Table of Contents

Introduction 1

Objective of the project 1

MCMIS Crash file 2

Data files used in the evaluation 3

Overall reporting to MCMIS Crash file 4

Missing data 12

Summary and conclusions 16

Evaluations reported in future reports 17

Tables

Table 1 Cases Reported and Expected, MCMIS Crash File 1994-2000 4

Table 2 Cases Reported and Expected by Crash Severity, MCMIS Crash File 1994-2000 6

Table 3 States Reporting Levels by Year, 1994-2000 12

Table 4 Percent Unrecorded (Unknown) for Selected MCMIS Crash File Variables, 1994-2000 12

Table 5 Percent Unrecorded (Unknown) for MCMIS Crash File Hazmat Variables, where Hazmat Placard = "Y", 1994-2000 15

Figures

Figure 1 Percentage of Expected Cases Reported to MCMIS Crash File for Trucks, Buses, and All 5

Figure 2 Percentage of Reported MCMIS Crash File Involvements by Crash Severity, 1994-2000 8

Figure 3 Percentage of Truck Crash Involvements Reported to MCMIS by Crash Severity, 1994-2000 9

Figure 4 Percentage of Bus Crash Involvements Reported to MCMIS by Crash Severity, 1994-2000 10

Evaluation of the Motor Carrier Management Information System Crash File, Phase One

Daniel Blower
Anne Matteson

Center for National Truck Statistics
University of Michigan Transportation Research Institute

Introduction

This is the first in a series of papers presenting the results of an evaluation of the Motor Carrier Management Information System (MCMIS) Crash file undertaken by the Center for National Truck Statistics at the University of Michigan Transportation Research Institute. In this paper we cover overall reporting levels to the MCMIS file; sources of underreporting by state and crash severity; and the problem of missing data. Later reports will evaluate data consistency, whether the correct cases are reported, and the accuracy of the reported data. The focus of the present report is at the national level, a "top-down" analysis. But the MCMIS Crash file is fundamentally a compilation of cases reported by individual states, so later reports will provide more of a state-by-state analysis. This process will identify states that are providing accurate and complete data, as well as states whose data are incomplete.

Objective of the project

The purpose of the work is to understand the strengths and weaknesses of the MCMIS Crash file, to identify problems with the data, and to propose solutions. The ultimate goal of the work is to assist in the continuing effort to improve the MCMIS file, validating it for use in crash analysis, and to support other objectives of the Federal Motor Carrier Safety Administration.

This project will evaluate the MCMIS Crash file in terms of completeness, accuracy, and consistency. Completeness has two components: 1) Are all the cases reported that should be? 2) Are the data complete for each record or is there substantial missing data? Consistency refers to the internal consistency of the data for each record. Are the data describing a crash consistent or are there contradictions? Judging consistency is difficult because, with only twenty-two reported data items, there are not many comparisons to be made among variables. Accuracy is measured against other sources, when possible. Accuracy is really a measure of the consistency of MCMIS data with those other sources. Whether the data in the MCMIS

Crash file is "accurate" when it conflicts with information from other sources is really a judgment of the relative quality of the two sources, based on knowledge of how the data are compiled, quality-control measures, and experience with the sources.

It is important to remember that the MCMIS file is a compilation of data files reported separately by the 50 states and the District of Columbia. While the data elements are specified and defined at the national level, each state develops its own method to collect and report the data. Therefore, data quality and completeness issues are state-level problems and must be addressed first at that level. The MCMIS Crash file aggregates data collected by the individual states. The strengths of the Crash file are a result of the potential to provide a national census of trucks and buses involved in traffic accidents. The Crash file's weaknesses are a result of the varying quality of the data uploaded from the states, including inconsistent, inaccurate, and missing data.

MCMIS Crash file

The Motor Carrier Management Information System (MCMIS) Crash file was developed by the Federal Motor Carrier Safety Administration (FMCSA), an agency of the U.S. Department of Transportation (DOT) responsible for monitoring and developing safety standards for commercial motor vehicles operating in interstate commerce. It was designed to replace the older MCS-50T data, which was not comprehensive enough to allow for research on motor carrier safety problems. A major virtue of the Crash file is that it contains records on all trucks and buses involved in a reportable crash, not just the reportable crashes of interstate carriers. A second advantage of the Crash file over the old MCS-50T data is that the data are not self-reported by carriers, but instead provided directly by the states. The combination of these two changes means that the MCMIS Crash file has the potential to provide a census of all trucks and buses involved in a traffic crash. The MCMIS Crash file thus is potentially a very valuable resource for FMCSA. Crash file data are used in the SAFER system to evaluate and compare the safety status of carriers. The Crash file also may serve as a census file of traffic crashes involving trucks and buses used for targeted research purposes.

