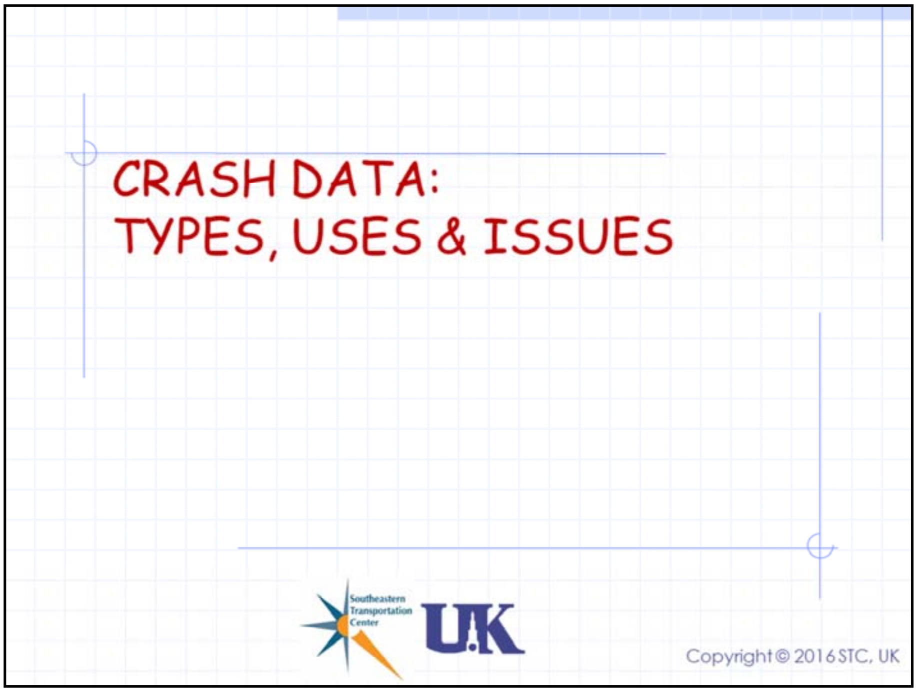
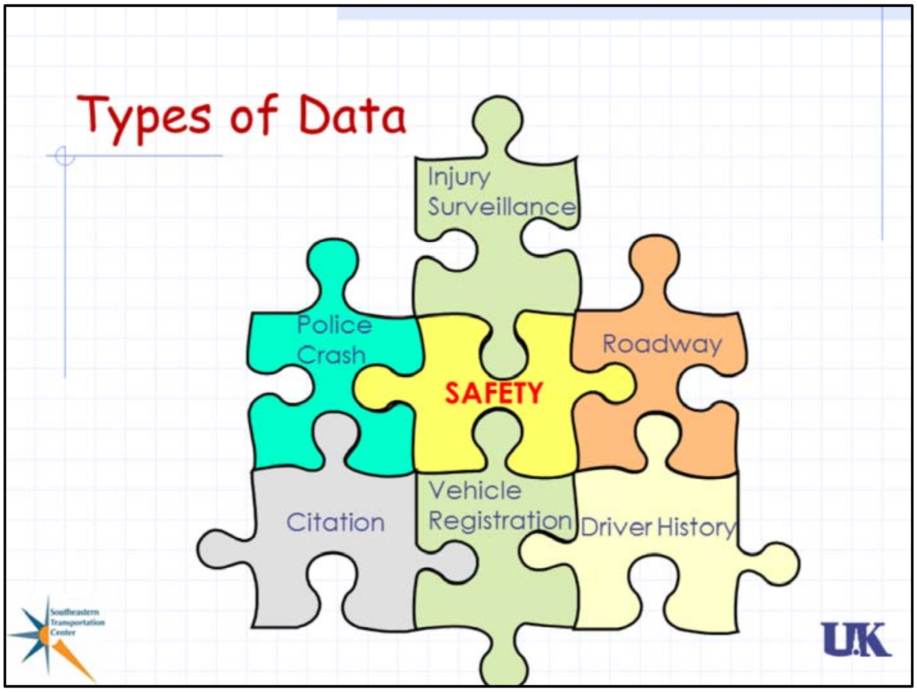


This module discusses the types of available crash data and identifies their uses in highway safety in order to improve safety management and decision making. It also provides an overview of data deficiencies and explains the limitations of crash data.

Question: How do we measure safety?

Answers: The idea here is to generate discussion and focus on the need for data in order to analyze, evaluate, compare and determine the issues surrounding highway safety. Any study undertaken requires some type of data in order for the researcher to be able to reach conclusions. The data could be simply crash data in order to identify potential trends and contributing factors, it could a combination of crash and roadway data in order to identify roadway related factors contributing to crashes, or it could be crash and socioeconomic data to evaluate their relationships. The availability of crash data becomes central to all such analyses. It is therefore imperative to ensure that data is first available and then reliable.



