.

4.3.4 Types of Hazmat Data Included

The FARS crash data file includes limited information regarding hazardous materials. Since 2005, the vehicle-related variables (up to two responses allowed) include a level that captures hazmat cargo releases as a result of a crash. The vehicle-related variables record pre-existing vehicle defects or special conditions related to the vehicle.

The vehicle configuration variable, added in 2001, identifies light trucks or passenger cars that display a hazmat placard. The driver-related variable (up to four responses allowed) includes an entry for "carrying hazardous cargo improperly." Finally, the hazardous cargo variable records if a vehicle was transporting hazardous material and if it was placarded.

Comparison of the identification of hazmat cargo in FARS and TIFA over a five-year period showed a large discrepancy. The FARS file identified hazardous material in the cargo in 1,257 cases over that period, while hazardous material was identified in only 1,049 cases in TIFA (see Figure 4-3). Surprisingly, when the comparison was made on a case-by-case basis, there was a large amount of disagreement between the files. As shown in Table 4-17, hazmat cargo was identified in both FARS and TIFA in only 706 cases over the observation period. FARS coded 551 trucks with hazardous material when the TIFA survey did not identify hazmat cargo, but in 343 cases, the TIFA survey showed that the truck had hazmat cargo and FARS did not.

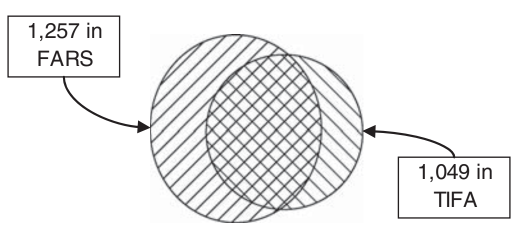

Figure 4-3. Comparison of trucks with hazardous materials in FARS and TIFA, 1999–2004.

Table 4-17. Hazardous materials in FARS and TIFA, 1999–2004.

Hazmat in cargo	Number of cases
Coded as Hazmat in FARS, not in TIFA	551
Coded as Hazmat in both FARS and TIFA	706
Coded as Hazmat in TIFA, not in FARS	343

Only a small part of the discrepancy is explained by differences in the ability to determine if the vehicle was transporting hazardous materials. In eight of the cases marked in FARS as carrying hazardous material, the TIFA survey was unable to determine if the vehicle held hazmat, while there were 33 cases in FARS where the analyst left the hazmat variable unknown, but the TIFA survey showed that the vehicle was carrying hazardous material. In most of the cases (880), the coding of hazardous material was directly contradictory (i.e., coded as hazmat in one and as not hazmat in the other).

Generally, there are a number of reasons to believe that the identification of hazardous material in the cargo is more accurate in TIFA than FARS. First, it is difficult in FARS to perform direct consistency checks on the hazmat variables, since FARS does not capture any other information about the cargo. Moreover, the FARS data collection protocol does not include direct contact with the carrier, driver, reporting officer, or other potential source, but relies primarily on the police report and other investigative documents. Moreover, the TIFA data collection protocol is based on a telephone survey of the motor carrier, driver, dispatcher, or safety director of the truck involved in the crash, as well as the reporting officer, so those sources are questioned directly. In addition, the TIFA data include other information about the cargo, so it is possible to perform basic checks on the accuracy of the hazmat coding, such as whether the truck was carrying cargo at all. The comparison of TIFA and FARS records showed that 264 of the cases coded in FARS were loaded with hazardous materials, while the TIFA survey showed that those trucks were empty at the time of the crash.

Finally, it should be noted that the TIFA survey specializes in trucks, while the FARS file covers all vehicle types. This focus on trucks allows the TIFA survey to go into greater depth and to develop more expertise in the details and varieties of truck operations. FARS analysts cover all vehicle types, and while the FARS file is a quality crash data source, it is not reasonable to expect FARS to have a higher degree of detail and accuracy than a file that has the advantage of narrowly focusing on only one vehicle type. In sum, while no doubt there are errors in the TIFA file, it is likely to be more reliable for analyzing truck crashes than FARS.

4.3.5 Usefulness of the Data for Determining Root Causes

Because the TIFA file incorporates the relevant data, a discussion of usefulness is deferred to Section 4.4 on the TIFA file.

4.3.6 Data Quality

FARS includes multiple layers of quality control. Cases are entered using computer software that includes validity and consistency checks. The validity checks ensure that the values entered are possible for the field. For example, if a field has valid values for one through seven, but an eight is entered, the eight would be rejected; similarly, alphas would be rejected if the field were numeric. There are also extensive computer consistency checks for each field, in which the value in one field is compared with the values in other fields to ensure that the fields are consistent. For example, if the first harmful event is a collision with a non-motorist, the field for the number of non-motorists involved must not be zero. There are multiple consistency checks for each field. Some of the checks are prescriptive, that is, certain values must be registered, while others flag unusual situations that should be reviewed. The consistency checks are documented in the *Coding and Validation Manual* (NHTSA 2002) for the FARS system. In addition, analysts receive annual training at a national meeting.

Similar quality control procedures are implemented as the cases are aggregated into the final file. The records are reviewed for timeliness and completeness. Statistical control charts are used to monitor the coding of key data over time, to see if distributions are wandering according to past experience. Typically, three versions of the FARS file are released. The "early assessment" file is released as a partial file that provides an initial look at the data for a year. Next there is a "complete" version that is typically released in the fall of the year following the data year. A "final" version, which includes all corrections and additional cases that have accumulated, is subsequently released, typically 18 or more months after the data year.

In some ways, the FARS file is the gold standard for data on fatal crashes. It is the product of considerable care over time, and is produced by a system that incorporates many checks for consistency and accuracy. On stable, well-understood data elements such as the environment of the crash, it is assumed to be of high quality and accurate. Items such as weather, time of day, light condition, and road type are coded from police reports and are dependent on the accuracy of the PAR. Although the accuracy of this information is unknown, these conditions are relatively stable and should be identified on the PAR with acceptable accuracy.

Similarly, it is assumed that the FARS file is acceptably complete (that is, virtually all vehicles involved in a fatal crash are included). However, given the sheer number of fatal crashes annually (about 40,000) and vehicles involved (about 55,000), it is not possible that every crash and every vehicle involved is included. In the process of compiling the annual TIFA file, each year, a small number of trucks appear on a police report for a fatal crash, but the FARS file contains no record for that vehicle. It is also possible that some fatal crashes are missed because the fatality occurs toward the end of the 30-day window. These few omissions are, however, undoubtedly very small in number and inconsequential.

The accuracy of the FARS data with respect to the main concerns of this project is a more complex matter. The TIFA file, indeed the entire TIFA protocol, allows an independent check for the accuracy of those data elements in common between it and the FARS file.

The inconsistencies with respect to the hazmat variables have been noted above. There also are problems with the identification of trucks in FARS. Cases extracted for the TIFA survey include some categories of vehicles that are not identified as medium or heavy trucks in FARS. These include light vehicles coded with a GVWR of more than 10,000 pounds. Each year, the TIFA survey determines that 200 to 300 of these vehicles are, in fact, medium or heavy trucks. In addition, the TIFA survey determines that a number of the vehicles identified in FARS as trucks are actually light vehicles. Typically there are 60 to 70 vehicles identified in FARS as a medium or heavy truck that prove to be a light vehicle or something other than a truck.

4.3.7 Additional Fields

The FARS system is quite complete and includes valuable fields. However, the addition of other data fields to the protocol would contribute to the ability to analyze crash causation. Some could be very easy to include, and require little modification of the program. Others would take additional resources, but would fit well within NHTSA's crash data program. Adding the following data fields is suggested:

- *Right of way.* This data element would identify which vehicle, if any, within a crash had the right of way prior to the collision. This could be readily coded from the PAR in most cases. Some state crash reports include right of way on the report. Right of way would be very useful in most crashes in identifying the vehicle that primarily contributed to the crash.
- *Accident type.* The General Estimates System (GES) file and Crashworthiness Data System (CDS) file both include an accident type variable coded at the vehicle level that captures the relative position and movement of the vehicle prior to its first harmful event. The TIFA data adds this to trucks in fatal crashes, but capturing this within the FARS system would be a valuable addition. An accident type field can identify key relationships that describe how the crash occurred and suggest contribution (for example, by identifying the vehicle that crossed over the center line in a head-on collision).

The following two fields would be useful although this would take additional resources and possibly require some change in the management of the FARS file:

- *Critical event* is a field that would identify and describe the event that precipitated the vehicle crash. This field is included in both the GES and CDS files, so the agency is very familiar with (and, indeed, invented) its use.
- *Critical reason* captures the "reason" for the critical event, classified broadly as driver, vehicle, or environment, with detailed levels under each. The variable is useful for identifying the immediate failure that led to the crash and would shed considerable light on crash causation. The field was used in the LTCCS, conducted jointly by the FMCSA and NHTSA, and in the National Motor Vehicle Crash Causation Survey (NMVCCS), conducted by NHTSA. Thus, the agency already has developed coding procedures for both variables.

However, adding these fields might require some changes to the FARS protocol. Both are difficult to code consistently and require a high level of focus and analysis. Currently, virtually all FARS fields are coded by analysts located off-site, that is in the 50 states and District of Columbia. But the coding of both GES and CDS is more centralized. In the LTCCS, both critical event and critical reason were coded by a small number of analysts in two locations. The National Center for Statistics and Analysis (NCSA) could adopt a similar method for the FARS file, if these data elements were added.

4.3.8 Potential Measures to Improve Data Quality

The FARS quality control system is complete and mature. It is subject to annual review and adjustment, including continuous training of the coders.

FARS might be improved if the system could be adapted to take advantage of the additional information provided through the TIFA system. FARS has not engaged TIFA in this regard, although one problem has been that information from TIFA has not been available in a timely fashion. However, greater cooperation between the systems would be valuable for both.

4.3.9 Compatibility with Other Databases

The FARS file does not include case identifiers that can be used to uniquely link to other data systems, such as the PAR number. Including the PAR number would provide a hard link. (Note that the MCMIS Crash file report number field in the past was supposed to include the PAR number in one of the fields, and it is recommended that MCMIS require that again. Currently, many states use a random report number, rather than using the PAR number.)

FARS does include information about the time and geographic location of the crash, as well as vehicle descriptive information, that can be used to obtain a probable link to records in other files. This includes date, time, state, county, city (if applicable), and an alpha string with the road name and mile marker. The file also includes latitude and longitude, although its accuracy is unknown.

FARS is highly compatible in another important way—capturing crash information using fields and code levels that are consistent with standard accepted practices. The code levels are almost always consistent with those available in other databases with elements in common. This includes the national crash files such as GES and CDS, as well as most state crash data systems.

4.3.10 Data Uses

NCSA uses FARS data along with GES data to produce an annual publication called *Traffic Safety Facts* (NHTSA 2008). This publication tracks annual trends for many crash factors of interest, such as vehicle involvements, deaths, injuries, restraint use, and drug or alcohol use. NHTSA and other analysts use FARS data in virtually all traffic safety analyses that require data on fatal crashes.

4.4 Trucks Involved in Fatal Accidents (TIFA)

4.4.1 Agencies/Organizations Responsible for Collecting and Entering Data into Database

The TIFA file is produced by the Center for National Truck and Bus Statistics at the University of Michigan Transportation Research Institute (UMTRI).

4.4.2 Database Years of Coverage

The TIFA file was initiated in 1980 and has been in continuous operation to the present.

4.4.3 Criteria for Reporting and Inclusion of Data

All cases in TIFA are also found in the FARS file, so TIFA shares the FARS reporting threshold as follows for crashes in which:

- A fatality occurs as a result of the crash, or
- A fatality occurs within 30 days of the crash, and
- At least one motor vehicle in transport on a public trafficway.

Additionally, the TIFA file includes only medium or heavy trucks, defined as trucks with a GVWR of more than 10,000 pounds.

4.4.4 Types of Hazmat Data Included

From its inception to 2004, TIFA data captured whether the cargo contained a quantity of hazardous material requiring a placard for each cargo body. Whether the hazmat cargo spilled as a consequence of the crash also was captured. The quantity of cargo is captured in terms of weight.

Package type information is not collected in detail. Cargo in TIFA is classified into general categories, such as general freight, solids in bulk, liquids in bulk, or gases in bulk. The specific cargo type, however, is recorded in an alphanumeric string. Examples of the type of information recorded in this field include gasoline, anhydrous ammonia, or jet fuel. More detail, such as the Motor Carrier (MC) type of liquid or gas tank, is not considered.

In the 2005 data year, the amount of data recorded for hazardous material was both expanded and restricted. Now, whether the truck was hauling hazardous cargo is recorded at the vehicle level, not separately for each cargo body. In addition, TIFA no longer collects cargo weight. Concurrently, more detail about the hazardous material is now recorded, including the hazmat class and UN number.

4.4.5 Usefulness of the Data for Determining Root Causes

This section provides a discussion of the variables in the TIFA Crash file that relate to understanding the factors that contributed to the crash or affected the severity of its consequences. Most of the variables are extracted from the FARS file and are added to the TIFA file without change. Of the contributing factors identified in Table 4-18, the TIFA file provides the information on truck configuration, cargo body, gross vehicle weight (GVW) (and gross cargo weight [GCW]), and accident type.

Virtually all of the factors listed in Table 4-18 are captured in a single variable. For example, roadway surface condition is described in one variable with several possible levels. Driver condition, however, is captured across several variables. There is no single variable that consolidates the information about the driver's condition at the time of the crash. Instead there are separate variables that identify whether the driver had been drinking or using illegal drugs. Driver fatigue or illness is captured in a "driver-related-factors" variable that includes a variety of driver state, action, and other conditions.

In this discussion, it is useful to recognize that there are two types of parameters: conditions present and conditions contributing to producing the crash. Both types can be contributing or root causes of accidents. For example, the weather at the time of the accident is a condition and it may or may not be a contributing or root cause of an accident. Clearly, if it is not captured, there will be no information to implicate weather as a contributing or root cause. The parameters contributing to producing a crash include parameters like driver condition pre-crash, conditions such as a car cutting off a truck, forcing it to take evasive action to prevent the accident. The important point is that to identify contributing and root causes of accidents, which parameters serve a specific function as to recognize that both types of parameters are needed. Descriptive information

Table 4-18. Contributing factors present in TIFA/FARS.

Vehicle	Driver	Packaging	Infrastructure	Situational
Configuration	Age	Package Type	Road Surface	Pre-Crash Condition
Cargo Body	Experience	Quantity Shipped	Road Condition	Event Type
GVW	Condition	Quantity Lost	Road Type	Vehicle Speed
	Valid License	Age (Cargo Tank)	Traffic Way	Impact Location
	Citation Issued	Rollover Protection	Access Control	Primary Reason
			Speed Limit	Accident Type
		Inspection History	No. of Lanes	Weather Condition
		Design Specification		Light Condition
				Time of Day

Key:
Factor obtained
Partially met
Not captured

characterizes conditions at the time of the crash. These include items such as light condition, roadway surface condition, time of day, and even road type. These pieces of data describe the circumstances and many can be related to changes in crash risk. That is, the condition *may* have contributed to producing the crash, but coding the condition does not tell us whether it did, just that it was present. For example, wet or icy roads increase the possibility that a vehicle will skid while maneuvering, but the fact that the road was coded as wet does not imply a judgment that the lower road friction contributed to producing the crash. However, the TIFA file also includes a set of data elements that identify actions and conditions that are either identified as contributing to the crash or that strongly imply they contributed to producing the crash. There are separate sets of roadway, vehicle, and driver "related factors" that identify defects, conditions, and actions that contributed to the crash. Charged traffic violations are also coded for the driver. These two types of variables—conditions present and conditions producing—will be discussed separately, with the producing factors discussed first.

There are four variables that directly identify factors that represent a judgment of contribution to a crash:

- Driver-related factors,
- Vehicle-related factors,
- Accident-related factors, and
- Violations charged.

Each of these factor variables are coded by the FARS analysts from the PAR or other investigative reports that the analysts may be able to obtain. That is, the analysts capture the recorded judgment of the reporting officers, rather than applying their own judgment or inferring conditions and actions from the available evidence. Accordingly, information coded in these factor variables has been explicitly stated by the original crash investigators.

The variables for driver-related factors record driver conditions or actions that may have contributed to producing the crash. Up to four responses may be recorded in these variables. The variables for driver-related factors that can be coded are divided into five different subsections that capture different types of influences on the driver. The first subsection includes driver conditions, such as fatigue, asleep, ill, blackout, or some other condition that impaired the ability of the driver to control the vehicle. The second subsection is called "miscellaneous factors" in the coding manual, but really each is a driver error of some sort. Examples of these errors are running off the road, driving with a suspended license, speeding, following improperly, and failure to yield. The next subsection of variables for driver-related factors identifies different conditions that obscured the driver's vision, including reflected glare, fog, trees, other vehicles, and buildings. There also is a subsection for conditions or events that caused the vehicle to swerve or skid, including wind, slippery road, objects or holes in the road, and swerving to avoid animals or pedestrians. Finally, there is a subsection to capture the presence of possible distractions, including cell phones, computers, and navigation systems.

