Chapter 3

Use of Probabilistic Linkage for an Analysis of the Effectiveness of Safety Belts and Helmets

Dennis Utter, National Highway Traffic Safety Administration

Abstract

This presentation will describe the use of linked data by the National Highway Traffic Safety Administration to generate population-based crash and injury state data that include the medical and financial outcome for specific crash, vehicle, and behavior characteristics. The linked data were used by NHTSA for a Report to Congress as mandated by the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991. Benefits were to be measured in terms of about their impact on mortality, morbidity, severity, and costs.

Hawaii, Maine, Missouri, New York, Pennsylvania, Utah, and Wisconsin, states with the most complete statewide crash and injury data, were funded by NHTSA to implement Crash Outcome Data Evaluation Systems (CODES). The states linked crash to hospital and EMS and/or emergency department data using their most recent data year available at the time, ranging from 1990-1992. Implementation of a uniform research model by the seven CODES states was successful because of the linked data. The presentation will discuss how the linked data were used to standardize non-uniform data and expand existing data for analysis.

Introduction

Motor vehicle traffic crashes continue to be a significant problem in the United States. Each year there are more than 6 million crashes investigated by police agencies. In these crashes 3.5 million people are injured, 450,000 of them severely, and nearly 42,000 are killed. Crashes produce a staggering economic toll, too. Nearly $151 billion are lost due to medical costs, property damage, legal costs, productivity losses, and other factors. Clearly, reducing the number of crashes and their severity is a necessity.

The National Highway Traffic Safety Administration (NHTSA) was created to reduce the number of deaths, the severity of injuries, and other damage resulting from motor vehicle traffic crashes. It does so through a variety of programs aimed at making vehicles safer, therefore mitigating the results of crashes, and by getting vehicle drivers and occupants to do things that would either prevent crashes or mitigate their outcomes. Evaluation of these programs requires a significant amount of data. Data linkage provides NHTSA, and the traffic safety community at large, with a source of population-based crash and injury state data that include the medical and financial outcome for specific crash, vehicle, and behavior characteristics.

Data files created from police reported motor vehicle crash data alone do not include medical outcome information for everyone involved in a motor vehicle crash. Thus, linking data became necessary when NHTSA was required by the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 to report to the Congress about the benefits of safety belts and motorcycle helmets. Benefits were to be defined in terms of mortality, morbidity, severity, and costs. Statewide crash data files were determined by NHTSA to be the only source of population-based information about the successes (those who use the countermeasure and receive no or a less serious injury), the failures (those who do use the countermeasure and receive an injury), those not affected (those who do not use the countermeasure and receive no injury) and those who were not injured as seriously as they might have been because of the safety device.

CODES

Hawaii, Maine, Missouri, New York, Pennsylvania, Utah, and Wisconsin, states with the most complete statewide crash and injury data, were funded by NHTSA to implement Crash Outcome Data Evaluation Systems (CODES). The states linked crash to hospital and EMS and/or emergency department data using their most recent data year available at the time, ranging from 1990-1992. The study population was chosen from police reported data because of the importance of the safety belt and helmet utilization data contained in the crash file. The safety belt analysis included 879,670 drivers of passenger cars, light trucks and vans and the motorcycle analysis consisted of 10,353 riders of motorcycles. This presentation will describe how linked data made it possible for NHTSA to conduct a medical and financial outcome study of the benefits of safety belt and motorcycle helmets using routinely collected, population-based, person-specific state data.

Use of Linked Data to Standardized Non-Uniform Data for Analysis

Outcome Analysis Using "As Reported" Data

Measuring outcome is complicated when using "as reported" utilization data. Using this type of data, the CODES results indicated that although each state was different, all safety-belt odds ratios from all states agreed that safety belts are highly effective at all analysis levels at less than the .001 significance level. The non-adjusted effectiveness rates indicated that safety belts were 89% effective for preventing mortality and 52% effective for preventing any injury. The downward shift in severity was demonstrated by the decreasing effectiveness rates ranging from 89% for victims who die to 75% for those who die or are inpatients and to 54% for those who die, are inpatients, or are transported by EMS. But these results are inflated. When safety belt usage is mandated, human beings being human have a tendency to exaggerate their use of a safety belt, particularly when crash evidence or their injury type and severity are not likely to indicate otherwise. Over reporting of belt use moves large numbers of unbelted uninjured persons into the belted uninjured column thus inflating belt effectiveness. NHTSA repeated the research model to incorporate observed safety belt utilization rates into the analysis. Adjustments were made based on the assumption that 35 percent of the belted who were uninjured or slightly injured may have misreported their belt use at the time of the crash. These adjustments obtained the more realistic effectiveness rates of 60 percent for preventing mortality and 20-45 percent for preventing morbidity. In the future, as state injury data systems are improved to include safety utilization and external cause of injury information, linkage will make it possible to use the injury data to confirm utilization of the safety device.

