Updated Review of Crash Data Elements for Motorcycles

Gary L. Winn, Ph.D., CHST
Mike Carr, MS
Department of Industrial and Management
Systems Eng.
College of Engineering and Mineral Resources
West Virginia University
P.O. Box 6070
Morgantown, WV 26506 - 6070
e-mail: gwinn@wvu.edu

Gary L. Winn

Updated Review of Crash Data Elements for Motorcycles

ABSTRACT

The purpose of this paper is to examine state police traffic accident reports (PARs) as they relate to motorcycle and motorcyclist post-crash data. Using NHTSA's 1995, 1997 and 1999 compendia of state crash forms, we counted direct and indirect references to motorcycles and operators. Because we found so few motorcycle oriented data elements, we argue that the creation of secondary data bases may be of limited reliability using PARs alone. We propose five alternatives to using PARs at face value. These alternatives include the use of "intelligent programmers", heuristic softwares and the completion of a motorcycle and motorcycle operator data needs study.

KEYWORDS: motorcycle safety; crash data; data elements

INTRODUCTION:

The issue of motorcycle safety is rarely absent from national attention. State legislatures and behavioral scientists seek to improve traffic safety with a variety of program initiatives. Traffic engineers seek to improve safety via highway construction and through methods such as roadway maintenance and even traffic calming. As electronic improvements in data collection, transfer and linkage occur, users of secondary, or complied, statistics seek primary data sets as their starting points; consumers of secondary data sets include insurance organizations, public health advocates and long-term care operators. Data flow models are available which maximize a particular state's data resources, efficiencies and existing infrastructure (Carr and Winn, 1996). However, much public motorcycle policy is based upon secondary data sets created principally using PARs.

In West Virginia, nearly a decade of research work has gone into the improvement of basic tools of crash data collection and database linkages (Winn, 1997a; Winn et. al, 1997b). Many improvements in data collection simply tightened up existing definitions of data elements and added in a few more for clarity. In a large, federally-funded project which ended in 1997, we found that many fundamental improvements in crash data quality and subsequent speed of collection and transfer can even be had by *technica*l improvements in the form itself (for example, making the crash form scannable, eliminating at least some errors due to key punching error or by making definitions intuitive so time consuming reference books are not necessary). We made some *cultural* improvements (fine-tuning officer training on special data elements which are used only rarely, for example). Many improvements in data quality were made by just asking the officers for suggestions. For example, we simply asked how to make the data flow easier and logical, making sure the most important data were gathered first, because officers are called upon to direct traffic and assist victims during data collection. We attempted to make enforcement feel ownership in the crash data system and it paid off: we made over a hundred improvements in the crash form and empirically-improved levels of user satisfaction. As we were obligated to "benchmark" our work (model our new form on best practice outside West Virginia), we used the 1995, and then the 1997 compendia of PARs from the 50 states and US territories published by the National Highway Traffic Safety Administration (NHTSA). A 1999 version is also available.

In researching the state PARs, we noted a dearth of information about low-registration vehicles, particularly commercial trucks and motorcycles. The motorcycle PAR comparison reported here began with the benchmarking study of state PARs undertaken in 1996 to glean best practice in the creation of a completely new West Virginia crash form for all vehicles and persons.

PROBLEM STATEMENT:

The objective of this project was to compare state PARs for references to motorcycle operators and motorcycles as vehicles and to recommend alternatives to using PARs at face value.

RESEARCH METHODS AND DATA COLLECTION:

The available data was considered to be the 1995 and 1997 version of state PARs compiled by the National Highway Traffic Safety Administration (US DoT, 1995; USDoT, 1997). We counted *direct references* to motorcycles or operators or motorcycle passengers if the reference was immediately obvious to our panel of three reviewers. An example of a direct reference might be "cc" or helmet used". An *indirect reference* to motorcycles or operators or passengers was counted if the data element could be inferred or otherwise passed the majority of our review panel of three. An example of an indirect reference to motorcycles or operators or passengers would be "VIN", which stands for vehicle identification number. The VIN can be decoded into four or five data elements unique to that vehicle which then available for subsequent analysis.

At the macro level, we first noticed that there are very few visual similarities among state PARs for motorcycles or anything any sort else. We then became aware that there were very few references of any kind to motorcycles. Table 1 illustrates motorcycle references which are shared by at least 40 percent of the states and the District of Columbia.