With the exception of the codes that merely specify possible distractions, the information captured in the variables for driver-related factors record the original police investigator's judgment on conditions influencing, or errors made by, the driver that helped to produce the crash. All the others point to driving errors, unsafe driver conditions (e.g., fatigue), or conditions that contribute to a high-risk situation, such as view obstructions or causes of skidding or swerving. Each of these may be regarded as identifying a causal factor in the collision.

Overall, the variables appear to be both useful and reasonably consistent. Missing data rates are low, with the variable left unknown in just over 1% of cases. Previous analysis has shown that the coding is reasonable. In two-vehicle collisions, typically only one driver will be coded with a factor, although there are some instances where both drivers are coded as having contributed to the crash. This makes some intuitive sense, because we generally think of crashes being primarily due to poor actions by one of the drivers. In addition, the codes have been shown to be consistent with the general configuration of the crash. For more on this subject, see *The Relative Contribution of Truck Drivers and Passenger Car Drivers to Two-Vehicle, Truck-Car Traffic Crashes* (Blower 1998).

The variables for vehicle-related factors capture vehicle defects that may have contributed to the crash. The FARS coding manual indicates that the variable is used to record pre-existing vehicle defects (i.e., defects not caused by the crash itself). The FARS manual (NHTSA 2002, p 344) also states that "The vehicle condition(s) noted only indicate the existence of the condition(s). They may or may not have played a role in the accident." Thus, according to the manual, defects noted may not have contributed to the crash. However, in light of how the data are collected, it is likely that the defects played a role. The source for the data is ultimately the PAR or other crash investigation document. Crash investigators typically do not routinely perform a vehicle inspection, but primarily detect vehicle defects if they contributed to the crash. Reviews of the PARs of crashes in which tire defects were recorded in the vehicle-related factors variable showed that in every case the tire defect was mentioned as contributing to producing the crash.

Missing data rates for the vehicle-related factors variable are low, averaging just under 2%. However, it is very likely, at least for trucks, that vehicle factors are greatly underreported. Typically, the vehicle-related factors record about 2% of the trucks in fatal accidents with brake problems, and much lower percentages with light system defects. However, roadside inspections of trucks—not crash-involved trucks but just trucks operating on the road—show much higher rates, with about 25% having one or more brake defects and similar percentages with light system problems. For additional discussion of this topic, see "Vehicle Condition and Heavy Truck Accident Involvement" (Blower 2002). It appears that the variables for vehicle-related factors significantly underreport the true incidence of vehicle defects in crashes. This is not surprising, since police officers are not trained to do vehicle inspections. Such underreporting, however, limits the use of the data.

The variables for accident-related factors record conditions at the accident level that may have contributed to the crash. These include roadway design problems and roadway defects such as worn or missing pavement markings, inadequate warnings, and roadway washouts. Additionally, the variable captures special circumstances that may have affected the crash, such as previous crashes nearby and police pursuits. Again, the variable is coded from the original PAR, with the requirement that the coded items were specifically mentioned. They are coded as factors that either existed or contributed to producing the crash.

Overall, missing data rates for the variables for accident-related factors are also low, and are not coded in only 0.3% of cases. However, in more than 93% of the cases, no accident-related factor is coded. Given the very slight incidence of roadway defects recorded in the variables, road design and condition play a small role or the source of the information is unable to capture its contribution to crashes.

Turning to the variables that capture conditions present or characteristics of the vehicle, driver, or environment, each may contribute to a specific crash, but the coding of the variable does not directly establish contribution. Instead, these factors are associated with crash risk—they may, in a sense, establish the preconditions for a crash. Table 4-19 lists all of the contributing factors for each analytical level within an accident, and provides a brief comment characterizing how the information is available in the TIFA database.

For the variables relating to infrastructure, all are taken from the FARS file and are coded in formats (i.e., code levels) that match national standards. For example, the same code levels are used in MCMIS, GES (General Estimates System), CDS (Crashworthiness Data System), and

Table 4-19. Coverage of contributing factors in TIFA.

Level	Variable	Comment
Vehicle	Configuration	Captures the exact configuration of the vehicle
	Cargo Body	Distinguishes 16 types of cargo bodies, including liquid, dry bulk, and gaseous tank types
	GVW	Prior to 2005, captures both gross vehicle weight and gross combination weight; weight variables dropped for 2005 and later
Driver	Age	Captured
	Experience	Not captured
	Condition	Captured in separate variables that identify alcohol use, drug use, and a set of multiple-response variables that code fatigue, asleep, ill, emotional, distracted, etc.
	Valid License	Captured
	Citation Issued	Captured in a multiple-response variable; up to four citations may be recorded
Packaging	Package Type	Captured only to the level of cargo body type
	Quantity Shipped	Weight only, prior to 2005; not captured in 2005 and later
	Quantity Lost	Spill/no spill captured only, not quantity spilled
	Tank Age	Not captured
	Rollover Protection	Not captured
	Inspection History	Not captured
	Design Specification	Not captured
Infrastructure	Road Surface	Captured, in standard format
	Road Condition	Captured, in standard format
	Road Type	Captured, in standard format
	Traffic Way	Captured, in standard format
	Access Control	Not captured directly; instead, inferred from roadway function class
	Speed Limit	Captured, in standard format
	No. of Lanes	Captured, in standard format
Situational	Pre-Crash Condition	Captured as pre-crash maneuver; standard format
	Event Type	Captured as first harmful event; standard format
	Vehicle Speed	Captured in standard format
	Impact Location	Captured as relation to roadway of the first harmful event; standard format
	Primary Reason	Inferred from the vehicle-, crash-, and driver-related factors discussed above
	Accident Type	Coded in the TIFA survey to match similar variable in GES and CDS
	Weather Condition	Captured in standard format
	Light Condition	Captured in standard format
	Time of Day	Captured in standard format

most state files. Characteristics of the vehicle are available in greater detail than other crash files, but code levels could be aggregated to match other crash data files if necessary. With respect to drivers, all of the information is available, except driver experience. The factors listed as "situational" also are available in formats that match national standards. Pre-crash condition identifies the vehicle's maneuver prior to the initiation of the crash sequence, and the format used has been widely adopted. The accident-type variable, which captures the relative motion and position of the vehicles prior to the collision, is very useful in understanding how a crash occurred. The format used in the TIFA file is also used in GES and CDS.

The primary area in which the TIFA file falls short relates to the details of packaging. TIFA includes only the general cargo body type with no information about the design specification, inspection history, or rollover protection. Moreover, quantity shipped is partially captured for 1999-2004, but not captured after that. For all years prior to 2005, quantity shipped was partially captured in the TIFA file as cargo weight, in pounds, for all cargo, but capturing this information was aban-

Table 4-20. Completeness of contributing factors in TIFA/FARS.

Vehicle	Driver	Packaging	Infrastructure	Situational
Configuration	Age	Package Type	Road Surface	Pre-Crash Condition
Cargo Body	Experience	Quantity Shipped	Road Condition	Event Type
GVW	Condition	Quantity Lost	Road Type	Vehicle Speed
	Valid License	Age (Cargo Tank)	Traffic Way	Impact Location
	Citation Issued	Rollover Protection	Access Control	Primary Reason
		Inspection History	Speed Limit	Accident Type
		Design Specification	No. of Lanes	Weather Condition
				Light Condition
				Time of Day

Key:
> 95%
50% to 95%
< 50%
Not captured

doned after 2005. The cargo weight variable was known precisely for about 84%, unknown for about 4%, and partially known (light load or full load) for the remaining 12% of the cases.

Generally, the fields derived from the TIFA survey or extracted from FARS are populated fairly completely. Table 4-20 shows that missing data rates for the fields captured are quite uniformly low. If the field is present, for the most part, it is complete in more than 95% of the cases. GVW and quantity shipped are exceptions, because it can be difficult to determine precise values after the fact, but even for those variables there is some information for 80% to 90% of the cases. Missing data rates are somewhat higher for vehicle speed prior to the crash, as that is even more difficult to determine. That information is taken from PAR and is available for about 75% of the cases. Complete rates of missing data, averaged over five years, are provided for all TIFA variables in Appendix D (available on the TRB website at www.TRB.org by searching for HMCRP Report 1).

4.4.6 Data Quality

The TIFA system includes multiple layers of quality control. The survey is administered by means of a telephone interview. Each case record is reviewed by an editor for accuracy, consistency, and completeness. The vehicle identification number of the power unit is decoded, and the vehicle description from the survey is compared with the original specifications. Information about cargo and operations are similarly compared with a library of information that has been accumulated from over 25 years during which the TIFA survey has been conducted. Survey information also is compared with the information on the original PAR. Any discrepancies are discussed with the interviewer, who may be required to make additional calls for information.

Once discrepancies are resolved, the data are then entered with verification. At that point, there is a computerized check of each batch of keypunched cases for consistency and to identify any invalid codes or responses that are outside of the usual range. Unusual responses are reviewed by the data editors. Finally, when a data year is complete, there is a computerized check of all cases for invalid codes, inconsistent data, or unusual responses.

4.4.7 Additional Fields

Although the TIFA survey adds valuable detail to the FARS data, additional data fields could add important detail about hazmat packaging and also enhance the information available on crash causation. Some data fields could be added very easily, with little modification of the program. Including other data fields would take additional resources. Recommended additional fields are as follows:

- The right-of-way data element could identify which vehicle, if any, within a crash had the right of way prior to the collision. This could be coded readily from the PAR in almost all cases. Some state crash reports include right of way on the report. Right of way would be useful in most crashes in identifying the vehicle that primarily contributed to the crash.
- Critical event is a field that identifies and describes the event that precipitated the crash for the vehicle. This field is included in both the GES and CDS files. Coding manuals are available and could be used to ensure that coding is consistent with NCSA standards.
- The managers of the TIFA survey might also consider adding a field for critical reason. Critical reason captures the "reason" for the critical event, classified broadly as driver, vehicle, or environment, with detailed levels under each. The variable is useful for identifying the immediate failure that led to the crash and would shed light on crash causation. The field was used in the Large Truck Crash Causation Study (LTCCS), conducted jointly by FMCSA and NHTSA, and in the National Motor Vehicle Crash Causation Survey (NMVCCS), conducted by NHTSA. Therefore, coding procedures are available. The suggestion to add a critical reason data field to TIFA, however, is subject to whether the information is available within the TIFA protocol, which relies primarily on PARs, to code this field.

The TIFA program could add the following additional information about hazmat cargo:

- MC number of the cargo tank, which has been collected in the past as part of a special data collection effort and, therefore, the feasibility of collecting this information has been demonstrated.
- Quantity of hazardous material transported, which would entail adding cargo weight data fields back to the survey. The program could consider capturing the quantity in terms of liquid measure, where appropriate.

4.4.8 Potential Measures to Improve Data Quality

The TIFA system is complete and mature. It is subject to annual review and adjustment, including continuing training of the coders. However, greater cooperation with the FARS program might help increase the accuracy with which trucks are identified.

4.4.9 Compatibility with Other Databases

The comments in Section 4.3.9 on the FARS data apply equally to TIFA, since TIFA uses the same case number system. However, there is an additional constraint on linking the TIFA file to other data systems, if that raises a risk of identifying specific individuals or organizations. The TIFA program is bound by commitments to respondents to protect their identity, and by the terms of its operations under the University of Michigan's Institutional Review Board (IRB). Thus, any effort to link the data to other data systems would be unlikely to be allowed.

4.4.10 Data Uses

FMCSA uses TIFA data for a variety of research purposes. In addition, the TIFA data are used by researchers at UMTRI and other universities for traffic safety research. The research is designed to identify the scope of traffic safety problems related to trucks and to identify risk factors in truck crashes, whether related to the vehicle, driver, operation, or environment. This information is used by government entities—including FMCSA, NHTSA, and certain states—for regulatory purposes.

In terms of uses relating specifically to the hazmat data in TIFA, over time, FMCSA has published data about the scope of hazmat crashes. In addition, FMCSA has requested data about fatal hazmat crashes to monitor trends.

4.5 Large Truck Crash Causation Study (LTCCS)

LTCCS was designed as a one-time study to compile a comprehensive set of accident data for approximately 1,000 large truck accidents. The data compilation began in 2001 and was completed in 2003, although analysis of the data is still ongoing.

4.5.1 Database Description

The LTCCS database is a series of tables that can be run in several relational databases. For purposes of this assessment, the tables were imported into the Microsoft Access database. The master table is the Crash file and contains 1,070 records representing the number of discrete accidents that were investigated. There were 107 of these accidents that were eliminated for one reason or another because they did not meet the analysis criteria. Thus, the LTCCS evaluates 963 discrete accidents. The next table is the *CRASHASSESSMENT* Table, containing a total of 2,284 records representing the data for the number of vehicles involved in the 1,070 accidents. Although all of these accidents involved at least one heavy truck, the second vehicle in many of the accidents was a passenger vehicle. Because the purpose of the study was to collect sufficient data to identify the causes of the accidents as well as to collect information on the drivers, vehicles, and environment present at the time of the accident, a lot of supporting information is provided.

The dataset contains a total of 43 separate tables summarizing the many factors that, in total, represent a comprehensive set of data for each accident (see Table 4-21). Table 4-21 includes

Table 4-21. LTCCS tabular structure arranged by category.

Overall	Driver	Vehicle or Packaging	Carrier	Environment or Situational
Crash	Driver Assessment	Air Bags	SAFER Authority Status	Environment
Crash Assessment	Driver Decision Aggression	Brakes	SAFER Carrier	Factor Assessment
Events	Driver Drugs	Cargo Shift	SAFER Insurance	
Injuries	Health	Crush	SAFER Review	
SAFER Inspection Summary	Driver Recognition Distraction	General Vehicle	SAFESTAT Rating	
Jackknife	PAR Violations	Hazardous Material	Carrier Interviews (1 Table)	
Non-Motorist	Sleep	Hazmat Inspection		
Occupants	Driver Interviews (14 Tables)	Truck Exterior		
	SAFER Crash Report	Truck Inspection		
	SAFER Driver Inspections	Overview		
	SAFER Driver Violations	Truck Units		
	MCMIS Driver	Vehicle Events		
	MCMIS Violations	Vehicle Exterior		

Note: Safety and Fitness Electronic Records (SAFER) System; Safety Status (SAFESTAT) Management System.

references to tables that summarize the results of driver and carrier interviews. Interview tables contain data on driver condition (aggressive driving, attention, condition, license status, fatigue, health, perception, and sleep). Interview tables also are provided for cargo shift, fire, jackknife, rollover, and trip.

Most of the database tables are related using two parameters, the case number and the vehicle number. Driver information is collected for all 2,284 vehicles, but it was possible to collect information on driver health for only 1,839 vehicles, about 80% of all vehicles involved in the accidents. Other significant data are contained in the *ENVIRONMENT* and *FACTOR_ASSESSMENT* Tables, conveying information on the *GENERAL_VEHICLE, VEHICLE,* and *TRUCK_UNIT* Tables that provide additional details.

As might be expected, although there are driver interview data for most of the 2,284 vehicles, the cargo shift data are not listed for all vehicles because cargo shift is not applicable to passenger vehicles. In total, cargo shift data are provided for 1,071 vehicles. Although very close to the number of accidents investigated, 1,070, there are multiple trucks involved in many of these accidents, so the data are actually provided for about 90% of the heavy trucks involved in the 1,070 accidents. The database contains descriptions for 1,207 heavy trucks and 29 bobtails (power units without a semitrailer). The *PAR_VIOLATION* and *BRAKES* Tables list defects found at the accident scene. The latter comes from a Commercial Vehicle Safety Alliance (CVSA) Level 1 inspection of the vehicles following the crash. The *CDCRUSH* Table lists the type and extent of vehicle damage, some coming from accident reconstruction analyses. Although not many vehicles were carrying hazardous material, there are two tables, *HAZMAT* and *HAZMAT_INSP*, that provide information specific to the packaging and hazardous material involved in the accident. The LTCCS included 57 vehicles carrying hazardous material with 77 material types, and provided detailed event descriptions for 30 of the vehicles.

4.5.2 Purpose and Function

The purpose of the LTCCS was to determine the causes of, and contributing factors to, crashes involving commercial motor vehicles. The study was mandated by the Motor Carrier Safety Improvement Act of 1999, P.L. 106-159.

4.5.3 Data Collection

A nationally representative sample of large-truck fatal and injury crashes was investigated during 2001 to 2003 at 24 sites in 17 states. Each crash involved at least one large truck and resulted in at least one injury or fatality. Data were collected on up to 1,000 parameters in each crash. The total sample, after non-qualifying accidents were eliminated, involved 967 crashes, which included 1,127 large trucks, 959 non-truck motor vehicles, 251 fatalities, and 1,408 injuries.

The data for each accident were collected from a wide variety of sources. These included the General Vehicle Form, the police accident report, medical reports, scene photographs, and post-accident inspections, including the CVSA Level 1 inspection of the truck involved in the accident. In addition, interviews were conducted with the drivers, witnesses, vehicle occupants, and carrier personnel. Local weather station data was used to describe weather conditions and the driver's log was used to determine hours of service. AASHTO documents provided the criteria used to determine the values that should be assigned to parameters such as the driver's line of sight at the time of the accident. Time-stamped toll and fuel receipts were also collected. Relevant data were also obtained from federally maintained databases. These databases included the MCMIS Registration File, and the Safety and Fitness Electronic Records (SAFER) system and Safety Status (SAFESTAT) Management System databases. Although the data collection involved using trained investigators to visit the site of each accident, the data collection was performed in phases so the highway where the crash occurred would not be unduly blocked while all the data were collected. The CVSA Level 1 inspection was conducted at the repair facility, not at the accident scene. Advanced photographic techniques were used to enable the compilation of scaled schematics and scene measurement logs. The typical approach was to mark key points in the accident progression while the vehicles were present and then go back later and take more extensive measurements. In addition, the accidents that were selected occurred close to the 17 locations where trained investigators resided. No effort was made to require that the investigators travel great distances, thereby forcing long-term closure of the highway. Simply stated, the data collection and compilation were designed to minimize disruption yet, at the same time, collect data on many relevant parameters.

