Traffic Safety Facts
Research Note
www.nhtsa.gov
DOT HS 811 057 December 2008

The National Motor Vehicle Crash Causation Survey

Introduction

To substantially reduce the high number of traffic fatalities and injuries, more needs to be done to prevent crashes by understanding the events leading up to a crash. The automotive industry has already applied significant resources into the research and development of crash avoidance features in vehicles. Many of the new features (ESC, traction control, lane-departure warning systems, etc.) are starting to appear in the fleet of newer model vehicles. NHTSA and other safety researchers are currently evaluating the effectiveness of these new technologies. Available databases, such as the National Automotive Sampling System (NASS) Crashworthiness Data System (CDS) do not provide information that can specifically serve the purpose of identifying pre-crash scenarios and the reason underlying the critical pre-crash events — information crucial to the evaluation and development of emerging crash avoidance technologies. Additional data are needed to identify factors associated with crash causation. With this objective, in 2005 NHTSA was authorized under Section 2003(c) of the Safe, Accountable, Flexible, Efficient, Transportation Equity Act: A Legacy for Users (SAFETEA-LU) to conduct a national survey to collect on-scene data pertaining to events and associated factors related to a crash. NHTSA's National Center for Statistics and Analysis (NCSA) has conducted the National Motor Vehicle Crash Causation Survey (NMVCCS) of crashes with focus on the factors related to pre-crash events involving light passenger vehicles.

This research note provides a brief description of the NMVCCS survey design, scope, the nature of information collected, and data limitations. Also, some statistics related to the driver, vehicle, roadway, and environment that often play a role in the crash, as collected through NMVCCS have been presented in this research note. This note also serves as a companion publication to the report to Congress that NHTSA submitted in September 2008.

Background and Objective

Nearly 30 years have passed since the last on-scene crash causation study was conducted (the Indiana Tri-Level Study in 1979). The information from the Indiana Tri-Level Study is outdated due to the changing nature of the vehicle fleet and vehicle technologies. Also, since the last study, driver behavior has changed due to a variety of dashboard electronics, also called telematics, pertaining to entertainment, navigation, and communication. Furthermore, the Tri-Level Study was not nationally representative in that it was only conducted in one small part of the country and was not based upon a statistical design. In 2006, NHTSA concluded a 100-car naturalistic study that was an instrumented-vehicle study undertaken with the primary purpose of collecting large-scale, naturalistic driving data. While this study captured information on overall driving behavior in crashes, near-crashes, and other incidents, it was not designed to conduct in-depth, on-scene investigations of crashes that are necessary to determine the factors related to pre-crash events. Crash-avoidance technology (e.g., collision-avoidance systems) continues to be introduced, and data are needed to evaluate these systems, as well as establish priorities among investments in emerging technologies. Recognizing the need for such data, Congress asked NHTSA to conduct NMVCCS – the first nationally representative survey aimed at providing information on the pre-crash environment in crashes involving light vehicles and building up a national database containing detailed information on events and factors leading up to a crash.

The objective of NMVCCS was to collect on-scene information on the events and associated factors leading up to crashes that involve light vehicles. This information will facilitate the statistical and clinical analyses that would help identify, develop, as well as evaluate current and emerging crash avoidance technologies for the improvement of highway safety.

Scope

Like any well-designed sample, NMVCCS had strict guidelines for a crash to qualify for an on-scene investigation. Only crashes occurring between 6 a.m. and midnight were considered for possible investigation. Taking into consideration the operational and statistical issues, NHTSA set the following criteria that a crash must meet in order to qualify for an investigation:

- The crash must have resulted in a harmful event associated with a vehicle in transport on a trafficway.
- EMS must have been dispatched to the crash scene.
- At least one of the first three crash-involved vehicles must be present at the crash scene when the NMVCCS researcher arrives.
- The police must be present at the scene of the crash when the NMVCCS researcher arrives.
- One of the first three vehicles involved in the crash was a light passenger vehicle that was towed.
- A completed police accident report for this crash must be available.

