ROAD SAFETY MANUAL
A GUIDE FOR PRACTITIONERS!

ROAD SAFETY MANAGEMENT

SAFETY DATA

Introduction
Identify Data Requirements
Crash Data System
Non-Crash Data System
Quality and Under-Reporting
Data Analysis
Integrating Data
References

World Road Association (PIARC)

Version 1 - 20/10/2015
5. EFFECTIVE MANAGEMENT AND USE OF SAFETY DATA

KEY MESSAGES

- The establishment and ongoing support of road safety data systems is critical to the effective management of road safety in all countries.
- A lack of accurate data has a serious impact on the effective management and delivery of positive road safety outcomes. The issue of under-reporting in many countries needs to be addressed. Guidance is available on how to improve this accuracy.
- Safety data is used for identifying the size of the road safety problem, analysis of problems, development of strategy, identification of solutions, for advocacy, and monitoring and evaluation.
- Key safety data includes exposure data (traffic volume, population data), final outcome data (deaths and injuries), and intermediate outcome data (average speeds, protective equipment fitment and use, level of drinking and driving, network and vehicle safety quality).
- Countries that do not collect this data need to commence this process as a priority. In the meantime crash injury and survey data should be collected on high risk routes (e.g. high volume roads with a recognised fatal and serious injury problem).
- Detailed guidance is available on how to establish and improve crash data systems, including the assessment of existing data sources, development of a crash report form, engagement with key stakeholders, development of a crash data system, and ensuring the quality of this data.
- Other data is also important for assessing and addressing road risk. There is an increasing awareness of the importance of road asset data (information on safety related road and roadside design elements and features). Where high quality crash data does not exist, this information can be used to identify and treat high risk locations.
- Integration of different sources of safety data can lead to a more comprehensive understanding of the road safety problem, and a greater ability to put effective actions in place to address this problem.
- Accurate and comprehensive road safety data can be used by many stakeholders to help improve road safety.
5.1 INTRODUCTION

Comprehensive safety data is required for effective road safety management. Safety data is essential for an evidence-based approach, particularly in producing results-focused strategies, action programmes and projects; identifying key crash types and locations; diagnosing the causes of serious and fatal injury in road traffic crashes; selecting treatments; and monitoring and evaluating progress. The establishment and support of data systems is specifically identified as part of the Global Plan for the Decade of Action, with Pillar 1 (Road Safety Management) highlighting the importance of this activity (UNRSC, 2011).

Crash data is a key type of safety data, which can provide a valuable source of information to assist in road safety management. However, this is only one type of data required for the effective management of road safety. Crash data needs to be supplemented by other information, including road inventory and survey data of key behaviours, enforcement data, road network and vehicle fleet safety, and emergency and medical system quality. This data is important in providing intermediate measures of safety. In LMICs where crash injury databases are not fully established or operational, such survey data is particularly important for the measurement and targeting of safety problems. Road safety management capacity reviews in LMICs indicate weak capacity for identification and measurement of road safety problems. As such, there is a need to build capacity to improve road safety data collection, storage, analysis and sharing.

This chapter provides information on the types of data required to effectively manage safety; the establishment of data systems; and the collection and use of this data. It also provides guidance on combining the different types of safety data to manage road safety more effectively.

Uses of this data can be found throughout this manual, including on:

- target setting and performance indicators (Road Safety Targets, Investment Strategies Plans and Projects);
- assessing risk and potential solutions (Assessing Potential Risks and Identifying Issues);
- intervention selection and prioritisation (Intervention Selection and Prioritisation);
- monitoring and evaluation (Monitoring and Evaluation of Road Safety).

Much of this chapter discusses the effective management of safety data at network level (e.g. for the whole country). However, it is recognised that establishment of a national data source (although essential) may be some way off for some countries. At the very least, the information in this chapter should be used to commence the collection of data on high risk routes, including through corridor and area demonstration projects (see Road Safety Targets, Investment Strategies Plans and Projects).

As indicated elsewhere in this manual, the focus for effective road safety management is on the elimination of death and serious injury (both of which are defined in Identifying Data Requirements), and this is also where greatest efforts should be made in the collection of safety data. Information on fatal and serious injuries, and the crash types (such as those identified in The Safe System Approach) and factors that lead to such injuries, should form the highest priority in data collection. However, there are also important uses for data on minor injuries and even non-injury crashes, and such information should also be collected where possible.
HOW DO I GET STARTED?

Countries must assess what safety-related information they already collect, who the key stakeholders (collectors and users) are; how this data is used; and what further information is required. Identifying Data Requirements and Establishing and Maintaining Crash Data Systems discuss these issues.

Countries must commence collection of ‘final outcome’ injury data (especially fatal and serious injury data). Initially this may come from high risk routes or corridors (usually high volume national roads). This is discussed in Establishing and Maintaining Crash Data Systems.

Countries must also start collecting ‘intermediate outcome’ data or information on performance indicators (see Identifying Data Requirements). Information on road and roadside elements is a high priority, and can be used to identify problems and solutions; even in the absence of detailed crash data (see Non Crash Data and Recording Systems). Other intermediate data includes compliance data (such as speed, drink-driving and helmet wearing rates; Non Crash Data and Recording Systems). This data can be used to identify issues and solutions, as well as in the broader management of road safety outcomes.
5.2 IDENTIFYING DATA REQUIREMENTS

There are a wide variety of uses for safety data, with many different users. As identified later in Analysis of Data and Using Data to Improve Safety, safety data can be used by policy-makers, traffic engineers, police, the health sector, the research community, insurance companies, prosecutors, vehicle manufacturers and others. Although summary data (particularly on crash fatalities) is available in most countries, more detailed information is required to fulfil the requirements of these users. Without this collection of data, it is not possible to take an evidence based approach to the management of road safety.

WHO (2010) provides discussion on the use of data for a public health approach to road safety. This document provides a comprehensive account of crash data systems, including their place in effective road safety management, and their establishment and use. It is essential reading on this topic, particularly for those working within LMICs who wish to establish or improve upon a crash data system. This document suggests a cyclic approach of:

- using data to define problems;
- identifying risk factors and priorities;
- developing strategy;
- setting targets and monitoring performance.

This process is then repeated.

WHO (2010) also provides guidance on the linkage between safety data and effective safety management (Figure ), giving a framework for the collection and use of this data. The WHO document makes it clear that crash data alone is not sufficient to manage safety, but rather it must be used in combination with other sources of information. This additional information is required to better interpret risks, thereby assisting in the monitoring of performance and achievement of results.
As identified in Figure 5.1 and Box 5.1 (and further discussed in WHO, 2010; GRSF, 2009, 2013), the desired results or outcomes of road safety management are expressed as goals and targets, and occur at a number of different but related levels. These include institutional outputs from the policies, programmes and projects that have been implemented, which influence a range of intermediate outcomes. These intermediate outcomes subsequently influence final outcomes. Ultimately, these should reduce fatal and serious injury, in alignment with Safe System outcomes.

**Final outcomes:** Outcome indicators may include the number of fatalities and serious injuries, crashes relating to certain road users (e.g. pedestrians, motorcyclists) or types (e.g. intersection, head-on), crash rates (e.g. crashes per population, vehicle registrations, or amount of travel)

**Intermediate outcomes:** Safety performance indicators may include behavioural measures such as average vehicle speed, drink-driving, helmet wearing rates, seatbelt wearing, attitude survey information; vehicle safety ratings; infrastructure measures, including road safety ratings, % of high volume high speed roads that are divided by a median, % of roads where pedestrians are present with adequate footpaths; and post-crash care indicators such as emergency vehicle response times.

**Outputs:** Process/implementation indicators may include the policies, plans or programmes that have been implemented and details of this implementation (e.g. campaigns to promote seatbelt use, hours of additional speed enforcement, investment in safe road infrastructure, number of new ambulances).
For example, analysis of data may have identified vehicle speed as a risk factor. A policy to improve compliance with speed limits will require an increase in speed enforcement. The output results of this intervention would be evidence of this increase in enforcement. Intermediate outcome measures might include the percentage of drivers exceeding the speed limit at selected locations. Changes in this measure (i.e. a reduction in speeding motorists) would help identify whether the intervention is having the desired effect. The final outcome indicators would include total deaths and serious injuries (ideally including a record of those that were identified as being speed-related), proving the ultimate benefit of this intervention.

Although crash data is a primary source of safety-related information, other data sources also serve a very important role. There is growing recognition of the use of asset data (including road design features) in road safety, and in many cases this information may already be collected and available for use. As identified later in this chapter, many countries do not have accurate information on crashes, and until such data is available, information about road design features and key safety behaviours provides an important means of identifying high risk locations and ways to address them.

Often different sources of information will be available on similar issues. Although multiple sources of information can be useful to help understand road safety issues more fully, it can also lead to confusion if the sources provide conflicting information. Differences can result from inaccuracies in data or differences in how the data is collected (see Quality and Under-reporting for further discussion of these issues). Where there is potential for confusion from the use of multiple sources of information, it is important to select a ‘single source of truth’ from a data source that will ultimately inform decision-making. Once this source of information is selected, justification needs to be provided as to why this source is preferred.