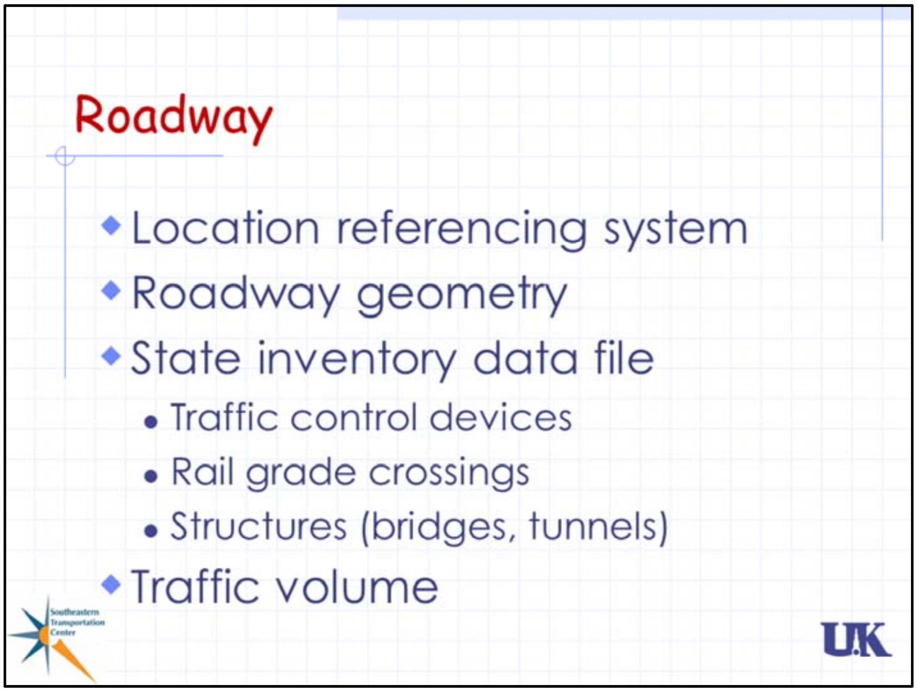
There is variety of data that is available and could be used in a safety analysis. Each one fulfills a specific need and addresses areas of needs. Transportation officials typically utilize a variety of data depending on their needs. Police crash data is typically available at the state and local level. There are also national efforts to ensure data transferability and these will be discussed later in the module. Roadway and traffic volume data is also often linked to crash data to allow for investigating the effects roadway geometry and features on crash occurrence. Vehicle registration and driver history could also provide additional information when specific human and vehicular issues are examined. Follow up of injuries and Emergency Management Systems (EMS) is another area where data is collected and tied to crash records to determine the level of crash severity aiming to address potential police information errors in such assessments (also to be discussed later in the module).

Police-Recorded Crash Data

- Weather conditions
- Time of crash
- Person type (driver, occupant, pedestrian, etc.)
- Number and severity of injuries
- Violations
- Location latitude and longitude
- Manner of collision
- Number of vehicles
- Alcohol or drug impairment
- Direction of travel
- Crash diagram
- Narrative crash description

Over the years, the National Highway Traffic Safety Administration (NHTSA) has made concerted efforts to ensure that complete, accurate, and timely traffic safety data are collected, analyzed, and made available for decision-making at the national, state, and local levels. To ensure this, typical data to be collected includes the time, location, environment, and characteristics (sequence of events) of a crash. In addition to this basic information, data identifying the people and vehicles involved and the consequences of the crash are also provided. Finally, violations and citations are included to identify any infractions.

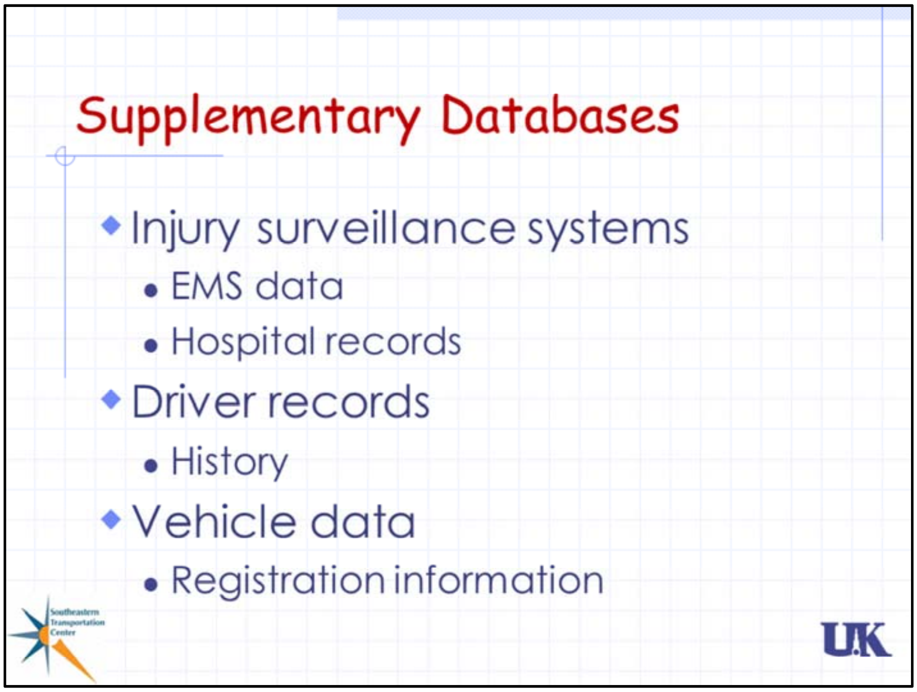
The crash data system is generally maintained by state departments of transportation, departments of motor vehicles or departments of public safety. Crash data are the most widely used type of data in road safety management. They provide guidance to decision makers within agencies and a powerful tool in support of safety legislation.



The main data provided in the roadway information is the location of the facility and its geometry. A variety of location reference systems are used throughout the country and most are dependent upon GPS information for accurately locating facilities.

The roadway geometry data typically provides cross section information on the roadway, e.g., number of lanes, lane width, shoulder type and width, median descriptors, pavement types, horizontal curvature, grades, etc. Most states also have supplemental files describing bridges (as part of the National Bridge Inventory) and railroad grade crossings (as part of the Federal Railroad Administration's Railroad Grade Crossing Inventory) that can usually, but not always, be linked to the basic roadway inventory file.