4.5.4 Data Compilation

The data from interviews, photographs, accident scene measurements, and vehicle inspections were used to populate many of the parameters in the database. For example, the interviews with the carrier and driver were used to compile data on the driver's previous sleep interval, the hours of service recorded on the log, as well as data on the driver's mental and physical state. Similarly, the measurements taken at the scene were used to generate scaled schematics and the scene measurement log. Data from the interviews of the driver and carrier were used to identify the driver aggression and driver distraction factor number. Photographs and on-scene measurements were compiled into deformation logs to be placed in the database as deformation codes.

4.5.5 Accuracy and Completeness of Data

Every effort was made to obtain a comprehensive set of data for more than 1,000 parameters. Many of the parameters were estimated from multiple sources of data and apparent or real inconsistencies could be resolved normally, thereby producing a consistent dataset for each vehicle involved in the accident. In addition to the 43 tables in the database, there are numerous supporting tables that define code numbers to be used instead of phrases or words. This increased the accuracy of entry among data compilers. In this regard, the LTCCS project generated a 512-page *Analytical Users Manual* to ensure that all of the parameters' codes were consistently entered into the database tables. The lengthiness of the manual is due primarily because a definition, source, cross reference, variable name, and attribute code ID was provided for each parameter.

4.5.6 Quality Control

Extensive quality control checks were performed to ensure the accuracy of the data put into the database. The use of attribute codes that are defined in the *Analytical Users Manual* greatly enhances the quality of the data. Based on the manual, codes are used for most parameters and since the manual defines the meaning for each of the numerical codes, there is little room for ambiguity. This minimizes the inconsistencies in the dataset.

4.5.7 Interconnectivity with Other Databases

It is not possible to connect the data in the LTCCS with other databases because the location and day of the month in which the accident occurred has been removed from the public version. The carrier is not named, the DOT number of the carrier is not given, and the vehicle registration number has been shortened. Any of these parameters would enable the datasets to be joined. There are only 57 crashes that involved hazardous material and, since HMIRS mainly reports spills, there are probably fewer than 10 crashes that might be reported in both datasets.

4.5.8 Analyses Using Database

The LTCCS raw dataset was presented to the analytical community in 2006 and numerous analyses have been performed on the dataset. The analyses presented here will focus on the hazmat truck accidents. Table 4-22 shows the breakdown of hazmat shipments included in the LTCCS. If a reportable quantity was being shipped, then the shipment would have to be placarded. There are only 40 placarded vehicles analyzed in the LTCCS. Table 4-22 shows that slightly more than 40% of the reportable shipments, 17, were Class 3 materials. Class 2 was the next most common, with 8 vehicles out of the 40, about 20% of the total. Thus, Class 3 and 2 shipments make up more than 60% of the total hazmat vehicles included in the LTCCS.

For the 40 reportable accidents, the database can be queried to look at health factors, as shown in Table 4-23.

Of the health factors listed in Table 4-23, other than requiring corrective lenses, almost all of the entries identify no health factors that might have contributed to the accident. To get better statistics for health issues that could affect safety would require a health assessment to be collected for at least 400 or more hazmat incidents. With 400 drivers, it might be possible to address the contribution from heart attacks. More than 1,000 would be required to get valid statistics on less common health conditions. This implies that if driver health is a contributing cause, it probably has to be captured in all hazmat truck accident records, as it was a few years ago in MCMIS, and for some reason has been left blank in MCMIS beginning in CY 2002.

Drug use by the driver also was tabulated. In the 40 drivers hauling hazardous materials, there were 10 drivers taking prescription drugs and 6 taking over-the-counter drugs. There were no instances where the driver was taking illegal drugs. In the 40 accidents involving a hazmat vehicle, in three of the accidents, the driver of the other vehicle was believed to be taking illegal drugs. A drug test verified the presence of the drug in one case and in the two others, the results of the drug test was unknown.

Table 4-22. Types of hazardous materials included in vehicles in LTCCS.

	Reportable Quantity Specified in 172.101 Hazardous Material Table [49 CFR 172.101]		
Material	**Yes**	**No**	**Unknown**
2.1 Flammable Gas	5	1	
2.1 LPG	1		
2.2 Nonflammable Gas	2	2	
3 Combustible Liquid	5		
3 Flammable	12	3	2
4.1 Flammable Solid	1		
4.3 Dangerous When Wet	1		
5.1 Oxidizer	1		
6.1 (Liquids)	1	1	
6.1 Zone A	1		
8 (PIH) Zone A	1		
8 Corrosive Material	6	1	
9 (Elev Temp Materl)	1		
9 (Hazardous Waste)	1		
9 Miscellaneous HM	1	1	
Total	**40**	**9**	**2**

Table 4-23. Listing of health factors present for vehicles in LTCCS containing hazardous materials.

Health Factor	Total	Health Factor	Total
Illness Factor Count	1	Astigmatic	2
Heart Attack	1	Other Vision	6
Epileptic Seizure	0	Unknown Vision	7
Diabetic Blackout	0	Other Factor Count	1
Other Blackout	0	No Factors	32
Cold Flu	0	Hearing Impairment	0
Other Illness	1	Prosthesis	0
Normal Vision	17	Paraplegia	0
Legally Blind	0	Strenuous Recreation	0
Myopic	5	Strenuous Non-Work	1
Hyperopic	4	Sleep Apnea	0
Glaucoma	0	Other Factor Physical	0
Color Blind	0		

The *ENVIRONMENT* Table provides a lot of information that is useful for defining the characteristics of the accident location. Table 4-24 shows the relationship between the *JUNCTION* and *INTERCHANGE* parameters for the 40 placarded trucks included in the LTCCS dataset.

As shown in Table 4-24, the data are internally consistent. There are no entries with *INTERCHANGE* = Yes that are not entered under a *JUNCTION* category related to an interchange. The data also show that of the hazmat truck accidents, more than 25% (11/40) occur at interchanges. Table 4-25 looks at the same data using the parameters *INTERCHANGE* and *FUNCTIONAL_CLASS.*

Table 4-24 shows that the 11 hazmat trucks in the LTCCS with *INTERCHANGE* = Yes are all associated with Interstate highways. Table 4-25 shows that there are also 16 additional accidents on Interstate highways not associated with interchanges. This means that of the 27 crashes involving hazmat trucks, 40% of the accidents are at interchanges. Given that interchanges occur only every few miles, on a per mile basis, accidents at interchanges on Interstate highways clearly dominate.

Although one might be tempted to look at the ratio of urban and rural freeway accidents from these data (2 rural and 25 urban), the fraction of the miles driven in urban and rural areas is not known. Thus, it is difficult to infer an accident rate from these data alone. In HMIRS, the origin and destination of the shipments is shown and if these data could be matched with HMIRS records, it would be possible to derive an estimate of these rates. There is another factor that could enter into the analysis as well. In the LTCCS, the accidents were not selected randomly; they had to be close to the location of 17 accident investigation teams and, as a result, the accidents selected could be biased toward urban accidents.

Table 4-24. Relationship between junction and interchange in LTCCS.

Junction	Interchange	
	No	Yes
Entrance/exit ramp related	1	7
Intersection	2	
Intersection related	2	
Non-junction	22	
Other location in interchange		4
Rail grade crossing	2	

Table 4-25. Relationship between functional class and interchange in LTCCS.

Functional Class	Interchange	
	No	Yes
Rural local	2	
Rural minor arterial	4	
Rural principal arterial – Interstate	1	1
Rural principal arterial – other	3	
Urban minor arterial	2	
Urban principal arterial – Interstate	15	10
Urban principal arterial – other	2	

In a typical data analysis, it is difficult to analyze just a few accidents. Thus, while it is possible to look at the decision factors associated with those 11 interchange accidents, the statistical uncertainty regarding the conclusion will be very high. Clearly, to identify significant differences would require more hazmat truck accidents in the dataset.

Just as the LTCCS targeted large truck accidents involving a serious injury or fatality, a comparable study that focused on hazmat accidents would provide a similar benefit. Rather than doing a two-year study of 1,000 truck accidents, there appears to be merit to doing a continual study of fewer truck accidents, perhaps 100 to 200 per year. To look for differences between hazmat truck accidents and regular truck accidents, it would be important to have data for both, perhaps a sample of 100 regular truck accidents and 100 hazmat truck accidents. If such a study were performed on an annual basis, it is important to have weighting factors to enable the findings from a limited sample of accidents to be related to the universe of accidents occurring annually. These can be developed as part of the sampling methodology or come from other databases such as MCMIS and HMIRS.

4.5.9 Summary and Potential Measures to Improve Root Cause Analysis

The analysis of the data from the LTCCS is still ongoing, so the following summary is based on its status as of the time of this report. The potential measures are prepared to focus on the objectives of this project.

4.5.9.1 Summary

The LTCCS represents a comprehensive analysis of serious, large truck crashes. The variables captured in the 967 accidents investigated by contributing cause category are shown in Table 4-26. As shown, all of the contributing factors listed under the categories for Vehicle and Situational and most of the contributing factors under the categories for Driver and Infrastructure are covered. The Infrastructure category's factors are actually known by the LTCCS analysts, but have been coded to prevent these data from being known by those outside the LTCCS program. Thus, the training and experience of the driver were the only contributing causes that are not captured under the Driver category. The Packaging category is not well captured, since package behavior was not the focus of the LTCCS.

4.5.9.2 Potential Improvements Based on the LTCCS Experience

Comprehensive studies, such as the LTCCS, are needed to obtain contributing and root causes of accidents. Similar to the LTCCS, these detailed analyses can be focused on a sample of all the accidents occurring in the United States, provided that the weighting of the sampling is known.

Table 4-26. Summary of variables captured in LTCCS.

Vehicle	Driver	Packaging	Infrastructure	Situational
Configuration	Age	Package Type	Road Surface	Pre-Crash Condition
Cargo Body	Experience	Quantity Shipped	Road Condition	Dangerous Event
GVW	Condition	Quantity Lost	Road Type	Vehicle Speed
Vehicle Defect	Valid License	Age (Cargo Tank)	Traffic Way	Impact Location
Vehicle Response	Citation Issued	Rollover Protection	Access Control	Primary Reason
	Driver Response	Inspection History	Speed Limit	Accident Type
	Training	Design Specification	No. of Lanes	Weather Condition
			Location	Light Condition
				Time of Day
				Health Consequences

Key:
Variable obtained
Partially met
Not captured

If a comprehensive study of selected classes of accidents is performed, there are significant advantages to performing a selected number of accident investigations annually rather than performing a larger intensive study over a one- to two-year period as was done for the LTCCS. The advantage of continuing studies is that the investment in training can be realized over many years, and trends in the data also can be evaluated.

Any program for identifying root and contributing causes could, for selected classes of accidents

1. Obtain data taken from interviews with drivers and other witnesses with information about the crash. There are many factors that cannot be obtained unless post-vehicle-inspections and reviews of driver qualifications are conducted.
2. Obtain data collected from SAFER, SAFESTAT, and the MCMIS Registration file. Such data would provide information on the programs to ensure that hazmat is being transported using well-trained drivers in safe vehicles.
3. Visit the scene of the accident to obtain first-hand observations, including photographs of the scene, vehicle damage, scaled measurements, and a scene measurement log. Such information is critical for accident reconstruction and also to resolve any data inconsistencies.
4. Consider the techniques developed in the LTCCS. It is not considered feasible to perform the level of analysis performed in the LTCCS on all hazmat accidents. However, just as the LTCCS looked at injury and fatality accidents involving heavy trucks in a selected area, so it would be possible to perform the LTCCS level of analysis for perhaps 50 to 100 hazmat accidents annually, perhaps those involving a specific type of hazardous material (e.g., TIH [toxic inhalation hazard] cargo tank shipments). Such analyses could be multimodal if accident investigations were coordinated.

4.6 Railroad Accident/Incident Reporting System (RAIRS)

Prevention of hazmat releases caused by railroad accidents differs from other modes in a variety of respects due to physical, operational, and institutional factors. Some of the important differences include

- Railroads typically operate trains rather than single vehicles;
- Operation is restricted to a fixed guideway or "single-degree-of-freedom" system;
- Railroad track age and other infrastructure is generally privately owned and maintained;
- Much of the operational control is either automated or controlled by individuals other than the train operating crew;
- Many railcars spend a substantial fraction of the time operating on and under the control of a railroad other than its owner; and
- These factors all have the effect of reducing certain general types of failure that can lead to a hazmat release, while elevating the importance of certain others.

Although human error is an important cause of railroad accidents, failure of infrastructure or vehicle components comprise a much larger percentage of hazmat accidents. Additionally, failures in the traffic control system sometimes cause accidents. These may be mechanical or electrical in nature, or they could be caused by human errors committed by personnel other than those operating the train, and who may be located hundreds of miles from the scene of the accident.

To support this type of approach, FRA and the railroads have a comprehensive accident reporting system that has roots dating back to 1910 and was implemented in its present form beginning in 1975. FRA regulations require that all accidents in which damage to track and equipment exceeds a specified monetary threshold (adjusted periodically for inflation) must be reported using Form FRA F 6180.54, the Rail Equipment Accident/Incident Report, which records 52 different variables regarding the circumstances and cause of the accident. Beyond this, major railroads maintain their own internal databases. These typically contain all of the information necessary to comply with FRA reporting requirements, and may contain additional data that individual railroads believe is useful for their own safety analysis purposes.

These efforts are significant to root cause analysis in several respects. At the most proximate level, the FRA reporting requirements ensure that all accidents of consequence are subjected to an analysis of the circumstances of the accident, and that both primary (and if applicable, secondary) causes of the accident be determined and reported to FRA. In some cases, these may require fairly intensive analysis of the accident scene if there is some uncertainty about the cause, and major railroads employ specially trained individuals responsible for this function. Understanding all of these aspects is pertinent to root cause analysis of railroad-accident-caused hazmat releases. In the following section, the nature and character of basic elements are described.

FRA divides accident causes into five broad categories:

- Track, roadbed, and structures,
- Signal and communication,
- Mechanical and electrical failures,
- Train operation—human factors, and
- Miscellaneous causes not otherwise listed.

Within each of these categories, there may be sub-categories and then, at the most detailed level, specific cause codes.

4.6.1 Track, Roadbed, and Structures

The most frequent cause of railroad accidents, and accidents resulting in a hazmat release are failures of the track system, especially rail failure due to various forms of fatigue-induced fracture. Railroads conduct frequent inspection of rails to find and remove defects; however, certain types are difficult or impossible to detect using existing technology. A number of other infrastructure failures also can cause serious accidents. The second most common track-caused hazmat release accident is track geometry, followed by roadbed problems, and switch and frog problems. Railroads use a combination of manual and automated inspection technologies to detect problems before they become critical, but some are not found and derailments can occur as a result. Overall, the FRA has more than 65 different cause codes for railroad-track-caused accidents. This enables a very fine-grained ability to analyze which causes are the most important contributors to hazmat accidents.

For both rolling stock and infrastructure, the American Association of Railroads (AAR), FRA, and Class 1 railroads are conducting or sponsoring research and development of better designs, materials, and operational practices that will be more resistant to failure. In parallel, they are also conducting research and development on an array of technologies intended to improve the inspection capability for a wide range of possible defects.

4.6.2 Signal and Communication

Accidents caused by signal and communication failure rank last among major categories of accidents and as a cause of hazmat releases. Unlike highways, virtually all railroad operations take place in a highly controlled environment. Specific rules and protocols apply to operation on all portions of the railroad. Communications and signals (C&S) are an essential element of these systems whose purpose is to ensure safe and efficient operation of the railroad. If some element of these systems malfunctions, it may result in incorrect or incomplete information being transferred to or from the train, thereby creating the potential for conflicting track occupancy authorities or excessive speed. Under these conditions, the consequences may often be a collision or derailment. Railroad C&S systems are thus designed to be extremely robust and embody extensive fail-safe elements in their design (i.e., if they fail, it results in a "safe" condition, indication, or message). Consequently, railroad accidents attributable to C&S failures are rare. In a recent study, they accounted for only 3/10ths of 1% of all the U.S. railroad mainline accidents. Nevertheless, when such failures do occur, the resultant accidents tend to have high consequences because the outcome will often be a collision or overspeed derailment, thereby resulting in relatively large impact forces. If hazardous materials are involved, there is a reasonably high potential to breach the car transporting them and cause a release.

4.6.3 Mechanical and Electrical Failures

Accidents caused by mechanical and electrical failure are the second most common major categories of accident cause, and third overall in causing hazmat releases, However, when one considers only mainline-accident-caused hazmat releases, they rank second. Railroads operate trains with hazardous materials in the consist, which ranges greatly in number of vehicles and length. The consist is defined as the group of rail vehicles that make up a train. These trains may have less than a dozen cars or more than 150, ranging in length from a few hundred feet to nearly two miles. This has a variety of implications in terms of the occurrence of accidents, and the consequent approaches to root cause analysis.