Different terms for injury severity are included throughout this manual. Definitions for different types of injury are provided in Box 5.2.

**BOX 5.2: DEFINITIONS FOR DIFFERENT INJURY LEVELS**

**Fatal injury:** any person killed immediately or dying within 30 days as a result of a road traffic injury accident, excluding suicides.

**Serious injury:** Injury that requires admission to hospital for at least 24 hours, or specialist attention, such as fractures, concussions, severe shock and severe lacerations. Some countries have adopted the Maximum Abbreviated Injury Scale (MAIS), and suggested that serious injury be defined as MAIS3+.

**Other/minor injury:** Injury that requires little or no medical attention (e.g. sprains, bruises, superficial cuts and scratches).

**Property damage/non-injury:** No injury is sustained as a result of the crash but there is damage to vehicles and/or property.

**Source:** WHO, (2010).
5.3 ESTABLISHING AND MAINTAINING CRASH DATA SYSTEMS

This section provides guidance on the establishment and maintenance of crash data systems. Information on the collection and use of other sources of data can be found in the following sections. Full details on the establishment and maintenance of crash data systems can be found in WHO (2010). The following is a summary of key issues.

SOURCES OF CRASH DATA

Identified effective practice acknowledges that no single crash injury database will provide enough information to give a complete picture of road traffic injuries or to fully understand the underlying injury mechanisms (IRTAD, 2011). A number of countries which have improved their road safety performance use both crash injury data collected by the police as well as health sector data.

National crash data are typically collected by police (WHO (2013) reports that over 70% of countries use police data as the primary source), and entered into crash database systems for easy analysis and annual reporting. In some circumstances, data are collected from hospitals, or from both of these sources. The use of health sector data for meaningful injury classification at country level is necessary to complement police data and to provide an optimal means of defining serious injury. IRTAD (2011) recommends that police data should remain the primary source of road crash statistics, but that this should be complemented by hospital data due to data quality issues and to identify levels of under-reporting (see Section 5.4). Furthermore, in-depth data is needed from crash injury research to lead to meaningful conclusions concerning crash and injury causation.

NATIONAL POLICE-REPORTED CRASH DATA

The police are well placed to collect information on crashes as they are often called to the scene. Alternatively, they may receive information about the crash following the event. Attendance at the crash scene allows for collection of detailed information that is useful for identifying crash causes and possible solutions.

A crash report form is typically completed (traditionally a paper-based form, although recently computer-based systems have been used), allowing collection of quite detailed information on the crash. Key variables typically collected include:

- crash location (including geographic coordinates);
- time of day, day of week, month of year, year;
- information on those who were involved (including road user type, age, gender, injury sustained);
- details regarding the road (whether at an intersection, speed limit, curvature, traffic control, markings);
- details on the environment (light conditions, weather, road surface wet or dry);
- information regarding what happened in the crash (vehicle movement types, objects struck (including off-road), and contributory factors such as speed, alcohol use or driver distraction);
- vehicle factors (type of vehicles involved).

Examples of crash report forms, including the types of detail that should be collected as a minimum can be found in WHO (2010). The advice provided in the WHO document is based on the European Common Accident Dataset (CADaS). In addition, a number of countries have developed their own minimum criteria. For example, the US has a Model Minimum Uniform Crash Criteria (further information is available on a dedicated website at http://www.mmucc.us/).
A balance needs to be reached between collecting the required information, and the time it takes to perform this task. If too much burden is placed on the police, it is less likely that the crash report form will be completed. Police are key stakeholders in the establishment and continued collection and use of crash data, and should be included at each stage of the process.

**HOSPITAL DATA**

Hospital data is used to identify levels of under-reporting or to obtain better injury information, particularly when police report data is not available or is inadequate. IRTAD (2011) suggests that because of under-reporting of crash data, hospital data should also be collected, and is the next most useful source of information for crash statistics.

Encouraged by the WHO and other institutions, medical authorities have established international recording systems that include road traffic injury. In particular, the International Classification of Diseases and related Health Problems (ICD) and the Abbreviated Injury Scale (AIS) coding systems are used widely. IRTAD (2011) recommend that an internationally agreed definition of ‘serious’ injury be developed, and that the Maximum Abbreviated Injury Scale (MAIS) be used as the basis for defining crash injury severity. This scale is based on maximum injury severity for any of nine parts of the body. A score of 3 or greater for one or more regions of the body (MAIS3+) is recommended as the point at which an injury is considered to be serious.

There are examples where enhanced data has been collected in hospitals thereby allowing more detailed analysis, as in the case study from Thailand below (Box 5.3).

**BOX 5.3 : CASE STUDY – GIS SPATIAL ANALYSIS OF HOSPITAL DATA, KHON KAEN PROVINCE, THAILAND**

The problem: Although information on crashes existed from other sources, it was not sufficient to adequately manage safety in Khon Kaen province, Thailand. An alternative source of information existed from hospital data, but this did not provide information on crash locations.

The solution: Khon Kaen Regional Hospital in the northeast of Thailand has had its own Trauma Registry since 1989. Additional information on road crashes has been added to this database through the use of an ‘Accident Report’ form. This form includes information on the circumstances of the crash, factors that may have influenced the occurrence of the crash, and descriptive information regarding the location of the crash. The development of the GIS spatial analysis of hospital data involved revising the accident report form and the injury surveillance forms. The accident report form was modified to exclude information on property damage as this required site inspections and the injury surveillance form was modified to reflect all injured person information e.g. the use of seatbelts etc. In addition to basic crash data, all crash locations were spatially coded using a Geographical Information System (GIS). A link recognition technique on GIS was developed to obtain road geometry properties. Using GIS mapping allowed for the visualisation of crash location features e.g. intersections, level crossings and roundabouts.

The outcome: The addition of this information has allowed detailed analysis of crash data, including the identification of high risk locations. A number of detailed studies on specific road safety issues have been undertaken over recent years based on this rich source of information.

Further details can be found in Ruengsorn et al., (2001).
A further example is provided in Box 5.4 from Egypt, demonstrating the integration of data from different sources, as well as the use of this data by various key stakeholders.

BOX 5.4 : CASE STUDY – EU TWINNING PROJECT ON ENHANCING ROAD SAFETY IN EGYPT

The problem: The need for a centralised and accurate accident data system.

The solution: The National Accident Management system is intended to be built as a new system for both investigating and analysing road accident data and it is intended as a shared ‘space’ for all parties responding to different kinds of road accidents, such as medical rescue services, police forces, fire brigades, hospital personnel and the road administration. The procurement started in the last quarter of 2010 and included the hardware and software of GPS based mobile terminals, to be used at accident sites for determining the location, documentation, reporting and information by police officers and ambulance personnel.

The new system will install the unified process for accident statistics and investigation, match the data from different sources and provide the different researchers on road safety with the necessary accident data. The handheld devices for the accident registration will contain the accident report sheets in the form of multiple choice tables, easily to be filled in at the accident sites and organized in a way which will prohibit personal data from being given to authorities other than the court, insurances and hospitals.

The outcome: The main outcome was that underreporting was successfully diminished and a close cooperation between the key stakeholders and a detailed system of information for accident investigations was achieved.

A key lesson in this exercise was that GPS based digital accident data management is a crucial tool for road safety policy and the infrastructure safety management as well. The data management system needs to overcome police restrictions on accessing accident data.

Source: Hans-Joachim Vollpracht, World Road Association (PIARC)
REGISTRATION SYSTEMS

‘Vital registration’ data can be used as a source of information on road deaths. This information comes from death certificates completed by doctors which state the cause of death. WHO (2010) reports that around 40% of WHO member countries collect vital registration data of the detail required for monitoring road traffic deaths. WHO and other organizations have instituted an international registration system that includes those injured in traffic crashes.

OTHER CRASH DATA SOURCES

Other sources of data on crashes can come from emergency services, tow truck drivers, members of the public, insurance companies, etc. However, it is important to recognise that the quality and extent of this information may be limited when compared to police and hospital reported data.

ESTABLISHING OR IMPROVING CRASH DATABASES

Before establishing a new crash database system (or improving a current system), it is recommended that a situational assessment be undertaken (WHO, 2010). This involves:

- stakeholder analysis;
- assessment of data sources and existing systems/quality;
- end user needs assessment;
- environmental analysis.

These same steps are also required when establishing or improving on the collection of non-crash data.
A stakeholder analysis involves identifying organisations and individuals who have (or should have) a role in the collection and use of road safety data. Critical stakeholders will include police, transport agencies and health departments, but there are likely to be many others.

An assessment of data sources is required to determine what information is already collected, and the quality of the data. This is often a significant problem in many countries.

An end user assessment involves understanding who the key users are and, how these key stakeholders use the information. This knowledge will help improve the usability of the data.

An environmental analysis involves understanding the political environment and critical partnerships required for the successful collection, analysis and use of the data. Without this understanding and appropriate collaboration it is likely that collection and use of crash data will be severely hindered. There are many examples where expensive crash data systems have been established, but data has not been entered into the system due to inadequate communications and poor cooperation.

