



## CRASH OUTCOME DATA EVALUATION SYSTEM (CODES)

**Coordinator and  
Data Contact:**

Michael Singleton, MS  
Kentucky Injury Prevention and Research Center  
University of Kentucky  
333 Waller Ave, Suite 206  
Lexington, KY 40504  
(859) 257-5809  
[msingle@email.uky.edu](mailto:msingle@email.uky.edu)

**State Web Site:  
National Web Site:**

None currently  
[http://www.nhtsa.gov/Data/  
State+Data+Program+&+CODES](http://www.nhtsa.gov/Data/State+Data+Program+&+CODES)

### Sources of Information for the Database

The Crash Outcome Data Evaluation System (CODES) is funded by the National Highway Traffic Safety Administration (NHTSA) for the purpose of linking state motor vehicle traffic crash report databases to administratively unrelated databases containing medical and economic information pertaining to persons involved in crashes. At the center of this effort is the Kentucky motor vehicle traffic crash reporting system, called Collision Reporting and Analysis for Safer Highways (CRASH). To date, CRASH has been linked with the state inpatient Hospital Discharge Database (HDD) for years 2000 through 2010, and with both inpatient and outpatient databases for years 2008 through 2010. The linkage is accomplished using a probabilistic methodology based upon research by Fellegi and Sunter (1969) and Jaro (1985, 1995), and implemented in software provided by NHTSA. This linked database enables the discovery of relationships between crash characteristics and injury outcomes for persons hospitalized as a result of motor vehicle crashes, and the assessment of the inpatient acute care charges associated with their treatment. There is no federal or state mandate requiring that this surveillance be conducted.

---

## Description of the Data Collected

CRASH reports are mandated in Kentucky for crashes occurring on public roadways involving an injury or property damage in the amount of \$500 or more. Officers collect information on all persons involved in the crash, including data on individuals (age, gender, date of birth, seating position, safety belt and helmet use, human contributing factors and more); vehicles (type, make, model, Vehicle Information Number, extent of damage, vehicular contributing factors and more); crash event (date, time, and location of crash, manner of collision, first and second collision events, most harmful event, and more); and environment (weather, light conditions, roadway conditions and characteristics, environmental contributing factors and more).

Hospital discharge reports are mandatory for all discharges of inpatients and emergency department (ED) patients from hospitals operating in Kentucky. The HDD database includes personal and medical information for each patient, including demographics, diagnosis and procedure codes, external cause of injury, monetary charges and payment sources billed and more.

## Strengths of the Data

The combination of these three population-based data sources through probabilistic linkage yields a data source on persons hospitalized or treated in EDs as a result of crashes on Kentucky's roadways. It thus enables analyses that would be impossible using either source alone. Crash reports lack reliable information about the type, severity, cost and treatment of injuries to crash participants. Hospital discharge data lack information about the many factors and circumstances that led to the crash and influenced its severity, and about the use of safety devices. Using the CODES linked database, we can discover relationships between risk and protective factors and medical outcomes.

## Specific Uses of Information

- Fact sheets on motor vehicle traffic safety topics
- Peer-reviewed research on traffic safety and injury prevention
- Data requests from NHTSA and from state and local users

### Data Limitations

There are two main limitations to the CODES data.

**Representativeness:** Some persons who are involved in crashes in Kentucky are hospitalized outside of Kentucky, and some who crash outside of Kentucky are hospitalized in Kentucky. Our data sources do not capture out-of-state events, so such cases will not be represented in our linked database. As a result, it is a significant challenge to determine how well the CODES database represents the population of all persons hospitalized as a result of crashes that occur in Kentucky. A more tractable question is how well it represents the population of persons who both crashed and were hospitalized in Kentucky, since these are the cases covered by our data sources. This question has been the focus of our evaluation efforts.

**Misclassification:** Some data elements on the CRASH reports are inherently difficult to capture reliably. For example, from comparing the reported seat belt use rate on CRASH with results of observational studies, we know that the latter is significantly over reported. This is because the vast majority of persons involved in crashes are not severely injured. By the time police arrive on the scene it is usually impossible to know whether such occupants were wearing seat belts, so the officer has to rely on self-reporting. The more severe the injury, the more likely the officer can directly observe belt use.

### System Evaluation

Both the CRASH and HDD systems perform computerized edit checks at the time reports are entered. Our evaluation efforts have focused on the predictive value positive (PVP) and sensitivity of the linkage process. We conducted an evaluation of the linked CODES database among persons hospitalized at the University of Kentucky Chandler Medical Center (UKMC) in order to determine the percentage of UKMC patients admitted for motor vehicle crash-related injuries who were matched incorrectly to a crash record. We found this type of error in less than 5% of cases. Our conclusion is that the linkage process has a very high PVP for persons who were hospitalized in Kentucky. A second study estimated the system sensitivity. This was accomplished by reviewing medical records for persons admitted to UKMC with an external cause of injury code indicating involvement in a motor vehicle crash, but whom we were unable to link to a CRASH record with a high degree of certainty. We estimate that about 15% of persons who crashed and were hospitalized in Kentucky are not represented in the CODES database, for a variety of reasons including crashes not being reported to police or failure of record linkage.

### Data Set Availability

A public-use Kentucky CODES data set is not currently available. Aggregated (tabular) data may be requested by contacting the project coordinator. Requests from researchers for access to the linked database will be referred to the project's data owners and Board of Directors for case-by-case consideration.

---

## Data Release Policy

Only aggregate data is released at this time.

## Data Publications

Data publications can be found at <http://www-nrd.nhtsa.dot.gov/Pubs/811181.pdf>.

## Suggested Data Citation

Kentucky Injury Prevention and Research Center (KIPRC). *Crash Outcome Data Evaluation System*. Lexington, Kentucky: University of Kentucky [data year].

## References

Fellegi IP, Sunter AB. *A theory for record linkage*. Journal of the American Statistical Association, 1969; 64(328):1183-1210.

Jaro M. *Advances in record linkage methodology as applied to matching the 1985 census of Tampa, Florida*. Journal of the American Statistical Association, 1989; 84(406):16-21.

Jaro M. *Probabilistic linkage of large public health data files*. Statistics in Medicine, 1995; 14:491-498.

## Contributing Author

Michael Singleton, MS, Kentucky Injury Prevention and Research Center