

3

**How to design, improve and
implement data systems**

3.1 Establishing a working group	49
3.2 Choosing a course of action	51
3.3 Recommended minimum data elements and definitions	56
3.4 Improving an existing system	60
3.4.1 Strategies to improve data quality	61
3.4.2 Strategies to improve data system performance	68
3.5 Designing and implementing a new system	77
3.6 Considerations for non-fatal data	90
Summary	92
Minimum data elements: full description	94
Crash data elements	94
Crash data elements derived from collected data	98
Road data elements	98
Vehicle data elements	103
Person data elements	107
Person data elements derived from collected data	113
References	113

MODULES 1 and 2 described the importance of reliable data for road traffic injury prevention and monitoring, and how to assess the road safety data situation in a country or region. This module describes how to use this information to strengthen existing road crash data systems, or design and implement new ones.

The sections in this module are structured as follows:

- **3.1 Establishing a working group:** This section shows how to develop a multisectoral working group to define objectives and choose the best approach. This is an essential step to ensure overall coordination of the system, with input from all stakeholders.
- **3.2 Choosing a course of action:** The appropriate course of action depends on objectives, the content and quality of available data, the features and functioning of available systems, and available resources.
- **3.3 Recommended minimum data elements and definitions:** This section presents an overview of minimum data elements for a common road crash dataset that can be used for national analysis (a full list of definitions and data values appears at the end of the module).
- **3.4 Improving an existing system:** This section covers strategies for strengthening an existing road crash data system (based on police data) through improving data quality and system performance. Topics include in-depth evaluation of existing systems, data collection tools, reporting requirements, training, quality assurance, database system features, linkage and data management plans.
- **3.5 Designing and implementing a new system:** This section describes the steps to take if there is no system, or if existing data systems cannot be modified to fulfil the key objectives (assuming there is the political will, resources and capacity to implement a new system).
- **3.6 Considerations for non-fatal data:** This section summarizes actions that can be taken to improve the availability and reliability of data on non-fatal road traffic injuries.
- **Minimum data elements:** full list of definitions and data values.

3.1 Establishing a working group

Decisions about strengthening road safety data systems should be made in consultation with a multisectoral working group. This should comprise agencies and individuals who have been identified as collaborating partners in the situational assessment (see Module 2). If there is a lead agency for road safety, it should be represented in the group. Working group members will have technical and practical responsibility for implementing changes to road safety data collection systems. Their regular duties will need to accommodate their new responsibilities as part of the working group.

The working group should identify a coordinator who will have overall responsibility for the group's work. The coordinator, whether paid or not, should have clearly defined responsibilities. These include overseeing the activities of the working group, monitoring progress, and ensuring that all those directly involved (as well as other key stakeholders) are kept well informed. The coordinator should have the necessary authority, resources and support to carry out these tasks. For this reason, the role is best filled by someone whose work already includes some of these responsibilities, for example, the chief technical officer within the transport department, the person in charge of the traffic police database, or the person responsible for data in the national road safety lead agency. In many cases it makes sense for the individual/agency that will have responsibility for the road crash data system to act as the coordinator.

In its first meeting, the working group should agree its purpose and the roles and responsibilities of each member. Key stakeholders and other parties who are interested in road safety data developments (but who will not have technical responsibilities) can be consulted through an advisory group, which meets less frequently than the working group.

Setting goals

One of the first tasks of the working group is to define goals for road safety data systems. If the national lead agency and/or national strategy has specified general road safety goals, these should be reviewed and discussed in terms of the data requirements for achieving and monitoring each one. Group members should then be given the opportunity to present their perspective on the primary goals for road safety data systems. Eventually the group must agree on common goals and key system requirements. This will help with the selection of common data elements and help identify which goals can be met with existing data sources and systems, and which require something new.

Module 1 described how reliable road safety data are needed to accurately identify problems, risk factors and priority areas, and to formulate strategy, set targets and monitor performance. Data are also needed for effective advocacy. This requires data that describe the social costs of road traffic injuries, i.e. deaths, non-fatal injuries and crashes (*final outcomes*); exposure to risk, e.g. population size, number of vehicles, traffic volumes; safety performance indicators such as helmet use or speeding (*intermediate outcomes*); and interventions implemented (*outputs*).

Though current good practice suggests countries set goals and targets for each of these areas, few have done so. Most countries and jurisdictions focus on reductions of deaths and non-fatal injuries (*final outcomes*). In practice this means policy-makers usually need data on absolute numbers, rates, trends, severity, and costs of road traffic injuries by geographic area, age group, crash type and road user type/mode of transport. This module therefore focuses on strengthening data systems for final outcome data, with a particular emphasis on road crash data systems based on police data.

Regardless of decisions about the short-term focus of a new data system (e.g. one based on final outcome data), the working group should develop a long-term strategy for meeting the full range of data needs required for effective road traffic injury prevention and performance monitoring (see Module 1).