Following this situational assessment, the recommended process for establishing a crash database system is to:

- establish a working group (typically drawn from the stakeholder analysis);
- choose a course of action;
- recommend minimum data elements and definitions;
- design and implement the system (or improve the current system) (WHO, 2010).

Crash location is a key element in collecting and analysing data, particular for road engineers. Without this information, it is not possible to determine what locations to treat in the future. In addition, if the crash location is known (whether from police reports or other sources of data), there is potential to link this crash data to asset or other data sources. This information may be of use in identifying other road-based elements that may have contributed to the crash risk.

Several methods are available for the accurate location of crashes, including the use of global positioning systems, reference to a local landmark (e.g. a link-node system), or reference to a route kilometre marker post (a linear referencing system).

Historically, crash data records were kept in paper-based filing systems, but now computerised database systems are used to store information on crashes. This allows relatively easy analysis of data, and is particularly useful in the identification of trends, high risk locations or areas, key crash types, etc. There are a number of computer software packages available for this task. At a minimum, such a system should have the capacity to:

- produce quick summaries on key crash variables (e.g. total crashes; crashes by different severity; crashes by different user groups, such as pedestrians; trends over time);
- allow cross-tabulations (i.e. analysis of two or more variables at a time (e.g. change in fatal pedestrian crashes over time);
- identify hazardous locations (e.g. locations, routes or areas with concentrations of crashes, or crashes of a particular type).

Crash data systems have become very advanced in recent years, with features added that allow quicker and more useful analysis. WHO (2010) and Turner and Hore-Lacy (2010) provide a list of other desirable features of crash data systems. These include:
• built-in quality checks (algorithms and logic checks);
• GIS linkage to allow accurate identification of crash location;
• ability to add new data fields without re-developing the database;
• database navigation features such as drop-down menus, clickable maps;
• pre-defined queries and reports;
• options for customised, user-defined queries and reports;
• mapping ability for data entry, crash selection and presentation of aggregated crash information;
• ability to export data to third-party applications (e.g. Microsoft Excel, Statistical Analysis Software (SAS)) for further statistical analysis;
• inclusion of crash narrative, sketches of crash scene, photographs and videos linked to crash;
• automatically generated collision diagrams;
• site ranking based on crash rates, numbers, costs;
• ability to monitor sites of interest, i.e. before and after treatments.

An example of the successful implementation of a crash data system is provided in the case study below from Cambodia (Box 5.5)
The problem: The database system for Cambodia was inadequate to adequately manage effective road safety outcomes. Since 2002, three different ministries were involved in road crash data collection in Cambodia (Ministry of Public Works and Transport, Ministry of Interior (MoI), and Ministry of Health (MoH)). Although the databases developed by the three ministries provided some relevant information on the road safety situation in the country, a need for improvement was observed as the databases were not compatible, there were important discrepancies between them, they were under-reporting the real situation, and they were limited in their scope.

The solution: In 2004, Handicap International Belgium and the Cambodian Red Cross proposed a new system, based on a standardised and more detailed data collection form. The objective of the new system is to provide accurate, continuous and comprehensive information on road crashes and victims for the purposes of increasing understanding of the road safety situation, planning appropriate responses, and evaluating impact of the road safety initiatives. The system was initially piloted in the capital city of Phnom Penh in 2004. The pilot’s success resulted in expansion to all provinces in 2006.

Handicap International (HI), an international aid organisation, played a key role in the development and implementation of the RCVIS and was initially responsible for overseeing the entire system. HI has slowly transitioned the work to the Cambodian government, moving data collection responsibilities to the MoH and the MoI in 2008–2009 and transferring system oversight to the National Road Safety Committee (NRSC) in 2010.

The system combines data from the police and health sectors. The structure of RCVIS (from 2009 onwards) is shown below:

![Future Structure of the Systems](image-url)

Figure 5.3 - Source: Chariya Ear, Handicap International, Cambodia.

The outcome: Combining both data sources has enabled the system to cover more elements on detailed crash information (from police) and casualty information (from hospitals). This has led to higher levels of reporting. There were several factors that helped with the success of the RCVIS. One factor was the development of the two RCVIS forms at the same time — one for traffic police and another for the health sector. This provided a common understanding on the technical aspects of the system, which required common questions (variables) in both forms. The main parties were all involved in this development stage, which also meant that they participated and approved the common sections of the two forms. Secondly, the Ministry of Public Works and Transport was empowered to play the
coordination role, with the understanding since the beginning that they would combine the two data sources. They were supported to organise meetings, lead discussion, chair RCVIS annual workshops, and become a co-author of the RCVIS publication, etc. Thirdly, the findings and data from the pilot phase helped all parties understand key issues and refine the process. It was clear that the reports from the pilot phase were highly appreciated by key government stakeholders, media, and the working group members.

Although there have been significant advancements through the RCVIS, there are still issues with the system. Data reporting is inconsistent, particularly from the health facilities. To address the issue of non-reporting, it is recommended that the NRSC and the Ministry of Health take steps such as increasing buy-in from senior level administrators in health facilities, formalising and standardising RCVIS reporting procedures, and showing health facilities how RCVIS data can be useful to them.

Source: Chariya Ear, Handicap International, Cambodia.

The Swedish Strada system is a unique database that integrates police and hospital data. It is important to recognise that although this linkage provides valuable additional information, it does occur at additional cost. Further details are provided in Box 5.6.

**BOX 5.6: CASE STUDY – SWEDISH TRAFFIC ACCIDENT DATA ACQUISITION (STRADA)**

The problem: Lack of reliable information on crashes, including information on injury outcomes.

The solution: In October 1996 the Swedish Road Administration was commissioned by the Swedish government to initiate a new information system covering injuries and accidents in the entire road traffic system. This was accomplished in co-operation with the Swedish Police, the Swedish National Board of Health and Welfare, the Swedish Institute for Transport and Communications Analysis, Statistics Sweden and the Swedish Association of Local Authorities and Regions.

Strada is based on information reported from two sources; the Swedish police and hospitals. All police districts, since 2003, report to Strada on a national scale. Strada also receives information from an increasing number of hospitals. The incorporation of hospital data makes this method very different from earlier methods of registration of injuries and accidents in the road transport system.

The outcomes: By bringing together data from two sources – the police and the hospitals – Strada provides more detailed information, thus increasing the knowledge of road traffic injuries and accidents. Incorporating hospital data decreases underreporting since the police have limited knowledge about some road traffic accidents (mainly involving unprotected road users: pedestrians, cyclists and moped drivers). In addition, the hospitals’ reporting of diagnoses broadens the knowledge of the injuries and their degree of seriousness.

The data are entered by the local police and hospitals into Strada, where a match immediately takes place.

To date, the level of accuracy in data matching is very high. Matching means the data was recorded by both the police and the hospitals. A more complete description of accident circumstances and injuries involved is obtained from matched data. This results in more accurate information on the real traffic safety problems which in turn facilitates the planning and prioritisation of traffic safety measures.
In 2013, the police recorded 15,000 accidents with 20,600 injured persons. They were matched with 32,700 persons who sought treatment in hospital. Of these, 9,800 were matched, that is known to both sources. It also means that 48% of those known to the police were also known to the hospitals; also 30% of those registered by the hospitals were registered by the police. One conclusion from this match is that the police reporting is far from comprehensive.

Since a number of hospitals do not yet report to Strada, the existing official statistics are based exclusively on accidents reported by the police. The information derived from the hospitals is shown in a supplement containing medical statistics.

Some countries have undertaken in-depth studies of serious crashes to provide a more thorough understanding of crash causation factors and possible solutions. Such studies typically investigate a sample of high severity crashes. As an example, in the UK, the ‘On the Spot’ project collected detailed and high-quality crash information over two regions. More than 2000 variables were collected for each incident based on scene investigation soon after the crash occurred, as well as follow-up communication with medical services and local government. The information was analysed to provide insight about human involvement, vehicle design, and highway design in crash and injury causation. Mansfield et al. (2008) provide an initial analysis of around 2000 incidents from this program. Such investigation can provide far more detail than what is typically available through a crash report, and to a higher degree of reliability.

Similar examples can be found in many other countries, including the USA, Germany, France, Malaysia, India and Australia. Some of these programmes have been in place for many years, and have produced large amounts of valuable information. One of the key outputs from the EU DaCoTA project (which collected and analysed data from European countries on various road safety topics) was guidance on the collection of such data, as well as standardised procedures (Thomas et al., 2013). A Pan-European In-depth Accident Investigation Network has been established, and tools such as an online manual for in-depth road accident investigations have been developed (see http://dacota-investigation-manual.eu).

The US has established the second Strategic Highway Research Program (SHRP2). SHRP2 is perhaps the most comprehensive database of information on factors occurring before and during crashes and near-crash events. The information collected includes data from the Naturalistic Driving Study (NDS) database. This dataset includes information from over 2300 drivers, collected through equipment installed in their own vehicles, and through normal driving. The massive amount of data collected through the NDS is supplemented through the Roadway Information Database (RID) which includes comprehensive information on road elements in the study areas as well as other relevant data (including crash data). This globally significant database is expected to provide the research basis for studies on driver performance and behaviour. More information can be found at http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/Pages/Safety_153.aspx.
DATA SHARING

Sharing of data from different sources is required for the comprehensive collection, analysis and integration of data. Efficient data sharing, particularly between the police and the highway authority, is essential for good practice road safety management.