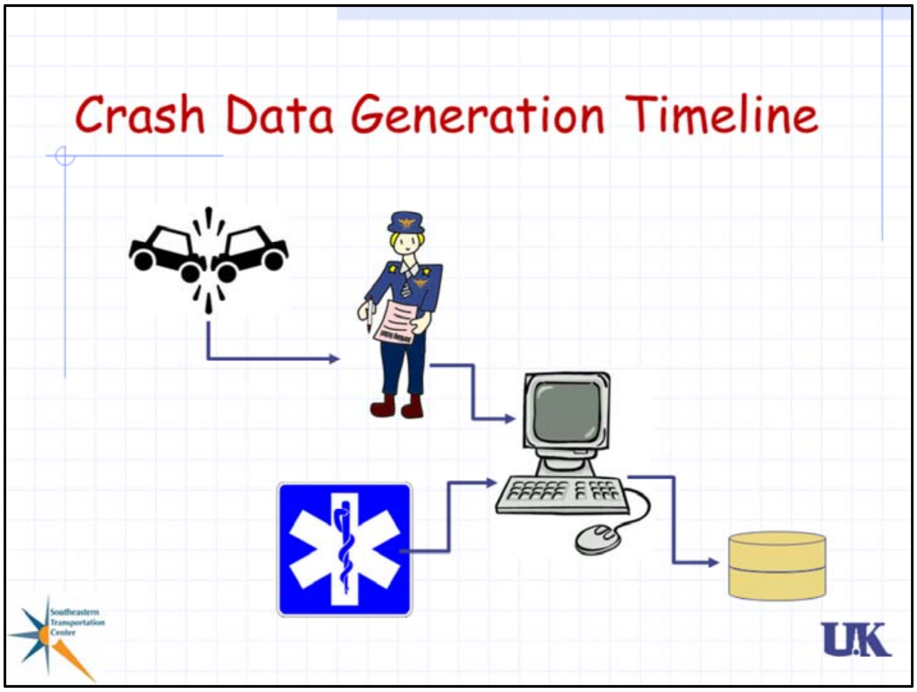
Another variable that is of interest in highway safety is the traffic volume of the facility and state highway agencies collect and maintain data on traffic volumes (Average Annual Daily Traffic – AADT). Traffic volume data may also include truck percentages. In general, traffic volume data are collected on the state-maintained system and traffic data in other roads may be obtained from local agencies. Most states conduct sample traffic counts on a periodic basis in addition to some permanent locations. The accuracy of traffic volume data should be ensured, especially if the Highway Safety Manual procedures are utilized for safety analysis.



Additional databases that can assist safety evaluations include the following:
Injury surveillance systems can be used to further assess the severity of the crash as well as costs associated with the consequences of the crash (medical bills, length of recuperation and healing, long-term rehabilitation). These systems rely on EMS and emergency department data and can be of great significance in bridging the gap between traditional traffic safety and public health issues.

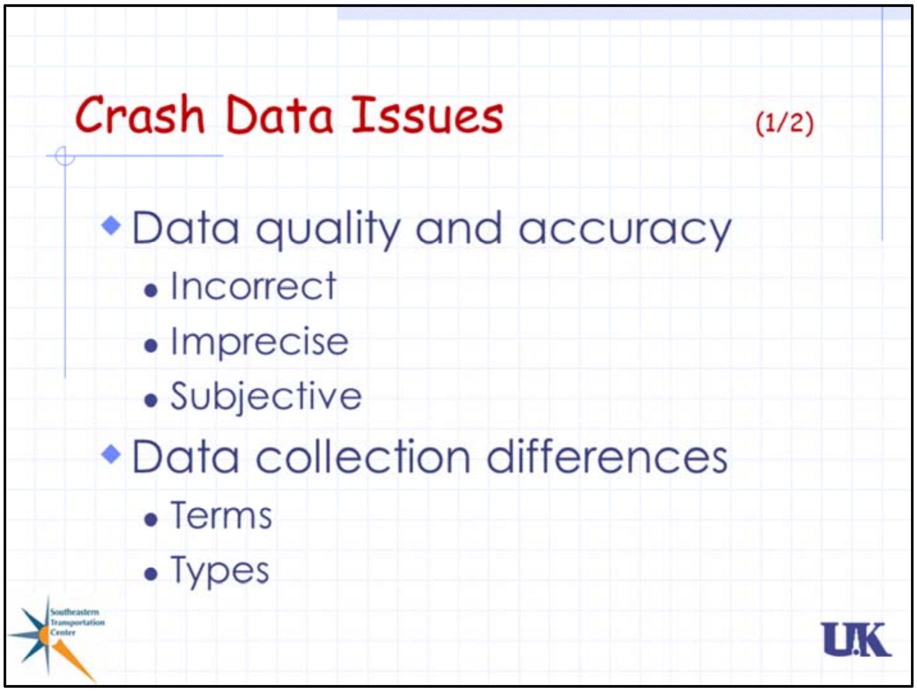
Driver records are maintained through department's of motor vehicles and include data on all licensed drivers in the state. The driver history file contains basic identifiers (e.g., name, address, driver license number), demographic information on the driver (e.g., age, birth date, gender), and information relevant to license and driver improvement actions (e.g., license issuance and expiration/renewal dates, license class, violation dates, suspension periods,, crash involvement).

Vehicle data provides registration information including owner information, vehicle license plate number, and vehicle make, model and year of manufacture. Vehicle information systems also contain information regarding commercial vehicles and carriers which may be registered in one state but are licensed to travel in other states. This information includes the U.S. DOT number, the carrier information, and any inspection or out-of-service information.