Definition of the Occurrence of an Injury and Injury Severity

Although the study population was defined from the crash report, the linked data were used to define the occurrence of an injury and the various levels of injury severity. This standardization was necessary to compensate for inconsistent implementation of the police reported KABCO severity scale by the different states. For example New York classified one-third of the occupants involved in crashes as suffering "possible" injuries compared to about 10 percent in the other CODES states. For CODES, injury and the severity levels were defined by combining "injury severity" on the crash report with "treatment given" on the injury records to create five levels -- died, inpatient, transported by EMS or treated in the ED, slightly injured or no injury. Police reported "possible" injuries were classified as non-injured unless the crash report linked to an injury or claims record. The severity levels were used to define the outcome measures (mortality, morbidity, injury severity, and cost) for the uniform research models for both the belt and helmet analyses as follows:

Mortality:	Died versus all other crash-involved victims.
Morbidity:	Any injured compared to those not injured.
Shift in Severity:	Separate effectiveness rates for each severity level were calculated and then compared to measure the downward shift in injury severity
Cost:	Defined as inpatient charges because non-inpatient charges were not comparable among the seven states.

Use of Linked Data to Expand Existing Data

Identifying Injuries Not Documented by the Police

Police are required to document only those crashes and injuries that occur on public roads and meet mandated reporting thresholds. In addition, some reportable injuries are not documented because of non-compliance with the requirements. CODES excluded cases not documented by the police because of the need for standardized safety device utilization information. But using only crash reports to document the injuries understates the total injuries. The CODES states used the linked crash and injury records to identify those injuries not documented by the police.

Identifying Financial Outcome

Data linkage provides highway safety with access to financial outcome information related to specific characteristics of the crash event. Lack of uniformity in the documentation of EMS and emergency department charges limited the CODES analysis to inpatient billed charges as indicated in the hospital data. These data were used to calculate average charges for inpatient drivers and all crash involved drivers. The analysis indicated that the average inpatient charge for unbelted drivers admitted to a hospital was 55% higher than for the belted, $13,937 compared to $9,004. If all drivers involved in police-reported crashes in the CODES states had been wearing a safety belt, costs would be reduced 41 percent (approximately $68 million in reduced inpatient charges or $47 million in actual costs). This type of information is powerful in the political arena and is unavailable to highway safety except through data linkage.

Identifying the Type of Injury

Linked data were crucial for the helmet analysis. By using only the level of severity NHTSA found that the effectiveness rates were low for helmets, 35% effective in preventing mortality, and only 9% effective in preventing morbidity. The downward shift in injury severity was much less than for safety belts. The linked data enabled NHTSA to redirect the analysis to brain injuries which the helmet is designed to prevent and found that helmets were 67 percent effective in preventing brain injury. That means 67% of the unhelmeted brain injured would not have been so injured if they had been helmeted. Looking at the costs for the brain injuries also justified focusing the analysis. Average inpatient charges for the brain injured were twice as high. Approximately $15,000 in inpatient charges would be saved during the first 12 months for every motorcycle rider who, by wearing the helmet, did not sustain a brain injury. Again, this type of information is more powerful than the overall effectiveness rate for helmets.

Barriers to Linkage of Crash and Injury Data

Probabilistic linkage requires computerized data. Unfortunately, not all states have crash and injury data that are statewide and computerized. Almost all of the states have computerized crash data statewide. Half of the states have developed state EMS data systems, but only a few have state emergency department data systems. A majority of the states have computerized state hospital discharge data systems. All of the states have computerized Medicaid and Medicare data systems, but few states have statewide computerized data files for private vehicle or health insurance claims data. Access to data for the less seriously injured victims, a group that includes many of the successes for highway safety, is difficult to obtain because the data may not be computerized. Or if computerized, they are computerized by provider or by insurance group and rarely statewide. Injury data are particularly useful to highway safety because they document what happens to all victims injured in motor vehicle crashes, regardless of whether the crash itself meets police reporting thresholds.

Benefits of Data Linkage

Data linkage provides documentation, generated from a state's own linked data, that is more credible among local decision makers who may be tempted to repeal the safety mandates, such as helmet legislation. And the data linkage process itself has the added benefit of making data owners and users more aware of the quality, or lack thereof, of the data being linked. The CODES states found that important identifiers that should have been computerized uniformly statewide were not; or if the identifiers were computerized, some of the attribute values were missing or inaccurate. All of the states became adept in discovering errors in the data and were motivated to revise their edits and logic checks. Thus annual linkage of the crash and injury state data provides the states, NHTSA, public health and injury control, with a permanent and routine source of outcome information about the consequences of motor vehicle crashes at the same time that the quality of state data are improved for their originally intended purposes.