However, it is important to note that some organisations may be reluctant to share certain data, particularly personal identifiers, due to the issues it poses surrounding the privacy and anonymity of those involved. One response is to collect the personal details on a separate page of the crash report form (e.g. name and address information). This page can then be removed before sending the remaining pages on to partner agencies. In some cases, it may be necessary to develop appropriate privacy policies to ensure this issue is addressed, or for certain variables to be removed to prevent the identification of individuals.

Crash data on its own is a valuable source of information on crash risk, and when combined with other sources of data, this value can be greatly increased. The following section discusses some of the other data sources, while Analysis of Data and Using Data to Improve Safety discusses combining these sources.
5.4 NON-CRASH DATA AND RECORDING SYSTEMS

Crash data is generally considered a key source of information when assessing and treating risk. However, in some countries, particularly LMICs, crash data may not be reliable or available at all. Additional surveys and data sources other than crash data may be the only reliable source of safety data available. As discussed in Identifying Data Requirements, this additional information (safety performance indicators) is also important in road safety management. These enable assessment of different policies, programmes and projects to identify their effect on road safety outcomes. This occurs through the collection and assessment of details on the interventions implemented and the intermediate outcomes.

A variety of other data sources are available, including information on road design and features, traffic data, survey data and exposure data.

ADDITIONAL SOURCES OF INFORMATION

ROAD INVENTORY

Road inventory data are a major source of information that can assist in assessing safety. Because the impact of different road elements is well known, different elements or combinations of elements can provide vital insight into crash problems, including the key crash types contributing to fatal and serious injury outcomes (see The Safe System Approach). The following data are of particular use:

- road class – e.g. freeway, arterial, local road;
- road width and type – divided (plus width and type of median) or undivided;
- adjacent land use – e.g. residential, commercial;
- lane number and widths – in each direction;
- intersection or crossing type – cross or ‘T’ intersection, roundabout, railway or other;
- traffic control devices – signals, stop signs, give way signs;
- road alignment – horizontal and vertical alignment;
- street lighting;
- road surface – type (asphalt, concrete, brick, unpaved) and condition (roughness, potholes, skid resistance);
- shoulders – width, type (paved, unpaved);
- speed limit;
- roadside hazards – including distance to hazard and type;
- pedestrian facilities – including presence of footpath, pedestrian crossing and type.

This presents a basic list of relevant road element types, but there are many other factors that may have an influence on safety outcomes. The International Road Assessment Programme (iRAP) collects around 70 attributes (see http://www.irap.org/about-irap-3/methodology for details of these attributes, and Proactive Identification for more information on iRAP, and Box for examples of the data collection undertaken in Mexico). In the United States, the Model Inventory of Roadway Elements (MIRE) provides a list of 202 elements that may be needed for making road safety decisions. Further information on MIRE can be found at http://www.mireinfo.org/about.html.
BOX 5.7: CASE STUDY – ROAD SAFETY DATA COLLECTION IN MEXICO

The problem: Nationally, Mexico has about 18,000 deaths in road accidents per year, with an annual average growth of 1.9%. Of these deaths, 25% occur in the Federal Highway Network, which constitutes the backbone of the road and national transportation system.

The solution: In order to improve road safety within the country and thereby the quality of life of the population, as well as safety conditions along the major national and international trade corridors, Mexico launched a preventative project aimed at assessing 46,000 km of Federal Highways during a first stage in 2012 and 19,000 km of collector roads during a second stage in 2013. The start of this project coincided with the launch of the UN Decade of Action for Road Safety, to which Mexico adhered, so that this project was thought as a breakthrough to improve road safety and as the start of a periodic road assessment program.

The outcome: The project, in each of its two stages, made it possible to generate a safety rating for each type of road user and the investment plan at 20 years for safer roads. For the 46,000 km assessed during the first stage, the evaluation in both stages showed about 50% of the total length assessed with one or two stars for vehicle occupants and higher percentages for the other more vulnerable road users. The project has made it possible to address significant problems throughout the system by implementing large-scale treatments (i.e. installation of median and roadside barriers), as well as to incorporate asset management measures (i.e. signs and pavement markings) within the annual program of the maintenance and rehabilitation agency, and to apply the results in the updating of the road design standards. It is also making it possible to count on an updated inventory of the roads assessed. The adoption of this technology as a periodic program will make it possible to monitor the effectiveness of the measures and programs implemented.

Case study provided by Dr Alberto Mendoza, IMT, Mexico.

Road features need to be spatially located (ideally through a GIS-based system) to allow effective analysis and cross-linkage.

Road inventory data relevant to road safety may already exist (e.g. through road asset database systems) or it might need to be collected. A situational assessment should be performed to see whether this data exists (see Establishing and Maintaining Crash Data Systems). Road inventory data has traditionally been used for road safety audit or inspection (see Assessing Potential Risks and Identifying Issues), but in more recent times, methods have been developed to quantify the likely safety outcomes based on these elements.

The collection of this data can be based on ad hoc approaches (e.g. through periodic inspection, public complaints, etc.), but should ideally be through a comprehensive programme conducted on a regular basis. The most common approach involves the collection of data through video images, and subsequent rating or coding of this data by trained experts. This information is then fed into a database or asset register (see Box 5.8).
Although the collection of road inventory data is highly useful, it needs to be done in a way that minimises costs (i.e. it must be quick), it must be accurate, and it must be safe. Several methods for data collection are available. In the most basic technique, data can be recorded on data collection sheets while travelling along the road of interest. This approach can only really be used on relatively short lengths, and it can be difficult to collect all relevant road variables when travelling at normal traffic speed.

For more extensive data gathering, computer-aided collection can be undertaken using a tablet or laptop, or information can be collected via video and coded safely back in the office. With a tablet, information is added to a database while travelling along the road of interest. Touch screen technology is typically used to select relevant road variables. Different symbols may be displayed on screen to facilitate quick data entry. As mentioned earlier, it is often difficult to enter all relevant variables when travelling at high speeds or in busy environments, and so video is often recorded to assist in later data entry and checking.

Another option involves the desktop assessment of video data. One or more video cameras can be used to collect data along the network of interest. A single camera can be used to gather information in a forward direction. Alternatively, more cameras can be used to allow better collection of road and roadside information. This video imagery is then used to code the variables of interest. The video images can be calibrated to allow measurement for more accurate collection of information (such as road width or distance to roadside hazards) and to ensure accurate spatial location of assets.

Video images are assessed, and can be paused to study more complex environments. Information from the images is added to a database for later analysis. This data entry may be through a dropdown menu system or manual population of a database. Data is typically collected for a discrete length of road (e.g. a 10 m section).

Figure 5.4: Populating a database with safety-related inventory data.
New technologies are being developed that will assist in more automated collection of road and roadside data. As an example, it is possible to collect information on features such as road width, horizontal and vertical alignment, and road surface condition using Light detection and ranging (LiDAR) and other vehicle sensors.

**TRAFFIC DATA**

Traffic data is important to collect and analyse, particularly traffic volumes (or Average Annual Daily Traffic, AADT). This data can be used to generate crash rates, which provide a good indication of safety performance, including the safety performance of specific routes, road types or even infrastructure elements. Other types of traffic data include the traffic mix (e.g. percentage of different vehicle type; motorcyclists, bicyclists and pedestrians) and vehicle speeds (mean and 85th percentile speeds, compliance with speed limits). Traffic data can be collected using manual traffic counts or through automated means (e.g. pneumatic tubes or permanent data collection devices installed in the pavement).

**OTHER EXPOSURE DATA**

Aside from traffic data, other sources of exposure data include population data (total number of people; number by each age group) for an area or country. This data is typically available from national census data. Vehicle registration data is also often collected and used.

**ATTITUDE SURVEY DATA**

Attitude surveys collect information on the views of drivers, other road users and residents. This information is considered an important source of feedback for assessing the effectiveness of a new programme or treatment, and can provide insight into driver behaviour (for example, low compliance levels with the posted speed limit).

**ENFORCEMENT DATA**

Information on the number of police checks (e.g. for speed, alcohol, restraint use), number of violations (e.g. number of vehicles speeding; motorcycle riders without helmets), and number of drivers punished (e.g. fined, penalties provided or imprisoned) are all useful measures. These will help assess the impact of new policies or actions on safety outcomes.

Other sources

In addition to the data sources mentioned above, other useful information can be gathered from the following sources, where available:

- hospital files – an additional source of information on crashes other than police files, which provides more detailed information on the injuries sustained, the length of stay in hospital, etc. (see Sources Data in Establishing and Maintaining Crash Data System for further detail);
- maintenance and operations files – information on the type and timing of any maintenance work;
- expenditure data – information relating to spending on safety-related initiatives;
- project history files – information relating to any previous major corrective works, including safety infrastructure improvements;
- insurance company data – crash history of the driver and the car;
- weather reports;
- vehicle data, including information from periodic vehicle inspection.