Table 1
Top Four Direct or Indirect Crash Form References

Reference	Frequency
Make, model or style of motorcycle	50/51 states
Helmet use	43/51 states
Vehicle Identification Number	31/51 states
Motorcycle Operator License	20/51 states

Altogether, there were only eleven direct or indirect references which could be classified as related to motorcycles. The remaining seven direct or indirect references included, endorsement type (18 states), operator endorsement (16 states), motorcycle damaged areas (5 states), motorcycle-only supplement (3 states), total number of occupants (3 states), crash diagram for motorcycle (2 states), motorcycle cc's (2 states).

Some PARs are extremely hard to read even under good light: most have 5 and 6 point type and one had 4 point type. We wondered how about data quality problems stemming from an officer to trying to read the form at night under a dome light in a police cruiser. Some PARs have instructions and some require the use of a manual, not always available with the PAR..

Arkansas had a motorcycle-only supplement in 1995 and it was what the review panel would call exemplary. It included 14 separate crash data elements, but the motorcycle-only supplement is not included in the 1997 or 1999 NHTSA state crash form compendia nor is the supplementary form on the NHTSA web site showing the state forms. That URL is: http://www.nhtsa.dot.gov/people/perform/trafrecords/forms/in.pdf.

We hypothesized that size of a state's motorcycle population might influence the number of motorcycle data elements. Texas, one of the nation's largest motorcycle registration state, has only four of the eleven identifiable motorcycle crash data elements. California, always near the top in registrations, has only five of the eleven motorcycle elements.

We hypothesized further that there would be a positive relationship between the length of state PAR and the number of motorcycle crash data elements. Tennessee's crash form runs to 12 pages with one whole page listing up to twelve other people involved. However, Tennessee's motorcycle data elements are very oblique: "body type" could, we presume, mean motorcycle but that's about as close as we get after Tennessee's four useful data elements. There are no instructions supplied to assist the reader. North Dakota's crash form is nine pages long but only includes two motorcycle data elements: a crash silhouette and driver/passenger seating position.

We also examined very short crash forms to look for a relationship, but again there was none. California's crash form is only three pages long and has five of the most frequent data elements. Ohio's PAR is only two pages long and also includes three of the same eleven.

CONCLUSIONS AND DISCUSSION

Our review of state PARs suggests that there is very little similarity among state crash forms for motorcycle data. There are only four data elements that even twenty or more states collect and only a total eleven motorcycle elements collected which can even indirectly be associated with motorcycles or their operators or passengers. The useful data elements across state PARs are limited to make/model, safety helmet use, VIN, and driver license.

While it is desirable and fairly easy to collect multiple PARs for analyses of motorcycle programs, we must conclude that there is precious little actual motorcycle data in those PARs when using them at face value. This is of even greater concern because data sets are electronically transferred so easily. Are the data elements comparable across states? Are data definitions comparable? Just because it is easy to transfer data sets quickly does not mean the sets are of any reliability.

Some alternatives to improving reliability of analyses when using PARs are now

suggested. One alternative when using multiple state-generated crash form data sets is to funnel state PARs through an "intelligent programmer" who knows how to decode and disaggregate PARs information:

- "motorcycle size" may decode into the more useful "displacement"
- "VIN" may decode into "displacement"
- "number of occupants" may decode into "passenger use"
- "seating position" may decode into passenger use or "number of occupants"
- " seating position" may decode into motorcycle or *not motorcycle* (ATV)
- "motorcycle damaged areas" may decode into motorcycle or *not motorcycle*

  (an ATV represented on the PAR may actually be a motorcycle or vice-versa).

A second alternative to improving reliability is the use of novel data sets particularly focusing on the crash diagram and the narrative, many of which are highly detailed. Because they do not lend themselves readily to scanning or conversion to optical disk, some electronic PARs data sets do not include diagram or narrative. However, crash diagrams may be supplied upon request. The diagram and narrative may have further detail which can help the same "intelligent programmer" make decoding which enlarge or clarify the data set. Many narratives will have notations which help add clarity such as "bike" which would help distinguish a motorcycle from an ATV, for example.

A third alternative is the use of software which establish probabalistic matches between PARs and other hard data such as hospital records, and thus create an artificial data set which are smaller overall with many more data elements which have been matched from one PAR to another type of record, even though the element itself may be different. The National Highway Traffic Safety Administration has worked for the better part of the last decade on its CODES software (Crash Outcome Date Evaluation System). CODES attempts to link a PAR crash outcome record to its crash injury record which is completed by the Emergency Medical Service

(EMS) and hospital personnel. Another probabilistic software has been used with some success and is called AUTOMATCH, which can be used to match any two files such as PAS to trauma registry or even death certificate.