NOTE

When road crash data systems are improved, the reported number of injuries can rise, sometimes dramatically, because the system has become more effective at capturing events. Political concerns about the impact of this apparent increase can underlie resistance to changes in road safety data systems. Open dialogue with policy-makers allows people and agencies to express their concerns about how such results might affect their funding, performance review, press coverage, and to consider together how these concerns might be addressed. The press should be sufficiently briefed before any new figures are released, to help them correctly report the change in trend. A media/information strategy is essential for highlighting these issues and minimizing political and community concerns.

3.2 Choosing a course of action

When improving road safety data systems, there is no single course of action that will be right for every country or jurisdiction, at all times. The following sections therefore describe a range of strategies for strengthening road safety data systems and provide examples of good practice. The working group must consider these possible strategies in the context of their own situational assessment (for example, see Case study 3.1), and identify what is likely to be the most effective and feasible approach. The working group must then decide the most appropriate course of action.

NOTE

Political support for investment of human and financial resources in road safety data systems is a critical piece of the jigsaw. Without this support it will be difficult to implement the kinds of changes needed to build good road safety data systems. If the situational assessment reveals major problems with – or a lack of – road safety data systems but there is not yet support to address this, then use the data gathered in the assessment to advocate for investment in data systems. Use the data to estimate the magnitude of the road traffic injury problem, and use information about data quality and under-reporting to argue the need for greater attention to road safety data.



CASE STUDY 3.1: **Using situational assessment to choose a course of action, Bengaluru (Bangalore), India**

The city of Bengaluru (Bangalore), India, carried out a situational assessment of the growing problem of road traffic injuries in the city. Data available suggested that there were increasing numbers of crashes involving pedestrians, cyclists and users of motorized two-wheelers.

Coordinated by the WHO Collaborating Centre for Injury Prevention and Safety Promotion at the National Institute of Mental Health and Neurosciences, the process began with a stakeholder consultation with the city's police, transport department and representatives from 25 hospitals. The goals and objectives of the assessment included reviewing the availability, quality and usefulness of existing data, defining what new data were needed, deciding on mechanisms to be used to collect this information and planning how the data would be used and disseminated.

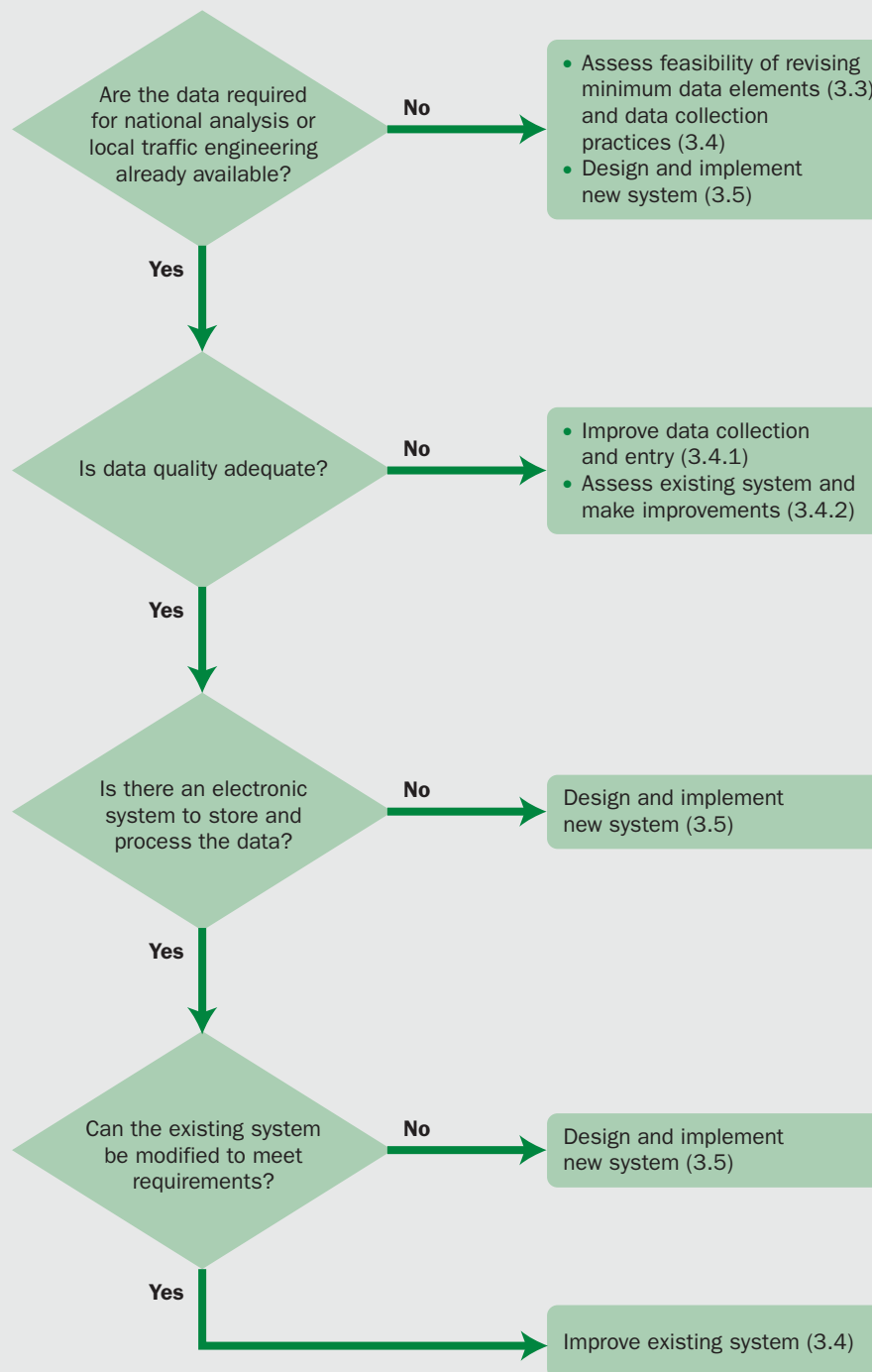
The situational assessment included a one-year data collection phase. This involved identifying and selecting centres that would collect the data; a preparatory phase to develop the tools to collect data, and training 300 police and hospital professionals to use them. Information gathered over the assessment period provided an estimate of the extent of the problem of fatalities, the extent and severity of non-fatal injuries, as well as information on the geographical distribution of injuries, socio-demographic characteristics of those most affected, and involvement of risk factors in road traffic crashes.

