U.S. Department
of Transportation
National Highway
Traffic Safety
Administration

www.nhtsa.gov

DOT HS 812 128

March 2015

NHTSA's Review of the National Automotive Sampling System: Report to Congress

DISCLAIMER

Suggested APA Format Citation:

National Highway Traffic Safety Administration (2015, March). *NHTSA's review of the National Automotive Sampling System: Report to Congress.* (Report No. DOT HS 812 128). Washington, DC: Author.

Contents

List of Tables and Figures

List of Frequently Used Abbreviations

CDS	Crashworthiness Data System
CISS	Crash Investigation Sampling System
CRSS	Crash Report Sampling System
EDR	event data recorder
ESC	electronic stability control
FMVSS	Federal Motor Vehicle Safety Standard
FRIA	Federal Regulatory Impact Analysis
GES	General Estimates System
NASS	National Automotive Sampling System
NCSA	National Center for Statistics and Analysis
NHTSA	National Highway Traffic Safety Administration
ODI	Office of Defects Investigation
PAR	police accident report
PSU	primary sampling unit
SCI	special crash investigation
VIN	Vehicle Identification Number

1. Executive Summary

The National Highway Traffic Safety Administration's traffic crash data provides the underpinning for informed highway safety decision-making at the Federal, State, and local levels. Accurate, accessible, timely, and standardized data allows decision-makers to:

- determine the primary factors related to the sources of crashes and their outcomes;
- develop and evaluate effective safety countermeasures;
- support traffic safety operations;
- measure progress in reducing crashes and their severity;
- design effective vehicle safety regulations;
- target safety funding; and,
- support defects identification and investigation.

NHTSA's National Automotive Sampling System (NASS) has provided nationally representative traffic crash information to the highway safety community for over 30 years. However, the data needs of the traffic safety community have increased and significantly changed since NASS was initially designed. In addition, the population demographics of the United States have changed over the last three decades, affecting how nationally representative the NASS data collection sites are. NHTSA recently undertook a thorough review of the NASS Crashworthiness Data System (CDS) and the NASS General Estimates System (GES), evaluating the sample design, the data collected, and the underlying information technology.

In the 2012 appropriation, NHTSA received funding to modernize NASS. To ensure that the new NASS best meets the current and future needs of the highway safety community, NHTSA sought user input from government, academia, and industry on both NASS GES and NASS CDS.

Based on its internal review and input from outside stakeholders, NHTSA is designing a new system that is flexible and scalable to efficiently and effectively accommodate these many requirements. The new system will combine the aspects of the current system that will be needed in the future with the additional needs not met with the current system. The new system will be composed of multiple components to accommodate these needs. There will be an annual records-based sample of motor vehicle police-reported crashes, the Crash Report Sampling System (CRSS), and an annual investigation-based sample of motor vehicle police-reported towed passenger vehicle crashes, the Crash Investigation Sampling System (CISS). The new system will also include periodic special studies on motorcycle, medium and heavy truck, or pedestrian and bicycle crashes. Operating the new system, including the periodic special studies, will result in better quality data now and for the future, but will require additional operating funds.

As discussed in this report, it is important to note that the Congress asked NHTSA to review several of its current data systems and describe what an optimal system would look like. This report meets the letter of that requirement, but NHTSA, like other Federal agencies, operates in a resource-constrained environment. NHTSA prioritizes its resources based on safety data and continually reviews its activities to ensure its funding is deployed in the most effective manner possible. As such, the optimal system described in this report may not be constructed precisely as detailed in this report.

NHTSA's design for the modernized system will:

- Support future analytical needs by conducting 15,000 passenger vehicle crash investigations per year in 73 newly selected CISS sites.
- Enhance the collection of pre-crash data and information on crash avoidance technologies, while continuing to collect crashworthiness data.
- Build a flexible system that could be adapted for periodic special crash studies, such as analyses of crashes involving pedestrians, motorcycles, or large trucks.
- Support analyses of all vehicle types and all crash severities by coding 65,000 police accident reports in 75 newly selected CRSS sites.
- Amplify the analytical potential of CRSS through linkage to other data sources.
- Make data more accessible and more secure, and provide users with better tools for analyzing the data.
- Process data in a more streamlined and consolidated information technology environment.

2. The Congressional Requests That This Report Fulfills

Congress made two requests for reports from NHTSA associated with the NASS modernization effort.

In Senate Commerce Report 112-83, the Senate instructed:

> "NHTSA must also undertake a comprehensive review of the data elements to be collected from each crash; solicit input from interested parties—including suppliers, automakers, safety advocates, the medical community and research organizations; and assess the need for more data from the pre-crash, crash, and post-crash phases. The agency should consider including the following factors as part of an enhanced data collection initiative: vehicle velocities; vehicle acceleration/deceleration; departure from the roadway; presence of crash avoidance or driver assistance systems in the vehicle(s); and road surfaces and conditions. … NHTSA shall provide a report on the results of the data element review and recommendations for revision."

In MAP-21, Congress instructed:

> "(a) IN GENERAL.—Not later than 1 year after the date of enactment of this Act, the Secretary shall submit a report to the Committee on Commerce, Science, and Transportation of the Senate and the Committee on Energy and Commerce of the House of Representatives regarding the quality of data collected through the National Automotive Sampling System, including the Special Crash Investigations Program.
>
> (b) REVIEW.—The Administrator of the National Highway Traffic Safety Administration (referred to in this section as the ''Administration'') shall conduct a comprehensive review of the data elements collected from each crash to determine if additional data should be collected. The review under this subsection shall include input from interested parties, including suppliers, automakers, safety advocates, the medical community, and research organizations.
>
> (c) CONTENTS.—the report issued under this section shall include—
>
> (1) The analysis and conclusions the Administration can reach from the amount of motor vehicle crash data collected in a given year;
>
> (2) The additional analysis and conclusions the Administration could reach if more crash investigations were conducted each year;
>
> (3) The number of investigations per year that would allow for optimal data analysis and crash information;
>
> (4) The results of the comprehensive review conducted pursuant to subsection (b);
>
> (5) The incremental costs of collecting and analyzing additional data, as well as data from additional crashes;

(6) The potential for obtaining private funding for all or a portion of the costs under paragraph (5); H. R. 4348—367

(7) The potential for recovering any additional costs from high volume users of the data, while continuing to make the data available to the general public free of charge;

(8) The advantages or disadvantages of expanding collection of non-crash data instead of crash data;

(9) Recommendations for improvements to the Administration's data collection program; and

(10) The resources needed by the Administration to implement such recommendations."

This report fulfills both requests.

3. Introduction

The National Highway Traffic Safety Administration is a data-driven agency. The National Automotive Sampling System (NASS), which includes the General Estimates System (NASS GES) and the Crashworthiness Data System (NASS CDS), along with Special Crash Investigations (SCI), provide NHTSA with real-world information on motor vehicle crashes[1] in the United States.

The data collected through NASS and SCI are used for a wide variety of analyses within the agency. Some of these are periodic and predictable, such as the use of GES data in the publication of the annual *Traffic Safety Facts* report. Many others are unique and difficult to predict, such as the use of CDS and SCI data in analyses conducted by the Office of Defects Investigations (ODI). Due to the unpredictable nature of many of the analytical uses of the NASS data, the system is designed to be a general, all-purpose data system capable of handling unforeseen analytical needs as they arise. NASS GES, NASS CDS, and SCI each have a separate design and function in the current system. They are described briefly below.

3.1. GES

General Estimates System data come from a nationally representative sample of approximately 50,000 police-reported motor vehicle crashes of all severities, from minor to fatal in 60 data collection sites. The system began in 1988, and was created to identify traffic safety problem areas, monitor large-scale trends, inform regulatory and consumer initiatives, and form the basis for cost and benefit safety analyses of traffic safety initiatives. The information in GES records is taken from police accident reports (PARs) and is therefore restricted to the data collected by different police jurisdictions. By restricting the scope to police-reported crashes, GES concentrates on crashes of higher interest and the greatest concern to the highway safety community and the general public. GES data are used in traffic safety analyses by NHTSA as well as other DOT agencies. GES data are also used to answer motor vehicle safety questions from other Federal and State agencies, Congress, researchers, and many other users, including the general public.

3.2. CDS

The Crashworthiness Data System has detailed data on a nationally representative, random sample of minor, serious, and fatal crashes. CDS began in 1977 and was redesigned in 1988 when GES launched. Currently field research teams, located at 24 data collection sites across the country, study about 3,500 crashes a year involving passenger cars, light trucks, vans and utility vehicles towed from the crash scene due to damage. Trained crash investigators obtain data from crash sites within days of the crash, studying evidence such as skid marks, fluid spills, broken glass, and bent guard rails. They locate the vehicles involved, photograph them, measure the crash damage, and identify interior locations that were struck by the occupants. These researchers also interview crash victims and review medical

[1] NHTSA also collects information on fatal crashes through the Fatality Analysis Reporting System (FARS), and the Crash Injury Research and Engineering Network (CIREN) is used to determine injury causation in severe motor vehicle crashes.

records to determine the nature and severity of injuries.

The detailed data collected by CDS research teams is used by NHTSA for evaluating:

- The overall state of traffic safety, and identifying existing and potential traffic safety problems.
- Crash performance, vehicle safety systems and designs.
- The nature of crash injuries as well as the relationship between the type and seriousness of a crash and the resulting injuries.
- Traffic safety standards and programs, which include alcohol and safety belt use programs.

In addition to NHTSA, CDS data is used by Congress, researchers, and the automobile and insurance industries.

3.3. SCI

Since 1972, NCSA's Special Crash Investigations (SCI) Program has provided NHTSA with in-depth and detailed crash investigation data. The data collected ranges from basic data maintained in routine police and insurance crash reports to comprehensive data from special reports by professional crash investigation teams. Hundreds of data elements relevant to the vehicle, occupants, injury mechanisms, roadway, and safety systems involved are collected for each crash. Approximately 100 crashes are designated for study annually.

SCI cases are intended to be an anecdotal data set useful for examining special crash circumstances or outcomes from an engineering perspective. The benefit of this program lies in its ability to locate unique real-world crashes anywhere in the country, and perform in-depth investigations in a timely manner. The data is used by the automotive safety community to improve the performance of state-of-the-art safety systems. Individual and select groups of cases have triggered both individual companies and the industry as a whole to improve the safety performance of motor vehicles, including passenger cars, light trucks, and school buses.

3.4. GES, CDS, and SCI Summary

Table 1 shows the purpose, scope, sources of data, and approximate annual sample sizes of NASS GES, NASS CDS, and SCI as they currently exist.

Table 1: Current Data Systems: GES, CDS, and SCI

	GES	*CDS*	*SCI*
Purpose	To monitor large scale crash trends and broad crash characteristics	To aid in the development, implementation and evaluation of motor vehicle and highway safety countermeasures	To examine the safety impact of new, emerging, and rapidly changing technology and explore alleged or potential vehicle defects
Design	Nationally representative	Nationally representative	Convenience sample
Scope of crashes	Police-reported crashes	Crashes involving an in-transport passenger vehicle towed from the crash scene due to damage	Crashes involving new, emerging, or rapidly changing technology; crashes that involve alleged or potential vehicle defects[2]
Data sources	PARs	PAR; scene inspection; vehicle inspections (for certain vehicles[3]); interviews and medical records (for certain persons[4]), interviews	PAR; scene inspection; vehicle inspections (for certain vehicles); interviews and medical records (for certain persons)
Annual # of sampled crashes	About 50,000	About 3,500	About 100

As with any survey, the types of questions that can be answered using the NASS or SCI data are a function of the variables that are collected. The sample size and survey design will determine the precision of the estimates that are calculated. Figure 1 shows the sampled cases or crashes and the number of data collection sites in CDS since 1988. In 1991, the number of sites was reduced from 36 to 24. Since 1991, CDS has continued to collect data from 24 sites, except for a three-year period from

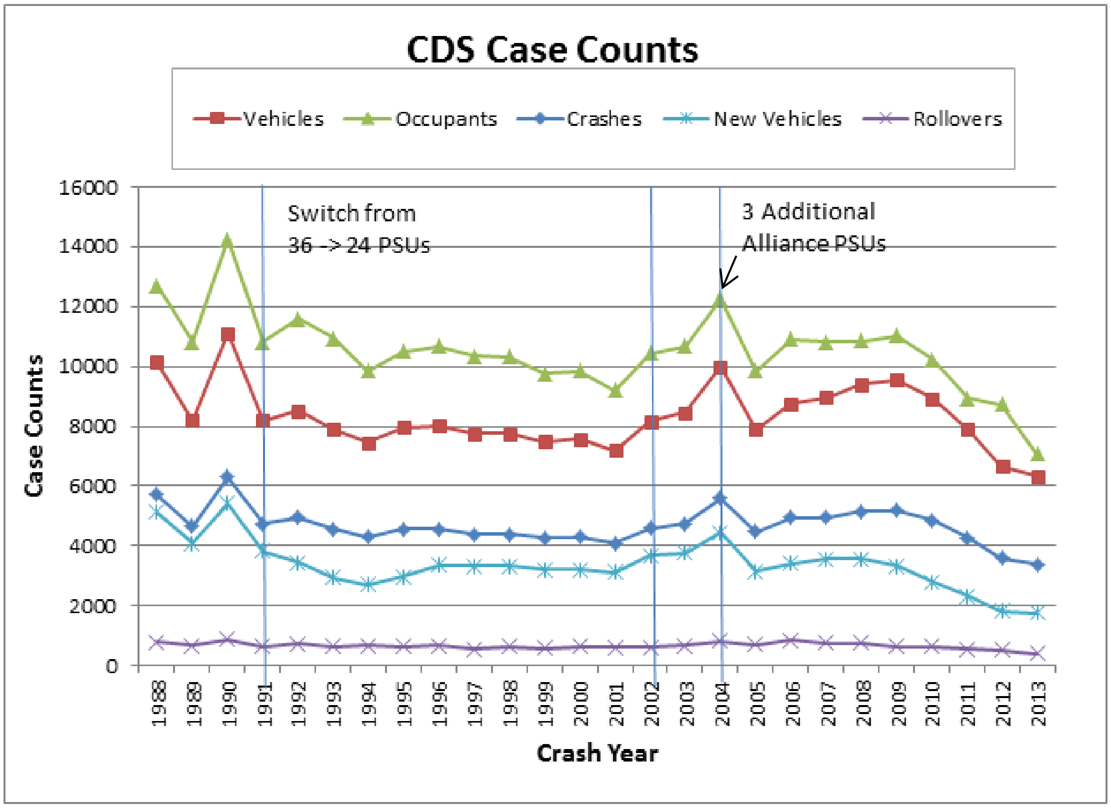
[2] The scope of SCI varies over time. At the time this report was written, SCI was investigating crashes involving: alternative fuel vehicles, unintended acceleration, child restraints, ambulances, non-traffic crashes, oblique/small overlap impacts, avoidance technologies, rollovers, quiet cars, heavy-truck rear impact guards, motorcoach fires, school buses, and advanced occupant protections.