With approximately 1.5 million railcars and approximately 800 different owners, railcars spend a great deal of time operating on railroads and by companies other than their owners. Frequently, repairs must be conducted on the road by someone other than the railcar's owner. Railcars have not generally been subject to programmed maintenance in North America. Instead, railroads and car owners have operated under a philosophy of run to (near) failure. The objective is to obtain as much life as possible from components without suffering failure. Due to the frequent and redundant inspections railcars receive as they move from terminal to terminal during their journeys, most failing components are indeed found before they cause a problem, but a few are missed and these sometimes result in an accident.

Each of the railcars in the train is subject to dozens of different failure modes with the potential to cause an accident. Although terminal personnel and the train crew are responsible for inspecting key attributes of the cars before departure and at certain intervals during a journey, many problems are difficult to detect or can become critical en route. The nature of trains is that crew members are separated from most vehicles by a considerable distance and will often be unaware of an incipient failure until it has already occurred. Consequently, railcars are necessarily robust and, wherever possible, designed in a fail-safe manner.

As mentioned above, railcars undergo human inspection in terminals and, in addition, railroads rely heavily on a variety of automated technologies to detect certain types of vehicle failure. The industry is aggressively developing new technologies to expand this capability to other components and failure modes.

Overall, FRA has more than 140 different cause codes for railroad-equipment-caused accidents.

4.6.4 Train Operation—Human Factors

Accidents caused by human-factors rank third among the major accident categories in terms of frequency but second in terms of causing hazmat releases. The majority of these are accidents in yards and industry tracks, although mainline collisions are a particularly important cause for the reasons discussed below. Human-factors accidents vary widely in their severity and potential to cause serious hazmat releases. Among the most common human-factors-caused accidents are various errors committed during switching, such as run through switches. These are generally low-speed incidents with little potential to cause sufficient damage to a hazmat car to produce a serious release. On the other hand, accidents caused by failure to obey signals on the mainline or other operating instructions can result in high-speed collisions with substantial potential to breach one or more hazmat cars. Several of the most serious hazmat release accidents in the past few years have been due to such failures. Both FRA and NTSB have placed a high priority on developing technology and implementing requirements for adoption that are intended to prevent certain types of human-factors accidents from happening. Notable among these is the recent rulemaking requiring implementation of Positive Train Control (PTC) on all rail lines handling toxic inhalation hazard (TIH) materials.

4.6.5 Summary of Causes and Impact

The points discussed are pertinent to the root cause analysis objectives of this project because there are a wide variety of possible causes, any one of which occurs relatively infrequently. These incidents are distributed over 150,000 miles of railroad lines and 1.5 million freight cars. Consequently, the rate of failure for any particular component or system is relatively low and dispersed across a large system. In order to understand the principal factors most likely to result in conditions that can lead to a hazmat release, a statistical approach is needed.

In the context of understanding the contribution of the current *FRA Guide for Preparing Accident/Incident Reports* (FRA 2003) to root cause analysis, it is worth reviewing a few of the categories that railroads are required to provide: Item 38—Primary Cause Code, Item 39—Contributing Cause Code, and Item 52—Narrative Description.

> **Item 38—Primary Cause Code** Proper entry of the correct primary cause code is of critical importance, not only for the accident being reported, but also for FRA's analyses conducted for accident prevention

> purposes. Because of the extensive use made of primary cause code entries, careful attention must be given to making correct entry for all accidents. (FRA 2003a, p 11)

As stated by FRA, this code is critically important to "accident prevention analysis," which is implicit in root cause analysis. There are several additional paragraphs providing more detail about the factors railroads should consider when identifying and possibly updating the Primary Cause Code as more information becomes available.

> **Item 39—Contributing Cause Code** If there were one or more contributing causes, enter the code for the foremost contributing cause. Otherwise, enter "N/A." An accident is frequently the culmination of a sequence of related events, and a variety of conditions or circumstances may contribute to its occurrence. A complete record of all of these would be beneficial in accident prevention analysis. However, it is not practical, even if it were possible, to develop forms and codes that would capture every detail that may be associated with the causes and resulting consequences of each accident. Therefore, the most appropriate combination of available codes that best identifies the likely primary and any contributing cause, and other factors, is to be used. Railroads are encouraged to use the Contributing Cause Code. When the events cannot be adequately describe[d] using the Primary and Contributing Cause, the railroad must use the Narrative Block to complete the causes of the accident. (FRA 2003a, p 13)

As discussed in this report, accidents are often the result of more than one factor. FRA explicitly recognizes this elsewhere in their discussion of the accident reporting and analysis process and, by providing Item 39—Contributing Cause Factor, they allow for one contributing cause to be identified. In the context of root cause analysis, this may be one area for improvement. FRA states that more than one contributing cause may be a factor and asks the railroad to enter the "foremost contributing cause," implying that only one be identified. It seems feasible that the process could be modified to allow for multiple (perhaps up to three) contributing causes to be identified, with a requirement that they be rank-ordered in importance (i.e., Contributing Cause 1, 2, and 3). However, this would often require a certain amount of subjectivity in ranking the causes and different individuals or railroads might use different criteria, thereby introducing additional variability. Furthermore, the notion of primary and contributing causes may have a tendency to over-simplify the process. It may be worthwhile to consider proximate and ultimate causes with more sophistication. Another potential measure could be to evaluate whether this aspect of the reporting system can be modified to enhance the value of the information.

In recognition of the potential complexity of factors contributing to an accident, FRA requires that railroads provide additional details in Item 52—Narrative Description. FRA's general instructions are as follows (FRA 2003a, p 15):

> A detailed narrative is basic to FRA's understanding of the factors leading to, and the consequences arising from, an accident. While many minor accidents can be described in a few brief comments, others are more complicated and require further clarification.

In addition, FRA specifically requests that information be provided on drug/alcohol involvement, cause, diesel fuel tank, hazardous materials, and other railroads. Of these, the following are of particular relevance to hazmat root cause analysis and the text for each is as follows (FRA 2003a, pp 15–16):

> Drug/alcohol involvement—Include a discussion of any drug/alcohol use connected with this accident. If positive tests were made, but usage/impairment was not determined to be a causal factor, explain the basis of this determination.
>
> Cause—Discuss any event(s) or circumstance(s) occurring prior to the accident that has relevance to the accident. Provide additional information concerning the reasons(s) for the accident when the causes found in Appendix C do not sufficiently explain why the accident occurred.
>
> Hazardous Materials—Identify the initial and number of any car releasing hazardous material. List the name and indicate the quantity of hazardous material released. Report the number of fatalities and injuries resulting from a direct exposure to the released substance. If there was an evacuation, estimate the size of the affected area and the length of the evacuation.

4.7 Marine Information for Safety and Law Enforcement (MISLE)

The MISLE database supports the Marine Safety and Operations Programs. MISLE contains vast amounts of data, including detailed vessel characteristics, cargo carriage authorities, involved party identities, bridges, facilities and waterways, and records of related Coast Guard activities. MISLE activities include law enforcement boardings and sightings, marine inspections and investigations, pollution and response incidents, and search and rescue operations. In addition, MISLE manages the information flow involving the administration of all of these activities, from the initial triggering event, to incident management and response, and the resulting follow-on actions. Its development was initiated in 1992 and it became fully operational in January 2002 when the Coast Guard transitioned from the Marine Safety Information Reporting System.

4.7.1 Database Description

The database is logically broken into a relational table structure that contains an activity table that includes all of the incidents reported to MISLE. As the example in Figure 4-4 shows, there are tables presenting an inventory of facilities and vessels that can be tied to the Activity and Events Tables. The activities are joined to the Facility Events and Vessel Events Tables, which provide additional information on the activity reported to MISLE. These, in turn, are joined to Facility and Vessel Pollution Tables that are also joined to an Injury Table that lists all of the reported injuries and fatalities associated with the activities.

The pollution activities are a very small portion of the activities reported to MISLE. Commercial, as well as pleasure boat, collisions and groundings are reported. If a reportable quantity of a hazardous substance is released (40 CFR Part 302), the National Response Center (NRC), which is administered by the Coast Guard, must be notified promptly, and the vessel operator must fill out Form CG-2692 and submit it to the Coast Guard to document the event. Note that the reportable quantity is determined using the EPA list of hazardous substances, which also includes marine pollutants.

4.7.2 Purpose and Function

The purpose of the MISLE database is to maintain a comprehensive record of vessel, facility, and Coast Guard activities related to commercial shipping. Incidents resulting in the loss of life to the public from private boating activities are also included in the database. The information system contains links to other resources so that Coast Guard personnel can respond quickly to any major incident. The database part, which is the focus of this discussion, reports all vessel or facility incidents related to commercial shipping. The documentation of pollution events, while significant, represents only a small fraction of the documented incidents. The MISLE system maintains a log documenting the status of all judicial activities associated with the documented incidents. The record of any incident with an ongoing judicial action is not available publicly until the case is closed. Since cases are commonly kept open for several years, a comprehensive picture of the number of pollution events occurring in a given year is difficult to identify from the publicly available file. The focus of much of the monitoring activities relates to efforts to speed up judicial actions so that cases can be closed more rapidly.

4.7.3 Data Collection

If a reportable event occurs, the vessel operator must fill out Form CG-2692 and submit it to the Coast Guard to document the event. There also are cases where Coast Guard personnel file an event report using CG-2692. Once filed, the Coast Guard accident investigators update the file as the investigation proceeds.

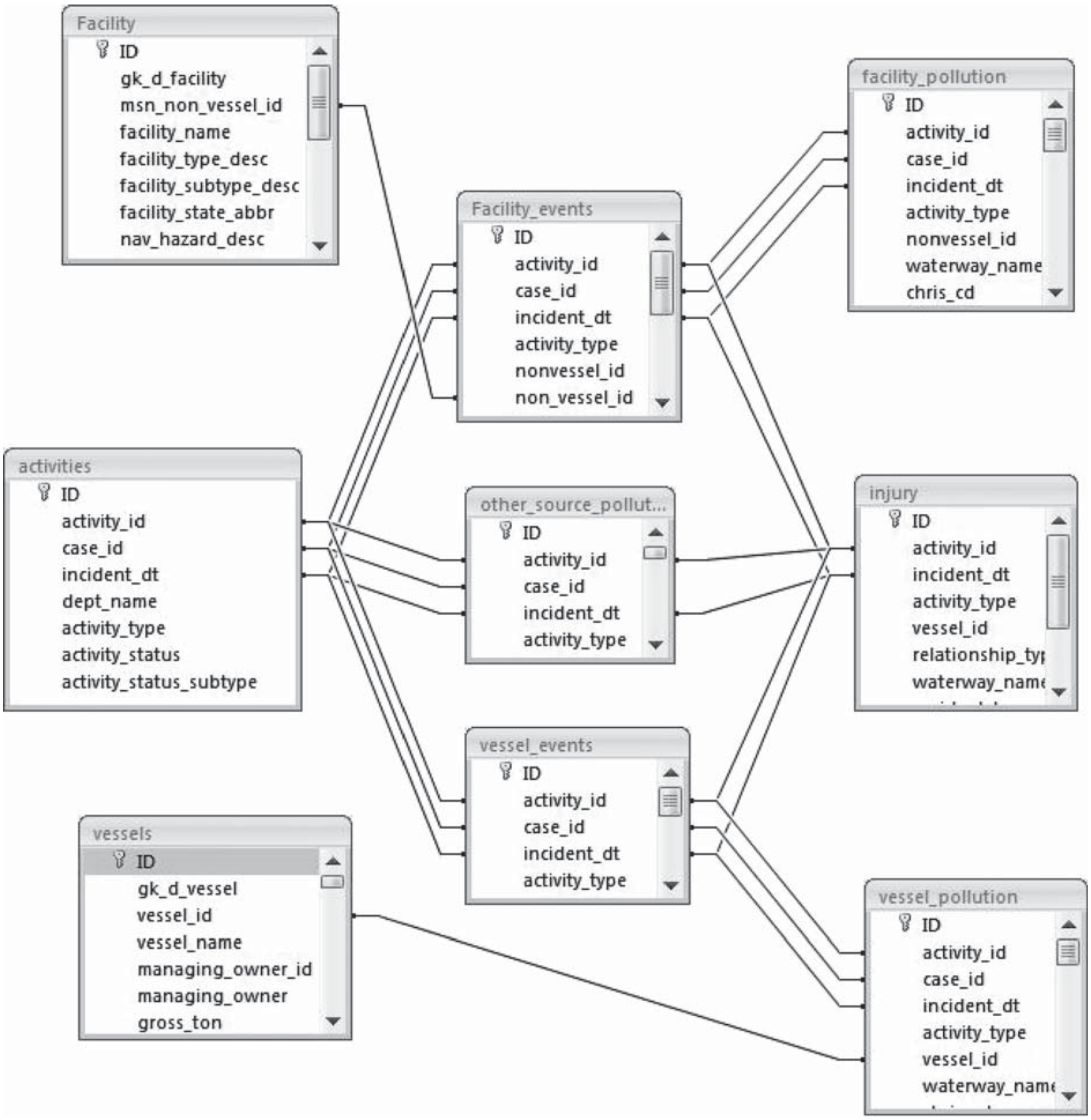

Figure 4-4. Example of tabular relationships in MISLE.

4.7.4 Data Compilation

Data from the accident form are transferred to the appropriate fields in the MISLE database. As the status of the investigation proceeds, the data are updated.

4.7.5 Accuracy and Completeness

Assessing the accuracy of the entries in the MISLE database is difficult for outside investigators. Since the majority of the events entered into MISLE involve legal action taken on the part of the Coast Guard, the accuracy of the entered data is likely to be very high. The completeness of the data entry is difficult to judge as well. Although all of the fields have entries, in many cases they are filled out with a standard term, such as "not noted" and it is not known if an entry should have been made in the field.

4.7.6 Quality Control

No assessment was made as to the use and effectiveness of MISLE quality control procedures.

4.7.7 Interconnectivity with Other Databases

There are no common fields that would enable this database to be connected easily to other databases. The MISLE record does contain the date, time, and location of the accident, expressed as the latitude and longitude, so if another database like HMIRS reported the same information, linkages could be made. Since the carrier name is a difficult field to join on (because the name must be identical right down to the spelling of the name, abbreviations, spaces, and periods) automatically connecting database fields would be very difficult. A brief attempt to connect the HMIRS reports listing mode as *water,* identified no events reported to both databases during a one-year period.

4.7.8 Analyses Using Database

Since the MISLE events could not be compared with the hazmat incidents reported in HMIRS with *water* as the mode, no analyses were performed.

4.7.9 Summary and Potential Measures for Improving Root Cause Analysis

Much of the MISLE database is accessible only to Coast Guard staff. Furthermore, the MISLE data become available to the general public only for closed cases and it can take several years to close many of the MISLE-reported incidents. This might be one of the reasons why it was not possible to find common events reported to both HMIRS and MISLE. Lack of timeliness, access, and interconnectivity are considered insurmountable barriers for MISLE use. Any database used for root cause analysis should provide timely accident reports, have unrestricted access, and be able to easily connect reports made to other databases.

4.8 NTSB Accident Investigations and Reports

4.8.1 Scope of Investigations

There are numerous NTSB investigations of individual accidents. While all commercial aircraft crashes are included, there are certain rail and truck accidents that are also selected for investigation by NTSB. This section will focus on one particular type of NTSB investigation—passive private grade-crossing accidents. A study (NTSB 1998) summarized the investigation of 60 passive grade-crossing accidents that occurred between December 1995 and August 1996. The accidents investigated were not selected on the basis of establishing statistical confidence. The criteria were that damage to the motor vehicle had to be serious enough for the vehicle to require towing from the scene and that it had to occur close enough to an NTSB regional office for an investigator to travel to the site before the vehicle was towed from the scene. The NTSB investigator recorded the types of signage present, as well as the characteristics of the grade crossing, and obtained witness statements.

4.8.2 Approach to Identifying Root Causes

The approach taken in all NTSB investigations is similar, although the scope of the individual pieces may vary. In all cases, a team of NTSB accident investigators is dispatched rapidly to the scene of the accident to collect evidence that might be usable in determining contributing causes. The evidence collected includes data describing the accident scene, the amount of damage to equipment, and the extent of injuries to individuals involved in the accident. Witness statements are always collected and it is pointed out in several investigations that it is important to get those statements quickly because witness memories fade. The NTSB then goes through an extensive analysis of the collected data and will frequently follow up with requests for additional information as the analysis proceeds. By going to the scene, the NTSB has all the contact information needed for follow-up purposes.

The NTSB investigator collects the following information:

- Location,
- Date and time,
- Lighting conditions,
- Type of motor vehicle (year and type),
- Train action reported (horn sounded/auxiliary lights on),
- Signs present (crossbuck, advance warning, and/or stop sign),
- Physical characteristics (limited sight distance, angle of intersection, road or track curve, and presence of a nearby road intersection), and
- Number of injuries and fatalities.