Other types of compliance data are discussed in the following section.
ADDITIONAL SURVEY METHODS

Often traffic data and driver behaviour data are not readily available. There is no set list of additional data that must be collected, and given the cost of any type of data collection, careful thought needs to be given to this task, regardless of whether this is conducted at national level or for specific locations. Additional data should be collected when there is a need for it, and collection of this data should be cost-effective.

The following section provides a brief description of some of the more common data surveys performed, as well as different methods that can be used. References to useful material is also provided.

**MEASURING SPEEDS AND TRAFFIC VOLUMES**

A spot speed survey involves the collection of a sample of speeds at a specific road location, or at a number of locations. This can then be used to determine the speed distribution of vehicles, which is useful for the following reasons:

- establishing the speed of vehicles (e.g. mean and 85th percentile speed) on the selected road;
- checking the proportion of vehicles that are travelling over the speed limit;
- assessing the likely contribution of vehicle speeds to crashes;
- checking for differences in speeds between different road users;
- tracking these speeds for monitoring purposes;
- evaluating the effectiveness of a new treatment installed to reduce speeds.

Vehicle speeds can be measured using manual methods (radar or laser guns, stop watch), or using automatic methods (loops or tubes). Automatic methods are better for studies that require a larger sample. Loops and tubes can also record more than just average speeds, such as traffic volumes, vehicle turning movements and traffic mix. These components are essential to understanding the safety issues that exist at a location. The GRSP Speed Manual (GRSP, 2008) and the UK government (DETR, 2001) provide in-depth guides to speed and volume measurements and how to manage speed-related safety
issues. See Box 5.9 for a case study on speed data collection in India.

**BOX 5.9: CASE STUDY - SPEED DATA COLLECTION**

The problem: data was required on vehicle speeds for a study in four states in India.

On-site traffic speed, traffic volume and crash data on four sample sites were selected from the road corridors that formed part of the 2011 Four States Project in India. The research team conducted traffic operating speed and volume studies for determining the 85th percentile speed and the traffic density on the corridors by type of vehicle. Researchers also collected available police crash data and conducted on-site crash investigations for a period of up to two months to obtain a better understanding of the accidents. These crash site visits were made possible through close liaison with the local police departments and also through regular routine patrols of the road sections.

The study was conducted on four state highways (two in Karnataka, and two in Gujarat) totalling around 300 km.

The outcome: It was found that the 85th percentile traffic speeds were well above the posted speed limit (in some cases the posted speed limit was not obvious). Large differences in speed were observed between types of vehicles. Cars had much higher 85th percentile speeds than other road users such as motorised two-wheelers, buses and trucks. The spread of speed for the same vehicle type was also quite high. For example, motorised two-wheelers travelled at speeds ranging from 25 to 80 km/h.

Some of the key learnings from this data collection were that:

1. Due to the lack of detailed data held by some road authorities and police departments in India, on-site speed and traffic volume data collection such as this is needed to truly understand how the road is being used.
2. With rapid urbanisation, traffic speed and volume studies need to be conducted on an ongoing basis.
3. Proper documentation needs to be maintained to ensure uniformity in data collection and, most importantly, data interpretation.
4. The activity can be recorded on video cameras to ensure that the data sample can be rechecked; this also acts as a safety measure.
5. Better equipment may help realise larger samples and faster data collection, but there must be a way to check the data collected to ensure reliability and accuracy.
6. The safety of the data collection team and other road users is of paramount importance. Careful planning and risk assessments are recommended prior to working in or adjacent to the live carriageway.

In addition to the production of safety ratings, the speed data led to recommendations for the setting of evidence-based speed limits as well as follow-up enforcement of these limits.

**Source:** Ravishankar Rajaraman, Manager - Safety Group, JPRI.

---

**MEASURING SEATBELT AND CRASH HELMET USAGE**

GRSP has developed two separate manuals, one dedicated to seatbelts and child restraints (GRSP, 2009) and the other dedicated to helmets (GRSP, 2006). Each of the manuals provides information on how to assess the extent of non-seatbelt and non-helmet use in a project region, as well as how to design, implement and evaluate a programme to target these issues. With regard to measuring seatbelt and
helmet usage, the guides list possible sources of this information, as well as how to collect the data through conducting community surveys and observational studies.

### MEASURING DRINK-DRIVING LEVELS

Much like measuring helmet and seatbelt usage, GRSP has developed a road safety manual for drink-driving (GRSP, 2007). This provides information on how to assess the situation and choose priority actions, as well as how to design, implement and evaluate a drinking and driving programme.

The guide suggests collecting data from relevant authorities, such as the police, road authorities and health sectors, to understand the size of the problem. Data on the level of compliance with existing laws can be collected through a combination of crash data (i.e. crashes involving drivers and riders with Blood Alcohol Content (BAC) levels exceeding the legal limit), the number of alcohol offences detected by police, the percentage of drivers stopped with a BAC over the legal limit, and by performing driver surveys (GRSP, 2007).

There are a variety of other intermediate outcomes that could be measured, depending on the interventions implemented.

### DATA RECORDING SYSTEMS

As for crash data, it is important that survey data be recorded in a way that can be analysed easily. It is also beneficial for the system to be developed so that data can be linked with other data sources. This is particularly important for surveys that cover a broad geographic area (such as traffic volume, asset or population data). Such systems may already exist for this data. A common method is to link data by location using a Geographic Information System (GIS). These systems can typically store information that is linked geographically for future analysis. Different types of data can be added to such systems as a ‘layer’, allowing more powerful assessment of risk (see Analysis of Data and Using Data to Improve Safety).
5.5 DATA QUALITY AND UNDER-REPORTING

When collecting, managing or utilising road safety data, it is important to remember that data quality can be compromised at any stage of the data process. This can be due to:

- missing or incomplete data, or errors in data collection and entry;
- differences in the application and understanding of variable definitions;
- under-reporting (WHO, 2010).

There are a number of consequences associated with poor data quality and under-reporting of crash data (Austroads, 2005; OECD, 2007). Some include:

- Lower numbers of casualties will reduce road safety as a public health issue and therefore it will be less likely to attract funding;
- Misleading information may cause road authorities to make ineffective and faulty road safety decisions and set inappropriate priorities;
- The success rates of implemented countermeasures cannot be fully assessed;
- Comparisons between jurisdictions and countries cannot be accurately made.

This section will consider factors that affect data quality, as well as methods for studying inconsistencies in data and how to improve data quality. Although this section concentrates on crash data, quality issues are also relevant to non-crash data, and care needs to be taken in the collection and interpretation of this.

MISSING OR INCOMPLETE DATA AND ERRORS

Data can sometimes be recorded incorrectly by the police or data entry staff. A major issue to note is that the person who fills in the form at the scene is, in most instances, not the same person who enters the data into the database (OECD, 2007). Missing, incomplete and incorrect data is often unintentional and is the result of human error. Due to officer priorities and workloads, the police cannot always attend the scene of a crash or may not have the time to completely fill out the crash report (which can be made worse by unnecessarily long data collection forms). Unclear variable definitions, as discussed in the next section, can also result in incomplete or incorrect data entry. Similar issues can also occur with non-crash data. For example, road asset data can be coded incorrectly, or data entry errors made during the analysis of speed data.
DIFFERENCES IN VARIABLE DEFINITIONS

The definition of each variable (crash type, injury severity, location, etc.) can differ between data sources (for example, police crash files, hospital records, insurance claims), jurisdictions and countries. This can lead to complications in the identification of crashes of interest, the comparison of datasets, and the evaluation of data quality within a dataset. Common confusing definitions are discussed below.

INJURY SEVERITY

The most common categories of injury severity are fatal, serious/severe and slight/minor injury. However, the exact methods used by police and hospital staff to determine which injuries fit into which severity categories can be problematic.

A recurring issue when comparing datasets from different countries is the timeframe that applies to ‘fatal’ injuries and crashes. The 30-day rule defines a fatal crash as when any person is killed immediately or dies within 30 days as a result of a road crash injury, excluding suicides. The 30 day rule is the most common classification used around the world, particularly by high- and middle income countries (WHO, 2010). Other countries, particularly lower-income countries, use the definitions of ‘at the scene’ or ‘within 24 hours’ to classify fatalities, which can create inconsistencies between databases. Adjustment factors have been developed to account for this (WHO, 2010); however, this assumes that similar proportions of vulnerable road users exist in each system, which is not necessarily the case (WHO, 2010).

The 30-day rule also implies that there is some coordination between the police officers who attended the scene and hospital staff in order to check for updates on patient status after 30 days. This is often not the case due to different priorities and workloads of those involved (WHO, 2010). The same issue arises with regard to non-fatal injury classification: a serious/severe injury is often classified as ‘admission to hospital’; however, police often classify this as all people who leave the scene in an ambulance (Austroads, 2005). Similarly, there is variation in what hospitals consider to be a ‘serious injury’ (see IRTAD 2011 for a...
detailed discussion of this issue). An increasing number of patients are being referred to specialist clinics (e.g. fracture clinics) instead of being admitted to hospital. Therefore, in some databases it is difficult to tell whether trends showing fewer admissions are a result of a change in the severity of crashes or a change in the health care management system (Ward, Lyons & Thoreau, 2006). IRTAD (2011) recommends that serious injury should be determined by trained hospital staff and not the police at the scene of a crash. In reality, such checks on crash severity outcome are often not made, and it is to the police in attendance to determine the severity outcome.