[3] In addition to not being able to conduct inspection on some vehicles (due to, e.g., not obtaining a vehicle owner's permission), or to such inspections being pointless (e.g., if the damage has already been repaired), CDS does not inspect, or conducts less extensive inspections on, certain types of vehicles. For instance, if a motorcycle collides with a (towed) car that is less than 10 years old, CDS inspects the car for damage, but not the motorcycle. For passenger vehicles over 10 years old, CDS only inspects the exterior (but not the interior) of the vehicle. These choices were made to address the balance between the numbers of cases collected for a given cost versus the information collected on each case for the same cost, and this balance is worth revisiting (as we are doing).

[4] CDS does not conduct interviews or collect medical information on motorcycle riders, occupants of trucks over 10,000 pounds GVWR, bicyclists, or pedestrians.

2002 to 2004 when the Alliance of Automobile Manufactures funded three additional sites.[5] The number of cases ("crashes" in Figure 1 shown by the line marked with diamonds) was above 4,000 from CDS inception until 2011, but dropped below 4,000 in 2012. The number of crashes collected in 2013 was 3,385. The steady reduction in the number of cases investigated and in the number of sites has reduced the data available for analysis and weakened the precision of the estimates. The reduction in cases is due to a number of factors, including increased costs, especially labor costs, over time while funding for the NASS had not increased accordingly.

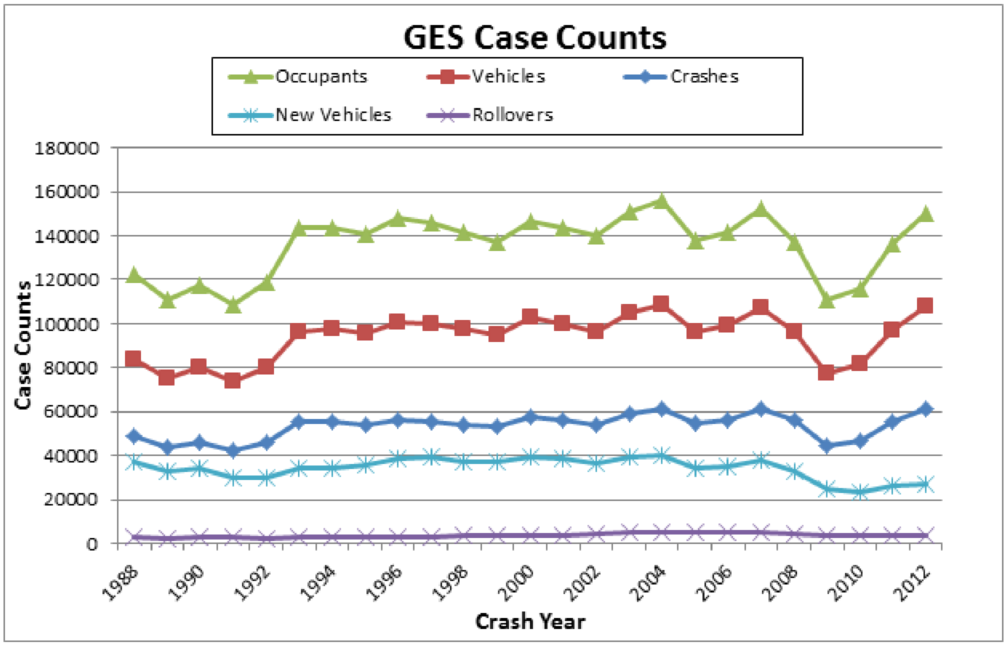
Figure 1: CDS Case Counts



Unlike CDS, the number of cases collected by GES has remained relatively constant and the number of sites in GES has been the same 60 sites since its inception in 1988. GES has averaged 53,430 crashes annually since 1988.

[5] The Alliance paid the NASS contractor to collect additional crashes from three additional data collection sites in a manner that the data from these crashes could be added to the annual file and weighted accordingly.

Figure 2: GES Case Counts



Both CDS and GES are experiencing a sharp decline in the cases involving new vehicles for a number of reasons. This is likely due either to changing demographics at collection sites or people keeping their vehicles longer (or both). New vehicles are defined here as those no more than four years old at the time of the crash. The percentage of new vehicles in CDS has declined from about 50 percent of vehicles in 1988 to 27 percent in 2012. In GES the new vehicles as a percentage of the total vehicles has declined from about 45 percent in 1988 to only 25 percent in 2012. Data for new vehicles is important for safety analyses because new vehicles are most likely to have the latest crashworthiness and crash avoidance technologies.

Figure 3: Percentage of Case Vehicles Less Than 5 Years Old

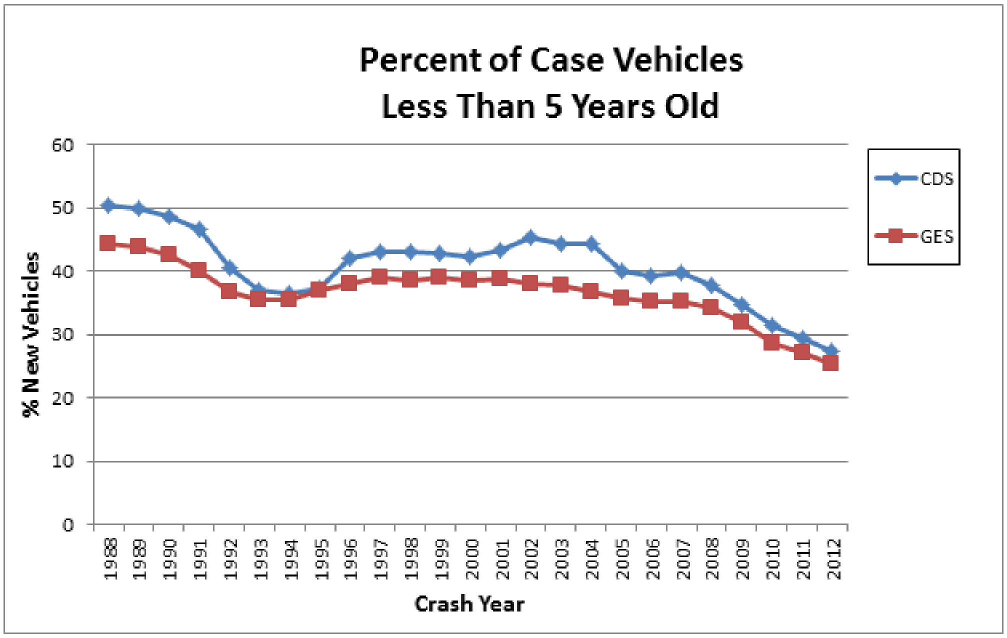


Table 2 shows the coefficient of variance, a measure of precision of estimates from data collected in a survey from a probability-based sample, in the estimation of the populations of interest for GES (all police-reported crashes) and CDS (all crashes where at least one passenger vehicle was towed due to damage). SCI is not included in this table because SCI cases are not selected randomly, making them unsuitable for the generation of nationally representative estimates with known estimates of precision.

Table 2: Precision of 2011 NASS

	Scope	Number of Cases	Weighted Estimate of Total	95% Confidence Interval	Coefficient of Variance
GES 2011	# of police-reported crashes	55,166	5,326,445	(4,773,872, 5,879,018)	0.0519
CDS 2011	# of tow-away crashes	4,278	1,888,859	(1,586,514, 2,191,204)	0.0735

Most analyses conducted using GES and CDS data are based on only a subset of the cases collected in a given year; for example, only rollover crashes or only cars equipped with electronic stability control (ESC). The precision of estimates in these analyses depends heavily on the size of the subset included.

4. Current Usage of NASS CDS, NASS GES, and SCI

The following sections discuss and give examples of how the data acquired through the NASS system is regularly used to support several of the key functions performed by NHTSA. Note that this is not an exhaustive list of the myriad of internal uses of NASS data; rather these examples serve to illustrate the capabilities of the NASS system given its current scope and volume of data.

4.1. NASS GES

NHTSA conducts a variety of public outreach services that both inform the public and attempt to influence and shape driving attitudes and behaviors. Because the GES sample is selected to be nationally representative of all police-reported crashes, it can be used to effectively monitor the state of traffic safety in the United States.

NASS GES data contributes to several annual publications, most notably the *Traffic Safety Facts* annual report, a compilation of the data collected by the GES and the FARS data systems. Also, GES and FARS data are combined in the Traffic Safety Facts Sheets, a series of publications that pertain to specific vehicle types, crash types, or passenger and driver demographics. These publications may include trends across years in addition to annual reporting of crashes, injuries and fatalities. GES contains enough detail to identify many important subpopulations, and has a sufficient sample size to allow estimates to be made on these subpopulations confidently. Topics covered by *Traffic Safety Facts* and the Traffic Safety Fact Sheets series include alcohol-related crashes, crashes involving pedestrians, bicyclists, motorcyclists, large trucks, and many other populations of interest.

The National Center for Statistics and Analysis also generates several publications each year (Research Notes and Technical Notes) that contribute to the body of vehicle safety research. These publications are in-depth examinations of specific areas of vehicle safety. The NASS GES data system allows these reports to be up-to-date, accurate, and relevant. Topics of research published in 2011 using NASS GES data include drowsy driving and electronic stability control.

NHTSA publications like Traffic Safety Facts and research notes are used by State Highway Safety Offices, independent research organizations, safety advocates, and vehicle manufacturers to inform their policies. Many of these entities also access the NASS data directly to perform their own analyses.[6]

4.2. NASS CDS

The NASS CDS is the only nationally representative survey of motor vehicle crashes that includes detailed information on passenger injuries and how they occur. CDS is used by academic institutions, vehicle safety advocates, vehicle manufacturers, medical researchers, and others. Internally, CDS data supports many of the key functions of NHTSA, including being fundamental to the rulemaking process.

[6] NHTSA makes all data (subject to the redaction or omission of personally identifiable information) collected in NASS available to the public, free of charge, at www.nhtsa.gov.

NHTSA issues and enforces Federal Motor Vehicle Safety Standards (FMVSSs). Crash avoidance features (such as brakes and lighting), occupant protection systems (safety belts and air bags), and fire protection (fuel tank integrity) are among the standards included in FMVSS. These standards establish performance criteria and test procedures that every new motor vehicle sold in the United States must meet. Vehicle manufacturers are required to certify that each new vehicle sold meets all of the applicable standards. NHTSA purchases vehicles on the open market and tests them. Should a vehicle fail any aspect of the standard, the manufacturer is required to recall the vehicle and fix the problem. Below, several examples illustrate the importance of CDS data in the rulemaking process.

4.2.1. Problem Identification

Because of the high level of detail on vehicle damage and occupant injuries in CDS, it can be used both clinically and statistically for problem identification. For example, in the course of the 2009 update to FMVSS No.216 (Roof Crush Resistance),[7] CDS data was used to investigate the observed patterns of roof intrusion resulting from real-world rollover crashes and to compare these patterns to the results of the current test procedures. Review of CDS cases showed that damage and intrusion in real-world crashes was considerably greater than the damage and intrusion shown in crash tests. Since compliance testing of several vehicle models found no compliance issues with the standard at the time, these cases provided important evidence for the contention that roof crush standards should be revisited.

The current scope of CDS data (all police-reported crashes in which at least one passenger vehicle is towed due to damage) ensures a wide variety of crashes will be investigated every year. This allows investigation of specific types of crashes (such as rollovers with significant roof damage) without the need for targeted data collection which can be costly and time-consuming. For the purpose of problem identification it is important to have an all-purpose data system that is rich enough with data on subpopulations of interest in such evaluations.

4.2.2. Regulatory Impact Analysis

When drafting a Final Regulatory Impact Analysis (FRIA) both GES and CDS are frequently employed to estimate the size of the population that might be affected by an FMVSS and to estimate any costs and benefits that the identified population might expect to experience.

In the FRIA of FMVSS No. 126, Electronic Stability Control Systems, CDS data were used to estimate the size of the crash population that was likely to be affected by the introduction of ESC. Specifically, it was estimated that there were 448,557 single-vehicle crashes annually that result in at least one occupant in each crash suffering an injury that has a Maximum Abbreviated Injury Score (MAIS) score of 1-5 (minor to critical injuries). Single-vehicle crashes are often the result of loss of control, making them likely to be affected by ESC. CDS estimates of affected population were then combined with estimates of effectiveness to create the benefit estimate.

[7] FMVSS 216 requires that passenger vehicle roofs be able to withstand a certain amount of force, so as to better protect occupants in rollover crashes.

4.2.3. Regulatory Evaluation

Both GES and CDS data are commonly used to evaluate the impact of regulations on real-world crashes and are often the sole sources of relevant crash data. For example, a regulatory evaluation of FMVSS No. 213 (child restraints) was published in 2010 and used CDS and State data to estimate the booster seat effectiveness in injury prevention for child passengers 4 to 8 years old. The analysis found that using booster seats rather than adult shoulder and lap belts resulted in a 14 percent reduction in risk of injury.[8]

4.3. SCI

SCI is able to capture rare or unusual cases that are of particular interest to the agency, but SCI data is not nationally representative like GES and CDS are. SCI data is commonly used internally by the offices of Defects Investigation, Rulemaking, and Research.

NHTSA conducts defect investigations and administers safety recalls to support its mission to improve safety on our Nation's highways. NHTSA is authorized to order manufacturers to recall and repair vehicles or items of motor vehicle equipment when investigations indicate they contain serious safety defects in their design, construction, or performance. NHTSA also monitors the adequacy of manufacturers' recall campaigns. Before initiating an investigation, NHTSA carefully reviews the body of consumer complaints and other available data to determine whether a defect trend may exist.

ODI frequently uses data from SCI and NASS CDS to support its efforts. When available, CDS data can be used to determine if a specific vehicle make/model is showing a defect. More often, ODI will contact the SCI team to request a detailed investigation of a particular case or to request surveillance of a specific crash scenario. SCI has three teams of investigators located across the United States that deploy to the scene of a crash once a case is initiated. Once on scene they conduct an in-depth investigation of the crash. SCI cases have the highest level of detail in any NHTSA data system and these investigations allow researchers to make a decision about whether or not a defect was a contributing factor in a crash.