The MCMIS Crash file contains data from state police crash reports involving drivers and vehicles of all motor carriers (interstate and intrastate) operating in the U.S. It now includes 22 data elements that the states are required to supply, along with several other data elements that are supplied by linking the state-supplied data to other files. The state-supplied data are based on a uniform set of crash data elements developed through the National Governors' Association (NGA). The data collected are entered by the states into a system called SAFETYNET, through which the data are transmitted to the FMCSA and entered into the MCMIS system.

Beginning January 1, 1994, states participating in the Motor Carrier Safety Assistance Program were required to report through the SAFETYNET system a standard set of data items on all trucks and buses involved in traffic crashes that met a specific severity threshold. Reporting requirements were designed to be simple and easily applied. Reportable crashes include one or more of the following vehicle types:

- A truck (used primarily for the transportation of property) having at least six tires in contact with the road surface
- A vehicle displaying a hazardous material placard
- A bus with seating for at least nine (15 before 2001) people, including the driver

The severity criteria for a reportable crash are equally straight-forward. Reportable crashes include one or more of the following factors:

- A fatality
- An injury requiring transport for immediate medical attention
- A vehicle towed from the scene as a result of disabling damage suffered in the crash

These straightforward definitions of trucks, buses, and reportable crashes facilitate uniform and comprehensive reporting by the states. Most states have implemented collection of the NGA data elements either by modifying their existing police accident reporting forms or by developing supplemental forms to be filled out for vehicles and crashes meeting the reporting criteria.

Data files used in the evaluation

To gauge levels of reporting to MCMIS, a comparison data set was developed from several applicable files. These files include the General Estimates System (GES) file and the Fatality Analysis Reporting System file, both compiled by the National Center for Statistics and Analysis in the National Highway Traffic Safety Administration (NHTSA); and the Trucks Involved in Fatal Accidents (TIFA) file and Buses Involved in Fatal Accidents (BIFA) file, both compiled by the University of Michigan Transportation Research Institute (UMTRI).

In each file, vehicles and crashes that would qualify for the MCMIS Crash file were selected. With respect to vehicles, all trucks and buses meeting the definitions used in the MCMIS Crash file were taken. For trucks, this includes trucks with at least two axles and six tires, or other vehicles placarded to carry hazardous materials. For buses, this involves buses with capacity for 15 or more passengers (prior to 2000) along with a driver, to the extent this could be determined in each data file used. Only crashes meeting the MCMIS crash severity

1995, the second year for which the states were required to provide full reporting. The level of reporting has stayed consistently low.