Sample Design

To make it cost-effective and yet achieve national representation, the NMVCCS used the existing NASS infrastructure to collect information about the events and factors that possibly contributed to crash occurrence. A multistage probability sample design was used to acquire a nationally representative sample of crashes. The probability of crash occurrence was taken into account by considering both the geographic location and day/time of the crash. This resulted into a two-dimensional sampling frame. While this frame was fixed in terms of the geographical area, day and time were dynamically overlaid on these areas. Using this sampling frame, crash selection for possible NMVCCS investigation was done sequentially through the following four stages:

- Selection of primary sampling unit (PSU) that is based on the geographical area and urbanization type across the entire United States;
- Selection of a time strip (a six-hour time interval between 6 a.m. and midnight) for each of the selected PSUs;
- Selection of days of the week for the selected time strip; and
- Selection of a crash within the selected time block, the combination of the selected day of the week and the time strip.

Due to operational issues, some local adjustments were made in certain PSUs. Similarly, in order to handle special situations, such as larger volume of transmissions or very large geographical area in a PSU, sub-sampling was implemented in certain PSUs. More detailed information is available in a NHTSA technical report.[1]

Sample Size

Based on the sample design detailed above, NMVCCS collected data on a total of 6,949 crashes over a span of 13,019 time blocks during a three-year period, January 2005 to December 2007. Of these, 5,470 crashes comprise a nationally representative sample. The remaining 1,479 crashes are meant for clinical studies.

Case Weights for National Estimates

To make the NMVCCS sample representative of all similar types of crashes for the whole of the United States, each of the 5,470 investigated crashes has been assigned a certain weight based on the probability of selecting a crash through the four stages of the sample design, together with some adjustments. The methodology used in determining design weights is discussed in detail in a NHTSA technical report.[1] National estimates related to several aspects of the crashes investigated in this survey can be obtained by using these weights.

In this complex sample design involving stratification, clustering, and missing adjustments, a computer-intensive variance estimation method[2] available in the SAS[3] software package can be used to compute the standard errors of the estimates.

Cases With Zero Weights

Of the 1,479 cases investigated between January 1, 2005, and July 2, 2005, 832 were investigated during the phase-in period and 647 cases failed to meet all the NMVCCS criteria. Due to insufficient information available for computing case weights, these cases have been assigned zero weights.

Data Collection Methodology

The NMVCCS aimed at collecting comprehensive and detailed information on a large number of variables related to vehicles, drivers, roadways, and environment. Therefore, every effort was made to prevent loss of information that could shed light on the pre-crash environment of a crash.

NHTSA's past experience in data collection has shown that the availability of crash data often diminishes with the passage of time. The loss of information may occur in several possible ways: the vehicle may be removed from the crash scene, the vehicle might have been altered, evidence may disappear, driver's memory of events may fade, etc. This makes it difficult to obtain an untarnished account of events. It is, therefore, imperative that the crash investigation begins as quickly as possible. Keeping these facts in view, every effort was made for the timely arrival of the NMVCCS researcher at the crash scene. These researchers constantly monitored crash occurrences and coordinated with EMS and police.

[1] Choi, Eun-Ha, et al. (April 2008) A Sampling Design Used in the National Motor Vehicle Crash Causation Survey, DOT HS 810 930. Washington, DC: National Highway Traffic Safety Administration.

[2] Lohr, S. L. Sampling: Design and Analysis. Duxbury Press, 1999.

[3] SAS/STAT 9.1 User's Guide, SAS Institute Inc., Cary, NC. 2004, pp. 4,185-4,240.

Once at the crash scene, the researcher confirmed if the crash satisfied the NMVCCS crash qualification criteria listed in the previous section discussing the scope of NMVCCS. The subsequent crash investigation aimed at acquiring the targeted information from all possible sources: the crash scene, police, drivers or surrogates of the drivers, passengers, vehicles, and witnesses. The priority at the crash scene was to conduct interviews of crash participants as well as surrogates of the drivers who, due to injuries or other reasons, could not be interviewed. The targeted information was collected using a set of field forms and a portable computer.