As well as collecting data, the assessment provided detailed information on the processes used to collect information. This suggested that while data on deaths could be collected from police (after some improvements were made to their data collection capacity), hospitals could also collect injury data using standard emergency trauma care forms. Stakeholders found that there were opportunities for integrating both police and hospital data, using particular technological tools. They also noted the critical need for administrative support, personnel training, monitoring and regular feedback to ensure the sustainability of a programme. Thus the assessment identified that road traffic injury surveillance was possible without building an entirely new system, but rather by strengthening existing systems to obtain relatively small amounts of good quality information.

Apart from collecting data, the process of conducting a situational assessment served to identify the need for capacity to be built for analysis and interpretation of data. Several limitations to the existing system were identified, and activities were initiated to address these, including cooperation by police and hospital staff, ongoing training, establishing mechanisms for data analysis, systematic feedback of data to stakeholders, resource allocation and assigning roles. The need to identify mechanisms to coordinate the programme, as well as to monitor and evaluate activities, was also considered. Following this assessment, attempts are being made to make continuous data available on a regular basis. The data have been used to strengthen a number of road safety and trauma care interventions. For more information see *1*.

The main strategies to strengthen final outcome data include:

- improving the data quality and system performance of road crash data systems that draw mainly on police data, or design and implement such a system if none exists (see Figure 3.1 and sections 3.3–3.5);
- improving health facility-based data on road traffic injuries. Consider implementing a surveillance system. This is especially important for improving non-fatal road traffic injury data (see section 3.6);

Figure 3.1 Using a situational assessment to choose a course of action

- improving the vital registration system, particularly death registration. Implementing required changes is usually beyond the mandate and capacity of road safety stakeholders, but you can advocate for change (for further guidance, see resources available from the Health Metrics Network, www.who.int/healthmetrics/en/);
- working with public health and road safety experts at an academic institution to combine existing data sources to generate more accurate estimates on the magnitude and impact of road traffic injuries (see Box 2.6 Module 2).

Sections 3.3, 3.4 and 3.5 provide guidance on strengthening road crash data systems that are based mainly on police data. Figure 3.1 illustrates how actions might be prioritized for strengthening or implementing road crash data systems. The flowchart assumes road safety is already recognized as a major health and development issue, and one requiring political backing and reliable data for effective planning and monitoring. Strategies for improving data on non-fatal road traffic injuries are discussed in section 3.6, since these involve measures that extend beyond a police-based road crash data system.

The case studies presented in this module provide examples of road crash data systems that are working well, mainly in low- and middle-income countries. These examples illustrate how a wide range of approaches can lead to successful data systems. There are many more valuable case studies in the World Bank Global Road Safety Facility guidelines for road safety management capacity review, which readers are strongly encouraged to view (2).

Police as key stakeholders

Police records are the primary source of road crash data in most jurisdictions, but police may not be responsible for the data systems that utilize this data, nor for initiating changes to such systems. It is easy to understand how resistance might arise in a situation where police have their own system for documenting crashes, and where their efforts in data collection are neither recognized nor appreciated. This can be compounded if transport or highway departments propose – perhaps without consultation – a new road crash data system that requires substantial changes in the way police officers work.