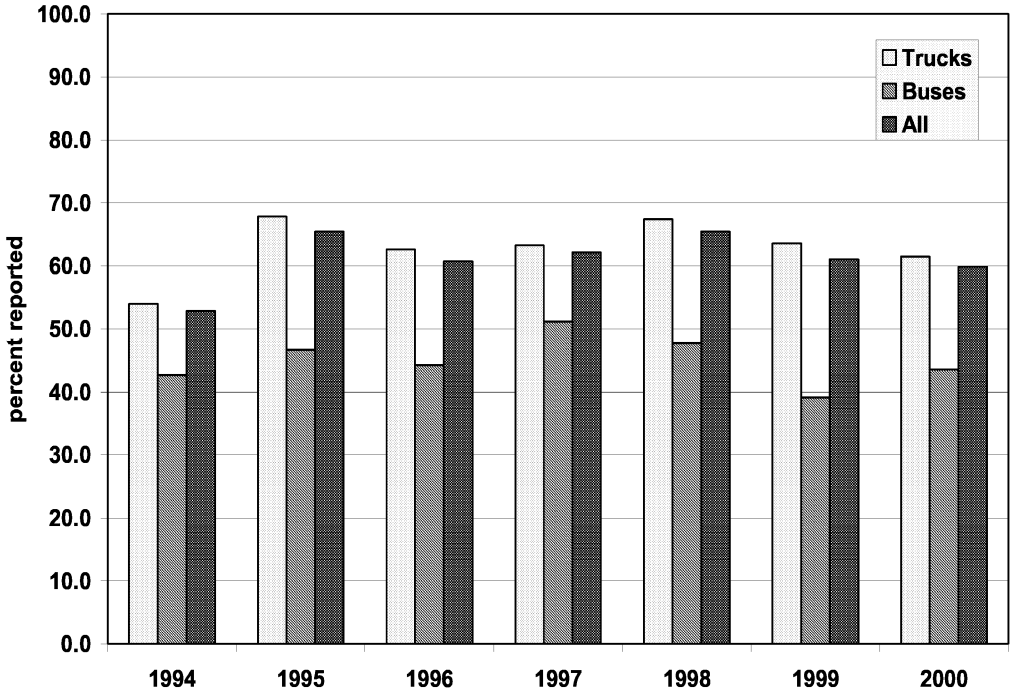

Figure 1 Percentage of Expected Cases Reported to MCMIS Crash File for Trucks, Buses, and All

Reporting of truck involvements is somewhat more complete than for buses. Again, only about 55% of expected truck cases were reported in the first year of reporting. The percentage increased to almost 68% in the second year, but never moved beyond that and in fact has only ranged between about 62% and 67% of expected cases. The results are even worse for reported bus crash involvements. Only 43% of expected bus involvements were reported in 1994. The percentage of expected bus involvements reported increased to about 51% in 1997 but then declined thereafter to less than 40% in 1999.

Overall, only about 44% of expected bus involvements are reported to the MCMIS Crash file. For trucks, the rate of reporting is somewhat better but still quite low at 63%. The combined reporting rate is 61%. Thus it appears that the MCMIS Crash file is still significantly below full reporting. Unfortunately it also appears that there is no trend toward fuller reporting. The reporting levels reached in the second year of the MCMIS file have stayed fairly constant since. Moreover, the level of reporting for bus crash involvements is significantly below that of truck involvements.

The completeness of crash involvement reporting varies by crash severity, and also differs for trucks and buses. Fatal and injury crash involvements tend to be reported more completely than towaway involvements. Truck crash involvements tend to be more

completely reported than bus involvements. In 2000, the number of reported MCMIS truck fatal involvements nearly equaled the number that actually occurred, according to the FARS file. Similarly, in 1997 and 1998, the number of truck involvements in crashes with an injury transported for immediate treatment was nearly equal to the expected number, though reporting trailed off in 1999 and 2000. However, in every year and for both trucks and buses, the reporting of towaway involvements to MCMIS is significantly lower than expected, in most years less than half. Bus reporting is more incomplete for every crash year and for every crash severity.

Table 2 Cases Reported and Expected by Crash Severity, MCMIS Crash File 1994-2000

	Fatal involvements					
	Trucks		Buses		Total	
	Reported	Expected	Reported	Expected	Reported	Expected
1994	2,810	4,801	82	266	2,892	5,067
1995	3,030	4,640	97	294	3,127	4,934
1996	3,368	5,001	139	350	3,507	5,351
1997	3,623	5,136	146	314	3,769	5,450
1998	3,963	5,202	137	308	4,100	5,510
1999	4,585	5,233	183	333	4,768	5,566
2000	4,923	5,298	232	357	5,155	5,655
Total	26,302	35,311	1,016	2,222	27,318	37,533
	Injury involvements					
	Trucks		Buses		Total	
	Reported	Expected	Reported	Expected	Reported	Expected
1994	42,000	57,000	4,000	10,000	46,000	66,000
1995	46,000	52,000	5,000	8,000	51,000	60,000
1996	53,000	59,000	6,000	10,000	59,000	69,000
1997	55,000	57,000	6,000	6,000	61,000	64,000
1998	53,000	52,000	5,000	8,000	58,000	60,000
1999	51,000	66,000	5,000	10,000	56,000	76,000
2000	49,000	62,000	6,000	10,000	54,000	72,000
Total	348,000	406,000	36,000	62,000	384,000	467,000
	Towaway involvements					
	Trucks		Buses		Total	
	Reported	Expected	Reported	Expected	Reported	Expected
1994	35,000	88,000	3,000	6,000	38,000	94,000
1995	38,000	70,000	3,000	8,000	40,000	79,000
1996	40,000	88,000	2,000	8,000	42,000	96,000
1997	41,000	95,000	2,000	9,000	44,000	105,000
1998	43,000	91,000	2,000	7,000	45,000	99,000
1999	49,000	94,000	3,000	9,000	52,000	103,000
2000	51,000	102,000	3,000	11,000	54,000	112,000
Total	298,000	629,000	18,000	59,000	315,000	687,000

Note: Reported and expected injury and towaway involvements rounded to nearest thousand.