Information Collected in NMVCCS

NMVCCS adopted the approach proposed by Perchonok[4] (1972). Accordingly, a crash in this survey is considered as a simplified linear chain of events ending with the critical event that precedes the "first harmful event" (i.e., the first event during the crash occurrence that caused injury or property damage). The researcher made an assessment of the crash based on this concept of the causal chain. Drivers were interviewed to obtain information about the drivers' perception of the pre-crash event environment and the events leading up to the crash. The targeted information was captured mainly through four data elements: critical pre-crash event, movement prior to critical crash envelope, critical reason for the critical pre-crash event, and the crash-associated factors. The details about these elements are given in the NMVCCS manuals.[5] Briefly speaking, the critical pre-crash event documents the circumstance that led to this vehicle's first impact in the crash sequence and made the crash imminent. The movement prior to critical crash envelope refers to movement of the vehicle immediately before the occurrence of the critical event. The critical reason is the immediate reason for the critical event and is often the last failure in the causal chain (i.e., closest in time to the critical pre-crash event.) The critical reason can be attributed to the driver (distraction, driving too fast, panic, etc.), vehicle (tires/wheels, brakes, etc.), roadway, or atmospheric condition (rain, snow, glare, etc.). In addition to the critical pre-crash event, movement prior to critical crash envelope, and the critical reason underlying the critical event, the researcher documented the presence of other factors associated with the crash. Identifying a critical pre-crash event and the critical reasons underlying that critical event are integral to the information sought in this survey. However, the critical event, the critical reason underlying the critical event, or the associated factors should not be interpreted as the cause of the crash.

NMVCCS data spans a set of at least 600 variables or factors related to drivers, vehicles, roadways, and environment. The NMVCCS researchers collected information that includes crash narratives, photographs, schematic diagrams, vehicle information, as well as data from the event data recorder (EDR) whenever available. The file (SAS datasets) containing the NMVCCS data is being released for public use, together with other NMVCCS manuals. This file contains data on 6,949 crashes: 5,470 cases with non-zero weights and 1,479 cases with zero weights.

Limitations of NMVCCS Data

The items necessary for data collection in NMVCCS were identified using various studies and resources. Thus, the data resulting from this survey contains abundant information that can provide in-depth knowledge about the causal chain of a crash: movement prior to the critical crash envelope, the critical pre-crash event, and the critical reason for the critical pre-crash event. However, the data has certain limitations in terms of the sample size, usage, and interpretation of results.

The data have been collected under certain restrictions, as discussed earlier. This limited the number of crashes on which information could actually be collected. Since NMVCCS collected data on crashes that meet certain criteria, the estimates obtained from the NMVCCS database should not be compared with those from other databases such as NHTSA's General Estimates System (GES) or the NASS CDS. Also, it is important to note that as in any sample survey, national estimates obtained from NMVCCS data are subject to sampling errors.

With regards to usage, NMVCCS data is best suited for analyses aimed at crash/risk assessment, identification of possible crash contributing factors, etc., and not merely estimating rates. However, it is important to note that the data covering a period of two and a half years (July 3, 2005, to December 31, 2007) has been weighted and is best suited for statistical analyses, whereas the first six months (January 1, 2005, to July 2, 2005) of data has been assigned zero weight and can be used for clinical studies. Last but not least, caution needs to be used in analysis and interpretation of results that use the data of subjective nature. For example, the critical pre-crash event, the critical reason underlying the critical pre-crash event, or the associated factors should not be interpreted as the cause of the crash.

[4] Perchonok, K. "Accident Cause Analysis," Cornell Aeronautical Laboratory, Inc., July 1972.

[5] NMVCCS SAS Analytical Users Manual and NMVCCS Field Manual.

Few Results

A total of 5,470 crashes investigated during the period July 3, 2005, to December 31, 2007, have been used as a sample to obtain national estimates reported in this research note. Based on the weights attached to a sample of 5,470 crashes, at the national level, this sample represented an estimated 2,188,969 crashes involving 3,944,621 drivers and 4,031,075 vehicles. The corresponding coefficients of variation of these estimates are 15, 13.9, and 13.6 percent, respectively. In NMVCCS, the targeted information is collected only on the first three vehicles that are referred to as case vehicles. Of the estimated total number 4,031,075 of vehicles involved in 2,188,969 crashes, 3,894,507 are treated as case vehicles. Additionally, NHTSA is releasing a databook[6] of all the variables which presents unweighted counts and weighted estimates of the attributes, as well as the corresponding percentages. The NMVCCS data provide detailed information about other aspects of the crashes: pre-crash movement, critical pre-crash event, critical reason, and associated factors.