Effective road crash data systems cannot succeed without the acceptance of the police force. The best way to make sure the system benefits police, and to foster a sense of police ownership, is to **involve police in all stages of project planning** for changes to the system (or implementation of a new system). This means seeking police participation in the situational assessment, the identification of goals, and the selection, implementation and evaluation of strategies to improve the system. Police participation in decisions about data collection procedures is particularly important in the design of data collection forms.

Frequently, extensive demands are made on the initial data collection process from the crash site, yet no information is passed back to the data providers. Police justifiably complain about the extra work required to collect data for other agencies when there is no productive benefit to their own work. From a practical perspective, therefore, it is critical that there is a structured feedback mechanism providing timely and useable information to traffic department managers, supervisors and, most importantly, crash investigators and patrol officers.

The provision of an annual report on crash statistics is not an effective feedback mechanism to help efficient law enforcement strategies to be put into place. The feedback must be provided on a regular basis to meet police user requirements.



Strategies for addressing reluctance or resistance among law enforcement stakeholders include (see also Case study 3.2):

- demonstrating to staff at various levels how the changes will ensure better information and analysis for positive road safety outcomes in terms of research, policing and engineering – and ultimately in saving lives. Procedures, paperwork and data entry would be streamlined to provide benefits to all;
- institutionalizing procedures through government order, or by defining Standard Operating Procedures (see Case study 3.6);
- fostering the commitment of senior police management to road safety in general, and crash data quality in particular. Top level political and management support can encourage wider acceptance;
- discussing the availability of personnel and vehicles for traffic policing with those highest in the police chain of command;
- considering the availability of essential equipment for police for accurate scene measurements and crash investigation (e.g. tape measures, maps);
- demonstrating how police data are being used by other stakeholders to improve road safety, e.g. distribute an annual report to police, including case studies of how police data were used to identify and improve hazardous locations;
- establishing mechanisms to provide relevant analyses to police in an accurate and timely manner, so they can use the data for intelligence-led enforcement (e.g. trend analysis of location, day of week and time for fatal and severe injury crashes).



CASE STUDY 3.2: Data-led enforcement strategy in Victoria, Australia

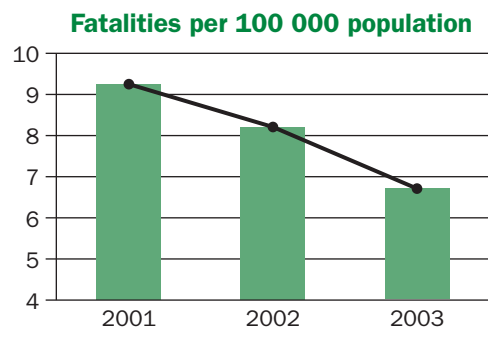
Over the past 30 years, Victoria Police and its road safety partners have consistently relied upon data to drive the road safety enforcement and education strategies.

In 2001, data analysis from fatal and serious crashes identified speed as a major factor in road trauma in approximately 30% of road crashes. Data from the VicRoads speed survey analysis indicated that in a 60 km/h speed zone, the majority of motorists drove at approximately the threshold speed, i.e. the speed at which enforcement action is taken. In many countries, this tolerance level (although never disclosed by police) has effectively become the speed limit.

The Transport Accident Commission, in collaboration with Victoria Police, undertook a massive and sustained education campaign called **Wipe off 5** – i.e. reduce your speed by 5 km/h. Research indicated that for every 1 km/h reduction in average vehicle speed, the crash frequency was reduced by 3%. Phase one began in August 2001, followed in November by a second phase with more intensive media coverage. The twin targets were careless and deliberate speeding, particularly in the low-level speeding range.

Phase three commenced in December 2001 with enhanced speed enforcement strategies, including additional mobile speed cameras and targeted police operations. The enforcement threshold was then reduced systematically from February 2002 by 1 km/h increments across the state each month for three months (i.e. overall the threshold speed was reduced by 3 km/h). The strong media message of *60 km/h means 60 km/h* was promoted in parallel. All data and actions were closely monitored, with an immediate effect in speed reduction.

The benefits of the threshold reduction strategy became clear from the substantial drop in numbers of Victorian road fatalities for the 2002 (16%) and 2003 (11%), and more importantly, a reduction in the five-year average. The enforcement threshold has been maintained to this date. Victoria Police continue to rely on data with intelligence-led, outcome-focused enforcement strategies.



3.3 Recommended minimum data elements and definitions

Whether you choose to modify an existing system or build something new, the common dataset composed of minimum data elements (variables) will be a key tool for ensuring the appropriate data are captured to enable analysis, and for maximizing consistency and compatibility of data collected across different jurisdictions.