A fourth alternative to using PARs at face value is the use of heuristic software which would glean key words from text-scanned PAR narratives. While we have said that narratives are not easily transferred electronically, recent studies show how useful heuristic softwares can be. The Highway Safety Research Institute at Chapel Hill did a study of cell phone use related to crash outcomes. While a researcher would never expect to find "cell phone" on any PAR, those words appeared frequently in narratives, where HSRI extracted them and created an artificial but very useful data set.

A fifth alternative for improving intra-state data quality is simply improved police training. In a recent study of West Virginia' commercial vehicles which was balanced for geographic representativeness, we found that enforcement often simply don't recall how to fill out the PAR for infrequent vehicle types such as commercial vehicles, and, by extension, perhaps motorcycles. In this study, police enforcement universally said that training and in-service refreshers would improve PAR completion rates and in turn, the accuracy and timeliness of submission. (Winn and Eck, 2000).

Our final alternative is to forget the PAR altogether as a primary data source. As surprising as this sounds, it means that PARs are probably not going to get better any time soon for infrequent vehicle types, including motorcycles. Instead, we suggest NHTSA undertake a special study of critical data elements for motorcycles. (A 1998 study by USDoT titled, "Transportation Statistics: Beyond ISTEA. Critical Gaps and Strategic Responses" (BTS, 1998)

does not mention motorcycles). This kind of study was promoted by the National Association of Governor's Highway Safety Representatives when it developed the recommended set of commercial vehicle data elements, now universal across states.. A similar data-needs study was published just this year for bicyclists and pedestrians (NHTSA, 2000). Perhaps the Motorcycle and Moped Committee of the Transportation Research Board would advocate such a study on critical motorcyclist data needs.

SUMMARY:

A study of state PARs reveals that they have a very small amount of motorcycle-oriented data available and that well under half of the states capture more than four basic data elements. We conclude that while PARS have limited reliability at face value, there are alternatives available for improving both reliability and utility, including "intelligent programmers" and the use of heuristic softwares to capture information in PAR narratives. Finally, a special study is suggested to examine critical motorcycle oriented data needs.

REFERENCES:

Carr, M.L. and Winn, G.L. *Accident/Injury Database Linking Assessment Report*. Final Report submitted to West Virginia Department of Transportation, Division of Highways. (Unpublished document: Morgantown, WV, August, 1996)

Winn, G.L. and Eck, R. and Gillespie, K. *Feasibility Study Concerning the Collection of Timely and Accurate Crash Data for Commercial Vehicles in West Virginia.* Final Report submitted to the Federal Motor Carrier Safety Administration and the West Virginia Public Service Commission. (Unpublished document: Morgantown, WV, June, 2000)

U. S. Department of Transportation, Bureau of Transportation Statistics. *Bicycle and Pedestrian Data: Sources, Needs and Gaps.* BTS-7-00 (Washington, D.C., June, 2000)

U. S. Department of Transportation, National Highway Traffic Safety Administration. *State Accident Report Forms Catalog: 1999 Update.* DOT 808- 854. Washington, D.C., February, 1999)

U. S. Department of transportation, Bureau of Transportation Statistics. *Transportation Statistics Beyond ISTEA. Critical Gaps and Strategic Responses.* BTS98-A-01 (Washington, D.C.: January, 1998).

U. S. Department of Transportation, National Highway Traffic Safety Administration. *State Accident Report Forms Catalog: 1997 Update.* DOT 808- 455. Washington, D.C., January, 1997).

Winn, G.L., Bucy, D.S. and Klishis, M.J. *Forms and Functions*. Security Management. pp. 93 - 96.(Arlington, VA, June, 1997a)

Winn, G.L. and Bucy, D.S. *Technology vs Culture: Improving the Efficiency of Traffic Crash Data Collection in West Virginia*. Proceedings of the Seventh International Conference: Traffic Safety on Two Continents. (Lisbon, Portugal, September. 22 - 24, 1997b)

U. S. Department of Transportation, National Highway Traffic Safety Administration. *State Accident Report Forms Catalog: 1995 Update*. DOT HS 808-322. Washington, D.C., December, 1995).