A recent example of NASS support of ODI is the unintended acceleration investigation that took place in 2010. SCI research staff provided surveillance for cases of unintended acceleration by monitoring media outlets and police accident reports collected by GES. SCI staff also provided 30 case reports of vehicles allegedly showing unintended acceleration. The data provided from these cases was derived from extensive interviews of vehicle operators and eyewitnesses, thorough scene inspections, crash reconstructions, and vehicle response data taken from the vehicles' event data recorders (EDRs). These case reports were an integral component of the final report that NHTSA published on unintended acceleration.[9]

[8] Sivinski, R., (2010, July). *Booster seat effectiveness estimates based on CDS and State data*. (Report No. DOT HS 811 338). Washington, DC: National Highway Traffic Safety Administration.
[9] NHTSA. (2011). *Technical assessment of Toyota electronic throttle control (ETC) systems*. Washington, DC: Author.

5. Comprehensive Review

In seeking to ensure that NASS CDS and NASS GES and SCI best meet the current and future needs of the highway safety community, NHTSA conducted a comprehensive review of the current systems, requesting input from users both internal and external to NHTSA. The three major components of the review were an internal review, a Federal Register Notice, and a Public Listening Session.

5.1. Internal Review

First, interviews were conducted of NHTSA researchers and policy makers to determine how data is currently used and how those data needs were likely to change in the future. The findings from these internal interviews were documented in the Report to Congress: NHTSA's NASS Data Needs.[10] In summary, NHTSA identified the following essential actions:

1. Add new data elements to better develop safety countermeasures, including ones related to crash avoidance and behavioral safety issues.
2. Expand the scope of vehicle types and crash types for greater analytical attention, such as: motorcycles, medium and heavy trucks, motorcoaches, bicyclists, school buses, and low-speed vehicles.
3. Enhance and improve the analysis of crash data, dissemination of data and the public's access to that data, and linkage to other safety information.
4. Obtain a sufficiently large database to identify emerging crash trends and occupant injury trends, including expanding the NASS data collection sites.

5.2. External Outreach

The second step taken by NHTSA was to gather input from its stakeholders – the highway safety community, including the industry, academia, and other government agencies. To reach the broadest possible group of stakeholders, NHTSA published a notice in the Federal Register announcing the survey modernization effort. This notice published on June 21, 2012, reflected NHTSA's intent to upgrade the information technology, research design, data elements, and data collection methods to meet the needs of government agencies, industry and academia in the United States and abroad. The notice solicited comments on the future utility of current data elements, recommendations for additional data elements and attributes, and a description of their anticipated data needs.

NHTSA sent the Federal Register Notice directly to all identified interested parties via e-mail and postal service. Over 525 letters and e-mails were distributed. Comments were submitted by the following.

- AAA Foundation
- Advocates for Highway and Auto Safety
- Association of Global Automakers, Inc.
- Center for Injury Research and Prevention at the Children's Hospital of Philadelphia
- Dynamic Research, Inc.

[10] August 2011, DOT HS 811 889.

- Exponent Failure Analysis Associates
- Ford Motor Company
- General Motors
- International Center for Automotive Medicine, University of Michigan
- Insurance Institute for Highway Safety
- Mercedes-Benz
- Motor and Equipment Manufacturers Association
- National Transportation Safety Board
- Pacific Institute of Research and Evaluation
- Robert Bosch, LLC
- Rubber Manufacturers Association
- Toyota Motor North America, Inc.
- University of Alabama
- University of Colorado – Denver
- University of Iowa
- University of Michigan Transportation Research Institute – general
- University of Michigan Transportation Research Institute – Biosciences Group
- Volkswagen Group of America, Inc.

5.3. Summary of Comments

Twenty-five organizations and individuals submitted comments in response to the Federal Register Notice. Most of the comments came from research organizations, automobile manufacturers, and their associations. Table 1 shows the summary of comments by stakeholder type.

Table 3: Comment Summary

Stakeholder Type	Number of Commenters
Automaker	6
Medical Community	4
Researcher	7
Safety Advocates	4
Supplier	3
(blank)	1
Total	25

The 25 organizations and individuals provided 313 specific comments. As expected, most of the specific comments were related to NASS CDS (292, or 93%). Twelve of the specific comments were related to GES, and the few remaining comments concerned other NHTSA databases or NHTSA data in general.

The most prevalent comments about GES dealt with increasing the availability and accuracy of Vehicle Identification Numbers (VINs). Twenty-five percent of vehicles in the 2010 GES data file have missing, unknown, or erroneous VINs. The other comments related to GES suggested including the crash narrative on the PAR verbatim and adding crash scene diagram from the PAR, as well as adding satellite images of the crash location or latitude and longitude of the crash. There was also interest in linking GES to other NHTSA databases and to other data sources such as driver license files, vehicle registration data, judicial records, citation files, or ZIP Code-based demographic data.

Specific comments for both GES and CDS fell into the common general topic areas (Table 4). The most common topic area was scene, followed by sample, data elements, and vehicle reconstruction.

Table 4: Comment Topic Areas

Topic Area	Number of comments	Topic Area	Number of comments
Scene	45	Manual	5
Sample	33	MMUCC	4
Vehicle Reconstruction	33	Workshop	4
Data Element	31	Child	3
Scope	21	On-Scene	3
Analysis file	17	Special	3
Injury	14	Traffic	3
Photo	14	VIN	3
Quality Control	12	Data Analysis	2
EDR	9	Funding	2
Pre-crash	9	Global Harmonization	2
Viewer	9	Data Element & Analysis File	1
Consistency	8	Scope & Vehicle Reconstruction	1
Crash Avoidance	7	Miscellaneous	8
Link file	7		

Once the comments were grouped into topic area, it was apparent that many of the comments were essentially the same or very similar. For example, of the 45 on-scene comments, only 18 were unique, with commenters requesting latitude and longitude data be added to both GES and CDS in the new system.

Comments that requested the same or similar information were grouped together. Table 5 shows there were 189 unique comments, and data element had the most unique comments.

Table 5: Unique Comments

Topic Area	Number of Comments	Topic Area	Number of Comments
Data Element	31	Manual	4
Vehicle Reconstruction	25	Child	3
Scene	18	VIN	3
Injury	12	On-Scene	2
Sample	12	Sample & Scope	2
Scope	10	Data Element & Analysis File	1
Quality Control	8	Funding	1
Photo	7	Global Harmonization	1
Viewer	7	MMUCC	1
Pre-Crash	6	Scope & Vehicle Reconstruction	1
Consistency	5	Special	1
Crash Avoidance	5	Traffic	1
EDR	5	Workshop	1
Analysis File	4	Miscellaneous	8
Link File	4		

The comments in the data element area were very specific and unique. Most were requests for additional data elements and some requested that we eliminate obsolete technologies such as automatic belts.

Most of the comments in the vehicle reconstruction, scene, pre-crash, and crash avoidance areas focused on adding variables related to crash avoidance technologies, that is, behavioral and vehicle systems that can enhance human performance and vehicle control. Requests also included additional pre-crash information on roadway geometry and environmental conditions, surface conditions, weather, and visibility, along with driver health status, behaviors, cognition, and response. In addition, requests included more detailed documentation of pre-crash maneuvers such as paths of travel, travel speeds, vehicle trajectories, and disposition. If collected, this information must be evaluated in concert with the identification of the availability and the performance of a broad array of crash avoidance technologies. It can be challenging to identify whether a particular vehicle is equipped with emerging technologies, especially when such devices are offered as optional equipment. Some commenters also suggested linking the VIN of a vehicle involved in a selected case with the manufacturer's "build sheet" that lists all of the options and features installed on a given vehicle during the manufacturing process. However, NHTSA has been unable to obtain this information from manufacturers.

Commenters wanted more detailed information on injuries and injury mechanisms, such as specific bones or pelvic region fractured, noting that medical imaging could be used for injury mechanism analysis.

Stakeholders provided a variety of comments related to the sample. Most requested an increase in the annual sample size, redistributed sample sites to better reflect current population distributions, and narrowing the range of national weights. Many wanted more cases with newer vehicles or more cases with moderately severe crashes or severe injuries. Some commenters wanted to increase the amount of VIN, EDR and Delta V (change in velocity) information by preferential selection of data collection sites or police jurisdictions that code VIN, or by selecting alternative cases when the VIN, EDR or Delta V information cannot be obtained. NHTSA is interested in increasing the amount of these data elements; however, some of the techniques suggested could bias or invalidate the sample.

Another significant recommendation from the stakeholders is to broaden the scope of CDS to include crashes involving motorcycles, commercial vehicles, pedestrians, bicycles, and other road users such as all-terrain vehicle (ATV) users. NASS included these crash types until the redesign in 1988. Since then, there have been short-term studies of large-truck crashes and of pedestrian-related crashes within NASS, and more limited studies of bus and of motorcycle crash causation outside of NASS. However, nationally representative crash-investigation-based data on this range of crash types is not currently available.

There were comments submitted by stakeholders about a lack of consistency in case contents and perceived decrease in case quality. Suggestions include having independent case reviews by a multidisciplinary team with expertise in crash investigation, vehicle design, biomechanics, and medicine to help ensure consistency and accuracy across all levels of cases. Users also note that missing or incorrectly coded data should be eliminated through improved efforts, more photographs, and imputation.

A topic that generated many comments was the photographic documentation of case evidence. Many users want the revised CDS to provide more comprehensive photographs using a camera or laser scanning device. Other items stakeholders mentioned as critical and needing to be photographed were the condition and final locations of the vehicles; landscape, roadside features, and traffic control devices; vehicle interiors including under the instrument panel, steering wheel position, seat locations, and positions; and instrumentation as displayed on the dashboards.

Several comments related to improving the quality of the photographs, as they can be poorly framed, out of focus, have poor exposure settings, and lack appropriate close-up images of restraint systems. Other photography recommendations include:

- providing cameras to police agencies to obtain on-scene photos;
- standardizing nomenclature for photos to easily identify specific images, including specifications for height and angles at which pictures are taken; and,

- enhancing photos by using cameras equipped with GPS and compass-enabled features to provide information on the location/distance between photos and the direction in which they were taken.

There were comments by stakeholders requesting improvement to the case viewer, the analysis files, and the users' manuals. Stakeholders requested enhancements to the search function, the ability to download query results as tables, and more functionality when viewing information. Several submissions requested that the analysis data be made available in an open source data format such as ASCII that does not require third-party software to access or read. The current software format presents an impediment for users who do not routinely use or have licenses for SAS software. The commenters recommended that the manuals need to have improved documentation on methodology used to determine national weights and mapping protocol to describe how to map new or changed elements to old data elements.

Several stakeholders agreed that collecting EDR data is important, but not a good substitute for on-scene crash data collection. Users also want the EDR data in an electronically available format other than in portable document format (PDF).

The third major step in the comprehensive review was a Public Listening Session on July 18, 2013, to solicit information and comments on: (a) adding, deleting, or changing the current NASS data elements for data modernization, and (b) recommendations for changing or improving the NASS data collection methodology. The listening session was announced in a Federal Register Notice on June 18, 2013. Interested parties were invited to come to NHTSA headquarters or join through the Internet to provide additional comment not previously submitted in response to the June 21, 2012 Federal Register Notice.

More than 150 participants or stakeholders joined the session in person or via Internet, with 7 stakeholders providing written input during the listening session. Most of the comments heard at the listening session mirrored ones submitted in response to the Federal Register Notice or identified through NHTSA's internal review. The stakeholders want a sufficient number of cases to identify traffic safety issues. Some stakeholders want to identify emerging trends in a timely manner and others want more cases involving children, serious crashes and/or pedestrians. Some stakeholders want to add crash avoidance information to GES and others want to expand types of crashes investigated in CDS. They want improvements to the analytical files available to the public and how the public can view the data. They want detailed crash causation data, injury information, more photographs, scaled scene diagrams, and the use and availability of crash avoidance technology on crash-involved vehicles.

Most of the stakeholders want both GES and CDS data but wanted to add other data. None of the stakeholders suggested we discontinue collecting of variables. However, one stakeholder did suggest replacing GES by using State data to create national estimates of motor vehicle crashes. Another stakeholder stated that without an increase to the NHTSA data collection budget additional outside sources of funding are needed to collect more data or more cases.

Besides soliciting information through the formal methods discussed above, NHTSA has also had many informal conversations and meetings with stakeholders, such as IIHS, Mercedes Benz, and Bosch, where improvements to NASS were discussed.

NHTSA assessed the feasibility of the suggestions received from the internal and external reviews for inclusion in the modernized NASS. Questions considered in the feasibility assessment were:

- Is the requested information currently collected?
- What is the best way to collect the information requested?
- What is the estimated time to collect?
- What is the estimated cost?
- What is the accuracy of the data?
- What is the utility of the data to all users?
- Can the information be released to the public?

Many of the system improvements requested are feasible and could be collected or achieved in the new system. Since we are redesigning the sample, most of the improvements to the sample will be achieved. Data that is currently collected but not available to the public could be made available to the public. Data elements that are relatively easy to collect, are not time consuming and/or do not require expensive equipment, could easily be implemented in the new system. The new system will be built with the flexibility to add and modify data elements. However, the operating budget will ultimately determine whether or not the requested data or improvements can be implemented in the new system.

5.4. Summary of Review

The internal NHTSA interviews and the two external outreach efforts did identify key areas for improvements:

1. Add new data elements to better focus on crash avoidance technologies, and behavioral and vehicle systems that can enhance human performance and vehicle control.
2. Expand the scope of CDS to include crashes involving motorcycles, medium and heavy trucks, pedestrians, and bicycles.
3. Improve the NASS sample design by (a) obtaining a sufficiently large database to identify emerging crash trends and occupant injury trends,(b) reselect sample sites to better reflect current population, and (c) narrow the range of national weights.
4. Enhance and improve the analysis of crash data, dissemination of data, and the public's access to that data and linkage to other safety information.

The next section explores in more depth expanding NASS in some of the key areas identified in the review.

6. Expanding NASS Analysis

Given the input from the review process, NHTSA recommends expanding the types of crashes covered by the new system and the data elements collected on these crashes. Doing so will address two key areas identified by the agency and its stakeholders as not adequately covered in the current NASS: crash avoidance, and a crash involving a motorcycle, a pedestrian, or a large truck. In addition, as many stakeholders requested, NHTSA recommends increasing the size of the NASS samples so as to improve the ability of the data from the system to support most data needs. By expanding NASS in these three areas the utility of NASS data for analyses is enhanced. Each analytical area is described below.

6.1. Crash Avoidance

Perhaps the greatest need for new analyses comes in the area of crash avoidance. New crash avoidance technology such as Forward Collision Warning, Lane Departure Warning, and other systems are designed to alert a driver to an impending crash threat or take actions such as automated or supplemental braking. NHTSA needs to assess the effectiveness of these crash avoidance systems, be prepared to investigate allegations of system malfunctions, and provide high-quality data to allow system designers to develop future systems.