The purpose of defining minimum data elements and specifying uniform definitions and criteria is to provide a dataset for describing road traffic crashes, and the resulting

injuries, that will generate the information necessary for national analysis and road safety improvements (3). Uniformity of crash data is especially important when combining sub-national datasets, and for international comparisons.

The common dataset concept is well-known and used in public health (4). The concept is used in some countries that have road safety data collected at state or provincial level, to facilitate uniform data collection, for example Australia's *Minimum Common Dataset for Reporting of Crashes on Australian Roads* and the USA's *Model Minimum Uniform Crash Criteria*, or MMUCC. In the *Community database on Accidents on the Road in Europe*, or CARE, a set of common data elements and definitions is accompanied by transformation rules so that national data collected according to various criteria can be transformed to meet the common data elements (*Common Accident Dataset*, or CADaS).

Minimum data elements should be selected considering the following criteria:

- **Data elements and values must be useful for road crash analysis.** These elements should be routinely collected when a road traffic crash occurs. Data that will not be used should not be collected.
- **Data elements and values should be comprehensive and concise.** Each variable must include a description and definitions of the possible data values (see minimum data elements at the end of this module).
- **Data which are very difficult to collect should not be included**, regardless of their value for road crash analysis.

When planning for the introduction of minimum data elements, try to minimize changes to the definitions and values of existing data elements, as these can cause problems with consistency and comparability of data over time. If definitions or data element changes are made, clearly note in official records the date of change and allow for some misclassification during the transition.

Minimum data elements proposed in this manual

The common dataset recommended in this manual is based on the Common Accident Dataset (CADaS), developed to provide a common framework for road crash data collection in Europe. The minimum data elements selected for CADaS were based on extensive research on data sources and systems available in 25 European countries, and stakeholders' needs and priorities for crash data analysis at the national level (5, 6). The data elements of CADaS were finalized after more than four years of consultation with road safety data experts and are currently being applied in the European CARE database.

For this manual, minimum data elements of CADaS were reviewed and selected according to the criteria listed above, bearing in mind the unique challenges faced in low- and middle-income countries. The resulting common dataset was reviewed by experts and practitioners in several low- and middle-income countries, and revised for relevance and feasibility. Implementation of this common dataset will help countries

strengthen their road safety data for planning and monitoring purposes, and will facilitate the consistency and comparability of road safety data internationally.

Like CADaS, the common dataset proposed in this manual can be adopted as a full set of criteria for a new data collection system, or adopted progressively in the course of improvements to an existing one. It is important to note that changes to an existing data collection system may not be required at all – depending on the practices and definitions in use, it may be possible to implement this common dataset by applying appropriate transformations to existing data so it can be analysed according to the format and definitions specified here.

Not all of these data will be collected at the crash scene. If other data sources are available and reliable, they should be drawn on for variables that are troublesome for police officers to record at the crash scene (e.g. road functional class, vehicle engine size).

The proposed dataset specifies **minimum** data elements, with an emphasis on those variables that will be useful for national analysis. Jurisdictions may need to collect additional variables to facilitate local analysis, law enforcement follow-up and in-depth crash studies. Extra variables can easily be added to this dataset, according to specific requirements and circumstances.

Before implementing the common dataset, the type of crashes that will be included in the database must be defined. The variables presented here are intended for use in documenting information about crashes which result in at least one injury – crashes resulting in property damage but no injury would be excluded from the database. If a country wishes to include damage-only crashes, the variables can be adapted to describe them.

Table 3.1 summarizes the minimum data elements, while Table 3.2 describes additional variables commonly collected. Box 3.1 explores one of these additional variables – purpose of journey – in greater detail. Detailed definitions and data values for each data element can be found at the end of this module.



Table 3.1 Minimum data elements: overview

Crash related	Road related	Vehicle related	Person related
<ul style="list-style-type: none"> • Crash identifier (unique reference number assigned to the crash, usually by police) • Crash data • Crash time • Crash municipality/place • Crash location • Crash type • Impact type • Weather conditions • Light conditions • Crash severity^o 	<ul style="list-style-type: none"> • Type of roadway* • Road functional class* • Speed limit* • Road obstacles • Road surface conditions* • Junction • Traffic control at junction* • Road curve* • Road segment grade* 	<ul style="list-style-type: none"> • Vehicle number • Vehicle type† • Vehicle make† • Vehicle model† • Vehicle model year† • Engine size† • Vehicle special function† • Vehicle manoeuvre (what the vehicle was doing at the time of the crash) 	<ul style="list-style-type: none"> • Person ID • Occupant's vehicle number • Pedestrian's linked vehicle number • Date of birth • Sex • Type of road user • Seating position • Injury severity • Safety equipment • Pedestrian manoeuvre • Alcohol use suspected • Alcohol test • Drug use • Driving licence issue date • Age^o

^o Derived or calculated from other data elements.