All of these items, except some of the physical characteristics, are included in the grade-crossing incident report submitted to FRA. The FRA accident database does not document the proximity of the grade crossing to other roads nor does it document whether the road or track is curved. The curvature of the road or rail track was not listed as a probable or contributing cause in any of the 60 cases (NTSB 1998), and there was only one case where the presence of traffic plus a nearby intersection was listed as a contributing cause. Thus, the absence of this information being captured in the FRA database is not considered to be a significant weakness. Whereas the NTSB findings are based on a site visit and witness statement, FRA does not require either of these. While there are narrative fields provided in the database, during the time period of the NTSB study, the narrative fields were left blank for all grade-crossing accidents for both active and passive grade crossings—more than 3,000 reports.

The analysis performed here included an additional task that was not performed by NTSB—a comparison of the data reported to FRA and the data reported by NTSB for the same accident. Although reports for all 60 of the accidents were found, the matching task, initially thought to be easy, turned out to be a challenge for the following reasons:

- NTSB did not include the FRA incident number, so it had to be discovered.
- NTSB listed the closest town to the grade crossing whereas FRA reports the nearest timetable station, county, and city (if the accident occurred within city boundaries).

For almost one-third of the cases for which matches could be found, it was necessary to refer to a map to find the location of the grade crossing, a time-consuming process. If both databases had provided the GPS coordinates (an option in the rail equipment accident database), then the month, day, and GPS coordinates would have made it much easier to match the accidents in the two databases.

4.8.3 Insights for Analyzing Root Cause

The insights are divided into three categories: (1) data quality, (2) probable cause findings, and (3) summary. These categories are described in the following subsections.

4.8.4 Data Quality

A comparison of the NTSB (1998) and FRA (2003) data revealed many significant differences. In five cases, there were differences in the number of fatal and non-fatal injuries. FRA establishes explicit reporting requirements and it is not known if NTSB followed the same requirements. For example, if a medical examiner determined that the driver committed suicide, the railroad would not have been required to report the fatality to FRA. Although FRA requires that the railroad file an amended report if the person dies within 180 days of the crash, if the railroad does not follow the progress of the injured person, then they would not know if the person died within that time period and this would not be reported to FRA. Whatever the reasons, the differences were not small. For Case 3, the FRA database stated no fatalities and NTSB reported 3 fatalities. For Case 7, the FRA database stated 1 fatality and 3 injuries whereas NTSB reported 3 fatalities. In Case 26, the FRA reported 2 injuries and NTSB reported 12 injuries. In Case 55, the FRA database stated no injuries and NTSB reported 6 injuries. In Case 60, the FRA database reported 1 injury and 1 fatality and NTSB reported 2 fatalities. Since the NTSB finding is always more severe, one can assume that NTSB did more follow-up in the period after the accident to more accurately reflect the number of fatal and non-fatal injuries.

There were 4 of the 60 cases where no corresponding FRA report could be found at the date and location listed in the NTSB report. In one case, the accident was found at the same time and location, three days later. For the other three, no report could be found. In one case, the vehicle attempted to cross at an abandoned grade crossing and that accident, while classified as a grade-crossing accident by the NTSB is classified as an obstruction accident by FRA for which a grade-crossing accident report need not be filed. For the remaining two, no FRA report corresponding to the NTSB report could be found.

There were large differences reported for other parameters as well. The view of oncoming trains in the vicinity of the grade crossing showed the largest discrepancy (see Table 4-27).

As shown in Table 4-27, NTSB separated its assessment of obstructed view into two parameters, the view as the motor vehicle approached the crossing and the view of the oncoming train

Table 4-27. Query on obstructed vision.

NTSB Findings		FRA Database View Parameter	Number of Cases
Sight Distance Limited on Approach	Sight Distance Limited at Stop Line		
Yes	Yes	Topography	1
Yes	Yes	Vegetation	1
Yes	Yes	No Obstruction	5
Yes	No	Permanent Structure	1
Yes	No	Vegetation	1
Yes	No	No Obstruction	24
No	Yes	No Obstruction	1
No	No	Passing Train	1
No	No	No Obstruction	22

at the stop line. The FRA report has only one parameter, *limited sight distance.* Based on the definition in the FRA reporting guide (FRA 2003), the view as the driver approaches the crossing, the first NTSB parameter, should be reported in the FRA database. The two assessments agreed in 27 cases, less than half the 57 for which a comparison was possible. In 22 of these cases, both assessments agreed that there were no obstructions. In one of five remaining cases, the obstruction was a passing train, a factor not considered to be an obstruction by NTSB. In the four remaining cases, both recorded that an obstruction to the driver's vision was present on approach. That leaves 30 cases where the two assessments disagree. In five of the cases, the NTSB investigators concluded that the driver's view was obstructed both on approach and at the stop line—at the stop line for one case, and for the remaining four on approach. It seems clear that the NTSB investigators and the employees filling out the FRA grade-crossing report differ on the definition of what constitutes an obstruction as a driver approaches the grade crossing. Given that the NTSB investigator visited the scene, the NTSB assessment is thought to be a more accurate assessment.

Differences also were observed in the types of signage at the grade crossings. In 10 of the 57 cases, the FRA database stated that no stop sign was present whereas the NTSB report stated one was present. In 4 of the 57 cases, the FRA database shows there was a stop sign when the NTSB investigator reported none was present. Being inaccurate in 25% of the cases makes it difficult to draw conclusions regarding the effectiveness of stop signs at passive grade crossings.

Because the NTSB investigators take statements from witnesses and observe the conditions at the accident scene at the time of the accident, they clearly have an advantage regarding reporting accuracy. There were cases where the NTSB reported a building, a large pile of rocks, overgrown vegetation, and cars on the tracks as blocking the vision of vehicles approaching the crossing. Clearly, conditions change at grade crossings and if there is no way to capture those changes in the data being used by the railroad employee filling out the FRA grade-crossing report, the FRA report will always be less accurate.

4.8.5 Probable Cause Findings

The NTSB (1998) report recognizes that determining the probable and contributing causes of passive grade-crossing accidents is a challenging task. It is even more difficult to summarize the findings based on only 60 cases. The following discussion summarizes the findings, recognizing that no finding is likely to be statistically significant.

There were 14 incidents at private grade crossings and 46 at public crossings. In 33 of the 60 cases, more than one-half, limited sight distance was listed as the primary cause. In 28% of the private crossing incidents, there was limited sight distance, while 63% of the public crossing incidents cited this as a factor.

Stop signs were present at 3 of the 4 private grade-crossing incidents with limited site distance and at 15 of the 29 public grade-crossing incidents with limited sight distance. Since one of the recommendations of an earlier NTSB report was that stop signs be installed at all passive grade crossings, it is interesting to note that stop signs were present at one-half of the private crossings and about one-third of the public crossings. Regarding the primary cause of these accidents at crossings with stop signs, the driver ran the stop sign in 3 of the 7 incidents at passive grade crossings (about half) and 13 of the 15 public grade crossings with stop signs (almost 90%). Clearly the recommendation to place stop signs at all grade crossings with limited sight distances will have limited effect until the compliance rate with the stop signs is improved.

Regarding injuries and fatalities, there were two cases where a person on the train was injured. Two crew members and 12 passengers were injured in these accidents, respectively. Fatalities or injuries occurred in 47 of the 60 incidents. Because the 60 cases were not chosen at random, no significance can be made regarding this finding. For the 14 private grade-crossing incidents in the database, there were 7 incidents in which there was an injury or fatality and in only 1 of those incidents was there limited visibility at the crossing. At the 46 public grade crossings in the database, there were 30 incidents where limited visibility was present and of those, 28 incidents resulted in one or more injuries or fatalities. The situation is similar with the 16 cases in which limited visibility was not an issue. In 12 of the 16, there was one or more injury or fatality. One can conclude from these data that while more than one-half of the passive grade-crossing accidents occurred at crossings with limited visibility, whether or not there is an injury or fatality is not a function of limited visibility. Rather, a large fraction of grade-crossing accidents result in one or more injury or fatality to the occupants of motor vehicles; this is just associated with the seriousness of the incident when it does occur.

The NTSB study clearly demonstrates the importance of site visits and driver and witness interviews if probable and contributing causes are to be identified. Witness statements appear to have the greatest value. Otherwise, it would be difficult to obtain data, such as whether the driver was talking to an occupant and never looked, or the driver was talking on a cell phone. It is clear that the NTSB conclusions were based on these interviews. Even in the seven cases where it could not be determined if the findings were based on witness statements, it is likely that the witness statements either influenced the determination of probable cause or validated the conclusion by providing supporting information. Two cases reported that an event data recorder was read to verify that the train engineer sounded the horn on approach of the crossing. Such collaborating evidence is also useful when attempting to identify probable and contributing causes.

In the FRA train accident database, there are no fields specifically designated for probable and contributing causes. There are narrative fields that could be used for statements provided by the railroad crew, motor vehicle occupants, and/or witnesses. The railroad train crew could be a valuable source of information. They could document contact information for witnesses and fill out a form describing the conditions at the crossing at the time of the accident. For example, they could note if the view of the train was blocked by overgrown vegetation, a recently constructed building, or something temporary (e.g., a pile of rocks or train equipment). This would enable the railroad officials to more accurately complete the FRA form. The railroad official could also obtain statements from the train crew and witnesses, placing their statements in the narrative fields of the current reporting form. The person filling out the FRA accident report would also have access to the readings from the train's event recorder at the time of the accident. A query of the FRA railroad crossing accident database for the period beginning in November 1995 and ending in September 1996 revealed that the narrative fields were left blank in all cases—nearly 4,000 records. The other alternative is to add fields where the person filling out the grade-crossing accident report can specifically list causes. This means that if the same person fills out both forms, he or she is already trained in determining accident causation.

4.8.6 Summary

The following potential measures apply to capturing root causes routinely in grade-crossing accidents:

1. Require the information that was captured by the NTSB investigator to be incorporated in the FRA grade-crossing accident form. While the railroad can be left to determine how this recommendation is to be implemented, it can probably be accomplished by either a site visit or by the train crew documenting the conditions at the accident scene and obtaining contact information for witnesses, including law enforcement personnel. It would be possible for the

railroad official filling out the grade-crossing incident form to follow up with these individuals as the form is being completed.

2. Require the narrative sections of the FRA grade-crossing accident form to be filled out with witness statements. Police officials, if present at the scene, are trained to provide accident details and their findings should be included in the narrative fields.
3. Readings from the train's event recorder at the time of the accident could be obtained and stored to verify some of the statements made by the train engineer regarding speed and whether or not the horn was sounded at the proper time.
4. The rail equipment incident database has fields entitled *primary cause* and *contributing cause.* Such fields could also be included in the grade-crossing accident reporting form.
5. Evaluate the need to provide additional guidance on the definition of restricted view of the railroad tracks. While both NTSB and FRA emphasize the importance of seeing the train as the motor vehicle approaches the crossing, it does not appear that the railroads were using this definition when filling out the FRA grade-crossing accident report. Railroads also could place maintenance cars, supplies, and equipment in a location where the view of oncoming trains is not obstructed at passive grade crossings. Although this occurred in only 1 of the 60 cases, it is a condition that can be addressed easily.
6. If the GPS location of all accidents was recorded in the database, by using these data along with month, day, and year, it could be possible to display the frequency of grade-crossing accidents in a region and also couple records in multiple databases, thereby expanding the amount of information available regarding an accident with no increase in the number of forms that have to be filed.

4.9 The Hazmat Serious Truck Crash Project Database

4.9.1 Introduction

The *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005), a project conducted from 2002 to 2005 for FMCSA, demonstrated two methods for improving the usefulness of a database for identifying the root causes of hazmat crashes. The project achieved this result by improving data quality (including comprehensiveness) and augmenting the database with additional fields.

The project demonstrates how, by adding specific data fields, checking data from the original source, and supplementing data with telephone calls, the user is able to develop insights into root cause analysis that could not be obtained without the application of these techniques.

This project had the following three basic objectives:

- Enhance the current methodology for identifying and characterizing serious hazmat truck crashes in the United States.
- Improve the capability to analyze causes and effects of selected serious hazmat crashes.
- Support the implementation of hazmat truck transportation risk reduction strategies for packagings, vehicles, and drivers.

The Hazmat Serious Truck Crash Project used the MCMIS Crash file for serious crashes occurring in 2002, extracted the crashes that involved hazardous materials and, for a sample of 1,000 hazmat crashes, supplemented the data in MCMIS with information from other sources. These sources included HMIRS, PARs filed by police from individual states, and direct correspondence with the involved carriers.

92 Hazardous Materials Transportation Incident Data for Root Cause Analysis

Sample crash information was input and stored in the Hazmat Accident Database, a specially designed database that enabled the aforementioned information sources concerning a particular crash to be assembled into as complete a record as possible. This included characteristics describing the crash event, as well as the accuracy of the information itself. Extensive database protocols and quality control checks were employed to accomplish this objective. Once database development was complete, analyses were performed on the database for the purpose of providing information that might support the development of more rigorous hazmat truck safety policy.

4.9.2 Adding Explanatory Variables to the Hazmat Accident Database

For the Hazmat Accident Database, the Hazmat Serious Truck Crash Project added a number of fields designed to capture the actions of the vehicle(s) before the crash. In addition, fields were added to provide more detail on the type and quantity of the hazardous materials, hazmat packaging, infrastructure, and such driver characteristics as age and experience.

Explanatory variables are crash characteristics that help explain cause and effect. Table 4-28 shows the five types of explanatory variables. The crash analysis process involved associating explanatory variables with impacts to determine how vehicle, driver, packaging, infrastructure, and situational characteristics influence crash occurrences in general, as well as those that result in spills. Appendix E (available on the TRB website at www.TRB.org by searching for HMCRP Report 1) describes selected analytical results from the project.

4.9.3 Crash Records Selection

Of the approximately 2,000 hazmat crashes in the MCMIS database in 2002 (out of a total 105,000 records), the Hazmat Serious Truck Crash Project selected about one-half of the hazmat crashes for a more in-depth analysis. These records were primarily selected on a random basis, with the exception that less common accidents involving hazmat Classes 1, 4, 5, 6, and 7 were all selected to obtain a large enough population. Before performing data analysis, weighting factors were applied to compensate for the non-random aspects of the selection.

Table 4-29 shows the hazmat crash classes selected for more detailed analysis.

4.9.4 Populating Records and Improving Data Quality

After the records were imported into the Hazmat Accident Database, HMIRS data were used to both fill in data for fields not included in MCMIS and for quality checking the existing data. Because the HMIRS fields are more fully populated, any fields in the database that were common to HMIRS and MCMIS were overwritten by the HMIRS information. The remaining HMIRS information was also incorporated into the database.

Table 4-28. Explanatory variables used in the Hazmat Accident Database.

Vehicle	Driver	Packaging	Infrastructure	Situational
Configuration Cargo Body GVW	Age Experience Condition	Package Type Quantity Shipped Quantity Lost Age (Cargo Tank) Rollover Protection Inspection History Design Specification	Road Surface Road Condition Road Type Trafficway Access Control Speed Limit No. of Lanes	Pre-Crash Condition Dangerous Event Vehicle Speed Impact Location Primary Reason Accident Type Weather Condition

Table 4-29. Sampled crashes by hazmat group.

Hazmat Group	Description	Analyzed Crashes		Estimated 2002 Totals	
		Crashes	Spills	Crashes	Spills
1.1 – 1.6	Explosives	19	2	21	2
2.1	Flammable gases	148	14	256	21
2.2	Non-flammable gases	60	8	102	12
2.3	Gaseous poisons	11	1	18	2
3.0	Flammable liquids	544	125	914	182
4.1 – 4.3	Flammable and reactive solids	7	2	8	2
5.1 – 5.2	Oxidizing materials	31	9	36	10
6.1 – 6.2	Poisonous and infectious substances	14	2	16	2
7.0	Radioactive materials	4	2	4	2
8.0	Corrosive liquids	75	16	139	23
9.0	Miscellaneous hazardous materials	57	23	86	27
Unknown	Hazmat group could not be determined	17	5	28	9

A similar process was used to input PAR data. As the information was being filled in from the PAR, the data entry form showed the default values for any parameters that were previously entered based on information supplied by MCMIS and HMIRS. Any inconsistencies were changed to reflect the information contained in the PAR. Frequently, the changes were not inconsistencies, but expansions of the data. For example, many PARs list the actual GVW of the vehicle and, in those cases, that number was input in place of a broad weight category.

The final step in populating the Hazmat Accident Database involved entering information that was obtained through direct telephone conversations with the involved carrier. These calls verified the accuracy of the entered information and provided specific information only the carrier could supply. This included such information as the amount of material being shipped; whether there was a spill and, if so, how much; the manufacturer and specification number of the packaging; and the year the packaging was fabricated. Valuable information on packaging characteristics was obtained from carriers who provided the DOT specification number for the tank, the year it was manufactured, the manufacturer, type of rollover protection on the cargo tank, and the inspection history. Many carriers could estimate the amount of material being shipped and if any was spilled. The type of damage to the cargo tank could sometimes be recalled, usually only if there was a spill. Most carriers also were willing to provide information on the driver's experience.

4.9.5 Quality Control Checks

Several quality control checks were built into the data collection process. Accuracy checks were performed at three critical junctures: (1) after the data from the PAR were entered for the crash, (2) after the carrier calls were completed, and (3) whenever a reviewer changed a pre-existing database entry. Special efforts also were made to identify and reconcile blank fields. In addition, error-trapping queries were run to identify reporting inconsistencies (e.g., Interstate highways that were not flagged as limited/controlled access). Finally, summary reports were generated of each recorded crash to use in reviewing the entered information or to use as a reference during carrier correspondence.