In some countries ‘property damage only’ or ‘non-injury’ crashes are required to be reported, and in others they are not. Sometimes the level of damage must exceed a certain monetary limit before it must be reported. Such additional information can be of use, especially in the identification of crash locations and likely causation, although it does entail a higher cost in terms of data collection and entry.

**ROAD TRAFFIC CRASHED**

The definition of a road traffic crash may incorporate or exclude crashes involving non-motorised vehicles. It may also exclude crashes that occur on private roadways or in off-road locations such as parks and parking lots. On the other hand, some countries collect information regardless of the location (WHO, 2010).

Another common issue is that hospital outpatient files often simply focus on the nature of the injury (e.g. broken femur) and sometimes neglect to mention the external cause of the injury. This can make it practically impossible to identify which cases are crash-related, and it also reduces the information available to identify and treat crash locations (WHO, 2010).

**LOCATION**

There are a number of different methods used to determine the location of a crash (as discussed in Establishing and Maintaining Crash Data System). Each of these methods can be subject to error, which can lead to inaccurate or non specific crash locations recorded by police. This can make it difficult to assess the significance of particular crash locations.
UNDER-REPORTING

Under-reporting can occur at any point in the data collection and data entry processes. WHO (2010) discusses the factors contributing to under-reporting in police data and health facility data in detail. Under-reporting often varies with crash severity, transport mode, road user types involved, victim age, and the crash location. Common findings are that (Austroads, 2005; Ward, Lyons, Thoreau, 2006):

- Crashes only involving one vehicle are less likely to be reported than multi-vehicle crashes.
- Reporting rates vary with hospital type (e.g. rural, private, etc.).
- The higher the injury severity, the higher the reporting rates.
- The older the victim, the higher the reporting rates.

This under-reporting issue can be a significant problem in all types of countries, but has been a particular issue in LMICs (see Box 5.10 and Box 5.11).

BOX 5.10: UNDER-REPORTING OF CRASH DATA

The Global Status Report (WHO, 2013) uses estimates based on a regression model for countries that do not report death registrations to the WHO in a specified format. In many cases, the WHO estimates differed considerably with the officially reported road deaths. A number of countries were estimated to record only around 15–20% of deaths, while in one case the estimate was just 2.5%. Obviously more needs to be done to improve reporting rates. The case study on RCVIS in Cambodia (see Section 5.3.2) highlights an approach that can be taken to improve reporting rates. This involved the collection of data from two main sources, an approach that has significantly improved the situation in Cambodia. A similar approach has been taken in Indonesia. In 2009, steps were taken to improve data collection, including combining the police data with insurance and hospital data. As indicated in the figure below, reporting of data improved substantially following this action.

Figure 5.5 - Source: WHO, (2013).

However, as indicated by WHO (2013), this activity had the unintended outcome of indicating a
substantial increase in road crashes for 2010. This apparent increase is not a result of an actual increase in road deaths, but rather an improvement in the recording of existing deaths. Several countries are experiencing similar apparent increases in road deaths, when in reality the level of data accuracy has improved. The improved data allows for better identification and management of road safety issues. However, the impression that crashes are increasing substantially is an issue that also needs to be managed.


**BOX 5.11: DIFFERENCES BETWEEN POLICE DATA AND HOSPITAL DATA ON ROAD DEATHS IN MEXICO**

In Mexico, the National Institute for Statistics and Geography (INEGI) generates the official statistics, based on the collection of road accident information from local police and prosecution agencies and from the Federal Police in the case of the accidents on Federal Highways. From this information source, deaths at the site of the accident are registered. For 2012, this source yielded a figure of 10,008 deaths. Furthermore, the Ministry of Health produces a database from the death certificates, registering the causes of deaths due to road accidents. From this second source, all deaths caused by traffic accidents are registered, yielding for 2012 a figure of 17,653. The difference between the figures from both sources are deaths that do not occur at the site of the accident but some time after the accident, i.e., by 2012 there were 7,645 of these deaths. Furthermore, the figure of 17,653 deaths for 2012 underestimates the total number of deaths due to road accidents, because for some of the deaths that occurred some time after the accident, in the death certificate a different cause than the road accident was recorded (e.g. cardiopulmonary arrest, infection, brain death, organ failure).

It is typically the case that higher levels of severity have better levels of reporting. Many countries (especially in HICs) record all fatal crashes, and have reasonable records of more serious injury (e.g. hospitalisation). Information for minor injury is typically less well reported. One quick way to determine the likely scale of under-reporting rates for non-fatal crashes is to compare the ratios for fatal crashes to other crash types between countries or regions. Although a number of factors need to be considered (e.g. road types, vehicle fleet, average speeds, etc.), the discrepancy in these ratios can indicate differences in reporting rates.

ASSESSING DATA QUALITY

Datasets can be assessed for under-reporting levels and data quality by comparison with other databases. A common comparison to make is between police crash data and hospital in-patient data. Another source is to use insurance claim data. Although these evaluations are very useful, it is not possible to determine the real number of total road crashes as there is no way to know the exact intersection of the two databases (OECD, 2007). There will be some crashes that are recorded in police crash report databases, but as victims are not always sent or admitted to hospital (i.e. in property damage only or minor injury crashes), they do not always appear in the hospital database. Conversely, there will undoubtedly be hospital injury records that are not crash related.

Matching hospital and police data allows cases to be checked for accuracy (ensuring the information provided in both databases is the same) and also provides a basis to estimate the proportion of under-reported cases in both the police and hospital files, as shown in the diagram below (OECD, 2007).
A common problem with this technique is that some countries do not allow the release of victim names and sometimes even personal identification codes. Cases can then only be linked by other characteristics, such as time, date and location (Austroads, 2005). Data can only be reliably maintained when the data quality is regularly monitored. WHO (2010) and IRTAD (2011) provide details on methods for assessing data quality and under-reporting rates.

**IMPROVING DATA QUALITY**

It is typically not possible to successfully collect data for every crash on a network, but not all crashes need to be reported to be able to draw conclusions and identify key priorities to improve road safety (Austroads, 2005). However, the more comprehensive the data set, the higher the reliability.

The main steps to improving data quality include:

- a review of variable definitions, ensuring they are simple to understand and apply;
- reinforcing the need to report crashes, e.g. by making it a legal requirement;
- improving data collection tools (e.g. crash report documents and apparatus, coding procedures);
- collecting accurate location information;
- improving training of police and data entry staff;
- ensuring the data collected is accurate and reliable through quality assurance measures.

Section 3.4.1 of WHO (2010) discusses in detail how the above steps can be put into action. It discusses effective solutions such as the benefits of data entry systems with built-in checks to minimise mistakes, and engaging with police so that they see the value and importance of this task and their role within it. It is also important to acknowledge that a balance must be found in the number of details the police must record at a crash scene. Too many questions will lead to incomplete or missing crash reports, whereas too few will limit essential details that are required for future analysis.
5.6 ANALYSIS OF DATA AND USING DATA TO IMPROVE SAFETY

Crash data can be extremely useful to a number of agencies and individuals, including:

- traffic engineers – in the identification, analysis and treatment of existing risks and the prevention of future risk problems;
- policy-makers – at national, regional and local levels in setting crash reduction targets, developing road safety action plans, and monitoring performance;
- police – in the identification of problem locations and times for enforcement;
- health sector – for resource planning, injury surveillance, health promotion and injury prevention interventions;
- research community – in preventative studies and in testing and improving the effectiveness of road safety treatments;
- insurance companies – in setting insurance rates and premiums;
- vehicle manufacturers – in the development of safer vehicles;
- prosecutors – in the use of data as evidence.

This section considers the availability of crash data, different users, and current international cooperative efforts to improve crash data.

AVAILABILITY OF CRASH DATA

Crash data is useless to organisations that cannot access it. Appropriate methods for distributing data should be developed for each agency that requires it, through the use of statistical reports, newsletters, websites and workshops (WHO, 2010). If the funding is available, an excellent way to make crash data available is through the use of online public searchable databases, which can provide customised reports based on location, injury type, or other crash characteristics (WHO, 2010). An example of such a system is provided in Box 5.12.

BOX 5.12: UK CRASH DATABASE ACCESS VIS THE INTERNET (CRASHMAP)

CrashMap (www.crashmap.co.uk) is a publicly available online tool that allows users to search for crashes (by severity) to see where and when they took place. Users may nominate whether to include all casualty types in the search results, or whether the results should focus on cyclist, pedestrian, child or motorcycle casualties. The results provide a mapped display of crash locations, graphed with colour tags to indicate the applicable severity level. Alternatively, tags may be uniquely coloured and display a number greater than one, to indicate that multiple crash reports exist for a particular location. Each tag provides an overview of the crash details, including the date of the crash, the severity level, and the number of vehicles and casualties involved. Further details of crashes are available upon registration with the site and at a small cost. This charge supports the ongoing maintenance of the service.