Currently NASS collects only limited information on the circumstances leading up to a crash, and does so long after the crash has been cleared, when key information (such as skid marks) may no longer be evident.[11]

NHTSA's National Motor Vehicle Crash Causation Survey (NMVCCS) demonstrated that key pre-crash evidence can be obtained by collecting data while vehicles and involved people are still on the scenes of the crashes. After a crash has been cleared, i.e., the vehicle and involved people leave the scene, evidence can quickly alter in nature or disappear entirely.[12] Thus it is important to collect such information before the crash has cleared, if at all possible. Crash data collected before the crash scene has been cleared is called "on-scene." Data collected subsequently (e.g., visiting the scene and/or the tow yards where the vehicles were towed to a few days after the crash) are referred to as "follow-on" data collection. NMVCCS has provided a rich source of information for better understanding key factors in crash occurrence and the potential for technology and behavioral change to reduce or mitigate crashes. However, NMVCCS was only conducted during a three-year period (2005 to 2007), thus presenting only a limited number of cases, and does not reflect any post-2007 technologies such as Forward Collision Warning and Lane Departure Warning, nor does it reflect any vehicle-to-vehicle (V2V) technologies or vehicle-to-infrastructure (V2I) technologies currently under development. To continue

[11] For instance, GES records whether a driver took an evasive action prior to impact (such as braking, or steering to avoid) if such information is noted on the PAR. CDS supplements such information with interviews of involved parties, and vehicle and scene inspections. SCI collects various information, depending on the nature of the topic being investigated. As with CDS, interviews and inspections are conducted after the crash has been cleared.

[12] EDRs will overwrite data recorded from a crash with subsequent events. Witness recollections diminish with time, and at-fault drivers may be less forthcoming after conducing activities (such as speaking with their insurance companies and possibly lawyers) that prompt concerns for their potential liability. Skid marks can lessen with the wear of subsequent vehicles traveling over them.

to fulfill NHTSA's responsibility to assess the safety and efficacy of crash avoidance technologies, NHTSA will need to collect more crash avoidance data in some manner[13] going forward.

The need to collect crash avoidance data was noted by several data users outside of NHTSA in the response to the Federal Register Notice seeking comment on the future NASS. Respondents in this area ran the spectrum of stakeholders, from automakers and suppliers to safety advocates and researchers.

Multiple respondents noted the importance of NHTSA providing high-quality data that would allow system developers to identify new arenas for crash avoidance technology as well as to improve their current systems. Volkswagen, Mercedes Benz, MEMA,[14] IIHS,[15] the NTSB,[16] and others cited collecting data on whether crash-involved vehicles were equipped with systems such as Electronic Stability Control, Crash Imminent Braking, Active Brake Assist, and Forward Collision Warning, and how these systems performed.[17] Others, including Toyota, Mercedes Benz, DRI,[18] and Global Automakers mentioned collecting quality data on pre-crash conditions that would help in designing and identifying the life-saving potential for the systems of tomorrow.

Driver distraction is one particular "pre-crash condition" that generated multiple comments and for which technology may offer countermeasures. Toyota, for example, would like more detailed information on the nature of the distraction, such as whether a driver was changing the radio station, texting, interacting with some other technology, or just not paying particular attention to driving. The NTSB conveyed a particular desire to note the presence of manufacturer-installed communication, entertainment, and navigation systems in crash-involved vehicles. Other respondents, such as Advocates,[19] expressed the general concern to collect quality data on the continued threat to traffic safety posed by driver distraction in general.

Respondents also noted the need to assess whether crash avoidance systems work properly.[20] IIHS would like to add questions to the interviews of crash-involved drivers whose vehicles had such systems to better ascertain the status of the system prior to the crash. The NTSB cited a need to evaluate the effects of crash avoidance systems, and AAA expressed a general need to have on-scene data.

---

[13] Recognizing that collecting on-scene data is expensive, NHTSA is designing the new system to address pre-crash data under different paradigms, each reflecting different costs and quality.

[14] The Motor and Equipment Manufacturer Association

[15] At the same time, IIHS advocated that on-scene data not replace the collection of crashworthiness data, as crash avoidance systems have not displaced the need to ensure the continued crashworthiness of vehicles. NHTSA concurs.

[16] National Transportation Safety Board

[17] Although NHTSA generally will not be able to determine with certainty in a given crash whether a particular system sounded a warning, supplemented driver braking, or executed another function designed to avoid or mitigate a crash, we can gather a variety of information from driver interviews, EDRs, and vehicle and scene inspections that can inform whether and how well a system performed.

[18] Dynamic Research, Inc.

[19] The Advocates for Highway and Auto Safety

[20] We expect that both on-scene studies and SCI (whose data rarely is collected before a crash is cleared) will provide key tools to assess the proper functioning of safety technologies.

NHTSA is addressing the need for more pre-crash and crash avoidance data in the new system in two ways: first, by enhancing the data captured a few days after the crash (in the follow-on investigation) to include available crash avoidance data; and second, by adding additional studies (modules) to NASS that focus on collecting high quality pre-crash data collected before the crash scenes have been cleared (through on-scene investigations). The on-scene studies are discussed in more detail in the next section.

In the follow-on investigations, NHTSA plans to collect and release information on whether vehicles were equipped with crash avoidance equipment, such as,

- Lane Departure Warning with Lane Keeping,
- Lane Departure Warning without Lane Keeping,
- Forward Collision Warning with Auto Braking,
- Forward Collision Warning without Auto Braking,
- Blind Spot Detection,
- Daytime Running Lights,
- Assisted Braking, and
- Automatic Crash Notification.

In addition, if through interviews we can determine if the technology worked or did not work, that information would be recorded. Additional improvements to data collection are discussed in Chapter 7. With more pre-crash and crash avoidance data collected through an enhanced follow-on and/or new on-scene study, NHTSA will[21] be better prepared to answer critical questions like the following:

- What is the effectiveness of a crash avoidance technology, such as Lane Departure Warning? How many crashes and injuries have been or might be avoided or mitigated, and at what societal cost savings?
- What is the potential benefit for V2V and V2I technologies (in terms of crashes and injures avoided or mitigated)?
- How much does distracted driving contribute to crash occurrence and severity?
- What other key factors contribute to crash occurrence and severity that could be used to design the next life-saving technology?

6.2. Motorcycle, Medium/Heavy Truck, and Pedestrian Crashes

With the recent substantial reductions in crash fatalities involving only passenger vehicles (cars, light trucks and vans [LTVs]), the proportion of fatal crashes involving a motorcycle, medium/heavy truck, or pedestrian has increased. In 2000, there were 27,784 fatalities in crashes involving exclusively cars and/or LTVs; that number declined dramatically to 18,553 by 2011. On the other hand, in 2000, there were 13,832 fatalities in crashes involving at least one motorcycle, medium/heavy truck, pedestrian, and/or other road user; the number of fatalities in these crashes also decreased, but less so, to 13,377 in 2011. In other words, in 2000 only 33 percent of the fatalities occurred in crashes that involved at least one motorcycle, medium/heavy truck, pedestrian, and/or other road user but that proportion increased to more than 40 percent in 2011.

[21] Subject, of course, to adequate funding.

New regulations or safety programs to reduce the occurrence or to mitigate the outcomes of crashes involving vehicles or road users other than passenger vehicles will require detailed crash data that cannot be obtained from NHTSA's current data systems. Only by expanding the scope to collect detailed information on such crashes can NHTSA address this need.

Public comment also supports this effort, with automakers, suppliers, safety advocates, researchers, and the other stakeholders voicing their desire for such data to be collected. DRI and the NTSB were among those advocating detailed crash data for motorcycles, while DRI, NTSB, MEMA, and IIHS specifically mentioned medium or heavy trucks. IIHS, AAA, NTSB, Toyota, Mercedes, PIRE, and Volkswagen mentioned pedestrians or pedalcyclists.

Consequently, NHTSA plans to accomplish this through the use of the additional studies or modules mentioned previously. NHTSA is proposing that the following on-scene data collection modules conducted periodically that will collect pre-crash data (like in the NMVCCS) and crash outcome data.

- *On-Scene Passenger Vehicle Module*: A data system of detailed investigations of passenger vehicle crashes that collects data before crash scenes have been cleared.

- *On-Scene Motorcycle Module*: A data system of detailed investigations of motorcycle crashes that collects data before crash scenes have been cleared.

- *On-Scene Pedestrian Module*: A data system of detailed investigations of crashes involving pedestrians or pedalcyclists that collects data before crash scenes have been cleared.

- *On-Scene Large-Truck Module*: A data system of detailed investigations of trucks weighing over 10,000 pounds GVWR that collects data before crash scenes have been cleared.

The types of highly specialized information needed on crashes involving motorcycles, pedestrians, or heavy trucks, combined with the fact that the crashes are concentrated in different areas,[22] make it substantially more efficient to collect the data in individual studies, separate from the GES and CDS successor components, the Crash Report Sampling System (CRSS) and the Crash Investigation Sampling System (CISS). In motorcycle, pedestrian, or heavy truck crashes the in-depth data needed often are specific to these vehicle types and include data that must be collected on-scene. Collection of these types of data falls outside the scope of CRSS and CISS data collection methods. Additionally, using the same data collection staff to gather all of these data poses risks to data quality. Thus, it makes more sense to collect these data under separate motorcycle, pedestrian, and truck data collections.

Through their detailed investigations, NHTSA will be able to answer, or answer better, critical questions like:

- How effective are technologies designed to avoid or mitigate pedestrian crashes?
- How effective are motorcycle antilock braking systems (ABS)?

[22] For instance, motorcycles are more common in warmer climates, while pedestrians are more common in cities.

- What are the key factors in motorcycle crash occurrence and whether a motorcyclist sustains a serious injury?
- How effective are crash avoidance and mitigation technologies in medium and heavy trucks?

6.3. Improved Analysis from Increasing the Sample Size

In general, increasing the number of crashes sampled in NASS will fortify the conclusions that can be reached. Although this assertion holds true throughout NASS, the greatest potential lies in increasing the sample sizes in the investigation-based CISS because it has a much smaller sample size than CRSS.

As indicated in Section 4 Current Usage of NASS, GES, CDS and SCI currently support many important analyses, and each of these could be answered more substantively and definitively if they are based on information from more crashes in the new system. [23] These include critical questions like:

- Is the number of motorcycle crashes increasing or decreasing and is the change significant?
- How effective (in crashes and injuries avoided or mitigated) are crash avoidance and mitigation technologies? If a new technology, do the benefits warrant mandating its incorporation in new vehicles, or adding it to the NCAP five-star rating program to better inform consumers? If a previously regulated technology, how effective was the regulation?
- Should a test procedure in the FMVSS be altered in some way to better reflect real-world crashes (such as NHTSA did with the roof crush standard)?
- Does an alleged motor vehicle defect exist? Although SCI provided only one piece of the unintended acceleration investigation (together with analysis of consumer complaints and other tools), it provided a critical piece. With greater resources for the SCI program, it is possible that NHTSA might have been able to resolve this issue of great concern to the driving public sooner.[24]

As discussed in Section 5, external stakeholders want NASS's sample size increased in order to provide a more substantive basis with which to develop and assess traffic safety countermeasures. The stakeholders who advocated increasing NASS's sample size include GM, Ford, AAA, NTSB, Advocates,

[23] NHTSA has designed the CRSS and the CISS in a way to maximize the precision of key estimates for a given operational cost. Such an optimization requires gathering and analyzing large amounts of data, and NHTSA is not yet in a position to quantify the impact of increased sample size on precision. SCI is a different matter, as it serves a different purpose. Serving the need to gather data quickly on time-relevant topics and potential emerging threats, increasing its caseload would generally benefit NHTSA's ability to draw appropriate, substantiated conclusions in timely manner. However, in order to provide more information quickly, it is currently, and is likely to remain, a "surveillance" system in nature, responding to as many relevant cases as it can acquire in a timely manner. As such, its data are not from a probability sample, and its precision is not quantifiable. Thus while increasing SCI's caseload is an important consideration in order to meet time-relevant topics and emerging threats, NHTSA cannot quantify the impact of additional SCI cases on the ability to substantiate conclusions.

[24] Increasing the caseload in CISS would in all likelihood better inform only some defects investigations, probabilistic systems like CISS are generally not effective vehicles for this purpose. Motor vehicle defects are typically (and thankfully) rare and unlikely to arise in appreciable number through probabilistic sampling, A multiplicative increase in CISS's operating budget would be needed to capture enough cases for this purpose, and we do not recommend this as a prudent use of funds.

UMTRI – Bioscience, ICAM,[25] and IIHS. NHTSA recommends sample sizes for the new NASS in the upcoming Section 8.

Increasing the number of data collection sites in NASS will boost the precision of the estimates. Over the years, the crash and injury prevention techniques have advanced dramatically. The high interest cases (crashes with severe injuries or crashes involving new model year vehicles) are becoming rarer or inherently rare. A small number of data collection sites either cannot produce enough high interest cases or does not reduce the overall error in the estimates enough to truly identify and provide reliable estimates that describe an issue. The most effective way to reduce the total sampling error is to increase the data collection sites. NHTSA identifies the number of data collection sites in the Section 8.

[25] The International Center for Automotive Medicine, at the University of Michigan

7. Future System Design

NHTSA is designing a system that is flexible and scalable to accommodate many different data needs and fiscal resources. This new system will affirm NHTSA's position as the leader in motor vehicle crash data collection and analysis, by collecting quality data to keep pace with emerging technology and evolving policy needs. This plan reflects our current thinking.

7.1. Analytic Objectives of the Future System

Beyond providing broad indicators of traffic safety, the original NASS was designed to address the most important traffic safety issue of its time – how well passenger vehicles protect their occupants in crashes (and thus inform possible protective improvements). Times have changed. Clearly, the new system needs to be designed to accommodate issues of crash avoidance and an expanded range of vehicle types. To address this, NHTSA is designing the new system to meet the following analytical objectives.

Table 6: Analytic Objectives

A1: To assess the state of traffic safety on an annual basis. Key traffic safety indicators include: the number of police-reported crashes nationwide, and the number of such crashes that involve "serious" injuries.[26]
A2: To assess the effectiveness of crash avoidance countermeasures for light vehicles, non-light vehicles, and motorcycles.
A3: To identify the key factors increasing crash risk in light vehicles, non-light vehicles, and motorcycles.
A4: To identify the key factors in whether an occupant, motorcyclist, pedestrian, or pedalcyclist sustains a "serious" injury in a crash.
A5: To assess the effectiveness of crash mitigation countermeasures for occupants, motorcyclists, pedestrians, and pedalcyclists.