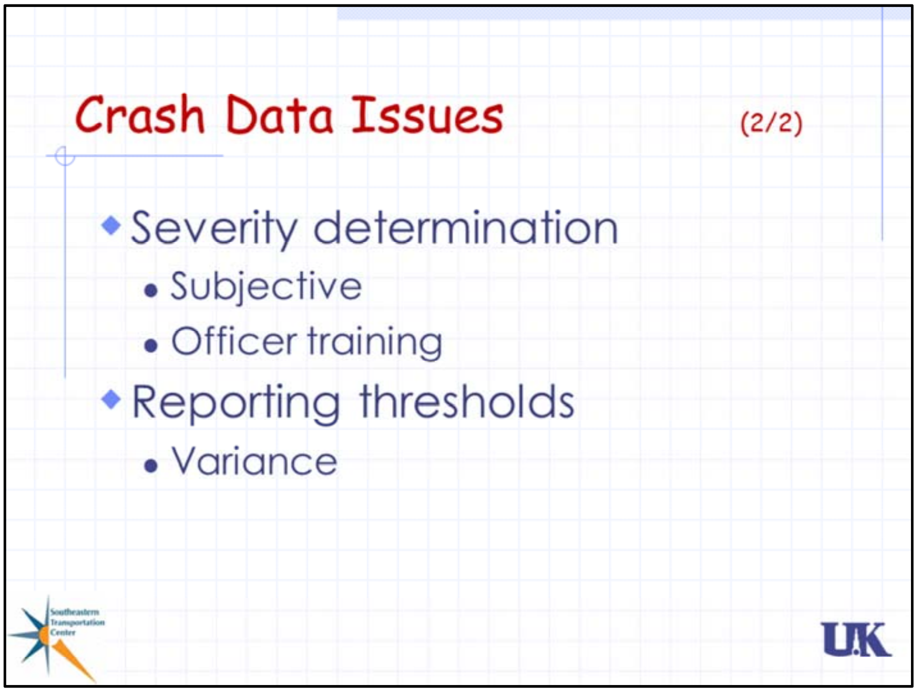
Crash data collection initiates when a crash occurs. At this point, the corresponding law enforcement unit (state patrol or local police) is notified and upon arrival at the location, the police officer completes short crash report documenting the crash specifics (location, time of concurrence, driver and vehicle data, pertinent violations and severity assessment). The officer then completes a more detailed report (full report or long form) in the office based on state- or local-specific requirements. This form is then coded and entered in the database by the law enforcement agency responsible for crash data collection. Typically all crash data (state or local level) is assembled into the state crash database that can then be used to data to identify and record the location of the crashes. Data is evaluated to ensure that it meets the state database requirements and once it is included in the database, it becomes available to the various agencies and entities for analysis and use. It is apparent that there is a time lapse between the crash occurrence and the time the data is available and this depends on state procedures and checks for the data. Typically, this time can be between 6 to 12 months. Crashes involving fatalities are reported to the NHTSA and investigated further for inclusion in the Fatality Analysis Reporting Systems (FARS).

It is apparent that there is a great potential for errors in this entire process from the point of collection to the point of analysis. The process is complicated, varies from state to state, and sometimes it even varies within states and among local governments.



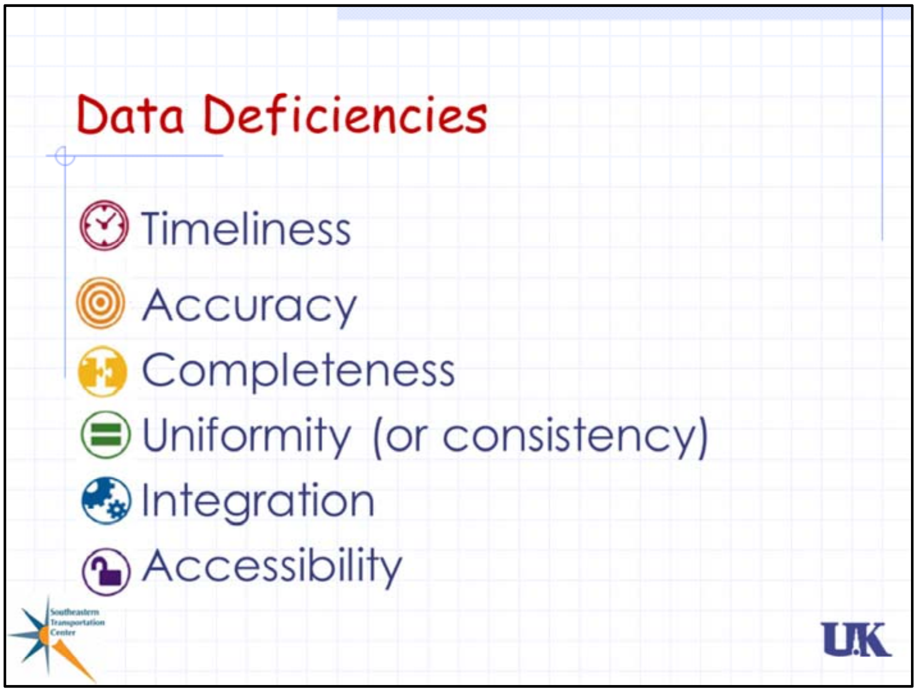
A number of crash data issues can affect the quality of the analysis. These issues can introduce bias and affect crash evaluation and analysis resulting in erroneous results and decisions. Data quality and accuracy is critical in any analysis undertaken in order to ensure that the reported crashes reflect the existing situation. The main sources of error for the data accuracy involve incorrect data entries, i.e., typographic errors or incorrect information regarding road designation, severity level, vehicle and driver information, etc., or imprecise data, such as use of generic terms to describe a location. Another form of error is the subjective nature of some data to be provided, such as level of severity of speeding (since he officer determines this after the fact as a potential contributing factor).

There may be different procedures followed between state and local agencies in crash reporting that could complicate the data analysis due to incompatibilities among the collected data. These procedures could be attributed to the use of different terms for the same data type or use of different types of data. For example, some agencies may record the Average Daily Traffic (ADT) as the volume while others may use the Annual Average Daily Traffic (AADT). It is possible that two agencies could simply call the same information differently, i.e., both use the AADT, while at other times each agency may be using a different metric. Such differences need to be identified and understood to avoid misuse of the data,



Additional issues that could affect the quality of the analysis include determination of severity levels and the threshold for reporting crashes. Severity assessment can be flawed, since police officers typically are not trained medical professionals to assess the level of injuries sustained in a crash. This could be a significant problem, since often countermeasures are evaluated not only on the changes they impact on the number of crashes but on the levels of severity as well. Studies indicate that crashes with high severity levels are reported more reliably (and more frequently) than crashes with lower levels, which may lead into bias towards countermeasures targeting improvements for crashes with high severity levels.

An issue of concern for comparative analysis across states and jurisdictions is the different reporting thresholds for a crash resulting in improper estimations if all are aggregated into one database. Changes in reporting thresholds to reflect inflation or cost of living adjustments need to be also acknowledged, since they will create different reporting levels between time periods. Another issue is also the number of unreported crashes. It is well recognized that not all reportable crashes are reported to eh police for a variety of reasons. This may also affect the development of countermeasures, since the true magnitude of the problem is not documented.