Another effective method to distribute data is through the media, which can act as an agent of change by influencing public and political opinions.

It is important to remember that regardless of the method of distribution, those responsible for crash data also hold the responsibility to protect the privacy of individuals involved. Steps can be taken to assist with...
this as outlined in WHO (2010).

USES OF ROAD SAFETY DATA

ADVOCACY PURPOSES

Data can be used to raise awareness about particular road safety issues, and to act as evidence and draw support for a certain policy, programme or allocation of resources (WHO, 2010). Common advocacy activities include workshops, news reports and campaigns. Advocacy is an important part of road safety – it can be the source of funding and public support. It is important to note that any advocacy material must take the target audience and the context of the recommendation/cause into account in order to have a desirable affect. WHO (2010) provides a number of tips for developing advocacy messages for policy-makers. Box 5.13 demonstrates the use of road safety data for advocacy purposes in Cambodia.
The problem: Although a developed database exists for the analysis of crash data, key information was not being provided to appropriate stakeholders.

The solution: Recognising that lengthy reports are not likely to be read by senior managers and politicians, the National Road Safety Committee in Cambodia revised the reporting from their crash system. Along with detailed reports, they produce a summary report that provides key headline analysis on crash outcomes. Clear graphs, tables and maps are provided that are easy to interpret, and can be quickly read and understood. This summary information is particularly useful for advocacy purposes and for informing senior management of key issues. Detailed information is still available for those who require this (e.g. technical staff).

The outcome: Information is provided in different ways according to end-user needs. This means that the information is now more accessible to key stakeholders.
IDENTIFYING ROAD SAFETY PROBLEMS

Road safety engineers are often the most common users of police-based crash databases for road safety work. Crash data is used to identify high crash risk sites, as well as possible identification of risk factors that are specific to the site. This is explained in further detail in Assessing Potential Risks and Identifying Issues.

In the identification of problematic crash locations, target groups or particular risk factors, policy makers use crash data to approximate the size of the problem in terms of counts, severity, trends and the costs of road traffic injuries (WHO, 2010). It is therefore important that these individuals have access to crash characteristics, such as age group, crash type and road user group, so that they can make informed decisions about which high-risk problems get priority and what solutions can be effectively implemented.

Police can also utilise crash data to target enforcement towards a particular issue or location. It is important that the police receive regular feedback so that they can see how their efforts in the collection of crash data, and in traffic enforcement, are having a positive impact (WHO, 2010).

Monitoring and evaluating the performance of initiatives

Crash data is essential to evaluate treatments and policies that have been introduced. Evaluations provide a knowledge base about the effectiveness of a given treatment, as well as ensuring that current programmes are providing the expected and desired results.

New analyses can strengthen the known effectiveness of an initiative, such as through the development of crash modification factors (CMFs). Further information is provided in Monitoring and Evaluation of Road Safety on the monitoring and evaluation of road safety countermeasures, including the effectiveness of treatments and development of CMFs.

INTERNATIONAL AND REGIONAL COLLABORATION

International cooperation is essential for data coordination and benchmarking. International assessments can help to identify and monitor national road safety issues, as well as to evaluate the effectiveness of any methods implemented on a wider scale. Benchmarking (through a comparison of safety performance with similar peer countries, regions, cities, etc.) can lead to the identification of road safety issues that need to be addressed. It is important to note that this cannot be achieved unless there is consistency across crash variable definitions. Coordination also helps countries and governments to improve their road safety data quality and collection systems (see Box 5.14).
In 1988, the Organisation for Economic Co-operation and Development (OECD) established the International Road Traffic and Accident Database (IRTAD). This database includes crash and traffic data from over 30 countries, which is continuously updated and analysed for trends.

The database includes data such as crash severity, road user group and road user age, and also includes relevant country details such as population, vehicle composition, road network length and seatbelt usage rates. This has allowed very useful benchmarking to occur, allowing comparison of fatality rates (e.g. road fatalities per 100,000 population) between countries.

The IRTAD Group is a working group consisting of road safety experts and statisticians from all over the world. Its main objective is to contribute to international cooperation on safety data and analysis. This is achieved through the exchange of data collection and reporting systems and trends in road safety policies, research and publications on key and emerging issues in road safety and through providing
advice on specific road safety issues to member countries.

The IRTAD Group is also in charge of the development of the IRTAD network and database coverage, twinning programmes to assist LMICs in improving their data collection and reporting systems, the IRTAD Conference, and publication of the Annual Report. It also provides standardised definition and methodologies for comparison purposes (e.g. defining injury and crash severities)

**Source:** OECD/ITF, (2014).

Within the framework of its outreach strategy in LMICs, IRTAD has launched a twinning programme to assist countries. IRTAD is working with a number of organisations in an effort to assist LMICs improve their data collection methods and database setup and management. Several such arrangements exist, including twinning between Cambodia and the Netherlands, Jamaica and the UK, and Argentina and Spain. Other partnerships are currently being developed. The case study in Box 5.15 provides information on the twinning arrangements between Argentina and Spain. Box 5.16 provides details of a broader regional observatory in Latin America. Box 5.17 provides details of the IRTAD/OISEVI Buenos Aires declaration on better safety data for better road safety outcomes.

**BOX 5.15: CASE STUDY – TWINNING ARRANGEMENT BETWEEN ARGENTINA AND SPAIN**

The problem: Lack of reliable and comprehensive information on accidents at a national level in Argentina.

The solution: In April 2010, the World Bank provided funding for a twinning arrangement between Argentina and Spain, as part of the IRTAD exchange. Spain assisted Argentina to improve its data collection and analysis systems, with the view to help the Agenzia Nazionale per la Sicurezza del Volo (ANSV, ‘National Agency for the Safety of Flight’, the lead agency in Argentina) become an IRTAD member.

The twinning program involved study tours to Spain and providing guidance on the management of road safety interventions. The main aspect of the twinning program was providing guidance on the development of a comprehensive and consistent accident data management system at a national level. It also included training practitioners from all the jurisdictions at the national, provincial and municipal levels, to ensure standardised data analysis and quality in preparing the diagnoses and reports.

The outcomes: By developing a national accident database, the data from all jurisdictions was collected using a standardised form to allow for the inclusion of the ANSV Database in the IRTAD group. The twinning program also created access to expert advice and access to technical information and research methodologies. Through the success of the twinning program, Argentina was included in the IRTAD database. The success of the twinning program also led to broader cooperation in Ibero-America. This led to the creation of the Ibero-American Road Safety Observatory (OISEVI).
BOX 5.16: IBERO-AMERICAN ROAD SAFETY OBSERVATORY

The problem: Lack of regional capacity in collection and analysis of road safety data.

The solution: The successful twinning between Spain and Argentina led to the establishment of a broader cooperation between countries in Latin America and the Caribbean region, called the Ibero-American Road Safety Observatory (OISEVI).

The OISEVI was created in 2011, and 18 countries have joined with the goal of sharing knowledge and best practice policy-making and planning. The main aim of OISEVI is to share road safety information, particularly best practices in policy formulation, planning, road safety strategies and data management. It is also aimed at improving expertise in road safety and knowledge sharing among practitioners and improving road safety outcomes in Ibero America.

OISEVI is also supported by a regional road safety database (hosted by ITF/OECD) based on the IRTAD model. This database is dedicated to supporting countries in Latin America and the Caribbean. The database uses the same standardised definitions and reporting as the core IRTAD system. The intention is that this linkage will result in progressive improvements in data quality over time. Given the success of this model, ITF/OECD is exploring opportunities to duplicate this approach to other regions of the world.

The IRTAD LAC Database is a kind of shadow database. It has the same structure as the main IRTAD database and is based on the same survey. The main difference is that it is a learning tool; countries can submit their national data, even when incomplete or when data are based on estimation. IRTAD LAC allows them to get acquainted with IRTAD requirements, adapt their data collection (if required), and start benchmarking with neighbouring countries. When data become stronger and more complete, they are then reviewed by the IRTAD Group for inclusion in the main IRTAD database.

Figure 5.9

The outcomes: The Observatory has produced three annual reports consolidating regional road safety data. Technical conferences have also been held on topics relevant to their region (e.g. a workshop on motorcycle safety). Baseline data has been collated to analyse behavioural changes in pilot countries. The data was collected using observational surveys on accident risk factors such as the use of seatbelts, driver distraction factors, use of child restraints and drunk driving.
In November 2013, 40 countries met at the Joint IRTAD/OISEVI Conference in Buenos Aires. The meeting agreed on 12 recommendations on better safety data for better road safety outcomes, including that:

- Reliable crash, contextual and exposure data are essential for understanding road safety issues, setting targets and implementing effective policy.
- A minimum set of data is required, including outcome data (including the number of persons killed and injured by type of road users, type of roads, time, etc.), output level data (including performance indicators) and contextual data (including exposure data such as population, the number of vehicle kilometres driven).
- Safety data should be aggregated at national level.
- A road safety observatory, under a lead road safety agency or a lead ministry, is a good institutional setting to raise the profile of road safety.
- Regular monitoring and analysis of key road safety risk factors should be undertaken.
- The international community should work towards harmonisation of data.
- Information on injury crashes is essential for a more complete picture of road safety, and MAIS3+ should be used to define a seriously injured road casualty.
- Police data will remain the main source for road crash statistics, but this should be supported by hospital data.
- Benchmarking between countries is a useful way to generate road safety improvement and learn from others.