These objectives are in NHTSA's view the most important the agency faces in the future that can be informed by crash data. The primacy of these questions is confirmed by NHTSA's stakeholders, with nearly all stakeholder needs being related to one or more these objectives.

Although these objectives address the spectrum of needs, it may be that a particular, highly specialized need might not "fit" under any of these objectives. Such a need might still happen to be met in the new system, although NHTSA did not design the system with it in mind. Although we would like to address every stakeholder's wish, no matter how specialized, this is clearly not practicable.

[26] The particular threshold constituting a "serious" injury can be different for different purposes. For the purpose of broad indicators, we mean a fatal or incapacitating injury.

Identifying and then investigating alleged motor vehicle defects is a key component of NHTSA's mission. Although primarily informed by other data sources, such as consumer complaints,[27] NASS and SCI have historically contributed to NHTSA's defect investigations, and the new system will better support this critical need.

7.2. The Components of the Future System

Many aspects of the NASS's current data collection system will continue to be needed in the future. However, NHTSA's and our stakeholder's data needs highlight some weaknesses in the current system: limitations in our current sample of data collection sites, not enough pre-crash and crash avoidance data, and relatively few detailed investigations of motorcycle, medium/heavy truck and pedestrian crashes. The new system will combine the aspects of the current system that will be needed in the future with changes to address the additional needs not met with the current system. As described in earlier sections, the new system will be composed of multiple components, or modules, to accommodate these needs.

The two major components of the new system will be a records-based data collection module similar to the current GES – the Crash Report Sampling System – and an investigation-based module similar to current CDS – the Crash Investigation Sampling System. CRSS will continue to provide the annual, nationally representative estimates of the number of police-reported motor vehicle crashes overall. In addition, CRSS will provide estimates by type of vehicle, and for a broad range of vehicle and crash characteristics that are needed to fully describe current highway safety and to trace motor vehicle crash trends.

CISS will collect accurate, detailed information about a nationally representative selection of passenger vehicle crashes that involve a passenger vehicle towed from the crash scene. Researchers will investigate crashes a few days after the crash (a "follow-on" investigation as described in earlier sections) gathering information from a variety of sources: crash site inspection, vehicle inspections, interviews, medical records and others. CISS will have enhanced pre-crash data and data on the presence and use of crash avoidance technologies.

CRSS and CISS will be designed to reflect characteristics of the current population and will be sufficiently large to identify emerging crash trends with improved precision. Both new samples will also oversample severe crashes and crashes involving newer model year vehicles such that the sample will include more of these types of crashes than would be expected if crashes were simply randomly selected. The CRSS and CISS samples were selected separately (the CISS data collection sites are not a subsample of the CRSS sites or vice versa) to better represent data users' needs. Reselection of the data collection sites in both samples will significantly improve the precision and relevance of the crash data.

In addition, as outlined in Section 6, NHTSA plans to conduct special studies periodically. These special studies will collect detailed data on motorcycle, medium/heavy truck or pedestrian crashes similar to past studies like the National Motor Vehicle Crash Causation Survey (NMVCCS) and the Large-Truck

[27] The public can report safety defects and other safety complaints to NHTSA at www.safercar.gov. These complaint data are also called Vehicle Owner Questionnaires.

Crash Causation Study (LTCCS). In these studies, the researchers gather time-sensitive information while the vehicles are at the crash scene (known as "on-scene" investigations). This study design results in more accurate pre-crash, crash causation, and crash avoidance data. These studies, an example of which is the On-Scene Passenger Vehicle Module, will also collect outcome data similar to the LTCCS.

A sufficient number of motorcycle, medium/heavy truck or pedestrian involved crashes will be collected in CRSS to track trends. However, crashes involving a motorcycle or pedestrian are unlikely to be found routinely in CISS. This is because crashes involving a passenger vehicle and motorcycle or pedestrian are unlikely to result in enough damage to the passenger vehicle to cause it to be towed from the scene. On the other hand, crashes between a passenger vehicle and medium/heavy truck are likely to result in a towed passenger vehicle and may be selected in the CISS. However, these investigations may lack information on the truck because the truck is unlikely to be damaged and thus hard to find a few days after the crash. Consequently for detailed assessments of these vehicles, the new system includes specialized modules that conduct investigations on just these vehicle types. All together, the "pieces" of the new system are shown in Table 7.

Table 7: Recommended Future System Components

Module	*Purpose*	*Analytic Objectives Addressed*	*Data Collection Sources*
Crash Report Sampling System	To provide a broad assessment of traffic safety and support the collection of NiTS data.[28]	A1	PAR only
Crash Investigation Sampling System	To inform crash mitigation and crash avoidance in passenger vehicles.	A2 – A5 for light vehicles	PAR + follow-on
On-Scene Passenger Vehicle Module	To better inform crash mitigation and avoidance in passenger vehicles by collecting data before crash scenes are cleared.	A2 – A5 for light vehicles	PAR + on-scene + follow-on
On-Scene Motorcycle Module	To better support crash mitigation and avoidance in motorcycles by collecting data before crash scenes are cleared.	A2 – A5 for motorcycles	PAR + on-scene + follow-on
On-Scene Pedestrian Module	To better support crash mitigation and avoidance for crashes involving a pedestrian or pedalcyclist by collecting data before crash scenes are cleared.	A2 – A5 for crashes with pedestrians and pedalcyclists	PAR + on-scene + follow-on
On-Scene Large-Truck Module	To support crash mitigation and avoidance in buses and trucks over 10,000 lbs GVWR.	A2 – A5 for buses and trucks over 10,000 lbs GVWR	PAR + on-scene + follow-on
Special Crash Investigations	To support crash mitigation and avoidance, defects investigations, and address needs not met by other components.	A2 – A5	PAR +follow-on

[28] Not in Traffic Surveillance's (NiTS) crash data is currently collected through GES, and would be collected under CRSS in the new system.

Conducting all seven studies on a continuous basis would be an expensive proposition. In NHTSA's view, CRSS and CISS need to continue as ongoing data collections, but the other modules are not needed every year. SCI, separate from CRSS and CISS, would continue on an ongoing basis.

CRSS needs to collect crash data on a continuous basis, as this PAR-based system forms the basis of our annual assessments of traffic safety (analytical objective A1). A continual CISS, with its detailed assessments of injuries and vehicle damage, provides a strong foundation for the crash avoidance and mitigation objectives articulated in A2-A5. A continual SCI is needed to support all NHTSA offices and address emerging issues in a near real-time manner.

The additional modules address gaps not covered by CRSS, CISS, and SCI. NHTSA proposes that these studies be carried out periodically.

Table 8: How Often Data Is Collected[29]

Component	*2016*	*2017*	*2018*	*2019*	*2020*	*2021*	*2022*	*2023*	*2024*
Crash Report Sampling System	■	■	■	■	■	■	■	■	■
Crash Investigation Sampling System	■	■	■	■	■	■	■	■	■
On-Scene Passenger Vehicle Module				■	■	■			
On-Scene Pedestrian Module				■	■	■			
On-Scene Motorcycle Module							■	■	■
On-Scene Large Truck Module							■	■	■
SCI	■	■	■	■	■	■	■	■	■

Table 8 shows CRSS and CISS data being collected every year, while the remaining studies are conducted periodically. For example, according to a "3-year-on, 3-year-off" schedule, the on-scene passenger vehicle and pedestrian modules are conducted simultaneously for 3 years followed by 3 years for the motorcycle and large truck modules), with 6-month periods in which collection is suspended in order to re-train for the next pair of studies.

7.3. Leveraging Technology for an Improved Data System

Perhaps the biggest opportunity for data system improvement comes from taking advantage of the substantial changes in technology that have occurred since NASS began over 30 years ago. The new system modernizes and consolidates multiple legacy and disparate information technology systems, using a combination of mobile and Web-based technologies to improve security and enhance data quality and overall reliability of NHTSA's crash data. Outdated server technology is being replaced with an integrated system that uses virtualization and more cost effective cloud-based technology. The new

[29] These reflect just one possible scenario, but are subject to operational budgets and identified priorities.

system complies with the Federal Data Center Consolidation Initiative and provides a centralized web-based platform for sample collection, case coding and quality control operations. Other specific technological improvement features of the new system – EDMI, improved EDR data collection, tablet computers, and the use of electronic PARs and virtual offices – will be discussed in the sections that follow.

7.3.1. Electronic Distance Measuring Instruments (EDMI) and Crash Reconstruction Software

Crash investigation involves taking lots of measurements, such as the length of skid marks, the depth of a dent in a bumper, and the extent of crush in a car's crumple zone. In the current system, crash investigators measure skid marks, and other key scene evidence by walking its length with a distance measuring wheel like those that used to be used by land surveyors.

Besides offering less accuracy than our data users need or would like for their analyses, this measurement method poses a risk to the crash investigation staff, as they must walk in the road, repeatedly, to collect data. Similar problems with measurement precision are encountered in the current method for assessing dents in vehicles and crush in crumple zones, which our investigators currently tackle with measuring tape.

Land surveyors, who face similar challenges in collecting multiple measurements of distance and depth, have long ago turned to technology. Electronic distance measuring instruments (EDMI) take digital "pictures" of a scene, producing point-to-point measures between objects using laser technology.

With such a device, our crash investigators can record all of the crash scene measurements our users clamor for without ever entering the roadway. This technology also saves time, taking these measurements in a fraction of the time required by the current methods. Plus, the measurements are highly and quantifiably accurate, measuring distances to within 0.06 inches from up to 4,900 feet away.

Having detailed and scalable scene diagrams was one of the most frequently expressed desires of our data users. By pairing the EDMI with software that automatically imports the distance measurements and draws the scene (known as "crash reconstruction software"), we will meet these needs.

As with scene measurements, EDMI also digitally measures crush on damaged vehicles, more quickly and accurately than our current manual methods do. Moreover, the same crash reconstruction software produces a digital, rotatable image of the vehicle.

Scene and vehicle measurements from real-world crashes play a large role in today's technological development cycle. Automakers and system designers feed the measurements from hundreds of crashes as input into sophisticated computer simulations to develop, test, and refine prototypes. Using EDMI and crash reconstruction software in the new NASS will give designers the cases they need to do their life-saving work.

7.3.2. Event Data Recorders

Most investigative tools involve some element of art as well as science, such as interpreting skid marks. Event data recorders (EDRs), commonly referred to in the lay world as automotive "black boxes," offer completely scientific measurements of key elements in traffic safety analysis: the force of an impact, vehicle speed, whether restraint systems were activated, and whether brakes were used.

Vehicles are not required to have EDRs, but most vehicles have them. NHTSA requires that, as of September 2012, data collected in EDRs follows a standardized format (CFR 49 Part 563).

Since the early 1990's NHTSA has collected some form of EDR data. This EDR data is typically stored in the air bag control module due to the processes it goes through to determine if air bag or pretensioner systems should deploy.

Although they contain such valuable information, the EDR data cannot be "read" without using a specialized tool. Beginning in the early 2000s commercially available tools were produced to access EDR data in certain model vehicles.

An example of this tool is Bosch's Crash Data Retrieval (CDR) tool, which is what NHTSA has used, and plans to continue to use, to access EDR data. When attached to the control module housing the EDR, the CDR "reads" the EDR data and outputs it to a file that can then be more easily read and stored. Although not every EDR can be "read" with the Bosch tool (or any tool for that matter), the Bosch tool supports a large range of makes and models.

A complete CDR kit generally costs about $10,000, plus about $1,000 per user per year for software and updates. Each new make and/or model that is released also requires a new cable, which costs on averages about $300. To minimize costs over the years, our NASS program only supported one complete CDR kit per PSU. Although NHTSA collected EDR information in the NASS CDS system, crash technicians are not fully equipped to handle all the newly supported vehicles.

Because EDR data is so vital to understanding crashes, NHTSA plans to equip all CISS crash technicians with the CDR tool hardware and software. This decision will not only better equip the crash technicians to collect EDR data more efficiently, we will be collecting more data overall.

The future of crash data collection relies on reliable data collected before, during and immediately after the crash. EDRs provide this information and are being used to understand the crash performance of advanced safety technologies in vehicles. As more vehicles are supported each year, our crash data collection systems should be equipped to record data from as many EDRs as possible.

7.3.3. Using Tablets to Collect CISS Data

The introduction of paperless technologies within the field data collection and investigative processes will allow for data to be collected quickly and accurately. One obvious improvement is the removal of duplicative data entry steps required in the current system that will be removed. In the current system an investigator fills out a paper form in the field and then manually enters the data into the system after returning to the office. Tablet computers in the new platform will be supported with an electronic form technology that will allow investigators to use a combination of inputs to increase accuracy and efficiencies in the field. The use of pen-based input is expected to enable users to add information to a form while standing next to the vehicle, thus making the process of capturing "real time" information more possible. Forms will be enhanced with standard drop-down, formatted fields and basic error handling to allow data to be captured more accurately at the location, versus having to fix common data errors when the investigator returns to the office.

In addition, linking of data extracts from EDRs and other supporting data from data capture devices will be done within the forms while the investigator is on the scene. These improvements are expected to remove duplicative steps that are currently in place for both at the scene and office locations.

The use of electronic data capture processes through tablets and supporting software will allow data to be transmitted through a synchronization process to a centrally managed server and database platform. The tablet-based data capture will also provide improved data security, as paper forms will no longer need to be carried to and from locations, and stored at multiple office locations. The use of tablet technologies and a mobile architecture will allow for centralization of field data processing to a single virtualized platform. The accessibility of the data will also be improved through centralization. For example, case data could be shared between investigative teams while in process to gain perspective and ideas on enhancing accurate interpretation of results. In the current system this cannot be done effectively due to the case data being stored physically in separate field server locations prior to completion.

7.3.4. Electronic PARs and Files

States are increasingly turning to electronic crash reports (e-PARs) to save time and money, to improve the organization and access of its records, and to better facilitate analyses. In our experience, an e-PAR usually takes the form of a PDF file created by scanning a paper document filled out by a responding police officer. Currently NASS staff access PARs both through electronic means[30] and in paper form, depending on the options available in and preferable to the sampled jurisdictions.[31]

Anticipating a greater shift towards e-PARs and wanting to be prepared to leverage the added efficiency afforded by this shift, NHTSA has partnered with a police jurisdiction to test our ability to capture NASS data elements from PDF crash files. The jurisdiction supplied us with a secure electronic file of crashes that includes the narrative and diagram.[32] From this file, we were able to successfully code some of the standard data elements automatically into NASS data elements.