Table 2 shows the reported and expected cases in the MCMIS Crash file separately for fatal, injury, and towaway involvements, and for trucks and buses. Fatal and injury crash

involvements tend to be reported more completely than towaway involvements, and truck crash involvements tend to be more completely reported than bus involvements. In 2000, the number of reported MCMIS truck fatal involvements nearly equaled the number that actually occurred, according to the FARS file. Similarly, in 1997 and 1998, the number of truck involvements in crashes with an injury transported for immediate treatment was nearly equal to the expected number, though reporting trailed off in 1999 and 2000. However, in every year and for both trucks and buses, the reporting of towaway involvements to MCMIS is significantly lower than expected, in most years less than half. Bus reporting is more incomplete for every crash year and for every crash severity.

Figure 2 displays the overall level of reporting to MCMIS by crash severity. Surprisingly, for the first five years (1994-1998) of required reporting, a higher percentage of injury crash involvements than fatal involvements was reported. One might expect that fatal involvements would be more likely to be reported, given their seriousness. But for each year from 1994 to 1998, a higher percentage of injury involvements was reported. However, the figure shows that the proportion of reported fatal involvements has increased each year, as would be expected, and that in the most recent full year of reporting, over 90% of fatal crash involvements, including both truck and bus, were reported. This trend of increasing compliance with reporting requirements is both expected and welcome. However, reporting completeness for injury crash involvements has actually decreased substantially in the last two years. And it should also be noted that towaway reporting started low and has not increased much. While the improvement of the reporting of fatal involvements is gratifying, a census file of fatal crash involvements is already available in the FARS, TIFA, and BIFA files. It is exactly the injury and towaway crashes that require a census file, which the MCMIS Crash file has failed to produce to date.

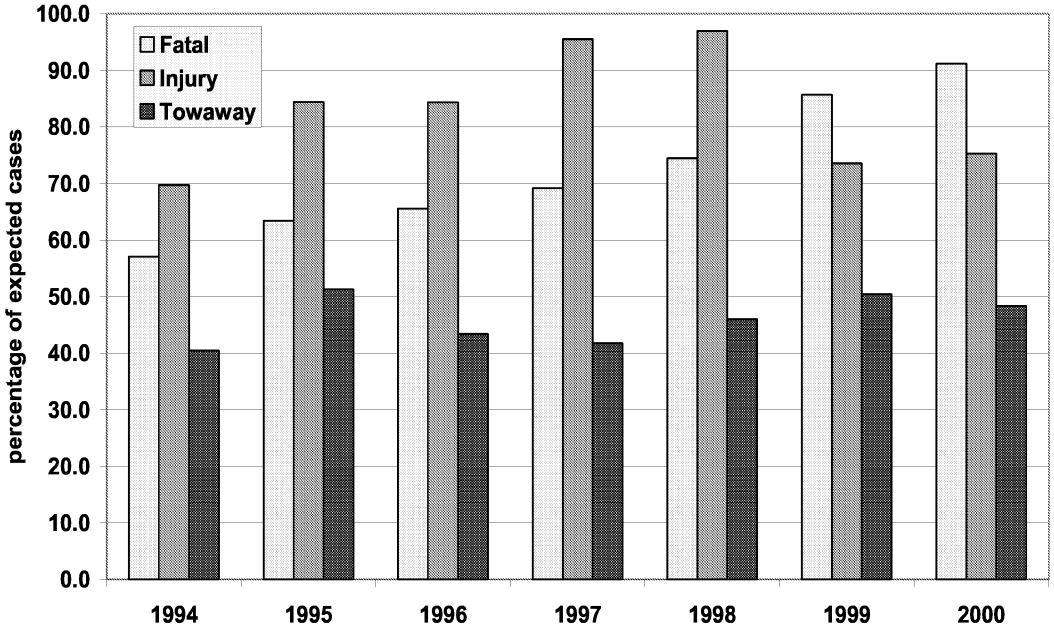

Figure 2 Percentage of Reported MCMIS Crash File Involvements by Crash Severity, 1994-2000