Full details can be found at the following website:
http://www.internationaltransportforum.org/jtrc/safety/Buenos-Aires-Declaration.html

In Europe, a centralised database of road crashes has been developed. The Community Road Accident Database (or CARE) is hosted by the European Commission and includes information on fatal and injury crashes. Details on individual crashes are retained (i.e. the information is not combined), thereby allowing for more powerful analyses to be conducted. A protocol for the collection of data has been developed, with common variables specified. The intention of the database is to provide the basis for analysis to:

- identify and quantify safety problems on European roads;
- evaluate road safety measures;
- determine the impact of actions;
- share experience.

Further information on CARE can be found on the European Commission website (http://ec.europa.eu/transport/road_safety/index_en.htm).
5.7 INTEGRATING DATA

The integration of safety data provides a large number of benefits, including:

- effective management of safety (e.g. linking intermediate and final outcome data; see Identifying Data requirements);
- identifying crash risk types and locations in a more comprehensive manner (e.g. through combined use of both reactive and proactive approaches, benchmarking, better information on injury outcomes, greater knowledge of crash contributors; see Assessing Potential Risks & Identifying Issues);
- better knowledge in order to identify cost-effective solutions (e.g. through better knowledge of the infrastructure that already exists; see Intervention Selection and Prioritisation).

In addition, linked data can be used to validate other sources of information. As an example, crash database systems can either draw information directly from asset data to provide additional information on road elements, or this linkage can be used to reduce the likelihood that data entry errors will occur by validating the presence of different road features or assets. It can also be used for research on specific topics.

Key linkages include combining crash data with:

- traffic data;
- road inventory data;
- vehicle registrations;
- vehicle inspection data;
- population statistics.

The linkage process involves several stages, and can be temporary (e.g. for a specific project or policy need) or permanent (e.g. for ongoing analysis and monitoring). Data needs to be collected in a format to facilitate linkage. This typically involves provision of a common data element, most usefully the spatial coordinates for road based elements (including crashes), while for non-spatial data, another unique identifier will be required for the datasets to be linked. A comprehensive safety information system may have a large number of component files.

Once the data has been linked, it can be analysed through merging of data files. For spatial data, a GIS software package is able to assist greatly in this task, and is particularly useful for mapping information from different sources.

Once the initial investment in collecting data has been made, it may be a relatively low cost task to join different sources of information together to meet a variety of needs, especially if a unique identifier has been used in each dataset. In other circumstances, especially where data is not in a compatible format, the task might be quite substantial involving considerable investment.

One of the more commonly used linkages is the calculation of crash rates to allow either benchmarking or identification of high risk locations. For example, crash data can be combined with population figures, traffic volumes, or vehicle registrations to provide a useful comparison of risk. Ideally, each of these would be presented as fatal and serious injury crash rates. Each of these measures is useful for different purposes, as outlined below:

- Crashes per 100,000 population: This measure reflects the direct impact of road crashes on a country, region or community. It provides a useful basis to compare road safety outcomes with other types of risk (e.g. risk of death from heart disease). It can also be used to identify risks for different subsets of the community (e.g. risk by age, gender, location). This is one of the most commonly used comparisons in
road safety, particularly because it is easy to collect relevant data to make this calculation.

- Crashes per vehicle kilometres travelled (VKT): This measure reflects the level of safety based on the amount of travel undertaken. It can be used to compare different travel modes (such as between car, bus or train), different road types (e.g. undivided compared with divided roads) or infrastructure (e.g. roundabouts compared with traffic signals). This requires good knowledge of traffic volumes and distances travelled, and this information is often difficult to collect.

- Crashes per 10,000 registered vehicles: This measure is often used as a proxy for the amount of travel, as it is much easier to collect than kilometres travelled. It can be useful for analysis of performance at a national level, but has limited applicability for more detailed analysis.

Crash data can also be combined with road inventory data. At a simple level, this can provide information about current road features that may be present, providing information about possible infrastructure safety improvements. For example, crash data of run-off-road crashes could be presented alongside information on current roadside barrier locations on a map to allow for a quick visual analysis of locations that might benefit from further barrier installation.

Combining data on crashes with roadway, asset, environment, and traffic volume data can lead to some important outcomes relating to safety performance of infrastructure. It is possible to compare the safety performance of different types of infrastructure for different traffic volumes. For example, the performance of divided and undivided roads can be compared for different traffic volumes. In addition, crash performance of different infrastructure can be compared at different levels of traffic volume, for different road user types, or for different environment types (e.g. low versus high speed environments). Box 5.18 provides information on the US Highway Safety Information System.

BOX 5.18: US HIGHWAY SAFETY INFORMATION SYSTEM

The US Highway Safety Information System (HSIS) is managed under contract to the Federal Highway Administration (FHWA). HSIS contains crash information (e.g. collision type, severity, vehicle information, sex and age of occupants, objects struck, and weather conditions), inventory (e.g. road type and function, cross-section, number of lanes, lane and median width, shoulder width and type), and traffic volume data for several States. Information on curve/grade and intersection variables is also available from some States. The combination of these different sources of data allows powerful analysis to be conducted on specific road safety issues.

A large number of studies have been conducted using this rich source of information. This has led to the production of various research reports, summaries and tools. Recent examples include a study that examined the safety effects of horizontal curves and grades on rural two-lane highways; a safety evaluation of lane and shoulder width combinations; an evaluation of the safety benefits of transverse rumble strips on approaches to stop-controlled intersections in rural areas; and a review of the safety benefits of ‘road diets’ (converting four lane arterial roads to two lanes, plus a central two-way turn lane).

Further information on the HSIS can be found at the following website: www.hsisinfo.org

Recent initiatives in integration have involved the combination of crash data and road risk assessment data. This provides a very powerful tool for identifying risk locations and possible solutions. Further information on the combination of this data can be found in Combining Crash Data and Road Data and Intervention Selection and Prioritisation.
**PATHWAY TO EFFECTIVE MANAGEMENT AND USE OF SAFETY DATA**

**GETTING STARTED**

- An assessment of data requirements should be made.
- For countries with no comprehensive crash data, information on final and intermediate outcomes should be collected for high risk routes (e.g. high volume roads) to allow measurement of safety problems and identification of measures. This collection could be undertaken as part of a corridor demonstration project.
- A crash data system should be put in place. The steps required for this include the need to: assess current data sources; engage with key stakeholders (the road agency, police and the health sector are especially important); develop a crash report form; develop a data system; and put in place a process to ensure data quality.
- The focus should be on the collection of the range of data needed to address fatal and serious injury crash outcomes which will include exposure data, final outcome data as well as intermediate outcome data.
- Road infrastructure/asset data collection should be considered to inform safety decisions, for instance through a road assessment programme. This can provide information on likely high risk crash locations as well as affordable treatments in the absence of comprehensive crash data.

**MAKING PROGRESS**

- A data collection strategy should be developed to ensure that essential information is collected.
- The crash data system should be routinely checked for accuracy and completeness (e.g. by comparing police and hospital data).
- The database should include basic features to allow comprehensive analysis of crash problem...
types, and be fit for use by the required stakeholders.

- Information on road assets relating to safety outcomes should be collected.
- Countries should be encouraged to aggregate data at national level, matching the IRTAD structure and progressively contributing to the IRTAD database.
- Other data relevant to the setting and monitoring of road safety targets and trends should be collected, and the accuracy of this data assessed.
- All outputs (such as reports) should be assessed to ensure that they are fit-for-purpose and address the needs of key stakeholders.

**CONSOLIDATING ACTIVITY**

- A comprehensive data collection strategy should be put in place and regularly monitored to ensure that it is fit-for-purpose, accurate and complete.
- A crash data base should be fully implemented that contains all crash data. Data should be spatially coded, and appropriate quality control checks should be put in place.
- Information on road assets relating to safety outcomes should be contained within a comprehensive roadway inventory database. This may require the development of a database, or linkage to an existing database.
- Linkages should be made between key sources of data, particularly between data collected by police and hospitals, and between crash and asset data.
5.8 REFERENCES

Austroads, (2005), Australasian road safety handbook volume 3, Why we continue to under-count the road toll, Austroads, Sydney, Australia.


Global Road Safety Facility GRSF (2013), Road Safety Management Capacity Reviews and Safe System Projects, Global Road Safety Facility, World Bank, Washington, DC.


GRSP (2009), Seat-belts and child restraints: a road safety manual for decision-makers and practitioners, FIA Foundation for the Automobile and Society, London, United Kingdom.