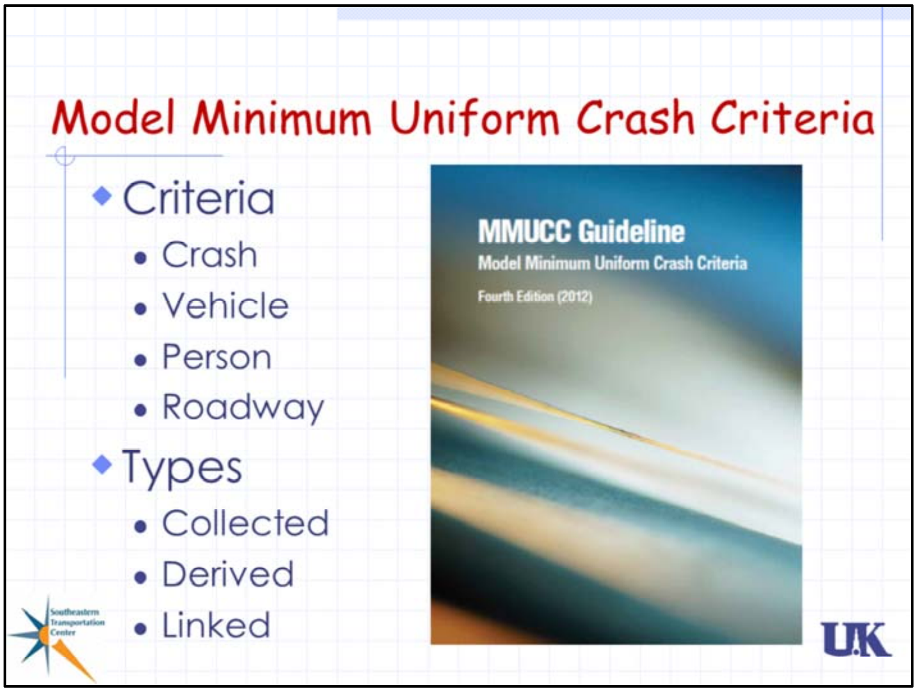


NHTSA identified the following six items as deficiencies that should be addressed to improve data quality. NHTSA identifies these as the data "six pack" to describe areas where data need to be improved.

1. Timeliness – Data needs to be provided in time in order to be relevant. This deficiency can result because the agency responsible for the crash data either does not receive the crash report forms in a timely manner from law enforcement, or the agency can not keep current with the volume of crashes being reported.
2. Accuracy – This probably the most critical element in the entire process. This can result because the information was incorrectly converted to electronic format; multiple people entered the data into an electronic database from a paper form; the form was hand written in poor conditions (lighting, weather, etc.) by an officer with other responsibilities at the time (e.g., attending to the victims, clearing the roadway, etc.).
3. Completeness – Uncertainty of codes during the coding process could result in leaving entries blank or incomplete and thus no information is entered into a required field on the crash report form.
4. Uniformity / Consistency – This could result due to lac of uniform codes and a single crash report form are not being used by all state and local police departments.
5. Integration – The crash database was created with incompatible versions of software and cannot be linked to spatial databases or other key fields such personal identifiers or vehicle data.
6. Accessibility – The agency does not provide access to the crash database fearing that "the wrong" person will get access to information that could possibly result in a misunderstanding or be used against the State in a court of law.

The National Safety Council developed an injury scale to allow for a better classification of motor vehicle crashes. The scale is called KABCO and each letter represents an injury level category defined as follows:

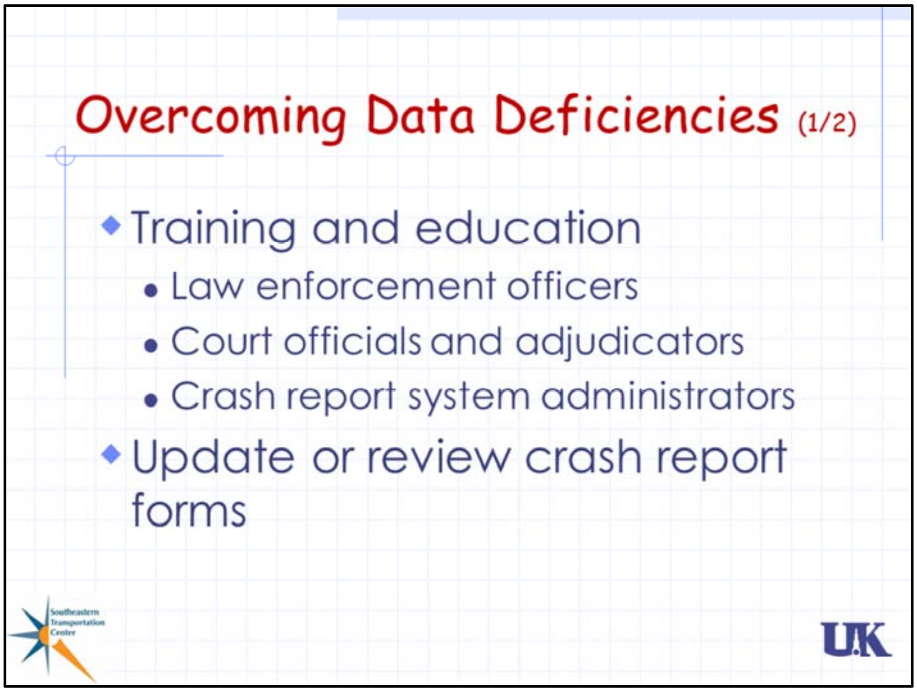
- K fatal injury is defined as an injury resulting in a death
- A incapacitating injury is defined as any injury other than fatal that prevents the injured person from walking, driving or normally conducting the activities the person was capable of performing before the injury
- B non-incapacitating injury is defined as any injury other than fatal or incapacitating injury evident to an observer at the crash scene.
- C possible injury is defined as any injury reported or claimed that is not fatal, incapacitating or non-incapacitating and it was immediately observable
- O no injury (or Property Damage Only - PDO) which is defined as a crash resulting in no juries and only property damage between vehicles.