Reporting levels for crash involvements of trucks are very similar to the overall level of reporting, because truck involvements are about 93% of all involvements reported to MCMIS. Nevertheless, it is useful to break out truck involvements separately. Figure 3 shows that the reporting of truck fatal involvements has improved in each of the years of MCMIS. In the 2000 crash year, almost 93% of truck involvements in fatal crashes were reported. On the other hand, while the proportion of injury crash involvements reported improved steadily until in 1998, when 103% of the expected number of involvements were reported, the proportion dropped below 80% in both 1999 and 2000. Towaway reporting fluctuated at a low level, breaking 50% in 1995, 1999, and 2000, but averaging under 48%.

Table A-2 in the appendix shows the level of reporting of bus fatal crash involvements for each of the states and the District of Columbia. The underreporting of buses involved in fatal crashes is much more extensive than for trucks. Over the three years represented in the table, almost half of fatal bus involvements were not reported. Thirty-three of the states and the District of Columbia misreported the number of fatal bus involvements by more than 20% each year on average. In contrast, Alabama, Delaware, Hawaii, Idaho, Nebraska, New Hampshire, North Dakota, Pennsylvania, South Dakota, and Vermont all reported within 5% of the correct number. It should be noted, however, that Vermont and South Dakota had no fatal bus crashes in the three years and Nebraska, New Hampshire, and North Dakota each had only one.

Judgments about reporting levels for nonfatal crashes are more subjective, since there is no file to compare to MCMIS data on a state-by-state basis.[1] However, by examining the level of reporting from year to year for each state, it is possible to get some sense of the direction the state is going. Table 3 tabulates judgments on the level of states' reporting efforts. These judgments were made by looking at total MCMIS case counts for each state by month and year from 1994-2001. The following definitions were used in determining each state's placement:

- Not reporting: The state reported no cases for that year.
- Nominal reporting: Annual case counts were less than 5% of the average number for other years.
- Partial reporting: Case counts were definitely below those for other years. In some cases, all months showed lower counts, and in other instances only a couple of months were responsible for the difference.
- Efforts at full reporting: Reporting looks fairly consistent with other years and monthly counts are stable. There could be some variation observed, but not enough to suggest that partial reporting is occurring.

As the reader will observe, all states have made some effort at reporting in each of the past five years. Moreover, the number of states just partially reporting has steadily decreased, so that in the 2000 data year, only nine states were judged to be only partially reporting. While some of the other trends reviewed thus far have been discouraging, at least insofar as participation in the MCMIS process is concerned, it appears that more states are participating.

[1] In a later phase of the project, however, we will compare reporting of nonfatal crashes for states for which we have the complete crash files.

Table 3 States Reporting Levels by Year, 1994-2000

Year	Not reporting	Nominal reporting	Partial reporting	Efforts at full reporting
1994	1	2	26	22
1995	0	1	26	24
1996	0	0	21	30
1997	0	0	18	33
1998	0	0	14	37
1999	0	0	7	44
2000	0	0	9	42

Missing data

Table 4 shows missing data rates for the most important variables in the MCMIS Crash file. The fact of missing data does not necessarily indicate a problem in every instance. For example, between 32% and 45% of the cases are missing a DOT number, but a DOT number is only issued to companies that operate trucks in interstate commerce or that carry hazardous materials (hazmat). Missing DOT numbers may simply indicate an intrastate carrier. Likewise, the lack of a crash city name could indicate that the crash took place outside of city limits.

In general, most of the baseline variables have reasonably low rates of missing data. Cargo body type, vehicle configuration, whether the vehicle was a truck or bus, number of fatalities, number of injuries, and number of vehicles in the crash all have very low rates of missing data. Data are complete for both number of fatalities and number of injuries, and only a few cases have obvious miscodes for those variables. Six cases are coded with more than 70 fatalities, including one with 970; and four cases are coded with more than 100 injuries, including one with 260 injuries, another with 630 injuries, and one with 998 injuries! But considering the roughly 727,000 cases reported over the time period, a handful of cases is insignificant. Other variables that are descriptive of the crash also have reasonably low rates. Driver license state, light condition, weather, road surface condition, and road access control all have rates of missing data under 10% in most years.