The conversion from paper to electronic PARs is currently in a transition period, with some jurisdictions using e-PARs, and some still using paper PARs, and the ones that use e-PARS do so in different formats. We will continue to ready ourselves for greater use of e-PARs, and expect to eventually see cost savings

---

[30] Some electronic access is conducted on-site (e.g., in accessing an internal database through using a jurisdiction's computer) and some is conducted remotely (e.g., by logging onto a website from a remote location).

[31] At present, the transition of a State or police jurisdiction to some form of electronic PAR can result in additional cost or delays in access. In addition to the types of time delays that frequently accompany a major structural transition, some States charge a per-PAR fee for NHTSA to collect e-PARs, and some institute of layers of review concurrently with the transition that result in an increased "lag time" between when the crash occurs and when it is made available to us for sampling. NHTSA is hopeful that we can work with States to reduce such obstacles, but how, when, and at what cost States choose to provide their PARs to NHTSA (if they choose to provide them at all) is up to them.

[32] The narrative and diagram are needed to determine if the crash should be included in the sample and to code some data elements such as pre-crash movements which are not commonly found on PARs but can be gleaned from the narrative and diagram.

and improved data quality as a result. In the meantime, we will continue to work with our law enforcement partners to collect PARs in whatever format (paper or electronic) can be most propitiously arranged.

7.3.5. Virtual Offices and Field Personnel

When NASS was designed 30 years ago, the idea of a "virtual office," in which a NASS field staff member could work from their home or another location, did not exist. Over the years, NHTSA has reduced field personnel and incorporated virtual offices, particularly in GES, and we plan even greater use of this cost savings measure in the new design.

Currently most GES sites are operated without any field staff. A centralized coding center can access the sampled jurisdiction's PARs electronically, usually in a PDF, eliminating the need to supply computer equipment or pay for travel time to and from home/office to sampled jurisdictions. Some field staff work part-time from their homes and other sites are close to CDS sites where the field staff share GES and CDS responsibilities in brick-and-mortar office space.

In CRSS, NHTSA thinks that more States, counties, and police jurisdictions will move toward e-PARs or electronic files so CRSS field staff will not be needed. In places where field staff are needed, they will use virtual offices.

Currently, all CDS sites have brick-and-mortar offices. Since in in both CDS and the future CISS, researchers go to the crash scene and tow yards for vehicle inspections in the sampled jurisdiction, field staff are and will be needed. However, NHTSA recommends using virtual offices whenever it is feasible and cost effective.

8. How Big Should the New System Be?

Meeting data needs with a crash system expected to last years, if not decades, will be a challenge. Our stakeholders have disparate needs, often concerning types of crashes that occur only rarely, and these are just the foreseeable needs. We will also face novel challenges involving technologies not yet developed, potential new sources of driver distraction, and unforeseen potential defects in automotive design.

To be blunt, there is no scientifically determined optimal figure for the number of cases or the number of sites in designing a system like NASS. Mathematics can only identify optimal solutions in the presence of completely specified constraints, that is, if we could articulate exactly the questions we wish to answer and the precision we require in answering them. However, neither is the case here.

Lacking the specificity needed to forge a mathematical solution to the question of what is optimal, we have determined a figure that represents our judgment of what is best, based on our historical experience and best anticipation of the future.

8.1. The Number of Crash Investigations in CISS – A "Rare Event" Approach

We envision CISS as primarily supporting the following analytical needs:

- provide nationally representative data with which to estimate the lives saved and injuries mitigated by proposed vehicle safety regulations;
- supply detailed crash data that automakers and others can use to design new crashworthiness technologies and improve crash protection;
- assess the real-world performance of crashworthiness technologies and effectiveness of crashworthiness regulations;
- supply certain basic information supporting crash avoidance and causation research;
- support NHTSA's Office Of Defects Investigations; and,[33]
- help identify emerging vehicle safety issues.[34]

In our view, the first of these (concerning vehicle regulations) requires that the crashes we investigate in CISS, as in the current CDS, be selected through probability sampling, in order that the estimates derived from them be nationally representative. The challenge then, and the driving element in determining the number of investigations, is to ensure that this selection mechanism generates adequate numbers of cases of various types to support the analytical needs.

In large part due to the sizable advances in traffic safety in recent decades, this challenge is not easily met. Many of the analytical needs involve types of crashes that occur only rarely. For instance, the

[33] Safety defects are virtually never anticipated in either their nature or in the makes, models, and model years in which they occur. Although we expect the crash data in the new system to have some utility for defects investigations, we do not recommend designing the type of massive and expensive data system that would be required to support defects investigations to an ideal degree. Instead, we recommend that the primary support for defects investigations continue to come from SCI and consumer complaint data.

[34] Likewise, we expect the new system to help identify some emerging threats, but a massive system to detect unforeseen threats is clearly prohibitive.

agency seeks to improve the protection to children in side impact crashes, which although they accounted for the death or injury of an estimated number of 6,500 children under age 15 in 2011, account for only 0.1 percent of all crashes. In the same year, 16,000 people were ejected from their vehicles in crashes, but only 0.3 percent of crashes involved ejections. Lithium ion battery safety, both for crash victims and first responders, is an agency priority, but only 1 percent of crashes in 2011 involved hybrid or electric vehicles.[35] Although some recent advances in automotive design and technology concern crash types that occur in large numbers (such as Forward Collision Avoidance and Crash Imminent Braking), automakers also look to our data in designing technologies that address less frequent crash types.

Selecting crashes at random would yield few cases of these rare crash types, absent high cost. For example, a random sample of even 50,000 crashes would only be expected to involve 50 children killed or injured in side-impact crashes, and 150 involving people ejected from vehicles. However, through design efficiencies, we expect to be able to capture adequate numbers of cases for most data needs with far fewer investigations. By designing a sample around key crashes of interest, user needs should be supported with a smaller number of investigations than a simple random sample.

8.2. Focus Areas for Crash Investigations

Crashes involving newer passenger vehicles

Crashes involving serious injury

Two themes that emerge from the analysis needs of the agency and our stakeholders are crashes of a newer passenger vehicle, and/or involve serious injury. For convenience, we'll refer to these two categories as "focus areas." We define a crash-involved vehicle to be "newer" if it is at most 4 years old at the time of the crash.[36] We define a crash to involve a "serious injury" if the police report indicates that a passenger vehicle occupant was killed or incapacitated.[37]

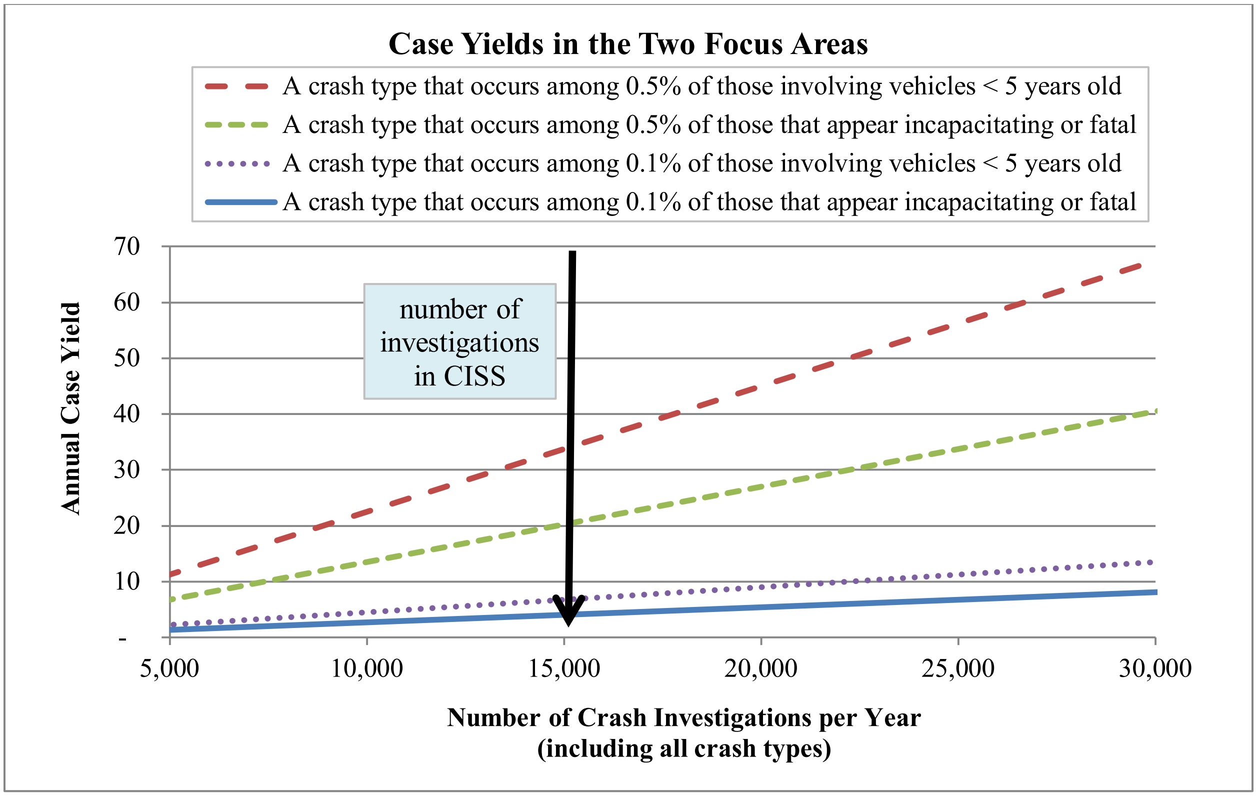
By incorporating information about the occurrence of these two broad crash types in the probabilistic selection of the data collection sites, and through judicious use of stratification and sample allocation in designing the crash selection mechanism, we greatly reduce the number of crash investigations required without sacrificing the national representation that comes from probability sampling.

The number of crash subpopulations involved in one or more analytical needs are too numerous to address individually. Rather, the chart below, Case Yields in the Two Focus Areas, illustrates our

[35] While vehicles that run solely on electric power commonly use lithium ion batteries, only some hybrid gas-electric vehicles do, so less than 1 percent of crashes in 2011 involved lithium ion batteries.
[36] That is, the difference between the year of the crash and the model year of the vehicle is at most 4.
[37] Police make their best assessment of injury severity based on the information available to them at the time. What looks to be an incapacitating injury on the scene of a crash may turn out to be relative minor, or vice versa. More than three-quarters of injuries indicated to be incapacitating on crash reports end up being "moderate" or "minor" as assessed by medical professionals with more information.

expectations for case yield based on the incidence rate of the particular crash type among the two focus areas. For instance, children killed or injured in side impacts account for about 0.5 percent of serious injury crashes. Figure 4 indicates that investigating 15,000 crashes per year, if selected using our design efficiencies, would yield about 20 annually that involve children in such crashes. At the same overall sample size (15,000 crashes per year), an analysis principally concerning crashes that account for 0.5 percent of crashes involving a car or light truck that is less than 5 years old would have slightly more than 30 cases per year on which to base its study.

Figure 4: Case Yields in the Two Focus Areas



Overall, the charts indicate that for crash types that account for of 0.5 percent of one of the two focus areas, a 15,000-crash sample (designed with our efficiencies) can be expected to generate about 20 to 35 cases per year. An extremely rare event, accounting for 0.1 percent of a focus area, would yield lesser numbers, about 3 to 8 crashes per year, and we are hopeful that combining data from multiple years of data would provide adequate analytical support.[38]

8.3. Recommended Number of Crash Investigations in CISS

Based on the computation from the previous section, we arrive at the figure of 15,000 crashes per year. It is our view that this is the number of investigations that should be conducted if the new system will be effective at meeting the evolving needs of our data users. This represents a slightly more than four-fold

[38] As a general rule, the agency does not like to combine more than 5 years of crash data to shed light on a problem that depends on the ever-changing crash environment. Five years of data should produce15 to 20 cases, even for very rare crash types (0.1% of one of the two focus areas).

increase in the cases from the current CDS. This number of cases is slightly less than what was recommended in the original NASS design.

8.4. The Number of Data Collection Sites in CISS

Increasing the number of investigations per year without modifying the other stages in a multi-stage sample is not an effective way to reduce the sampling error. When the numbers of crashes are estimated from multi-stage samples[39] instead of all crashes of the universe, then the analyst must also consider the variability in the estimates. These national estimates have variability, or sampling error, associated with the selected multistage samples. The size of the sampling error is determined by the sample sizes at each stage of the sample. Simply increasing the crash sample size without increasing the primary sampling unit (PSU) sample size will only reduce the sampling error a limited degree.

To reduce the total sampling error the number of data collection sites or PSUs should be increased. This is an important factor to consider when we decide the number of crash investigations. A small number of PSUs either cannot produce enough high-interest cases or does not reduce the overall error of the estimates enough. The new sample design must have a larger number of PSUs to produce precise estimates.

To reduce the sampling error, the new sample design will be a stratified sample because subgroups of crashes can vary greatly within the overall crash population. Crashes of new vehicles are more likely to be found in affluent areas of the county. The stratified design divides the country in 24 homogeneous subgroups except for Los Angeles County.[40] We believe the number of PSUs in the new CDS should be 73; selecting 3 PSUs from each subgroup (3 x 24 = 72) plus 1 additional PSU consisting of Los Angeles County. This is slightly more than 3 times the current number of PSUs at 24 and is very similar to the original NASS sample design completed in 1977, which called for 75 PSUs.

8.5. The Numbers of Data Collection Sites and Crash Reports for CRSS

Standard statistical formulas indicate that increasing the number of PSUs in GES from the current 60 to 75 PSUs will improve the margins of error by about 10 percent.[41] The figure below depicts the general relationship between the number of PSUs and the margin of error. Considering the relative inexpensiveness of collecting PAR-only data, and based on the types of analyses the agency and other users pursue, NHTSA recommends conducting the CRSS at 75 PSUs.