NHTSA, FHWA, Federal Motor Carrier Safety Administration and Research and Innovative technology Administration developed in 2009 the Model Minimum Uniform Crash Criteria (MMUCC). This provides voluntary guidelines for states in order to improve and standardize their state crash data. MMUCC provides a minimum set of data elements that are accurate, reliable, and credible within states, among states, and at the national level. States could achieve "MMUCC-compliance" by the addition of data elements and attributes to their crash report form in the following four major groups:

- Crash – This describes the overall characteristics of the crash and consistency among the minimally reported data is required (similar to those presented earlier).
- Vehicle - The motor vehicle data elements describe the characteristics, events, and consequences of the motor vehicle(s) involved in the crash.
- Person - The person data elements describe the characteristics, actions, and consequences to the persons involved in the crash.
- Roadway - Roadway data elements are generated by linking crash reports to the roadway inventory and hardware data files when these data files exist in the state. The data elements used for linkage include crash roadway location and others as necessary depending upon the type of roadway inventory system implemented by the state.

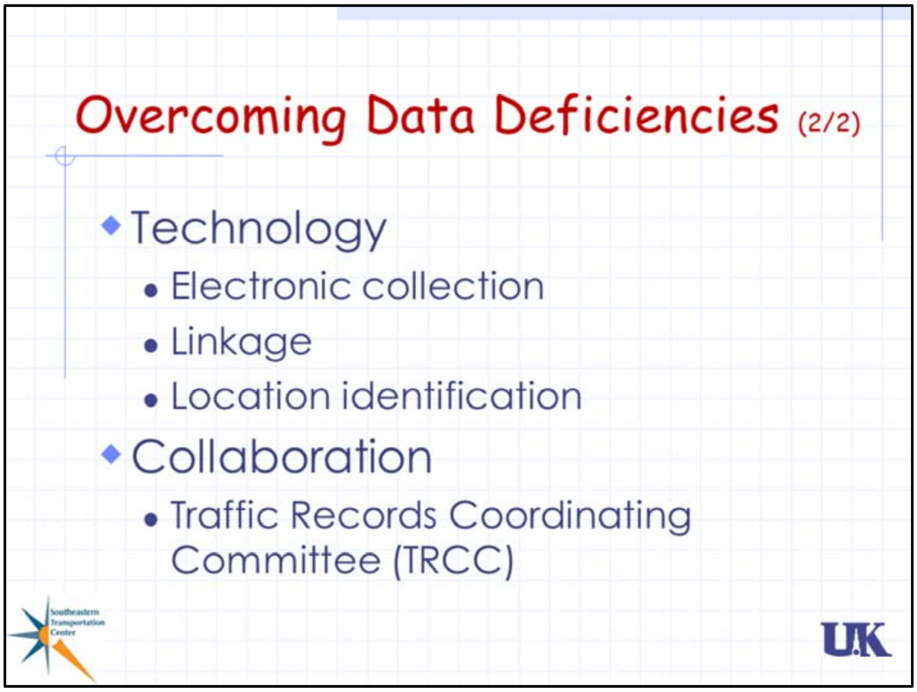
The data can be collected in the field (collected by police at the scene and recorded directly onto the crash report), derived (generated from computerized crash data) or linked (data generated when the crash data file is linked to injury, driver history, vehicle registration, roadway inventory, or other data files).



Efforts to address the data issues noted before have been undertaken in order to improve problem identification, needs assessment, priority setting, resource allocation, project selection, and countermeasure evaluation.

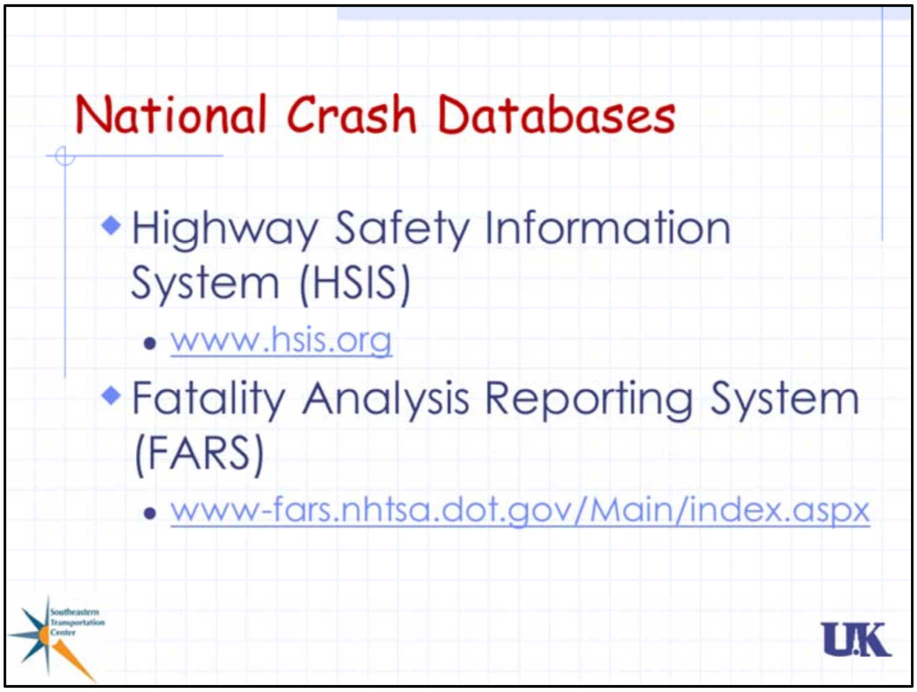
A basic effort to address some of the issues is focused on continuous training for police officers. Police reports completed at the crash site are the first source of potential errors and can create a variety of inaccuracies when data is coded and entered into the database. Such training efforts could also include information on the importance of the crash data collected and techniques to ensure accurate data collection. A good understanding of the various uses of the crash data is also important so police officers can see the value and importance of the crash data during the decision making process for safety improvements and investments. The notion that crash data is merely to serve and comply with insurance company needs should be dispelled. Training of court officials and adjudicators is also important as changes to safety legislation and penalties occur. Training crash report system administrators to properly handle reports with inaccurate or missing information can result in more accurate data.