4.9.6 Database Enhancements and Limitations

The fields in the Hazmat Accident Database reflect a list of parameters that are considered pertinent for safety analysis. While every effort was made to obtain relevant information, it was not expected that it would be possible to populate all of the fields. Nevertheless, significant improvements were made in the breadth and accuracy of hazmat crash information from which safety assessments and root cause analyses can be performed.

These improvements are evident by comparing initial MCMIS tables with the completed Hazmat Accident Database. In addition to broadening the selection of eligible entries to many of the descriptive tables, new tables also were created that are not present in either MCMIS or HMIRS, such as Pre-Crash Events, Primary Reasons, and Impact Location. Moreover, data collected from PARs and from carrier correspondence for nearly 1,000 MCMIS crash records enabled many MCMIS data fields that were initially blank to be populated.

Despite these improvements, some fields were largely blank. For example, no PAR captured information on evacuations. Only one state, Kentucky, captured information on road closures. The vehicle speed was captured in roughly 50% of the PARs, and the trailer dimensions, length, and width could be obtained in only one-quarter of the cases. The other fields were filled out for more than 80% of the selected crashes and in some states that figure was 100%. Some states, such as California, have extensive PARs that provide information on all of the key parameters as well as other parameters that might be of future interest. Roughly 60% of the states use a commercial vehicle supplement, designed to capture data required for the MCMIS Crash file. These supplements tend to have a uniform hazmat section that provides all of the information needed to fill out the five hazmat entries in MCMIS. Unfortunately, about 25% of the states that are known to have commercial vehicle supplements did not provide the supplemental form. Appendix E (available on the TRB website at www.TRB.org by searching for HMCRP Report 1) presents representative analyses that were conducted using the Hazmat Accident Database.

4.9.7 Summary

The *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005) convincingly demonstrated that by adding explanatory fields to MCMIS, selecting a sample of crashes for more detailed investigation, matching the same crash in HMIRS with the one in MCMIS, using PARs to check data quality and complete added data fields, and telephoning carriers to collect data on such elements as hazmat type and quantity, root cause analyses could be made more accurately and thoroughly. Consequently, the project team suggests that the following lessons learned from the Hazmat Serious Truck Crash Project be applied to the data collection process:

- Add selected explanatory fields to increase the type of data available for analysis.
- Select an annual population of hazmat crashes for more detailed investigation. This population could be selected based on a number of criteria such as type of crash or type of hazardous material. For example, for a particular year (or years), all rollover crashes resulting in a spill could be selected or all crashes involving a spill of Class 3 hazardous materials could be selected. The number of years selected would relate to the number of accidents available for analysis. For rarer events and hazmat classes, a larger number of years would be chosen.
- Match all applicable crashes to HMIRS. Use HMIRS data to supplement MCMIS data wherever possible.
- Use the PARs to supplement and check crash data. Collect PARs from the states and use them to ensure data quality and to complete data in the added fields.
- Telephone key contacts such as carriers to collect unique data. Carriers should be called to collect data on such characteristics as driver age, packaging type, and type and quantity of hazardous material shipped and/or spilled.

CHAPTER 5

Potential Measures for Improving the Identification of Root Causes for Hazardous Materials Crashes

5.1 Introduction

This chapter discusses potential measures to improve the capability of officials and researchers to identify the root causes of hazmat transportation accidents. As discussed in Chapter 1 of this report, the following definition of root cause was used to develop these measures:

> One or more contributing factors that lead to the occurrence of a transportation accident and/or affect the severity of its consequences.

As previously noted, root cause identification may depend on detailed and accurate information available for five major parameters of vehicle, driver, packaging, infrastructure, and situational. Inadequate information in any one of these parameters may result in the inability to accurately identify the root cause of the hazmat crash.

If a contributing factor can be mitigated, the likelihood of occurrence and corresponding impacts of an entire class of accidents could be significantly reduced. When focusing on one class of accidents, such as single-vehicle cargo tank rollovers, much can be learned when the data show that a contributing factor is present in a large fraction of the accidents. Consequently, policies can be developed and actions initiated to improve safety.

5.2 Information System Development

A key finding emanating from this study is the need to establish a root and contributing cause information system. The system would have the following two major components:

- Linking crash databases together so information in different databases can be easily retrieved for the same crash. This incorporates some of the same elements proposed by PHMSA for increasing the effectiveness of hazmat databases and techniques developed for FMCSA's Serious Hazmat Crash Project.
- Selecting a group of hazmat crashes annually for collecting additional information that will enable officials and researchers to identify the root and contributing causes of that class of hazmat transport accidents. This follows lessons learned from FRA's detailed examination of selected crashes, NTSB's focus on investigating a certain class of crashes, techniques used for the Serious Hazmat Crash Project, and the TIFA database for adding to information in the FARS database.

In order to move toward the identification of root and contributing causes, officials and researchers need to utilize all available data related to either a single hazmat crash or an entire population. Where crash information is collected in more than one database and by different parties, the data could be combined to provide a thorough accident portrait. For example, the same hazmat truck crash may appear in FMCSA's MCMIS database and in PHMSA's HMIRS database. The information in these databases could be linked to combine information from different sources on the same accident. Similar reasoning could apply for a hazmat rail crash found in FRA's RAIRS database and HMIRS, or a fatal hazmat truck crash in TIFA and MCMIS.

Beyond the inherent advantage in linking hazmat crashes in different databases, additional information is needed to more effectively identify root and contributing causes. Current accident data need to be supplemented by information about the circumstances and conditions that existed before the accident occurred, factors not presently captured. Since these additional pieces of information could come from a variety of sources, the term "information system" and not "database" is used to describe the components and structure of such a system.

An example of the value of this approach is an analysis that NTSB performed on grade-crossing accidents. Approximately 60 unprotected private grade-crossing accidents were selected for study, beginning with the information commonly recorded in RAIRS. NTSB gathered additional information by visiting each site, and collecting information that both supplemented RAIRS data and validated on-site conditions listed in RAIRS. NTSB also obtained witness statements from accident sites as soon as possible. None of the information in the witness statements is captured in RAIRS and it is those data that showed driver distractions to be a frequent contributing cause of accidents. The data collected by the NTSB showed that driver grade-crossing visibility was often more limited than documented in RAIRS tables. Had the analysis relied solely on RAIRS data, poor visibility would not have been cited as a major contributing cause for these passive grade-crossing accidents. Although field data collection may be too labor intensive and costly on a recurring basis, this example illustrates the advantages of supplementing data in the databases.

Another approach with considerable potential, which is currently being implemented by FRA, is more detailed investigation of a representative sample of accidents. FRA conducts approximately 100 detailed investigations of rail crashes annually. These investigations obtain additional data that are not captured in RAIRS.

Using the terminology in this report, approaches taken by NTSB and FRA comprise an information system for selected accidents. NTSB and FRA investigations illustrate the feasibility of supplementing information contained in current databases to address a specific class of accidents, improving the ability to identify contributing and root causes for these classes of accidents. Although site visits and witness statements might be difficult to obtain on a routine basis, clearly, if the FRA chose to target its 100 detailed investigations on a particular class of accidents, additional data could be obtained for those targeted accidents. In successive years, the focus of the detailed investigations could be switched to a different class of accidents. For example, if FRA targeted private grade-crossing accidents for 60 of the detailed investigations, it could have produced a report similar to the one produced by NTSB, concluding that the contributing cause, perhaps even the root cause, of many of the accidents was crossing visibility.

5.2.1 Develop Framework for Identifying Contributing Causes and Root Causes of Hazardous Material Accidents

This section focuses on developing an information system capable of capturing the data for thousands of hazmat accidents that occur each year. This information system would not reside in a single database. Rather, the system would use a number of relatable databases, analysis tools, and reports that can, in their totality, contain the information in sufficient detail and quality to identify root and contributing causes of accidents. This would include those databases that are currently used to collect information on hazmat crashes.

To identify the root and contributing causes of various classes of accidents, an analyst must be able to relate inventory information to the accident tables. Inventory information characterizes the hazmat information system, including driver characteristics (e.g., age and experience); hazmat package characteristics (e.g., tank type and age); and vehicle characteristics, carrier characteristics, and mileage traveled. With this information, an analyst can mine a dataset and search for the common contributing causes of various classes of hazmat accidents. For example, without information on the number of hazmat truck drivers in various age and experience categories, it is impossible to determine those age and experience categories that are over-represented in hazmat accidents.

Note that this kind of analysis is appropriate to identify coarse-grained factors that increase risk. For example, the result could be used to determine that older or younger drivers are at higher risk. This should be used with stable characteristics, such as age, vehicle type, and road type, but cannot be used to identify specific errors or transitory conditions that generated a crash, such as a distracted driver or tire failure.

Currently, PHMSA is completing Phase I of the Multi Modal Hazmat Intelligence Portal (HIP) system. The system is being designed to acquire hazmat information at a single location. Under HIP, data from FAA, FMCSA, NTSB, the Coast Guard, and PHMSA will be available by carrier, shipper, manufacturer, and packaging company. Only parts of the system, which is being designed for the enforcement staff, will be available to the public. Although FMCSA is supplying cargo tank and hazmat compliance reviews as well as inspection results, sharing hazmat crashes in the MCMIS database is not yet part of the system and PHMSA has no immediate plan to incorporate this information into HIP. The incorporation of crash files in the future could enhance system capability.

5.2.2 Availability of Carrier Characteristics Inventory Information for Analysis with Accident Data

FMCSA maintains both the MCMIS Census and Crash files. The MCMIS Census file contains inventory information that, if routinely updated and validated, could be useful for identifying which motor carrier characteristics are over-represented in hazmat crashes. A study that made such comparisons for a targeted group of 100 accidents over a one-year period might identify changes to the type of inventory information that should be collected. As a result, a program could be initiated requiring that the new information be obtained when existing carriers re-register and new carriers register for the first time.

PHMSA has an annual hazmat carrier registration requirement. These data also could be used to determine which carrier characteristics are being over-represented in hazmat crashes. It is likely that additional information would be required on the vehicle configuration, packaging, and driver if one wished to determine whether these characteristics are being over-represented in hazmat accidents. Clearly, changes to the data elements in the current HMIRS and Hazmat Registration file would be required to capture more than carrier inventory data. Note that although privacy issues arise when a driver's name is tied to age, experience, physical condition, or the extent of injuries, there is no privacy issue if the driver's name, license number, or Social Security number is not associated with the physical characteristics of the driver. Packaging could be handled in a similar fashion, without disclosing any business-sensitive information.

5.2.3 Add or Modify Inventory Data in Databases

This subsection includes specific potential measures for adding inventory data to the system in key locations such as in the major incident databases or supplemental databases such as PHMSA Registration Database or MCMIS Census file.

5.2.4 Link Data from HMIRS, MCMIS, RAIRS, and Other Information Sources

Existing fields suitable for linking individual hazmat crashes in different databases would be described. Where needed, fields that describe the event location, such as lat/long coordinates, street addresses, river and rail mile points, and FIPS codes, could be added and/or better quality controlled (using GIS technology) to facilitate the linking of databases. Common accident identifiers are suggested to encourage data integration, validation, and sharing.

For all hazmat truck crashes, the DOT number could be correctly reported and entered into the Crash file. The use of a police report number would be another possibility, for linking traffic accidents. To ensure this takes place, a copy of the police report could be submitted with the crash report. FIPS codes for all geographic entities (such as states, counties, cities) would be used as available. Time could be entered from the police report. MCMIS also could make sure that the report submitted to the crash file includes the police accident report number.

5.2.5 Develop a System for Each Database That Will Target About 5% of Hazmat Crashes for More Detailed Investigation

Ideally, the chosen accidents would be from a common class of accidents. As one potential example, a sample of hazmat crashes involving rollovers of hazmat tanker trucks could be selected for more detailed root cause investigation. Another potential application could be the selection of crashes involving a particular commodity, such as propane, to determine if these accidents are over-represented and, if so, what measures could be developed to decrease both their frequency and severity. A further option could be the selection of hazmat crashes occurring on a particular type of roadway or involving a certain category of driver based on age and experience. Finally, a random sample from all accidents could be selected for further investigation.

This approach is already being implemented, although not exclusively for crashes involving hazmat. FRA currently supplements selected rail accidents in the RAIRS data by about 100 detailed accident investigation reports published annually. The Coast Guard does a few detailed accident investigations, and NTSB investigates almost all air crashes as well as selected serious crashes for other modes. By examining 100 accidents in detail, FRA is able to obtain additional data for accidents of interest and thereby probe deeper into the root and contributing causes of those accidents. The FRA example provides a workable framework for investigating the root and contributing causes of hazmat accidents. Similar additional investigations could be undertaken by each agency responsible for a major database.

For trucks, a slightly different approach was taken in the Large Truck Crash Causation Study (LTCCS). This study obtained detailed parameters for approximately 1,000 heavy truck accidents. While no attempt was made to use the data in MCMIS or HMIRS, it is an example of a comprehensive approach whose only limitation is that it could not be performed on an annual basis. Although the LTCCS and NTSB reports provide potential models for these investigations, the TIFA framework that provides detailed analyses for all fatal large truck crashes is a more feasible model because it is performed annually, relies on telephone calls to check information, and is less expensive than the LTCCS approach. The *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005) applied a similar approach as that used in TIFA on a sample of hazmat crashes found in MCMIS for a particular year. In addition, the project used information for the same crash found in HMIRS wherever possible.

5.3 Improving the Effectiveness of All Databases Required to Identify Root Causes

5.3.1 Ensure Data Completeness and Accuracy

Data completeness includes accidents that are not reported as well as accident reports with incomplete data. If the system reports only one-half of the accidents that occurred, then it may take twice as long to obtain a good understanding of how the system responds to a set of changing conditions. Even then, this may not be possible because those accidents that were unreported may not be representative of those that are reported. For example, accident severity, driver, and/or carrier characteristics might differ. Typically, less severe accidents are underreported. This is true whether the cases are self-reported or not. In addition, carriers with poor records are less likely to self-report. This means less safe carriers are less likely to be identified in a timely fashion.

Incomplete reports can also have a negative effect on accident analysis since a complete record of cause and effect is not captured. Moreover, when one parameter is not filled out for one accident record and a different parameter is missing in another accident record, the situation is worse than if the same parameter value is not filled in for each accident. When the negative consequences of underreported accidents and incomplete reporting are combined, the ability of an analyst to draw conclusions from the data is significantly compromised. For example, for semitrailer truck accidents with hazmat releases, if only 70% of the accidents are reported and, for those that are reported, only 75% of the truck configurations are known and speed at the time of a crash is populated only 50% of the time, it is virtually impossible to have any confidence in an estimate of the annual number of semitrailer truck hazmat releases in which excess speed is a contributing cause. This is because there would be complete data for only about 25% of the accidents that occurred ($0.70 \times 0.75 \times 0.50 = 0.2625$). Moreover, there is no assurance that every accident with complete information is representative of the three that are missing.

5.3.2 Complete Values for All Parameters

No credible information system can operate using records in which many fields are blank. Parameters are specified in a database because they can be collected and are important to collect. Seemingly unimportant parameters are essential for specifying the causal chain of events or characterizing the circumstances of the event.

For any database incorporated into the proposed information system, required fields should not be left blank. In some instances, it may be necessary to provide an estimate of the appropriate data entry. Some databases allow for this, but maintain a code to distinguish between measured and estimated values. This type of information is valuable for parameters that change as a result of the accident sequence, such as speed. Although the code for "unknown" will always be an option, it should be used only if it is truly impossible to estimate a value for one of the required parameters. Often, an estimate is good enough, since categories are generally aggregated for analysis.

Techniques would need to be implemented for ensuring that information required for the data fields is both collected and correctly entered into the database. For example, the overwhelming majority of truck carriers surveyed believe that PHMSA should contact carriers to collect data for fields that are incomplete. Some techniques that could be implemented by database managers include the following:

- Computerize data collection to enforce and validate data coding.
- Create required fields that will not accept the record until those fields are complete.

- Develop incentives that will reward those that provide complete and accurate data into a database. For example, for states, a system could be developed to provide feedback in the form of crash data for the state. Or for carriers, data could be provided that defines how a company compares to its competitors.

5.3.2.1 Add Latitude and Longitude

Add latitude and longitude to all databases to provide the exact location of a hazmat crash and enable the identification of the same accident when it appears in more than one database. This would facilitate linkages between databases where information on the same crash is compared and also would aid in obtaining site-specific information associated with a particular location.

5.3.2.2 Add a Specific Description of the Hazmat

A specific hazmat description could be added to the database and would include the type or hazmat class identified on the placard and/or shipping papers, name of the chemical being transported, quantity being transported, and—if there is a spill—the quantity spilled. Adding this information will provide significantly more detail on the hazmat being shipped and make the database more valuable for identifying root causes of hazmat crashes.

5.3.2.3 Electronic Submission of All Crash Reports to the Major Databases Using the Web to Facilitate Accuracy

All crash reports could be submitted electronically using the web to facilitate accuracy. Currently, information is submitted from the states using the SafetyNet System. Although this approach is undoubtedly fixed for the near future, the electronic submission of the complete police accident reports (PARs), including truck and hazmat supplements, to the states could reduce the opportunity for error when the information is entered into SafetyNet. Similarly, if HMIRS reports were submitted using a web interface, an added source of errors would be removed by eliminating one data transfer step. In addition, an automated system could be developed to check the forms for completeness and identify incorrect coding.