* Depending on the quality and detail of road inventory and hardware data available, it may be possible to obtain this data element through linkage to other databases.

† Depending on the existence, quality and detail of a motor vehicle registration database, it may be possible to obtain this data element through linkage to motor vehicle registration files.

Table 3.2 Examples of additional variables that are commonly collected

Crash related	Road related	Vehicle related	Person related
<ul style="list-style-type: none"> • Location relative to roadway 	<ul style="list-style-type: none"> • Urban area • Tunnel • Bridge • Number of lanes • Markings • Work-zone related 	<ul style="list-style-type: none"> • Vehicle identification number (VIN, issued by manufacturer) • Registration place and year • Registration number • First point of impact • Insurance • Hazardous materials 	<ul style="list-style-type: none"> • Distracted by device • Driver licence class and jurisdiction • Driver manoeuvre • Trip/journey purpose (see Box 3.1)

BOX 3.1: 'Purpose of journey' data element

Information about crash victims' *purpose of journey* is an important building block for effective road safety intervention, particularly as it shows the contribution to the overall road toll of road traffic crashes related to different occupations, and suggests areas for intervention. Such data, however, are not collected in most jurisdictions around the world and this lack of surveillance has been identified as a major obstacle to improving occupational road safety. It has been proposed that road safety data systems include a *purpose of journey* data element.

Definitions for the *purpose of journey* field used in the UK and in Queensland, Australia include the following:

- Journey as part of work
- Commuting to/from work
- Taking child/student to/from school
- Pupil travelling to/from school
- Life and network necessities and social activities (e.g. grocery shopping, visiting friends)
- Life enhancement activities (sports, hobbies, driving for pleasure)
- Holidays and weekends away
- Other (to be specified)
- Unknown

For this data field to be effective, police officers must be sufficiently trained to understand the importance of the data, to identify the *purpose of journey* and record it appropriately.

Source: 7, 8.

3.4 Improving an existing system

Results of a situational assessment should suggest areas for improvement in an existing road crash data system. If an in-depth evaluation of the system was not conducted as part of the situational assessment, it should be done now. The US Centers for Disease Control and Prevention guidelines for planning and conducting an evaluation of a surveillance system is a useful tool to guide this process, and is summarized below (9).

Evaluation should begin with a flowchart of the process, description of the system, including its purpose, operation and resource requirements, building on the information gathered in the situational assessment (see Module 2, section 2.2). Stakeholders who use data generated by the system should be involved in defining questions to be answered by the evaluation.

The system should be assessed for its usefulness, e.g. the ability to detect road traffic injuries in a timely way, to allow counts of the number of road traffic deaths and injuries and description of their characteristics, to facilitate intelligence-led enforcement, and to facilitate assessment of the impact of interventions. Key questions to answer are whether the system meets its objectives, and whether the data are being used to improve road safety.

The CDC guidelines recommend assessing system performance with respect to several attributes (9). The importance of each attribute depends on the system's objectives.

- **Simplicity** of structure and ease of operation.
- **Flexibility** – can easily adapt to changes in operating conditions or information needs.
- **Data quality** – completeness, accuracy and validity.
- **Acceptability** – willingness of relevant agencies to participate in the system.
- **Sensitivity/reporting levels** – proportion of cases detected, and ability of system to monitor changes in trends.
- **Representativeness** – accuracy in describing the occurrence of road traffic injuries over time, and their distribution in the population by place and person.
- **Timeliness** of dissemination of information for intervention and programme planning. This is especially important for enforcement and public awareness strategies.

It will also be useful to have information on the features available in the crash database system, such as built-in quality checks during data entry, mechanisms to make navigating the system easier (e.g. drop-down menus, map-based selection), mapping, location-based analysis (e.g. site ranking by crash rates, numbers, costs or contributing factors).

Combined with information from the situational assessment, the results of this evaluation can be used to develop recommendations for improving the quality, efficiency and usefulness of the system (see also 10).



Improvements to existing data systems often focus on software solutions and database management (see section 3.4.2). Strategies in this area are appealing because they may be faster and easier to implement, and people may be excited about the possibilities of new technology. However, when prioritizing a course of action it is imperative to keep in mind that **even the best system cannot produce high-quality data if data collection and entry practices are poor**. Improving data quality through changes in data collection/entry may be more difficult to implement, but in the long run, reliable road safety data cannot be obtained without them (see section 3.4.1).

3.4.1 Strategies to improve data quality

The situational assessment and in-depth evaluation may reveal that the major problem with an existing system is poor data quality, related to data collection and entry practices. Several strategies that may be used to make improvements in these areas are discussed below:

- Review definitions
- Strengthen reporting requirements

- Improve data collection tools
- Collect accurate location information
- Improve training
- Quality assurance measures

Box 3.2 contains a checklist of questions to assess how data quality can be improved.