In the past few years, several states as a result on MMUCC have reviewed and updated their crash report form. This should be considered as a continuous improvement effort, since analysis of crash trends may require additional information. For example, the use of cell phone was not previously recorded as a contributing factor and this has been added in the crash forms of several states to reflect this recent change.



The use of technology can also improve data deficiencies. Electronic crash report systems, GPS location devices, and barcode or magnetic strip technologies that collect vehicle and license data can reduce data entry errors and improve data accuracy. NCHRP Synthesis 367 Technologies for Improving Safety Data provides a comprehensive summary of crash data collection innovations.

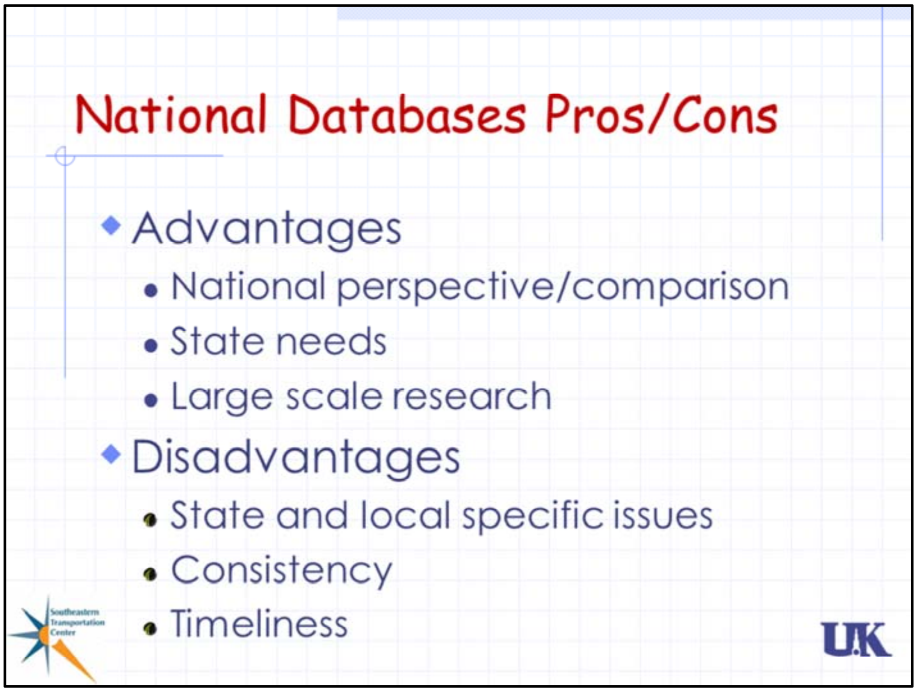
Safety data is collected through a variety of agencies and as such it requires collaboration among all parties involved. Many programs establish data collection task forces or committees to promote collaboration among safety stakeholders. The task force or committee holds regular meetings or workshops to highlight data sharing issues. In most states this function is performed by the Traffic Records Coordinating Committee or TRCC.



Various crash databases are maintained though a variety of national agencies that could provide for nation-wide analysis and research. The two most commonly used databases are:

The Highway Safety Information System (HSIS) database is maintained through FHWA and it provides the crash records for a number of states based on voluntary participation. The HSIS is a roadway-based system that provides quality data on a large number of crash, roadway, and traffic variables. The data are acquired annually from a select group of States, processed into a common computer format, documented, and prepared for analysis.

The Fatality Analysis Reporting System (FARS) database is maintained by NHTSA's National Center for Statistics and Analysis (NCSA) and contains data from 1978 through present. This database is updated annually and contains detailed information for crashes that resulted in at least one fatality. To be included in FARS, a crash must involve a motor vehicle traveling on a traffic way customarily open to the public, and result in the death of a person (either an occupant of a vehicle or a non-motorist) within 30 days of the crash. The FARS file contains descriptions of every fatal crash reported. Each case has more than 100 coded data elements that characterize the crash, the vehicles, and the people involved. General statistics are provided on the home page or through annual reports, and specific queries may be conducted from the FARS website. The queries allow the user to specify the variables of interest as well as the year of analysis to obtain various crash statistics.



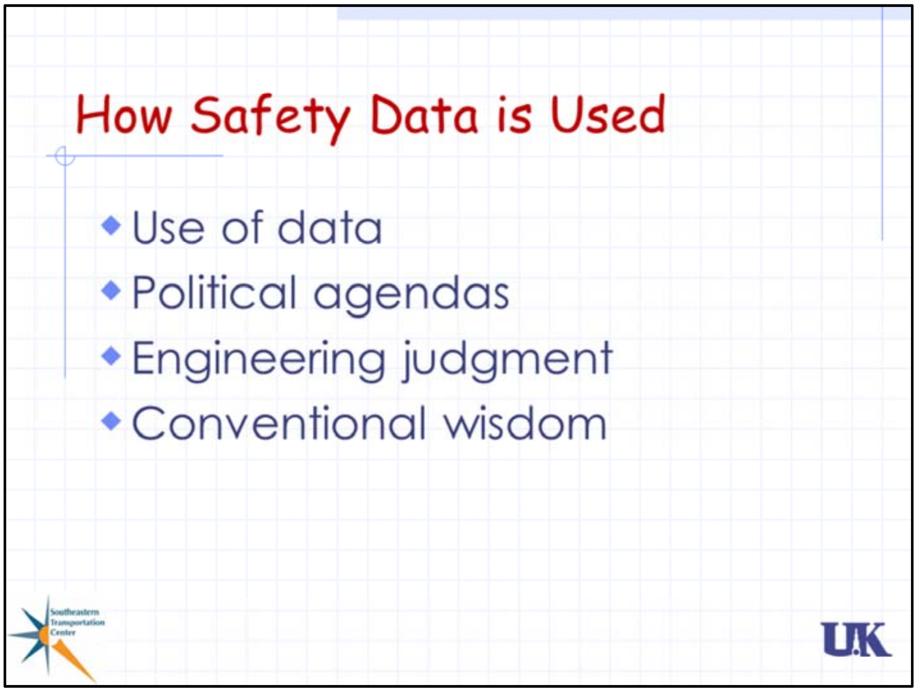
Federal databases can provide a representation of the nation status. However, data at this level could have distinct advantages and disadvantages.