5.3.2.4 Add "Error Trapping"

"Error trapping" could be applied to all databases in order to eliminate errors by applying a program that would test whether certain logical connections have been made within the same accident record. One example of an error that could be eliminated is a situation where a travel speed of 40 mph is listed for a crash associated with a truck backing into a loading dock. Another example of a "logical inconsistency" could involve an accident where a single-truck cargo tank is reported to have lost 20,000 gallons of a hazmat liquid. The error-trapping program would flag this spill size as erroneous because it is inconsistent with realistic truck cargo tank capacities.

5.3.2.5 Increase State Checks of the Quality of Hazmat Crash Submittals to FMCSA

As part of their reporting obligation, each state could quality-check reports for all hazmat crashes in MCMIS. States could read the reports, complete missing fields, and make any needed corrections before submittal to FMCSA. This step would increase the reliability of the MCMIS Crash file. As the use of electronic submissions becomes the norm at all levels, states could be encouraged to develop translational programs that automatically populate the fields in the MCMIS Crash file from the electronic entries in the PARs, thereby eliminating translation errors and the need to check for such translation errors between the PAR and the MCMIS Crash file. States with demonstrated higher quality submittals might be rewarded.

5.3.2.6 Ensure that the Databases' MCMIS Report Number for Serious Hazmat Crashes Can Be Linked to the PAR

Agencies checking the quality of MCMIS crash reports should be able to link the PARs with the MCMIS report. Including the PAR number in all MCMIS hazmat crash reports would allow the PAR associated with each crash to be easily identified and quickly retrieved. Currently, in many states this is a time-consuming manual process that could be easily remedied if the PAR number was referenced in the MCMIS Crash file and the PAR was electronically retrievable. Linking the MCMIS report number to the PAR could be a suggestion for police reports and require the following:

- States could be encouraged to automate collecting standard administrative data as much as possible since MCMIS data are extracted from police reports.
- Include non-spill hazmat crashes in HMIRS for all "serious crashes" involving placarded shipments.
- Include digital photos of the accident scene in all of the major databases.
- For truck carriers in MCMIS and HMIRS, include carrier out-of-service (OOS) rate per million miles. MCMIS inspection data could be able to determine the count of OOS vehicles for a carrier. Could then be normalized by the VMT estimate in the Census file.
- For HMIRS and MCMIS, include violations for drivers during the crash.
- Involve other parties such as emergency responders and insurance companies in data collection and reporting. In HMIRS, the *IEVENT* Table could be used to capture remarks made by multiple individuals.

5.3.2.7 Include All Applicable Hazmat Accidents

Each database should include a process for ensuring that all hazmat accidents that meet the specifications for inclusion are recorded. In order to accomplish this objective, agencies managing these databases could consider the following:

- Conduct training for carriers and police in how to complete accident reports. For truck accidents reported to MCMIS, because there are tens of thousands of different police officers filling out PARs each year and individuals from more than 50 states and territories filing crash reports, Web-based training is likely to be the most cost-effective method of delivering such training. Although the situation is simpler for other modes (e.g., rail and barge) because fewer individuals are responsible for compiling and reporting crash data, Web-based training might be cost effective for these individuals as well.
- Develop more efficient systems for entering data at the accident scene, such as greater dependence on electronic data entry and use of bar codes.
- Review news reports of hazmat accidents to identify those that should have been included in the database. Investigate the techniques used by financial news reporting services to automatically flag newsworthy statements made by company officials and reported in news reports.
- Check PARs to identify those accidents that should be included. For truck transport, this would be greatly facilitated if PARs were available electronically, thereby making the identification of reportable crashes more accurate and uniform.
- Provide incentives to states to verify that all hazmat truck crashes that should be reported have been included in crash databases.

One suggested project would be to establish a multi-agency task force to identify all hazmat accidents that meet the reporting threshold, considering all modes of transport. The project would perform the searches listed above and compare the results with HMIRS, FARS, RAIRS, and MCMIS, as appropriate for the mode. The outcome would come close to being a census of hazmat accidents. This census could then be used by HMIRS, FARS, MCMIS, and RAIRS as a measuring tool to determine their respective completeness. It could also provide the sampling frame for in-depth investigations as suggested in this subsection.

5.3.2.8 Ensure Accuracy and Consistency of All Data Entered into the Databases

Reporting incorrect data can make it impossible to identify causal relationships, even with an infinite amount of data. Thousands of vehicles might experience the same initiating event and correct for it without incident. The next vehicle might experience the same initiating event and, through a series of unique conditions, be unable to correct for the initiating event. If these unique conditions are not accurately recorded, then the true cause(s) will never be identified.

Potential measures for ensuring data accuracy fall into the following categories:

1. **Include parameters that are directly measurable** such as how far from the roadway a vehicle was after a crash or parameters that are directly observable, such as whether it was dark or light. Such parameters should be recorded with high accuracy.
2. **Include event data recorders on hazmat transporters.** Hazmat vehicles are a good test bed for these devices. Knowing vehicle parameters for the 5 seconds prior to a crash would be invaluable. Just knowing the vehicle speed, when the brakes were applied, and with what force the brakes were applied would be insightful.
3. **Record parameters that occurred before or during an accident** and could be objectively described but are no longer apparent to the accident reporter. An example of this type of parameter could be an accident that occurred as a result of a white-out condition that is no longer present at the accident scene.
4. **Include parameters that are based on highly subjective witness statements.** Police officers are highly trained to recognize unusual driver behavior and these subjective observations should be recorded. The identification of contributing causes falls into this class of recorded information.
5. **Record parameters that become available some time after the accident.** Investigating officers will frequently require a driver submit to a drug test when impaired performance is suspected. Even though this information is not available for several weeks, such information can be obtained and is routinely captured in the TIFA database. Other accident databases could adopt the TIFA procedures for capturing this information.

5.3.2.9 Place All Components of the Proposed Information System Under a Quality Assurance Program

A process should be implemented that provides periodic reports stating the accuracy of the data entry process. The program could include the following measures:

1. **A consistent data dictionary** that provides accurate definitions of the parameters such that two individuals, when assessing the same accident, would fill in the same cause codes (where multiple interpretations are possible, the question/answer format adopted in RAIRS provides an effective means to ensure more consistent reporting);
2. **Consistent PAR forms** that contain the minimum amount of data required to identify root causes;
3. **Training for individuals** responsible for completing PARs or reports for HMIRS and MCMIS. As stated previously, since an individual might only need to fill out a report once every few years, Web-based training that could be accessed to assist the reporting official might be a very cost-effective training tool; and
4. **A data checking procedure** to ensure quality control of information using such techniques as:
    - Checking a sample of all data submitted and entered into the database,
    - Making telephone calls to participants for a sample of accidents,
    - Comparing reports with press accounts of hazmat crashes, and
    - Carefully reviewing the database to identify inaccuracies and inconsistencies.

Table 5-1. Accident parameters.

Vehicle	Driver	Packaging Hazmat	Infrastructure	Situational
Configuration	Age	Package Type Hazardous Material	Road Surface	Pre-Crash Condition
Cargo Body	Experience	Quantity Shipped	Road Condition	Dangerous Event
GVW	Condition	Quantity Lost	Road Type	Vehicle Speed
Vehicle Defect	Valid License	Age (Cargo Tank)	Traffic Way	Impact Location
Vehicle Response	Citation Issued	Rollover Protection	Access Control	Primary Reason
	Response	Inspection History	Speed Limit	Accident Type
	Training	Design Specification	No. of Lanes	Weather Condition
			Location	Light Condition
				Time of Day
				Health Consequence

5.3.2.10 Data Breadth

This section covers the type of data breadth required to upgrade the system. This is difficult to specify because it depends on the type of analysis that must be performed to get to the root cause of a class of accidents. There is not, a priori, an assumption that can be made regarding what information is important since this relates to the contributing causes.

5.3.2.10.1 Data Breadth for Trucks. For trucks, basic data may be needed in one of five areas including vehicle, driver, packaging, infrastructure, and situational. Table 5-1, adapted from the *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005), can be used to insure that there is no missing data in each of the five areas.

5.3.2.10.2 Data Breadth for Trains. For trains, basic data may be needed in one of five areas including train consist, engineer/crew, packaging/hazmat, track type, and situational.

5.3.2.10.3 Data Breadth for Water Carriers. For water carriers, basic data may be needed in one of five areas including barge or vessel type, captain/crew, cargo configuration/hazmat, waterway, and situational.

5.4 Potential Measures for Improving Capability of Specific Databases to Identify Root Causes

5.4.1 Potential Measures for MCMIS

The following potential measures apply to enhancing the ability of MCMIS to identify root causes of hazmat accidents.

5.4.1.1 Provide Training in Completing Reports for Carriers and Police

The goal of this effort would be to improve the reliability of the MCMIS database by providing targeted training for those individuals responsible for submitting accident reports. The source of MCMIS data is the PARs that are completed by police officers. The most effective method for improving the quality of PARs would be to develop an online training package that provides police departments with guidance for investigating a crash and completing the PAR for serious hazmat crashes. The California Highway Patrol's *Collision Investigation Manual* could serve as a model for developing these materials.

5.4.1.2 Complete All Parameter Fields

Fully completed MCMIS parameter fields offer the potential for the single biggest improvement in MCMIS crash reporting. Some fields could be required to be filled out before an accident record can be uploaded, particularly those related to the vehicle, carrier, driver, route characteristics, and point of contact information.

- **Complete Driver Condition Field**
  Since 2004, the *DRIVER_CONDITION_CODE* field has been left blank. In *Hazardous Materials Serious Crash Analysis: Phase 2* (Battelle 2005), the code "Appeared Normal" was the common entry for about 94% of the vehicle crash records. Being able to flag those 6% for more detailed study might result in improved driver performance not only for the 6% identified but for some of the 94% that appeared normal but, in fact, were impaired. Since this is the only field that captures driver performance in MCMIS, it is suggested that this field be filled out again.
- **Complete Hazmat Fields**
  All five hazmat fields should be completely and accurately filled out for accidents involving trucks carrying hazmat. As described in Section 4.1, this is often not accomplished. If two fields must be filled out for a consistency check, this can occur in only 32% of the cases—the accidents where four or five of the fields are filled out. Furthermore, for the 32% of the cases where two or more descriptive fields are filled out, the entries are often inconsistent, making it difficult to accurately determine even the class of hazardous material being transported. Although it is normally possible to identify the name of the hazardous material from the data reported in the *VEHICLE_HAZMAT_MATERIAL* field, it should be noted that in either the recording of the information or in the electronic transmission of the data, the field is often being truncated.

  Bar codes could be used to supplement placards to supply the police officer with accurate data on the carrier, vehicle, driver, and type of hazardous material. These data could be read easily with an inexpensive hand-held bar code reader then transferred to a police officer's computer or printed and attached to the PAR.

  The use of radio frequency identification (RFID) tags on all large trucks transporting the most dangerous hazmat, such as TIH and explosives, should be considered. Information in the tags could include the driver, vehicle, hazmat cargo, carrier, and vehicle. This system would improve the accuracy of police reporting and also provide a valuable tool for emergency responders to identify hazmat remotely.
- **Specify the Location Accurately**
  The location field should be specified in a manner that enables the accident location to be found on a map. Presently, the accident location can be found on a map for about 30% of the crashes. Specifying the route number or street name followed by the longitude and latitude would appear to be a straight-forward way to register the location. The difficulty in identifying the accident location on a map is exasperated by truncation errors occurring somewhere in the recording or record transmission process, thereby eliminating key information in the *LOCATION* field.
- **Provide State Personnel Access to Other Key Data**
  State personnel entering the data into the MCMIS crash record system should have access to databases containing related information, such as the MCMIS Registration file and the 49 CFR Part 172 Hazardous Material Table. Having access to such files would enable state personnel to perform a quality control check on the hazmat entries and fill in any information missing from the PAR. Linking the data entry process with these, and other, files would make it easier to accurately populate fields in the MCMIS Crash file.
- **Ensure that the MCMIS Report Number Be Linked to the PAR**
  Agencies checking the quality of MCMIS crash reports should be able to easily link the PARs with the MCMIS report. Therefore, it is suggested that the PAR number be included

in all MCMIS hazmat crash reports. FMCSA could restore the rule for how the *REPORT_NUMBER* field is constructed. Prior to 2001, states were instructed to use the PAR number in the *REPORT_NUMBER* field. Although that rule is no longer required, some states embed the PAR number in the *REPORT_NUMBER* field. The PAR number of a crash would permit a hard link to a specific crash and not only provide more definite access to the PAR, but it would also facilitate linkages to crashes found in other databases such as HMIRS.

5.4.1.3 Add Data on Pre-Crash Conditions

Pre-crash data concerning the driver and the vehicle would be entered into SafetyNet by the states. Driver information would include the driver's safety, violation, and health records. Vehicle information would include vehicle defects discovered at the scene. At the very least, for a crash involving hazmat, a Level I inspection would be conducted at the scene or another location, with special attention to such defects as brake adjustment and tire condition. In addition, information about the vehicle's maintenance history could be provided. This may simply reflect that the maintenance records were current and required maintenance had been conducted. This is considered a long-term measure. At present, this information could be obtained for targeted accidents, say 100 per year as FRA does currently. If it is shown to be cost effective, the number of targeted accidents could be expanded over time.

5.4.1.4 Enhance the MCMIS Crash File Data Dictionary

The MCMIS data dictionary could be enhanced so that it contains not only the definition of a parameter and the format for the field in the database, but also the format of the data to be entered. Specifying the format in the database does not necessarily define the data entry format as evidenced by the current dataset.

A section answering some commonly asked questions would be valuable as well. For instance, if the PAR lists the carrier location as one of the carrier's freight depots, should that address be entered in the MCMIS Crash file or should the address of the carrier's home office, taken from the MCMIS Registration file, be entered? Another consideration might be related to the choice of entering a street address or a postal box number. Another example is whether an Interstate route should be designated as I-70, IR70, I070, I70, or some other format. If the use of longitude and latitude when specifying the location is adopted, then the format and accuracy also should be specified. If the coordinates were expressed in decimal degrees, then specifying the longitude and latitude to two decimal places would place the accident on a highway, yet if specified to three decimal points the location would be shown as either being on the left- or right-hand side of the right-of-way.

5.4.2 Potential Measures for HMIRS

The following potential measures apply to enhancing the ability of HMIRS to identify root causes of hazmat accidents.

- **Provide Training for Carriers**
  Training in completion of the 5800.1 report should be provided to carriers. Included in the course would be a unit on root cause identification. Some of the training elements given to NTSB inspectors could form the basis for this training unit. The course would be presented in a webinar format, comprised of two four-hour sessions.
- **Use Drop Down Lists**
  HMIRS could use drop down lists for such data as place, carrier name, vehicle type, container type, and hazmat type. This would prevent unneeded mistakes resulting from different interpretations of particular spelling.

- **Include a Copy of the PAR with Reports**
  Carriers would submit a copy of the PAR for any HMIRS reportable traffic accidents. This provides another tool for PHMSA to confirm the accuracy of material in the 5800.1 report. PHMSA could check the carrier's report against the PAR to identify inconsistencies. These inconsistencies could be sent to the carrier for confirmation or clarification. Changes to the 5800.1 report could be made where needed.
- **Ensure that Filers Fix Incorrect Data Before the Submission Is Accepted**
  In FY09, PHMSA will introduce an online Incident Reporting System that will require filers to fix incorrect data before the submission will be accepted. However, since the carriers will also be able to file the reports via other methods, the effectiveness of these checks will be limited to electronically submitted reports.
- **Include Non-Spill Hazmat Crashes in HMIRS**
  Most serious crashes are potential spills even if none occurs. Therefore by including all of these "serious crashes" involving placarded shipments, officials analyzing the crash data in HMIRS will be able to more effectively determine root causes. The definition FMCSA uses for the inclusion of a crash in the MCMIS database could be applied to all non-spill hazmat crashes that are not currently included among the incidents that must be reported to HMIRS. By including non-spills, PHMSA would be able to provide data on "successes"; that is, what worked well in the hazmat transportation system. For example, if a particular packaging was involved in multiple rollover crashes and resulted in fewer spills than another packaging, this information could provide valuable evidence for use of a particular packaging type. The recent requirement that Type C accidents be reported is a step in this direction, showing that it is considered feasible to include such non-spill accidents in the database.
- **Encourage Carriers to Enter Multiple Hazardous Materials in the 5800.1 Form**
  Carriers sometimes carry more than one hazardous material. This is true for less than truckload cargo shipments. Therefore, carriers can easily enter this information in the 5800.1 report sent to PHMSA. Many do not break the information down, making it impossible to distinguish good and poor package behavior.
- **Send All Reports to the Carrier for Confirmation**
  A report could be generated automatically after data are entered into HMIRS and a letter or email sent to the reporting carrier to confirm the data. A certain amount of time, such as one month, could be used to allow the carrier to check the accuracy of their submittal and to add information their internal investigations may have discovered. PHMSA could change the content of the 5800.1 report, if required.
- **Subject Crashes That Meet a Certain Threshold to Follow-Up Audits**
  All crashes involving certain classes of hazardous materials could be subject to a more detailed PHMSA audit. The audit would thoroughly check the accuracy and completeness of the accident description and collect additional information where required. Specific material types with a certain hazard threshold could be selected for this audit. For example, the hierarchy below could be followed with the top of the list having the highest priority for audit.
  - Class 1: explosives,
  - Class 7: radioactive,
  - Class 6.2: infectious substances,
  - Classes 2.3 and 6.1: toxic inhalation hazard (TIH),
  - Class 2.1: flammable gas,
  - Class 3: flammable liquid,
  - Class 4: dangerous when wet, and
  - Class 5: oxidizer.
- **Provide a DOT Number for All Reports**
  A DOT number should be reported in all HMIRS records for en route accidents. This will facilitate matching information on the same crash in other databases.