BOX 3.2: **Checklist for data quality improvement**

- Should variables be added or removed from the minimum dataset?
- Do current definitions need to be modified?
- Should reporting requirements be changed?
- Do current data collection tools allow simple and quick data capture? Do they need to be revised?
- How can paperwork requirements and data collection and entry procedures be simplified?
- Are police officers sufficiently trained for road crash data collection? Is further training needed for data collectors and/or those responsible for data entry?
- How can we improve identification of crash location?
- What quality assurance measures should be considered?

Review definitions

Module 2 discussed how definitions and criteria used to specify variables for data collection can affect data quality (by influencing which events are included or excluded) and also the likelihood of error in measurement or recording. The definitions of the minimum data elements provided above are intended to maximize data quality and comparability.



Every change to the variables in the common dataset for national analysis means a change to the data collection form and data collection procedures used by police officers at the crash scene, and possibly in follow-up to the crash. If the change requires additional training for all traffic police, do not underestimate the magnitude of this task.

In addition, most police officers are already responsible for a mass of paperwork for a multitude of competing priorities. There must be a balance between analysis-related data needs and the demands placed on the time and workload of police officers.

Special attention should be given to the definitions associated with data values for classifying injury severity, to ensure that they are clear and easy to apply without specialized medical training (see definitions provided in minimum data elements). The definitions and methods of measurement specified for crash location also require careful consideration, to make sure they are compatible with data collection procedures (e.g. if the police force is not even equipped with tape measures, let alone Global Positioning System (GPS) devices, do not define crash location by GPS/ Geographic Information System (GIS) coordinates).

Strengthen reporting requirements

Where there is a legal requirement to report to the police road traffic crashes that result in injury, and for the police to officially record/report such crashes, injury crashes are more likely to be documented and counted in a data collection system.

In contrast, legal requirements for drivers to call police to the scene of the crash even if no one is injured can involve the police in lengthy administrative processes for minor incidents. This may reduce the number of officers available to answer calls to crash scenes, and reduce the time officers have for proper data collection and follow-up.

Improve data collection tools

Most of the world's road crash data collection is done by hand, using paper forms. A standardized data collection form, along with training in its use, can therefore improve data quality. The structure and layout of the data collection form can have a significant impact on data quality. Data collection forms are often designed to match database structure and the logical grouping of data elements. While this may simplify transfer of data from the form to a database ('data entry'), it may not be the best structure for data collectors, causing them to record details incompletely or inaccurately.

If data forms need revising or creating from new (see Case study 3.3), the following recommendations will help (4):

- Seek and use the experience and expertise of people who will be responsible for recording information on the form – e.g. police officers who oversee crash sites in the case of primary data collection forms, or data clerks if the form will be completed by abstracting data from police reports.
- Seek advice from a statistician with experience in data collection instruments. This person can help with form design and structure to help ensure maximum speed and accuracy in data collection, and that recorded information is easy to collate and process.
- Make the form easy on the eye, easy to understand (e.g. when space allows, define abbreviations in the margin), as short as possible, and above all, easy to complete.
- Pre-code the forms (i.e. print the codes on the forms themselves next to the data to be entered) for as many fields as possible, using numbers rather than letters or symbols. Numeric codes are easier to process and less prone to data entry errors.

- A clear reference book containing detailed instructions for completing the form should be available to help answer questions and assist training.
- Test the form in real data collection scenarios, with people who will be completing it once it is adopted for general use. Document problems collectors have understanding fields or recording responses, and the length of time it takes to complete. Revise the form as necessary.

In addition to fields for recording information on the crash, road, vehicle and person variables, data collection forms should provide adequate space for a sketch of the collision, including measurements, and a summary narrative description of events (11).

Some software packages specially designed for the collection and processing of road traffic crash or injury data incorporate the option of computerized data collection forms to record data electronically at the crash scene. This saves manual data entry, but it can be expensive and requires high capacity for electronic processing. It also requires police officers who are comfortable and confident with the use of the handheld devices or portable computers. Data validation can be problematic when data are recorded electronically, as there is no paper record to verify the data. Electronic data collection should be field tested extensively before implementation, and evaluated for its effect on reporting rates.

Collect accurate location information

Identifying the exact crash location is not always easy for police, and often data on location are inaccurate or not specific enough to allow for detailed analysis by location. The specificity of location identification varies considerably between countries and also between different areas of the road network (e.g. inside and outside built-up areas).