Table 4 Percent Unrecorded (Unknown) for Selected MCMIS Crash File Variables, 1994-2000

	Year						
MCMIS variable	1994	1995	1996	1997	1998	1999	2000
Apparent driver condition	21.2	20.4	15.9	15.0	15.9	22.1	43.4
Axles	14.9	11.5	10.5	10.3	13.2	22.1	36.3
Cargo body type	0.0	0.0	0.3	0.1	0.1	0.6	2.1
Carrier address/state	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Carrier name source	30.7	36.3	33.2	32.1	34.6	38.3	56.0
Census number (DOT number)	45.0	45.3	37.9	36.2	33.6	32.9	34.5
Census state	97.3	95.4	79.8	39.7	38.9	40.2	45.1
Citation issued	99.0	99.6	97.9	97.1	96.7	93.8	69.6

The vehicle identification number (VIN) is also missing in about 20% of the cases. This is particularly unfortunate because VINs contain a wealth of information about the truck, including make, model, model year, gross vehicle weight rating, number of axles (of the power unit), and cab style.

Gross vehicle weight, gross vehicle weight rating, and the hazardous materials-related variables also have high rates of missing data. Gross vehicle weight and gross vehicle weight rating both are difficult to code and require some specialized knowledge on the part of the police officer. The gross vehicle weight rating of a truck is generally stamped on a plate fixed to the frame of the door or some other location, but without specialized training, a reporting officer may not be able to locate or interpret the codes. Gross vehicle weight is the actual weight of the truck at the time of the crash. Truck drivers and operators that either haul goods for hire or often operate close to gross weight limits often know this, since moving goods is how they make their living and the purpose for which they use the truck. Operators of trucks in other applications may not know and, unless the reporting police officer can locate a weigh station receipt or some other documentation, determining the weight is not generally feasible.

In addition to the gross vehicle weight missing from almost a quarter of the cases, the data coded is unreasonable or unlikely in a significant fraction of the cases. Over nine percent of the cases are assigned a gross weight over 80,000 pounds, while only around 3% of tractor-semitrailers involved in fatal crashes have gross weights that high. Almost 2% of the MCMIS crash cases are coded with gross weights over 150,000 pounds, which is extremely unlikely. It is more likely that such great weights are miscodes.

While coding weight variables may be inherently difficult, high rates of missing data on hazmat variables are less understandable. Hazmat placards are designed to be visible and are a well-known warning sign. As in the case of driver citations, the gross missing data rate is misleading, since it is likely that the police officer just left the question blank if a truck did not have a placard.

There are five variables that record the presence and type of hazmat cargo. To determine missing data rates, we used the "hazmat placard" to identify MCMIS vehicles carrying hazardous cargo. The hazmat placard variable, of course, is subject to missing data and inaccuracies, like all other variables. There were 406 cases over the period from 1994 to 2000 where the hazmat placard variable was either blank or "no," but there was valid data in the other variables on hazmat. These 406 cases amount to 1.4% of the 28,195 vehicles recorded with a hazmat placard, which, under the circumstances, is a reasonably low error rate. In any case, it is necessary to use one variable against which to measure missing data, and the hazmat placard variable appears to be the most appropriate.

Table 5 shows the percentages of cases with missing data on the hazmat detail variables, where the vehicle was coded as displaying a hazmat placard (hazmat placard="Y"). Missing data on all the variables is quite high. For the hazmat class, missing data averages around one-third of the cases. The more detailed identification of the material, the four-digit hazmat ID number, is missing in about 22% to 79% of the cases; there was no information about hazmat cargo spill in 13% to 73% of the cases; and the material's name was not recorded in 16% to 65% of the cases.

Table 5 Percent Unrecorded (Unknown) for MCMIS Crash File Hazmat Variables, where Hazmat Placard = "Y", 1994-2000

	Year						
MCMIS hazmat variable	1994	1995	1996	1997	1998	1999	2000
Hazmat class	37.8	30.3	25.2	25.5	38.3	81.3	36.4
Hazmat 4-digit ID number	28.2	26.8	27.4	22.3	47.5	79.0	27.4
Hazmat cargo spill	16.2	13.5	13.1	18.3	39.5	72.8	18.6
Hazmat name	61.7	58.9	60.9	65.1	42.3	15.6	50.1