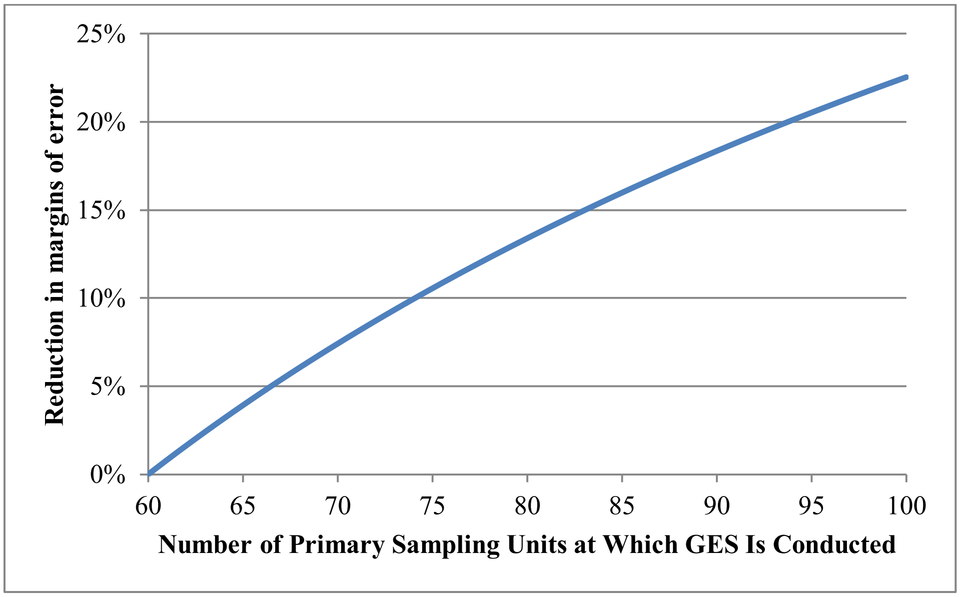
---

[39] For example, a county or group of counties are first selected at the first stage as PSUs, then within selected counties, police jurisdictions are selected at the second stage as secondary sampling units, and finally within selected police jurisdictions, crashes are selected at the third stage as tertiary sampling units.

[40] Los Angeles has an inordinately large number of crashes, so it is a certainty PSU and is considered its own stratum.

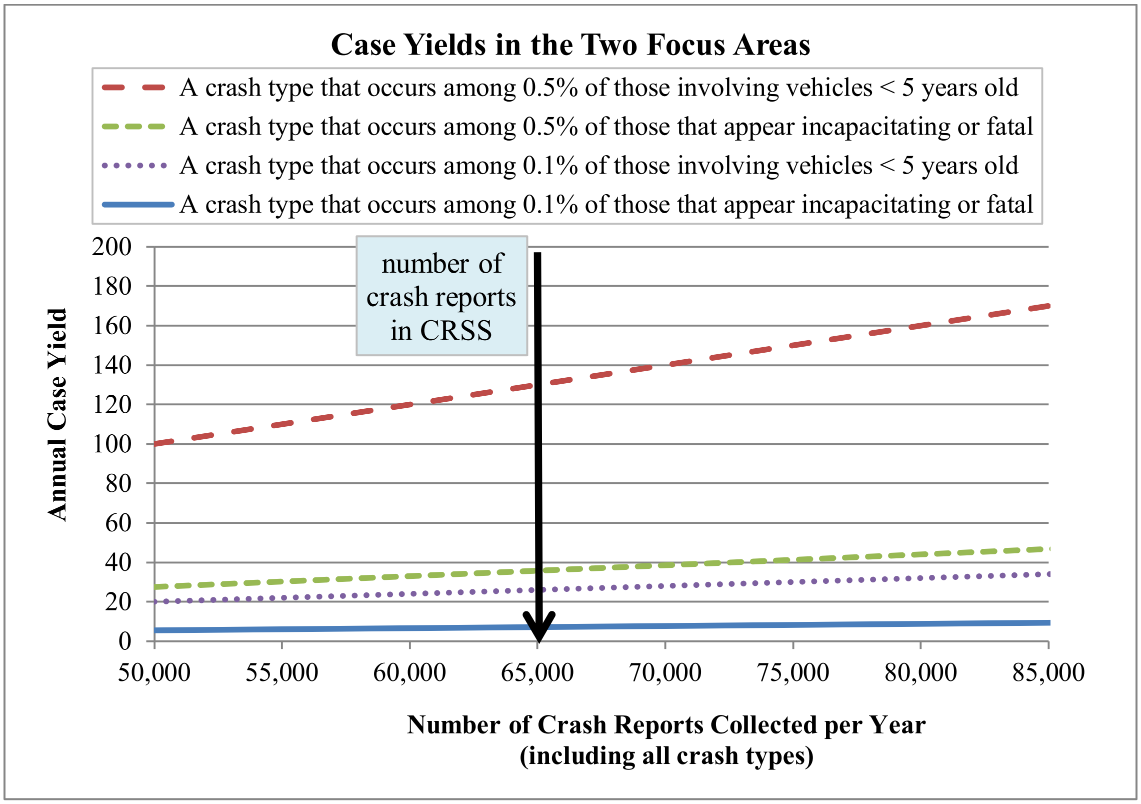
[41] All else held constant, doubling the number of PSUs cuts variances in half. See e.g., Cochran, W. G. (1977). *Sampling techniques*, 3rd ed. New York: Wiley.

Figure 5: Number of Primary Sampling Units at Which CRSS Is Conducted



In addition to producing reasonably precise estimates, it is also desirable for CRSS to contain adequate numbers of crashes of high interest for case studies (e.g., for automakers to use to design the next life-saving technology). Based on the data needs, we are designing CRSS to focus on the same two areas of foreseeable interest – crashes involving serious injury and crashes involving newer vehicles. As illustrated in the Figure 6, this should provide about 7 cases per year for crash types that occur among 0.1 percent of those that appear incapacitating or fatal, and about 25 per year for types that occur among 0.1 percent of those involving vehicles under 5 years old. We are hopeful that this will provide adequate case yields to support most data needs.

Figure 6: Case Yields in the Two Focus Areas



8.6. Summary of Recommended Samples

Based on NHTSA's review of the NASS, we have developed two new samples that:

- Reflect the current population of drivers, vehicles and crashes in the country,
- Target more severe crashes,
- Target newer vehicles,
- Have smaller margins of error for key estimates,
- Are flexible to increase the number of sites if additional funding is available without having to reselect the entire sample again,
- Are flexible to change the sample units if necessary at any stage of the sample, and
- Better represent data users' needs by selecting the CISS and CRSS samples separately from different sample frames.

NHTSA recommends that CISS have significantly more crash cases collected at significantly more sites than the current CDS. For CRSS, NHTSA recommends a slight increase in cases and collection sites compared to the GES. Current and recommended levels of collection sites and crash cases are shown on the following page in Table 9.

Table 9: Current and Recommended Sample Sizes for GES/CRSS and CDS/CISS

System	Collection Sites	Crash Cases
Current GES	60	50,000
CRSS	75	65,000
Current CDS	24	3,500
CISS	73	15,000

9. Advantages or Disadvantages of Expanding Non-Crash Data Instead of Crash Data

To properly address the advantages or disadvantages of expanding the collection of "non-crash data," the term must be defined. For sake of this discussion, non-crash data does not mean non-crash vehicle-related incidences like carbon monoxide poisonings; rather, non-crash data is exposure data from naturalistic driving experiments like SHRP 2 (Strategic Highway Research Program), in which most of the data is normal driving with no crashes, and near misses and other critical incidences are examined.

NHTSA estimates the benefits of three types of countermeasures: crashworthiness, crash avoidance, and driver behavior countermeasures. The basic building blocks of such analyses are target population estimates (how many crashes are involved, people injured and killed) and effectiveness of the countermeasure (percent reduction in crashes, injuries and fatalities). Non-crash data can provide valuable insights into driver behavior that can contribute to programs designed to affect that behavior.

9.1. Target Population

Crash data is essential to determining the target population for all three types of countermeasures.

Crashworthiness countermeasures demand very refined target population estimates. For example, to estimate the target population for countermeasures to decrease chest injuries in side impacts, annual data is needed on how many chest injuries occur by injury level, how many fatalities occur, and what are the interior contact points (side door, armrest, B-pillar, console, other occupant, etc.) within the vehicle that caused those injuries. Other important information include the severity of the crash (delta- V), whether these are caused by striking a tree or being struck by another vehicle, the age of the occupant injured, etc. Large amounts of data are needed to obtain a good statistical sample of all of these variables in this very refined target population. To provide the level of detail needed for this analysis to get injury levels and interior contact points, crash investigators must go to the crashed vehicle, get hospital records, etc. Non-crash data is of no value for this type of analysis.

For crash avoidance estimates, non-crash data may or may not be of value for determining the target population, depending on the specific countermeasure and how refined you try to make the target population. Different crash types were examined for NHTSA's electronic stability control analysis (single-vehicle run off road, rollovers, and multivehicle crashes). Non-crash data is of limited, if any, value for this analysis.

For driver behavior countermeasures, non-crash data may be the best way to refine a target population of crashes. For example, to create driver distraction guidelines, non-crash data could be used to estimate how many crashes (based on near-misses) are caused by driver distraction. Currently those estimates are derived from PARs or from NHTSA's investigated crashes, but the only sources of that data are the interviewed vehicle drivers. In some cases drivers will tell you they were reaching for an object, or handing something to a child in a rear seat. Non-crash data can provide information about what the driver was doing and where the driver was looking in near-misses.

Non-crash data can also provide information about how often people are tired or about to fall asleep. This information can be used to refine target population estimates to examine a countermeasure that tracks driver eye glances and warns drivers when their eyes were off the road for too long of a period.

9.2. Effectiveness Estimates

For crashworthiness effectiveness estimates of new countermeasures that are currently not in any or many vehicles, NHTSA compares test data of vehicles with and without countermeasures using dummies to determine the reduction in injury levels resulting from the countermeasures. When many vehicles in the fleet have a crashworthiness countermeasure NHTSA will evaluate the effectiveness of vehicles by comparing injury and fatality rates of vehicles before and after implementation of the countermeasure. Data from EDR or delta-V of the vehicles involved in the crash is being used more often in these analyses. Non-crash data is of little or no value in determining the effectiveness of crashworthiness countermeasures.

For crash avoidance countermeasures, effectiveness is determined by a variety of means, depending upon the availability of data. If a countermeasure is already available in the marketplace on a sufficient number of vehicles to obtain statistical significance, then the crash data collected is sufficient to make an estimate, but not as thorough and informative as if naturalistic driving data on that countermeasure was also available. For example, electronic stability control for light vehicles existed for several years before NHTSA evaluated its effectiveness. For this type of analysis, NHTSA typically uses a statistical method called paired comparison. The crash type affected by the countermeasure is compared to a control crash type that would not be affected by the countermeasure, for specific make/model/model years preceding introduction of ESC versus those matching make/model/model years after introduction of ESC. NHTSA does not need to collect non-crash data to do this type of analysis. However, naturalistic driving data could provide clues about what drivers were doing in situations in which ESC was initiated. This information could lead to consumer information to avoid those types of situations.

If a countermeasure is not already available in the marketplace, how is effectiveness estimated? There are several methods to collect information that can be helpful in estimating effectiveness. The agency has used field operational tests (FOT) in which NHTSA supplies select drivers with test vehicles having the technology, equips those vehicles with a number of cameras to see what the vehicle is doing, what the driver is doing, how the driver reacts to warnings given, etc. These tests are naturalistic driving tests that are focused on specific technologies. There are normally not enough crashes in these FOTs to estimate effectiveness directly, but they provide valuable information on timing of the warnings given and timing of driver reaction to the warnings.

A second method of developing information useful for effectiveness estimates is to test vehicles on a test track to determine the characteristics of the countermeasure, and test driver reactions to situations where the countermeasure would provide warnings, brake automatically, or otherwise function to help prevent the crash.

A third method is to use a simulator, set up to simulate crash conditions in which the countermeasure would provide a warning and to determine driver reactions (such as how they steer and what they perceive) to those conditions. This information is used in estimating effectiveness.

Naturalistic driving would provide a fourth set of data that could be used in estimating effectiveness of crash avoidance countermeasures. Drivers are the cause of over 90 percent of crashes, and data from a naturalistic driving study would provide many pieces of different information relating mostly to driver actions such as how often drivers are falling asleep or how often are they distracted. There are many situations that could be examined to determine how driver related factors will impact the estimate of effectiveness, including the timing and force of braking.

NHTSA has another naturalistic driving study underway in Ann Arbor, Michigan, on V2V communication. The agency will be using a combination of data from this study, an analysis of crashes at intersections where pre-crash EDR data are available on both vehicles in the crash, and driver simulator data for driving reaction/response times, to estimate the effectiveness of V2V in reducing intersection crashes.

Combining these pieces of data, NHTSA performs an engineering analysis of the timing of a collision (how much time before the collision is an alert given, how fast are the vehicles going, how will the driver react to the alert, is there enough time/distance to stop the vehicle, etc.) and all of these actions may be simulated and varied for different crash scenarios to estimate effectiveness.

The obvious conclusion from the previous discussion is that the more data from different sources is available, the better NHTSA's estimates of effectiveness.

The disadvantage of naturalistic driving information is that it is very expensive to collect, and costly to analyze and store for future analyses. This is why it is only collected as a field operational test on a small scale for specific technologies, or collected once every several years on a more general larger scale, rather than on an ongoing basis. For example, SHRP2 costs over $100 million annually to run, and we would expect its data to be used for many years.

In summary, while non-crash data has clear utility, it also comes with a very high cost. Until we reach the distant point in the future where technology has prevented crashes from ever occurring, crash data will continue to serve a vital role in making road travel as safe as it can be.

10. Costs of Collecting and Analyzing Additional Data

The original NASS sample design completed in 1977 called for 75 investigation-based data collection sites but never reached that level. In the early 1980s NASS data collection reached a high of 50 sites. However, in 1988 the NASS was redesigned into a crash record-based (GES) system at 60 sites to produce better overall crash estimates and an investigation based (CDS) system at 36 sites to ensure better crashworthiness estimates. Over the years the number of GES data collection sites has remained constant at 60, while CDS sites were reduced to 24 due to budget constraints.

The cost to operate and maintain the current NASS GES and CDS data collection process includes five major categories: training for program participants, data collection itself, data coding, quality assurance and information technology. For each of these areas, NHTSA must enter into contracts to complete the work. Labor costs, which increase annually, account for more than three-quarters of data collection costs. The following are brief descriptions of the data collection steps performed by contractors in the current GES and CDS:

GES – records based data collection:

1) Correctly select crashes according to the sample design.
2) Physically obtain a copy of the selected crash report.
3) Once the crash report is obtained from a specific police jurisdiction or State, the trained staff must accurately synthesize information from each State jurisdiction and code the required data elements into a nationally uniform reporting format.
4) Next, these data are entered into a relational database maintained by NHTSA's information technology contractor where they are electronically combined into a single national data set.
5) A contractor conducts a quality control process to ensure completeness and accuracy of the data set.
6) Once the quality control process is completed, the final data file is completed, reported upon, and made publicly available.