- **Verify Carrier Names**
  An additional quality assurance check could be performed to verify that the name being entered corresponds to the name provided on the annual PHMSA registration form. This may be already being done but was not mentioned at the time of interviews with officials.
- **Include the Number of Power Units and Drivers**
  For HMIRS, the number of power units and drivers could be included as data elements. All carriers reporting to HMIRS should be in the DOT Census file, so the number of power units and drivers could be extracted from there. This information provides an indication of carrier size that may reflect on the carrier's ability to complete the 5800.1 report.
- **Ensure Package Failure Entries Are Complete**
  PKGFAIL entries should be filled out for all reports submitted to PHMSA. If incomplete, the report would be returned to the carrier for completion of the information and submitted either by phone or e-mail.
- **Continue to Emphasize the Importance of New Reporting Requirements for Damage to Lading and Lading Protection Systems**
  The new requirement that carriers must file a 5800.1 form following an accident if there was damage to lading and lading protection systems on a cargo tank of 1,000 gallons or greater, even if there is no loss of hazardous material, could be emphasized. This is the new requirement to report Class C accidents. Such a notice might be given to carriers when they are informed that their annual hazmat registration application has been reviewed and approved. (This measure would be superseded if all non-spill hazmat accidents were reported in HMIRS.)

5.4.3 Potential Measures for TIFA

5.4.3.1 Potential Measures for Additional Fields

Although the TIFA survey adds valuable detail to the FARS data, additional data fields could add important detail about hazmat packaging and also enhance the information available on crash causation. Some elements could be added with little modification of the program. Others would take additional resources, but are nevertheless possible.

- **Right of way** could identify which vehicle, if any, within a crash had the right of way prior to the collision. This could be coded readily from the PAR in most cases. Some state crash reports already include right of way. Right of way would be useful in identifying the vehicle that primarily contributed to the crash.
- **Critical event** is a field that identifies and describes the event that precipitated the crash. This field is included in both the GES and CDS files. Coding manuals are available and could be used to ensure that coding is consistent with NCSA standards.
- The managers of the TIFA survey might also consider adding a **critical reason** field. Critical reason captures the "reason" for the critical event, classified broadly as driver, vehicle, or environment, with detailed levels under each. This variable is useful for identifying the immediate failure that led to the crash and would shed light on crash causation. The field was used in the Large Truck Crash Causation Study (LTCCS) and in the National Motor Vehicle Crash Causation Survey (NMVCCS). Therefore, coding procedures are available. The suggestion to add critical reason to TIFA is tentative, however, as it is not clear whether the TIFA protocol can uncover this information.
- The TIFA program could also **add additional information** about hazmat cargo, in particular
  - MC number of the cargo tank. This information has been collected in the past as part of a special data collection, so the feasibility of collecting this information has been demonstrated.
  - Quantity of hazmat transported. Cargo weight could be added back to the survey. The program should consider capturing the quantity in terms of liquid measure, where appropriate.

5.4.3.2 Potential Measures to Improve Data Quality

The TIFA system is complete and mature. It is subject to annual review and adjustment, including continuing training of the coders. However, greater cooperation with the FARS program could probably help to increase the accuracy with which trucks are identified.

5.4.3.3 Compatibility with Other Databases

The comments previously associated with the FARS data apply equally to TIFA, since TIFA uses the same case number system. However, there is an additional constraint on linking the TIFA file to other data systems, in that it raises a risk of identifying specific individuals or organizations. The TIFA program is bound by commitments to respondents to protect their identity, and by the terms of its operations under the University of Michigan's Institutional Review Board (IRB). Thus, any effort to link the data to other data systems would be unlikely to be allowed.

5.4.4 Potential Measures for RAIRS

For rail transport of hazardous materials, Item 1, and to a substantial extent Item 2, are fairly well developed, thanks to a combination of company, industry, and government programs.

The requirement that railroads provide the reporting mark and number of all cars involved in releasing hazmat and the quantity released is useful because this facilitates acquisition of more information about the design of the car via the AAR's Universal Machine Language Equipment Register (UMLER) database. An audit of the 2007 FRA data found that for all of the hazmat cars that the Class 1 railroads indicated had released hazardous materials, they provided the required information for all but one that they indicated had released product. The single discrepancy was a case in which a railroad indicated two cars had released product, but only one car was identified. It is not possible to determine which was incorrect, the number of cars that actually released hazardous materials in the accident, or if a second car's identity and information should have been provided. The same level of compliance was not evident for the non-Class 1 railroads. However, non-Class 1 railroads are responsible for a much smaller percentage of the hazmat cars that release in accidents so the impact of these on overall data completeness is much less. The value of this requirement could be considerably enhanced in terms of root cause analysis if the following changes were made with regard to reporting mark, car number, and identification of the commodity being transported:

1. Provision of the reporting mark and number of all derailed cars of any type;
2. Provision of the commodity, reporting mark, and car number for all derailed cars placarded as transporting hazardous materials;
3. Provision of the commodity, reporting mark, car number, and quantity released for all tank cars, whether or not they are transporting hazardous material; and
4. Provision of the same information called for in No. 3 for all intermodal, portable tanks being transported on cars that derail, along with the reporting mark and number of the railcar transporting them.

The reasoning for each is as follows:

- Knowledge of all derailing cars would provide the all-important "denominator" data needed to establish normalized rates of failure for various railcars and their critical design elements. It does little good to know if 10 times as many of one car type release compared to another if one does not know if there were 10 times as many cars of the former type being transported compared to the latter.
- Although tank cars transport the majority of hazardous materials, covered hoppers, boxcars, and intermodal cars transport significant quantities as well. It is useful to be able to distinguish the performance of different car types in accidents, which implementing these potential measures would allow.

- Tank cars used to transport some non-regulated materials are often identical to cars transporting certain regulated materials. Knowledge of their exposure to accidents and performance in accidents will substantially improve the robustness of the data, and consequent confidence in, and accuracy of, the statistics pertaining to tank cars of similar design that are used to transport hazardous materials.
- Development of information on intermodal portable tanks (isotainers) is needed to understand their performance in accidents and strengths and weaknesses in their damage-resistant design. This could be achieved in a manner analogous to the understanding that has developed during the past 38 years of studying railway tank cars. This mode of bulk transport is expanding, especially in the area of import and export of hazardous materials. Recording the information described in this subsection will provide a basis for development of such statistics.

Although the information presented in this subsection is not all that is needed for improved root cause analysis of rail transport of hazardous materials, in combination with detailed data on equipment design recorded by the railroads in the Universal Machine Language Equipment Register (UMLER) and the data on releases recorded by PHMSA in HMIRS, it would substantially strengthen our understanding of the factors affecting railcar performance and failure modes in accidents.

5.5 Conclusions

The research conducted under this project has demonstrated that there has been considerable progress during the past 20 years in the development and refinement of databases that include hazmat accidents. The project focused on identifying potential measures for improving the identification of the root causes of hazmat accidents using these databases. The project findings have provided researchers and officials with an overview and analysis of the individual databases and resulted in many potential measures for improving specific databases. However, implementation of the major measures—including establishing an information system, linking databases to take advantage of accident descriptions in more than one database, performing detailed sampling of a specific set of crashes to assemble more detailed information, and the adoption of techniques to improve data quality and completeness—could yield the greatest improvements in the ability of interested parties to conduct root cause analysis. This, in turn, would enable officials to identify problems for which a solution or mitigation will result in an improvement in hazmat shipment safety.

5.6 Follow-On Project

If implemented, the findings of this report could lead to the enhanced identification of root and contributing causes of hazmat crashes. The implementation of the potential measures that were identified will likely present both technical and institutional challenges. Consequently, the project team also suggests that in order to evaluate the feasibility, usefulness, and costs of implementation, a pilot program be implemented to demonstrate that the system will work effectively in identifying root causes. The team suggests that the pilot test focus on truck hazmat accidents and involve linking the HMIRS database, which provides excellent data on the hazmat material and package, with the MCMIS database, which provides superior data on the driver and accident environment. To supplement the data found in the two databases, the pilot program could link at least 100 crashes and supplement the data primarily by telephoning carriers and other key sources such as police officers and tow truck drivers. The pilot project would select the sample of hazmat crashes based on a set of consistent criteria for a similar group of crashes such as hazmat truck rollovers. The pilot study would also document the costs associated with linking the two databases, identifying the sample for more detailed data collection, and the actual collection of the additional data through telephone contacts and other methods. Finally, the pilot test could use the hazmat crash data to demonstrate the system's ability to identify root cause and analyze the results from a number of crashes to pinpoint areas where suggestions can be made to improve hazmat safety.

References

Battelle. 2001. *Comparative Risks of Hazardous Materials and Non-Hazardous Materials Truck Shipment Accidents/ Incidents*. Washington, D.C.: U.S. Department of Transportation, Federal Motor Carrier Safety Administration. http://www.fmcsa.dot.gov/documents/hazmatriskfinalreport.pdf.

Battelle. April 2005. *Hazardous Materials Serious Crash Analysis: Phase 2*. Washington, D.C.: U.S. Department of Transportation, Federal Motor Carrier Safety Administration. http://www.fmcsa.dot.gov/safety-security/ hazmat/Hazardous-Materials-Serious-Crash-Analysis-Phase2-April2005.pdf (accessed Aug, 2009).

Blower, D. 1998. *The Relative Contribution of Truck Drivers and Passenger Car Drivers to Two-Vehicle, Truck-Car Traffic Crashes*. UMTRI 98-25. Ann Arbor, Michigan: University of Michigan Transportation Research Institute.

Blower, D. 2002. "Vehicle Condition and Heavy Truck Accident Involvement." In *Proc. of the International Truck & Bus Safety & Policy Symposium*. Knoxville, TN: Center for Transportation Research, University of Tennessee, and National Safety Council, pp. 311–322.

Bureau of Transportation Statistics. *Project 5 Overview—Developing Common Data on Accident Circumstances*. Washington, D.C.: Bureau of Transportation Statistics. http://www.bts.gov/cgi-bin/breadcrumbs/Print Version.cgi?date=23151349 (accessed Sep 2008).

Bureau of Transportation Statistics. *Rail Equipment—Train Accident Data*. Washington, D.C.: Bureau of Transportation Statistics. http://www.bts.gov/cgi-bin/breadcrumbs/PrintVersion.cgi?date=23115859 (accessed Sep 2008).

Coray Gurnitz Consulting and Abacus Technology. 2007. *Comprehensive Safety Analysis 2010: 2006 Listening Session*. Washington, D.C.: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Dobbins, James and Mark Abkowitz. Jan 2009. "Use of Emerging Technologies for Marine Accident Data Analysis Visualization and Quality Control." In *2009 TRB 88th Annual Meeting: Compendium of Papers*. DVD. Washington, D.C.: Transportation Research Board.

Environmental Protection Agency. 1995. *User's Guide to Federal Accidental Release Databases*. Washington, D.C.: U.S. Environmental Protection Agency, Chemical Emergency Preparedness and Prevention Office.

Federal Motor Carrier Safety Administration. *Motor Carrier Safety Management Information System Code Book*, Washington, D.C.: U.S. Department of Transportation, Federal Motor Carrier Safety Administration. http://mcmiscatalog.fmcsa.dot.gov/beta/Catalogs&Documentation/Documentation/census/mcmis_doc.asp (accessed Aug 2009).

Federal Railroad Administration. April 2005. "Human Factors Root Cause Analysis of Accidents/Incidents Involving Remote Control Locomotive Operations." In *Research Results*, RR05-05. Washington, D.C.: Federal Railroad Administration.

Federal Railroad Administration. May 2003. *FRA Guide for Preparing Accident/Incident Reports*, DOT/FRA/RRS-22. Washington, D.C.: Federal Railroad Administration, Office of Safety, http://www.google.com/search? sourceid=navclient&aq=0h&oq=&ie=UTF-8&rlz=1T4SUNA_enUS316US218&q=fra+guide+for+preparing+ accident+%2f+incident+reports (accessed Aug 2009).

Federal Railroad Administration. May 2003a. "Chapter 7—Form FRA F 6180.54 Rail Equipment Accident/Incident Report." In *FRA Guide for Preparing Accident/Incident Reports*, DOT/FRA/RRS-22, Washington, D.C.: Federal Railroad Administration, Office of Safety. http://www.google.com/search?sourceid=navclient&aq=0h&oq= &ie=UTF-8&rlz=1T4SUNA_enUS316US218&q=fra+guide+for+preparing+accident+%2f+incident+reports (accessed Aug 2009).

Government Accountability Office. 2005. *Highway Safety: Further Opportunities Exist to Improve Data on Crashes Involving Commercial Motor Vehicles*. GAO-06-102. Washington, D.C.: U.S. Government Accountability Office.

Hedlund, J. and D. Blower. 2006. *Large Truck Crash Causation Study (LTCCS) Analysis Series: Using LTCCS Data for Statistical Analyses of Crash Risk.* FMCSA-RI-05-037. Washington, D.C.: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

"National Crash Data Bases Underestimate Underride Statistics." February 1999. *Road Management & Engineering Journal.* http://www.usroads.com/journals/rmej/9902/rm990202.htm.

National Highway Traffic Safety Administration. 2002. *2002 FARS Coding and Validation Manual.* Washington, D.C.: U.S. Department of Transportation, Federal Highway Traffic Safety Administration. http://www-nrd.nhtsa.dot.gov/Pubs/FARS02CVMAN.PDF (accessed Aug 2009).

National Highway Traffic and Safety Administration. August 2008. "Traffic Safety Facts 2007." DOT HS 811 017. Washington, D.C.: U.S. Department of Transportation, National Highway Traffic and Safety Administration. http://www-nrd.nhtsa.dot.gov/Pubs/811017.pdf (accessed Aug 2009).

National Transportation Safety Board. 2002. *Safety Report: Transportation Safety Databases.* NTSB/SR-02/02. Washington D.C.: National Transportation Safety Board.

National Transportation Safety Board. July 1998. *Safety Study, Safety at Passive Grade Crossings.* NTSB/SS-98/02 Vol 1 and 2. Washington D.C.: National Transportation Safety Board.

Paulsson, R. 2005. *Deliverable 5.2: In-Depth Accident Causation Data Study Methodology Development Report (SafetyNet).* Brussels: European Commission. http://www.erso.eu/safetynet/fixed/WP5/SN_Deliverable_5%202%20v3_051123_Final%20submission.pdf.

Scopatz, Robert A. 2006. *Illinois Department of Transportation Crash Data Process Audit.* College Station, TX: Data Nexus, Inc. http://www.dot.state.il.us/trafficsafety/ITRCC/Illinois%20Crash%20Data%20Processes_04102006.pdf.

Scopatz, Robert A. 2001. "Crashes Involving Long Combination Vehicles: Data Quality Problems and Recommendations for Improvement." *Transportation Research Record 1779,* pp. 162–172. Washington D.C.: Transportation Research Board, National Research Council.

Shappell, S. A. and D. A. Wiegmann. 2000. *The Human Factors Analysis and Classification System—HFACS.* DOT/FAA/AM-00/7. Washington D.C.: U.S. Department of Transportation, Federal Aviation Administration.

Shupe Consulting. 2001. *Unified Reporting of Commercial and Non-Commercial Traffic Accidents.* No. SD2000-14-F. Pierre, SD: South Dakota Department of Transportation.

Transportation Research Board. 1981. *Transportation Research Circular 231: Truck Accident Data Systems: State-of-the-Art Report.* Transportation Research Board of the National Academies, Washington, D.C.

Vogt, A. and J. G. Bared. 1998. *Accident Models for Two-Lane Rural Roads: Segments and Intersections.* Report No. FHWA-RD-98-133. Washington D.C.: U.S. Department of Transportation, Federal Highway Administration.

APPENDICES

Unpublished Material

Appendices A through E as submitted by the researchers are not published herein. They are available on the TRB website at www.TRB.org by searching for HMCRP Report 1. Titles of Appendices A through E follow:

Appendix A: Questionnaires
Appendix B: Questionnaire Results for Carriers and Database Administrators
Appendix C: Brief Summary of the 2005 MCMIS Crash Records
Appendix D: The Percent of Missing Data for Variables from TIFA/FARS, 1999–2004
Appendix E: Selected Analyses Performed with the Hazmat Accident Database

Abbreviations and acronyms used without definitions in TRB publications:

AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	Air Transport Association
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
HMCRP	Hazardous Materials Cooperative Research Program
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
PHMSA	Pipeline and Hazardous Materials Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S.DOT	United States Department of Transportation