Though missing data rates on the hazmat variables are generally quite high, they vary considerably from year to year. In fact, 1999 stands out both for high and low rates of missing data. In 1999, rates were unusually high for hazmat class, hazmat four-digit ID number, and hazmat cargo spill. But the rate of missing data on hazmat name was unusually low. What accounts for these anomalies? In a word, California. It appears that there was some problem with the data uploaded by California in 1999. The number of cases reported with a hazmat placard jumped by several thousand, and most of the new cases came from California. Examining the hazmat name variable provides a clue to the error. As Table 5 shows, missing data for hazmat name was unusually low in 1999. Most of the difference is explained by over 6,200 California cases in which the hazmat name was reported as "N". This indicates that data from California for the hazmat variables was probably reported in the wrong fields, resulting in a large number of cases incorrectly coded with hazmat placard="Y". Probably the "N" reported in the hazmat name field should have been recorded in the hazmat placard field, though this is just speculation, as we do not have information on the details of how fields are reported from California to the MCMIS file.

Nevertheless, even discounting California cases in 1999, rates of missing data on hazmat variables are very high, and severely limit the utility of these variables for analysis. In roughly one-third of cases, it is not possible to identify the type of hazmat involved in a crash, even in general terms. Moreover, the missing data rate for the hazmat four-digit ID number does not account for the additional 5.0% of cases with invalid ID numbers, so that, leaving out the anomalous year of 1999, only about 62% of hazmat placarded vehicles have valid information on the type of hazardous materials carried.

Summary and conclusions

Reporting to the MCMIS Crash file has stabilized at a relatively low level. There has been continuous improvement in the reporting of trucks involved in fatal crashes, but less improvement for fatal bus involvements. Recent years have seen about 90% of truck involvements in fatal crashes reported, though only about 65% of bus involvements. The upward trend, particularly for trucks, is welcome, but tempered by the fact that useful census files for fatal crashes already exist. Involvements in injury crashes is reported at a higher rate than for fatal crashes, but the relative overreporting of injuries could be due to failing to apply the injury severity threshold correctly. The injury reporting criterion is any injury transported for medical treatment. The relative overreporting of injury crash involvements could be due to reporting all injury crash involvements, rather than just those with injuries transported for treatment. The "transported for treatment" criterion may be too difficult to apply in practice. The relative underreporting of towaway crashes could just be a neglect of this non-serious accident type. More consistent reporting may be achieved by a simpler reporting threshold.

Buses are generally underreported, both overall and in comparison with trucks, at all severity thresholds. Most of the emphasis in crash reporting has been on trucks; so a heightened emphasis on buses may be necessary to improve reporting levels. The recent change in the definition of a reportable bus to eight passengers plus a driver will probably make reporting more difficult, because the definition of a bus in terms of passengers overlaps family vans, which may result in confusion.

It appears that an increasing proportion of states are reporting to the MCMIS crash file, but, seven years after inception, only 42 of the 50 states and the District of Columbia are apparently making efforts at full reporting. The number of states has been increasing, but still much progress remains to be made.

Missing data rates are reasonable for variables that provide simple descriptive information about the accident scene, but unacceptably high for details about the vehicle and driver. Most driver licensing and citation information is unavailable. The vehicle identification number, key to important physical details about the vehicle, is missing in about 20% of cases. Gross weight is missing or unreasonable in about 35% of cases. Considering variables on hazardous materials, it appears that the hazmat placard variable is generally reliable in identifying hazardous materials cargoes, but the variables that provide details about the cargo are missing between 20% and 38% of the time. Simple computer checks could flag cases with unreasonable or missing information at the state level. In fact, such checks are the only practical way to improve reporting.

Evaluations reported in future reports

In the next report, we will focus on the results of matching MCMIS Crash file cases with individual cases in other crash files, primarily UMTRI's Trucks Involved in Fatal Accidents (TIFA) and Buses Involved in Fatal Accidents (BIFA) files. Both the TIFA and BIFA files are census files (all cases) of the respective vehicle type's involvement in fatal traffic accidents.

The tables in the Appendix show reporting levels in terms of the gross number of cases reported, but not whether the correct cases were reported. Since the TIFA and BIFA files include all trucks and buses, respectively, involved in a fatal crash, the MCMIS Crash file should include, in theory, each case reported in TIFA and BIFA. The next paper will report on an effort to locate each individual TIFA and BIFA case in the MCMIS Crash file. The results will identify cases correctly reported in MCMIS, cases not found in MCMIS, and cases reported in MCMIS that do not appear in either the TIFA or the BIFA files. In addition, the matching process will allow us to evaluate the accuracy of MCMIS file variables, by comparing data on the case in the MCMIS file with comparable information in the TIFA and BIFA files. Patterns of underreporting and inaccurate reporting will be identified and suggestions will be made for improving reporting to the MCMIS Crash file.