Movement Prior to the Critical Crash Envelope: The movement prior to critical crash envelope is the movement of a vehicle immediately before the occurrence of the critical event that made the crash imminent. Table 1 presents the frequency distribution of the variable that defines this movement. The statistics in this table show that in about 46 percent of the estimated 2,188,969 crashes, the vehicles were going straight prior to the occurrence of critical pre-crash event. The other prominent types of movement prior to critical crash envelope included negotiating a curve (21.0%) and stopped in the traffic lane (16.4%). While no vehicle was coded as entering a parking position, in about 0.5 percent of the crashes, the vehicles were involved when leaving a parking position. Also, turning left was a more common (1.6%) movement prior to critical crash envelope as compared to turning right (0.6%).

Table 1: Movement Prior to Critical Crash Envelope of Vehicles With Critical Reason

Movement Prior to Critical Crash Envelope	Unweighted	Weighted	
		Estimate	Percentage
Going straight	2,643	998,623	45.6%
Negotiating a curve	886	460,686	21.0%
Stopped in traffic lane	900	360,038	16.4%
Decelerating in traffic lane	241	99,854	4.6%
Avoidance maneuver to a previous critical event	183	65,335	3.0%
Changing lanes	143	46,097	2.1%
Turning left	122	34,675	1.6%
Accelerating in traffic lane	95	32,281	1.5%
Passing or overtaking another vehicle	58	22,352	1.0%
Turning right	35	13,768	0.6%
Starting in traffic lane	42	12,806	0.6%
Leaving a parking position	33	11,192	0.5%
Merging	17	5,808	0.3%
Backing up (other than for parking) position)	11	5,270	0.2%
Making a U-turn	15	6,333	0.3%
Other	10	3,442	0.2%
Unknown (includes driver not present cases)	39	10,410	0.5%
Total	**5,470**	**2,188,969**	**100%**

Estimates may not add up to totals due to independent rounding.
Data source: NMVCCS

[6] NMVCCS Data Book.

Critical Pre-Crash Event: Another important piece of information contained in the NMVCCS database is the critical pre-crash event which refers to the action or the event that puts a vehicle on the course that makes the collision unavoidable, given reasonable driving skills. It could be associated with the vehicle that was assigned a critical reason or with one of the other case vehicles. Table 2 presents unweighted and weighted frequencies and weighted percentages of the critical pre-crash events, based on the vehicles with the critical reasons. Of all the critical pre-crash events coded in NMVCCS, in about 36 percent of the crashes, the vehicles with critical reason were turning or crossing intersection. Among other such critical events, about 22 percent went off the edge of the road, and about 11 percent were running over the lane line. Regarding the role of the other case vehicle, in about 12 percent of the crashes the critical reason was the other vehicle stopped prior to the critical pre-crash event, 4.8 percent had the other case vehicle traveling in the same direction, and only a small percentage (0.1) traveling in the opposite direction. In about 2 percent of the crashes, the poor road condition was the critical pre-crash event.

Table 2: Critical Pre-Crash Event for Vehicles With Critical Reason

Critical Pre-Crash Event	Unweighted	Weighted	
		Estimate	Percentage
Turning or crossing at intersection	2,183	787,236	36.0%
Off the edge of the road	1,083	481,139	22.0%
Stopped	641	267,780	12.2%
Over the lane line	567	239,339	10.9%
Traveling too fast	207	109,118	5.0%
Traveling in same direction	317	105,717	4.8%
Poor road condition	81	45,632	2.1%
Traveling in opposite direction	7	2,510	0.1%
Other	82	33,725	1.5%
Unknown	4	1,155	0.1%
Total	**5,470**	**2,188,969**	**100%**