CDS – investigation-based data collection:

1) Correctly select crashes according to the sample design (includes travel to jurisdictions that list PARS, completing a sample algorithm, and returning to the jurisdiction to obtain sample PAR).
2) A highly trained crash technician begins a detailed scene documentation, vehicle inspection, interview, and determining possible sources of injury.
3) These data are then synthesized into a uniform case format and entered into a relational database maintained by NHTSA's information technology contractor where they are electronically combined into a single national data set.

4) A contractor conducts a quality control process to ensure completeness and accuracy of the data set.
5) Once the quality control process is completed, the final data file is completed, reported upon and made publicly available.

NCSA maintains very high standards for the individuals who assist in data collection efforts and for the quality of the resulting data and analysis. To attract and retain the qualified staff necessary to maintain these high standards, we must assure the payment of competitive wages. These costs generally increase annually, at least to match the level of inflation, and NHTSA must keep pace with these increased operating costs.

In the past few years while the NASS budget has remained flat, NHTSA has implemented serious cost-cutting measures that have led to reductions in staff and data collected. Fewer cases being collected increases the time it takes to identify issues, create countermeasures and evaluate their effectiveness in the real world. To continue to operate under a flat budget, NHTSA has had to reduce the number of contract crash investigators and support staff in NASS CDS sites. In addition to diminished CDS staff, some case details are no longer collected. At the current funding levels, it is highly likely the case numbers in CDS will continue to diminish and the case quality will continue to degrade.

10.1. Costs of Data Modernization

In the 2012 appropriation, NHTSA received funding to perform a thorough review of the current sample design and data collection procedures in NASS CDS and GES. A contractor was hired to work with NHTSA's experts to create a scalable, nationally representative sample design for NASS.

As described in Section 8, NHTSA recommends increasing CISS (the new CDS) sites to 73 and CRSS (the new GES) sites to 75 while also increasing the number of cases in each system. NHTSA also recommends deploying staff at each CISS site with an average of two contract crash investigators and one support person. In addition, more extensive data collection would be conducted in CISS, such as for better scene documentation and additional crash avoidance data. Tablet computers and scene diagraming programs would be employed by the crash investigators, enabling more data to be collected as well as increasing data collection efficiency.

To accurately estimate what CRSS and CISS would cost when the final sample design and data collection protocols are not decided is quite difficult. For example, the use of e-PAR data collection and advanced scene diagraming technology could lead to some cost reductions. However, we do not know at how many sites we will be able to use e-PAR collection until we begin system implementation.

We have built a cost model based on our current data collection cost knowledge, but with the following assumptions:

- Each CISS site will be fully staffed with an average of 2.6 staff per site, two contract crash investigators and one part-time support person, compared to the current 2.4 average staff per site.

- Quality control will be enhanced at every stage of data collection and coding for CRSS and CISS cases.
- Additional data will be collected for every CISS case in response to NHTSA and stakeholder input.
- Better IT systems will allow cases to be processed more efficiently and securely, and users will be able to access more data more quickly.
- CISS crash investigators would be equipped with tablet computers and electronic measuring devices.
- Each CRSS site would at least initially require a contract data collection person, though we expect this requirement will dwindle over time as e-PARS and remote data collection become increasingly available.

These assumptions will lead to higher quality information in each case.

CISS data collection costs per case would increase slightly. This increase is due to more robust data collection, such as EDR, and enhanced quality control of the CISS data while gaining some efficiency from EDMI and tablets. The increased number of data collection sites will have a substantial impact on the annual operating costs.

CRSS data collection costs on a per case basis probably will increase when compared to GES because CRSS will be collected independently from CISS. In the current NASS, 24 sites are both GES and CDS sites and 14 sites are adjacent to CDS sites, so GES collection costs are minimal for 38 of the current GES sites. In the new CRSS, all 60 sites are independent and therefore may require a full or part-time data collector. The same amount of data will be coded in CRSS as is currently collected in GES, but additional data linkages and increased quality control of CRSS data will potentially lead to increased costs.

Implementation of the optimal CRSS and CISS would be phased in over several years due to the complexity of these systems, especially the very detailed CISS. Based on the best information NHTSA has at this time, the annual operational costs for the optimal CRSS and CISS could be as high was $50 million. The optimal CISS would also require up to $10 million in one-time costs for equipment, e.g., EDR retrieval kits, scene diagram equipment and tablet computers. Table 10 shows the estimated costs to achieve the optimal system in 2018.

Table 10: Costs of Current Systems and Estimated Costs of Phasing in the New Systems

Fiscal Year	System	Operational Costs (in millions)	One-Time Equipment Costs
2013	NASS GES and CDS	$12.5	-
2014	NASS GES and CDS	$17.5	-
2015	NASS GES and CDS	$17.5	-
2016	CRSS and CISS	$30.0	$2.0
2017	CRSS and CISS	$37.0 - $40.0	$4.0
2018	CRSS and CISS	$45.0 - $50.0	$4.0

Total costs for the new system would include additional costs for the other modules of the new system (such as for the on-scene passenger vehicle module). Although a vital component of the new system, these modules are very dependent on the amount and type of data collected and in how many sites. NHTSA would not propose beginning these modules until at least 2019 when the base CRSS and CISS systems are fully implemented.

11. Potential for Obtaining Private Funding

While NHTSA might be able to locate funding sources that would pay the costs of the agency to collect and analyze specific datasets, the NHTSA's ability to retain funds from providing such services is limited by certain statutes.

The "miscellaneous receipts" statute requires that an agency receiving funds from any source deposit those funds into the Treasury as soon as practicable without deduction for any charge or claim. Specifically, this statute requires an agency to deposit any funds obtained from a private source into the Treasury and not into the agency's own appropriations for further use. The Government Accountability Office (GAO) has suggested in its guidance that to retain such funds would constitute an improper augmentation of an agency's appropriation. The GAO has cited the Appropriations Clause of the U.S. Constitution and the miscellaneous receipts statute for this view. This underscores the general concept that executive branch agencies remain dependent upon the congressional appropriation process and that Congress retains control of the public purse under the separation of powers doctrine.

In order to overcome the requirement that NHTSA deposit any funds it would receive in the Treasury, the agency would require specific statutory authority to retain the funds. (NHTSA currently has one such authority allowing the agency to retain funds from the sale of crash-tested motor vehicles and motor vehicle equipment.) Any specific statutory authority would need to contain clear language allowing the agency to retain funds collected for data analysis work in its own accounts.

NHTSA does not plan to request this authority for several reasons. To maintain the integrity of an on-going data collection system and to collect the best quality data continuously, NHTSA must have a known budget well in advance of each data collection year. The potential variability in outside funding would create operational inefficiencies; for example, by starting up and shutting down collection sites continually. In addition, NHTSA maintains an unbiased data collection program and the perception of funding from the very industry it regulates could bring into question the integrity of the data collected.

12. Potential for Recovering Costs From High-Volume Users

NHTSA is permitted to establish user charges or fees for services it provides. Services subject to user charges or fees are described as Federal activities that convey special benefits to recipients beyond those accruing to the general public. In order to establish fees, the agency would be required to promulgate them through a regulation. NHTSA, however, would still require specific statutory authority to retain fees for services collected from the public. In the absence of such authority, the agency would be required to deposit any collected fees in the Treasury as miscellaneous receipts.

Government policy requires that a user charge be assessed against each identifiable recipient of a special benefit. Specifically, user charges apply when a government service allows a beneficiary to obtain more immediate or substantial gains or values than those that accrue to the general public; provides business stability or contributes to public confidence in the business activity of the beneficiary; or, is performed at the request or the convenience of the recipient, and is beyond the services regularly received by other members of the same industry or group or the general public. No charges are permitted when the identification of the specific beneficiary of the service is obscure, and the service can be considered primarily as benefiting the general public. Under these policy guidelines, NHTSA would not be able to establish fees for services it already provides or plans to provide to the general public, but only for those services that create a special benefit for a recipient.

The determination of an appropriate user charge would be based on the amounts necessary to recover full costs to the agency. Full costs would include an appropriate share of the costs covering direct costs (such as personnel costs) and other indirect costs, including applicable overhead. Using agency records to estimate costs, NHTSA would seek to establish a chargeable rate (as opposed to a fixed price) for providing services.

The implementation of any user charges would be through regulation (similar to the steps that the Department of Justice used to establish chargeable fees for FOIA processing) consistent with providing a notice to and receiving comments from the public. This process would also establish the basis for identifying individuals that receive special benefits and would be subject to fees. Specifically, the agency would establish a definition of user for services associated with collecting and analyzing data. NHTSA would not be able to apply fees against those determined to be high volume users unless the agency also determined that those individuals would be receiving a special benefit. In this way, the fees would apply to anyone receiving a special benefit from the work regardless of the volume of requests. (The established fees would be periodically adjusted to reflect changes in costs.)

In order to make use of any fees collected for services provided, NHTSA would require legislative authority to retain such fees in its accounts. NHTSA would need to seek such authority prior to making any effort to establish fees and to make services available to users that resulted in conveying a special benefit.

At this time NHTSA would not seek establishing user fees. NHTSA’s current data files are made available at no charge to users that range from the industry to public school students. Potentially inhibiting the use of data by imposing a fee would not seem consistent with NHTSA’s safety mission.

13. Conclusions

NHTSA's crash data systems play a fundamental role in identifying the size of traffic safety issues, tracking trends, providing data for developing countermeasures and evaluating outcomes. Additional cases and data collection sites increase the precision of safety estimates made from the data, reduce the time needed to determine and validate areas of specific interest to the Offices of Vehicle Safety Research, Rulemaking, Defects Investigation, and Behavioral Safety Research, and better assist researchers around the world in making informed decisions on vehicle design and safety policy.

Both the NASS GES and the NASS CDS currently serve many needs. The NASS crash record-based program, GES, tracks trends and serves to evaluate countermeasures. The NASS investigation-based program, CDS, is focused on the collection of vehicle crash and injury outcome data. Using CDS data, NHTSA is able to recognize that a problem exists and effectively evaluate the life-saving potential of new technologies. While these data requirements continue, traffic safety data needs have substantially changed since NASS was designed in the 1970s, so NHTSA conducted a thorough review of the system.

NHTSA recommends modernizing NASS by creating two, new, separate crash sample data collection systems:

13.1. Crash Investigation Sampling System

The CISS will:

- **A. Support future analytical needs by conducting 15,000 passenger vehicle crash investigations per year in 73 newly selected CISS sites.**
- **B. Enhance the collection of pre-crash data and information on crash avoidance technologies, while continuing to collect crashworthiness data.**
- **C. Be a flexible system that could be adapted for periodic special crash studies, such as analyses of crashes involving pedestrians, motorcycles or large trucks.**

As noted by Congress and data users, the number of passenger vehicle crash investigations has deteriorated to a point where the sample is barely adequate for some current analytical needs and unquestionably inadequate for future analytical needs. It is unfortunate that the agency is in a position where we must wait for enough cases to show up in our databases to spot an emerging threat or issue a life-saving regulation. As detailed in this report, we recommend increasing the number of investigations to 15,000 per year, conducted at 73 newly selected data collection sites that better reflect today's demographic profile and traffic patterns.

With more crash avoidance technologies entering the vehicle fleet and with driver distraction on the rise, NHTSA:

- Undertook a comprehensive review of CDS's data elements to bolster the information CISS would provide in these areas, without compromising our regulatory responsibilities in crashworthiness.
- Anticipates that the information collected from EDRs will be an integral part of the agency's data modernization efforts. The output data from the vehicle's EDR is a fundamental component for understanding the specific details for the events and factors related to the crash. Future EDR technology will further improve our vehicle research process by enabling the agency to collect and provide information on advance crash avoidance systems in real-world driving conditions.
- Identified new investigatory tools that, in addition to making data collection safer for data collection staff, will provide researchers with highly accurate scene diagrams with which to design the next life-saving technology.
- Built structural flexibilities into the new system so that it may easily accommodate the types of specialized studies that would best support the needs of pedestrians, motorcyclists, and large truck crash victims. These flexibilities will also improve our ability to conduct rapid response studies that would help identify emerging threats and supplement SCI in providing critical support to defects investigations.

13.2. Crash Report Sample System

The CRSS will:

A. Support analyses of all vehicle types and all crash severities by coding 65,000 PARs in 75 newly selected CRSS sites.

B. Amplify the analytical potential of CRSS through data linkage to other existing data sources.

As noted by Congress and data users, the current GES sites reflect U.S. demographics as they were when the sites were selected 30 years ago, and thus a fresh sample is required. Based on the information currently available and the fact that our population and driving environment are more heterogeneous than they were 30 years ago, we project that we will need to collect data on 65,000 police accident reports in 75 sites to obtain a representative picture of all severities of crashes involving all vehicle types. Additional cases will provide more nationally representative information on motorcycle, large truck and pedestrian crashes. Additional sites for CRSS (as well as CISS) will increase the size and geographic disparity of a key pool from which SCI selects cases, so that it can better support defects investigations for the agency. Further, the CRSS would be enhanced by data linkages, such as through linking crash data to vehicle registrations. This would greatly improve the percent of crashes with known VINs, presenting much more information than we currently have for crashes involving non-passenger vehicles and providing the Office of Defects Investigation with a larger pool of cases with reliable make and model information.

13.3. Improvements to Information Technology Infrastructure

A. Make CRSS and CISS data more accessible and more secure, and provide users with better tools for analyzing the data.

B. Process CRSS and CISS data in a more streamlined and consolidated environment.

The information technology infrastructure for the NASS CDS and GES are currently separate and do not have the capacity to be easily updated to reflect new data collection needs or to adapt to new IT security requirements. We are building a consolidated and flexible IT infrastructure for CRSS and CISS that will pay dividends in an agile system that can be changed to reflect changing user needs in the future. The new system will fully comply with personally identifiable information security requirements and provides full-system backup and recovery processes at a secondary data center. Responding to the needs of stakeholders, we are designing new user interfaces and data query tools, and will provide the data in more accessible formats.

13.4. Resources Needed to Implement System Modernization Improvements

Crash data is vital to the NHTSA's mission. The new systems, CRSS and CISS, will provide better quality and more timely data essential to reducing fatalities and injuries and preventing crashes that cost society $277 billion a year. To meet future information needs, NHTSA will need to dedicate additional resources to these activities – both operational funding and staff. Until the number of data collection sites is determined, it is difficult to precisely quantify the additional operational budget requirements and the specific number of new staff positions. NHTSA's crash data collection must evolve to keep pace with emerging technology and changing demographics and policy needs. Expanding the number of data collection sites and the number of cases collected while providing more accessible and better quality data will enable NHTSA to affirm its position as the world leader in motor vehicle crash data collection and analysis.

DOT HS 812 128
March 2015

U.S. Department
of Transportation
National Highway
Traffic Safety
Administration

www.nhtsa.gov

11342-031915-v2