The evaluation described above is limited to fatal crashes, because only vehicles involved in fatal crashes are included in TIFA/BIFA. Subsequent evaluations will match MCMIS Crash file cases with selected state crash files. UMTRI has a library of several state crash files that can be used to match with cases in MCMIS. Once again, differences between the content of the state files and the corresponding record in the MCMIS Crash file will suggest quality-control measures that can improve the consistency and accuracy of the file.

Appendix

Table A-2 Fatal MCMIS Bus Involvements by State Compared with TIFA and FARS, 1998-2000

State	1998 MCMIS	1998 FARS	1998 Under/ over	1999 MCMIS	1999 BIFA	1999 Under/ over	2000 MCMIS	2000 FARS	2000 Under/ over
Alabama	6	7	-1	2	2	0	3	3	0
Alaska	0	1	-1	0	0	0	0	3	-3
Arizona	0	4	-4	0	6	-6	2	7	-5
Arkansas	1	3	-2	0	3	-3	0	2	-2
California	7	38	-31	17	45	-28	33	38	-5
Colorado	4	9	-5	5	5	0	5	7	-2
Connecticut	2	3	-1	3	4	-1	2	2	0
Delaware	1	1	0	3	3	0	1	1	0
Dist of Col.	1	1	0	2	2	0	1	3	-2
Florida	9	23	-14	14	26	-12	11	38	-27
Georgia	0	14	-14	2	8	-6	6	13	-7
Hawaii	3	3	0	1	1	0	4	4	0
Idaho	0	0	0	2	2	0	0	0	0
Illinois	3	14	-11	0	12	-12	1	13	-12
Indiana	2	4	-2	3	6	-3	3	4	-1
Iowa	1	2	-1	1	1	0	1	5	-4
Kansas	1	1	0	2	4	-2	4	5	-1
Kentucky	3	6	-3	0	1	-1	2	3	-1
Louisiana	0	4	-4	5	4	1	1	2	-1
Maine	0	1	-1	0	1	-1	0	0	0
Maryland	0	9	-9	0	8	-8	0	5	-5
Massachusetts	2	3	-1	0	2	-2	2	3	-1
Michigan	2	8	-6	3	9	-6	0	17	-17
Minnesota	3	5	-2	2	5	-3	6	9	-3
Mississippi	0	1	-1	2	2	0	2	2	0
Missouri	3	3	0	5	7	-2	5	8	-3
Montana	1	2	-1	1	1	0	0	0	0
Nebraska	1	1	0	1	0	1	1	0	1
Nevada	0	4	-4	4	4	0	2	6	-4
N. Hampshire	0	0	0	0	0	0	1	1	0
New Jersey	8	14	-6	3	17	-14	4	13	-9
New Mexico	1	3	-2	1	5	-4	2	5	-3
New York	23	29	-6	32	36	-4	28	33	-5
North Carolina	3	5	-2	1	4	-3	0	7	-7
North Dakota	0	0	0	0	0	0	1	1	0
Ohio	0	11	-11	4	12	-8	30	9	21
Oklahoma	1	1	0	1	3	-2	5	8	-3
Oregon	0	2	-2	0	6	-6	0	0	0
Pennsylvania	13	13	0	22	23	-1	16	17	-1
Rhode Island	1	1	0	0	1	-1	0	2	-2
South Carolina	0	3	-3	0	6	-6	1	2	-1
South Dakota	0	0	0	0	0	0	0	0	0
Tennessee	1	8	-7	2	2	0	2	7	-5
Texas	19	22	-3	17	18	-1	29	28	1
Utah	2	2	0	3	3	0	2	3	-1
Vermont	0	0	0	0	0	0	0	0	0
Virginia	1	5	-4	3	7	-4	1	3	-2
Washington	0	4	-4	7	10	-3	3	4	-1
West Virginia	0	2	-2	2	2	0	1	1	0
Wisconsin	8	8	0	5	4	1	6	7	-1
Wyoming	0	0	0	0	0	0	2	3	-1
Total	137	308	-171	183	333	-150	232	357	-125