Estimates may not add up to totals due to independent rounding.
Data source: NMVCCS

Critical Reason for the Critical Pre-Crash Event: The third element in the causal chain is the critical reason which refers to the immediate reason for the critical pre-crash event and is often the last failure in the causal chain. The information about these critical reasons as collected in NMVCCS can be useful in evaluating the effectiveness of vehicle-based countermeasures in mitigating the effects of various driver performance, recognition, and decision errors. A critical reason can be attributed to a driver, vehicle, or environment. Normally, one critical reason is assigned per crash and as such, can be subjective in nature. Although the critical reason is an important element in the sequence of events leading up to a crash, it may not be the cause of the crash nor does it imply the assignment of fault to a vehicle, driver, or environment, in particular. The critical-reason-related statistics are presented in Table 3. It should be noted that in 110 crashes, due to various reasons, the critical reason could not be determined and hence not assigned to the driver, vehicle, roadway, or environment. About 41 percent of the driver-related critical reasons were recognition errors that include inattention, internal and external distractions, inadequate surveillance, etc. Of these, the most frequently occurring critical reason was inadequate surveillance that refers to a situation in which a driver failed to look, or looked but did not see, when it was essential to safely complete a vehicle maneuver. This critical reason was assigned to drivers in about 21 percent of crashes. Internal distraction as a critical reason was assigned to drivers in about 11 percent of the crashes.

About 33 percent of the driver-related critical reasons were decision errors that included too fast for conditions (8.4%), too fast for curve (4.9%), false assumption of others' actions (4.5%), illegal maneuver (3.8%), and misjudgment of gap or others' speed (3.2%). In about 10 percent of the crashes, the critical reason was a performance error, such as overcompensation (4.9%), poor directional control (4.7%), etc.

Among the nonperformance errors assigned as critical reasons to drivers in about 7 percent of the crashes, sleep was the most common critical reason (3.1%). The effectiveness of vehicle-based countermeasures used in mitigating the effects of various driver performance, recognition, and decision errors could be evaluated using this information.

Table 3: Critical Reasons for Critical Pre-Crash Event Attributed to Drivers

Movement Prior to Critical Crash Envelope	Unweighted	Weighted	
		Estimate	Percentage
Recognition Errors			
Inadequate surveillance	1,091	427,507	20.9%
Internal distraction	475	215,917	10.6%
External distraction	233	75,607	3.7%
Inattention (i.e., daydreaming, etc.)	217	73,059	3.6%
Other/unknown recognition error	99	51,717	2.5%
Subtotal	2,115	843,804	41.3%
Decision Errors			
Too fast for conditions	351	172,565	8.4%
Too fast for curve	176	100,155	4.9%
False assumption of other's action	265	91,762	4.5%
Illegal maneuver	212	78,112	3.8%
Misjudgment of gap or other's speed	217	64,547	3.2%
Following too closely	57	21,149	1.0%
Aggressive driving behavior	106	29,394	1.4%
Other/unknown decision error	333	126,449	6.2%
Subtotal	1,717	684,133	33.4%
Performance Errors			
Overcompensation	215	99,928	4.9%
Poor directional control	250	95,403	4.7%
Other/unknown performance error	27	7,766	0.4%
Panic/freezing	19	7,100	0.3%
Subtotal	511	210,197	10.3%
Non-Performance Errors			
Sleep, actually asleep	159	62,974	3.1%
Heart attack or other physical impairment	138	49,868	2.4%
Other/unknown critical non-performance	72	32,466	1.6%
Subtotal	369	145,844	7.1%
Other/Unknown	371	162,132	7.9%
Total	**5,083**	**2,045,577**	**100%**

Estimates may not add up to totals due to independent rounding.
Data source: NMVCCS