Advantages

- Nation-wide data could be used as representing the entire nation either through inclusion of the entire population or an adequately representative sample. Such data can address national-level policies such as safety belt use, alcohol-related crashes, and young driver fatalities and then be used to develop programs to address them at the national level. The data can then be also used to determine the efficacy of the policies implemented.
- State needs can be also identified from national databases due to the potential for comparisons across states. National data can be used to identify states with greater than average crash and fatality rates for total crashes or for specific areas of concern and then be targeted for improvement.
- National databases can also be used to identify trends and conduct statistical analyses that could not be completed utilizing the smaller size state databases.

Disadvantages

- National data may not be able to capture local issues. The national databases would be bale to distinguish of specific location issues that may not show up in such databases.
- The fact that national databases rely on input from various states creates a potential issue of consistency and accuracy of the data provided. Even though most states collect the same information, there is a lack of commonality among the specific information provided or the level of detail within an entry. As noted earlier, there are different levels of reporting thresholds for each state and this alone could create significant issues for comparisons among states.
- Timeliness is an issue at all levels of database management, but the issue is compounded at the federal level. Federal data are often received from state and local agencies, and lags at the local level can significantly impact the timeliness of the federal database.

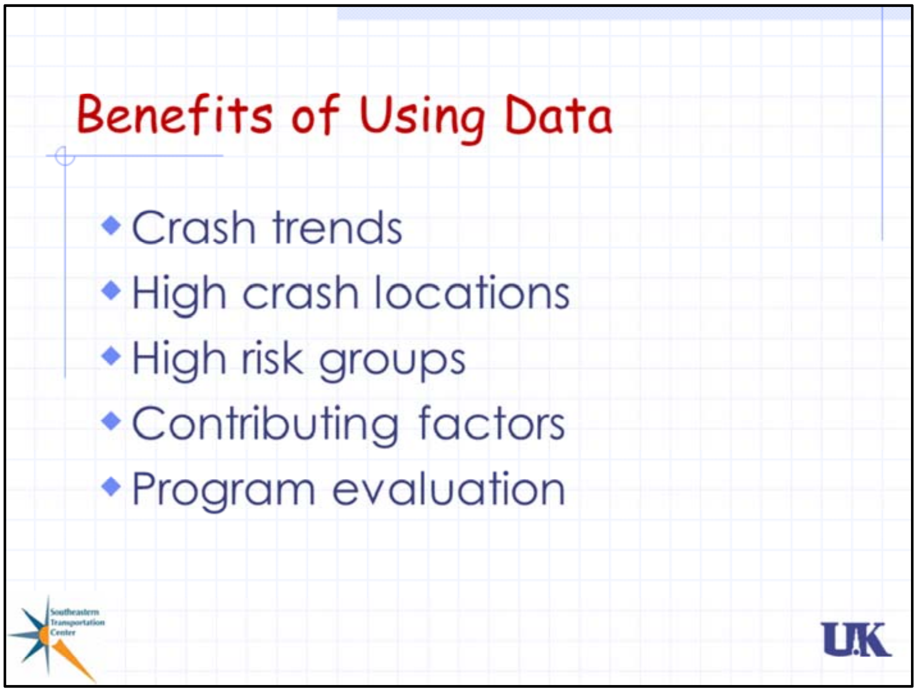


It is important to understand that safety data is an integral part of any decision-making effort. Decisions are improved and are more effective if they are based on comprehensive and accurate data. Frequently, decisions are not based on data due to complexity of safety issues, variety of agencies involved in the decisions, and historical crash data but rather formed based on political priorities, engineering judgment, and conventional wisdom.

Legislative bodies typically make the transportation investment decisions and therefore there is a large political influence on reaching them. Constituents can convince elected officials to press for investment in their areas and this can result in improper allocation of funds. A good example is congestion. The public often cites congestion as their number one transportation problem because it confronts them on a daily basis, when in fact, safety has a far greater impact.

Engineering decisions are often based on the implicit assumption that safety is built into the design guidelines and this can have safety implications that are unknown and most importantly, difficult to measure. Most engineers are not trained to consider the safety implications of their decisions outside of the traditional measures such as providing adequate stopping sight distance. However, recent tools (Highway Safety Manual, Interactive Highway Safety Design Model, Safety Analyst, etc.) can provide the tools to address this. And, by the way, the “judgment” of safety practitioners and law enforcement is often faulty as well.

Conventional wisdom can also affect safety especially since many safety countermeasures have not been evaluated. The use and potential effects of the implementation are often based on conventional wisdom. This means the effectiveness of the countermeasure is generally accepted as true, but no proof has been shown.



The uses of data can result in many benefits when safety issues are examined. Data can provide decision makers with systematic process to determine efforts to improve roadway safety. Data also allows for a quantitative description of the problem and provides means for evaluation of implemented countermeasures and interventions.

Data can assist in determining crash trends and can be used to determine crash rates in order to allow for comparisons among regions with varying environments. Data can also help managers identify high crash locations and identify high risk groups such as younger drivers, older drivers, impaired drivers, and motorcyclists. This can result in development of specific countermeasures targeting the high risk location or group.

Data can also identify contributing crash factors. This examination can assist decision makers in developing programs that could combat these factors and improve roadway safety and reduce crashes. Safety programs are typically developed with an aim to identify safety goals and evaluate strategies. Data is critical in such efforts, since it provides for a comparison of a before-and-after conditions in order to evaluate efforts undertaken and countermeasures implemented. Using data to identify road safety problems allows planners and engineers to effectively communicate safety needs to decision makers in the form of lives saved and injuries prevented.