Crash-Associated Factors

A crash-associated factor is the factor that is likely to add to the probability of crash occurrence and can be attributed to any of the crash elements: driver, vehicle, roadway, or environment, or even to a combination of them. Of the estimated 3,944,621 drivers of case vehicles, about 18 percent were engaged in at least one interior non-driving activity. In addition, about 59 percent of the drivers were not engaged in any non-driving activity and in the case of about 23 percent of the drivers, the non-driving activity was unknown. Also, among the drivers of case vehicles, the drivers between the ages of 16 to 25 had the highest percentage (6.7%) of being engaged in at least one interior non-driving activity. The percentages of drivers of other age groups who were engaged in at least one interior non-driving activity were comparatively much lower. Talking on a cell phone or with a passenger was the most frequent interior non-driving activity associated with about 12 percent of the drivers of case vehicles. The highest percentage (4.0%) of the drivers of case vehicles engaged in talking belonged to the 16-to-25 age group, while only about 2 percent were between 26 to 35 and about 2 percent between 36 to 45.

Conclusions

NMVCCS is the first nationally representative survey of events and associated factors leading up to crashes involving light passenger vehicles. This research note presents some results from a nationally representative sample of 5,470 crashes investigated during a two-and-a-half-year period from July 3, 2005, to December 31, 2007. National estimates obtained from the crashes that were sampled show that of all the vehicles assigned a critical reason, about 36 percent were turning or crossing at an intersection just prior to the crash – characterized as the critical pre-crash event. An additional 22 percent of such vehicles ran off the edge of the road, and 11 percent failed to stay in the proper lane. The information pertaining to the crashes at intersections can be used in the design of intersection collision avoidance technologies. The data from run-off-the-road crashes can be beneficial in evaluating the effectiveness of ESC systems. The design of the emerging lane-departure warning systems can be enhanced by analyzing the data pertaining to vehicles that failed to stay in the proper lane.

The critical reason underlying the critical event is assigned by the NMVCCS researcher after on-scene evaluation of the potential problems related to the vehicle, roadway, environment, and driver. This is achieved through prompt investigations, interviews with the drivers, assessment of the vehicle components, and an evaluation of the roadway condition and geometry. Through such multifaceted evaluations, the critical reason for the critical pre-crash event was attributed to the driver in a large proportion of the crashes. Many of these critical reasons included a failure to correctly recognize the situation (recognition errors), poor driving decisions (decision errors), or driver performance errors. The information on such crashes will be greatly beneficial in designing vehicle-based crash avoidance technologies that can address the driver-related critical reasons like distraction and inattention, or loss of control of the vehicle.

Among the critical reasons attributed to drivers, about 41 percent were recognition errors, about 33 percent were decision errors, about 10 percent were performance errors, and about 7 percent were nonperformance errors. About 18 percent of the drivers were involved in at least one non-driving activity, with the majority (about 12%) engaged in conversing either with other passengers or on a cell phone. The effectiveness of emerging crash avoidance technology that use existing vehicle systems such as adaptive cruise control, braking systems, seat belt pretensioners, motorized seats, sunroofs, etc., in mitigating the effects of various driver performance, recognition, and decision errors can be assessed using this information.

The researchers, through their assessment of the vehicles, also assigned critical reasons to the vehicles. In such cases, failure of the tires/wheels was the most frequent vehicle-related critical reason followed by the failure of the braking system. The design and refinement of dashboard warning systems monitoring the status of critical vehicle elements such as the brake system, tire pressure, tread depth, etc., will benefit from such information.

In some cases, the NMVCCS researchers also assigned critical reasons pertaining to the roadway or the environment through an assessment of the roadway design, environmental conditions, and participant interviews. Among such cases, roads slick with ice and other debris was the most frequent roadway-related critical reason, followed by an obstruction to the driver's vision as attributable to flawed highway designs, poor signage, and inadequate infrastructure maintenance. The information collected on such crashes can help in the development and evaluation of crash avoidance technologies that adapt to adverse weather and roadway conditions.

The nationally representative sample of crashes collected through NMVCCS will enable statisticians, automotive engineers and human-factors researchers to perform more in-depth analyses of various aspects of crash avoidance. NHTSA believes that this may enhance its capability, as well as that of the automotive industry and other private organizations, in designing and evaluating the effectiveness of emerging crash avoidance technologies. This will also aid in making refinements to existing crash avoidance systems thereby supporting NHTSA's mission of saving lives, preventing injuries, and reducing vehicle-related crashes.