Methods for recording crash location include the following (for more information see 11, 12):

- **Road name/route number and latitude/longitude (X, Y) coordinates.** This method records the crash location using the X, Y coordinates of a given location in a geographic coordinate system. The most reliable way is to take measurements with a mobile Global Positioning System (GPS) device at the crash scene, immediately following the crash. This uses satellites to provide actual geographic coordinates for the current location that can be converted to a local or national coordinate system. GPS may not work well in built-up areas where the satellite signal is blocked. Coordinates can also be read from a map, but this requires up-to-date maps and leaves greater room for error.
- **Linear Referencing System (LRS).** This provides a mechanism to describe the location of a point with unknown map coordinates by referencing it to a known point along a road network (e.g. kilometre posts along motorways). Using this method, unique route numbers are assigned to continuous sections of road. The route numbers may or may not correspond with those shown on official road maps



CASE STUDY 3.3: **Redesigning the statistical road crash data collection form, Spain**

Spain's Traffic General Directorate (DGT) identified the need to substantially redesign its statistical accident data questionnaire, thereby improving its usefulness, increasing its capacity to meet the growing demand for information, and reduce the time, effort and other resources invested in data collection.

The redesign process included:

- exploring the current state of crash data collection, using existing data from police crash data archives;
- establishing a Technical Work Commission, made up of representatives of the police in charge of crash data collection, the DGT, and universities and ministries involved in the collection and use of road traffic crash and injury data;
- the collection and review of information from the different crash data collection forms at the national and international level;
- carrying out a survey of police traffic-crash data management, to diagnose the state of information collection procedures, systems and quality;
- proposing contents for revised form, based on results obtained from the previous steps and distinguishing between crashes occurring on urban and inter-urban roads;
- organising workshops to enable the Technical Work Commission to reach consensus on which variables to include, the structure of the data collection form, and categories, codes and definitions. Achieve consensus on varying levels of information required, according to the severity of the crash. For each field of information, rate the usefulness and the level of difficulty for collecting the data. The following recommendations are worth mentioning:
 - Homogenization of criteria and the definitions for data collection by all police when an injury crash happens, as well as harmonization with European standards.
 - Improvement of the information fields and adaptation to new technologies, e.g. a proposal for monitoring injured victims at 30 days, inclusion of geographical coordinates to identify specific crash location, and alcohol and drug-use information.
 - Definition of an alternative method for classification of 'crash type' that overcomes current limitations, and allows a study of the sequence of events of a crash from a statistical perspective (METRAS method of sequencing events, see 13).
- evaluating the new form using consolidated data from a pilot field study;
- detailed specifications for database structure and hardware;
- using software applications and computer systems that are flexible and adaptable for the collection, management and analysis of crash data, and that include automated data quality checks. Development of interchange mechanisms with other databases currently in operation;
- developing an instruction handbook and implementing a training programme;
- implementing a new crash data collection system;
- evaluation.

This methodology has been applied successfully in the state of Catalonia to redesign the crash data collection form, and currently is being applied at the national level to introduce standard crash data collection throughout the country.

or road signs. A zero-km point is chosen on each road, and a particular location is measured by the distance from that point. Good km-post sign systems with appropriate frequency (200 metres on main roads and 500 metres on secondary roads) facilitate accurate identification of crash location. Missing, damaged or too few km-posts make the system less useful for police in locating a crash scene. The effectiveness of LRS in crash location also depends on the availability of sufficiently detailed and accurate maps for police officers recording location information.

- **Link-node system.** This method uses known points along the road network, usually intersections, identifying them as nodes with a unique number. The section of road connecting one node to another is called a link and is also assigned a unique reference number. Specific locations can be identified by specifying the distance from a node and the direction of travel. Like LRS, the effectiveness of a link-node system depends on easily recognizable reference marks along the road and the availability of sufficiently detailed and accurate maps for police officers recording location information.
- If none of the systems above is in use, which is often the case in low- and middle-income countries, the **road name and street number** can be used to identify crash location. This is the least accurate method, with particular problems in rural areas.

In cities or towns it is important that police officers attending a crash scene record the highway or street location and accurately measure the distance from the nearest cross intersection, junction or identifying landmark. The point of collision should be clearly identifiable from known or permanent features so that follow-up investigators can identify the crash location from the description provided. In country or rural areas, the same level of care is required to identify the exact location of the crash with reference to known landmarks, accurate distances from towns or villages, permanent road markings, mile or kilometre posts, road signs, property locations, curves or road intersection points.

In all situations, road and vehicle directions (compass point directions – particularly north) should be recorded accurately. The police vehicle odometer can be used check the distance of the crash from identifiable landmark locations. Even if GPS is available and used, physical measurements with a tape measure should still be undertaken to determine the point of impact. Photographs of the crash scene and surrounding environment, supported by accurate measurements, can also provide useful information.



At a minimum, officers required to attend crash sites, record particulars and investigate the incident should have a 100 metre tape measure, a 10 metre tape measure and a local map.

To make the best use of crash location data, the method for specifying location in the crash file should be compatible (or the results convertible) to the location information system used in corresponding files in other databases, such as road inventory or traffic files. Algorithms can be developed to match GPS/Geographic Information System (GIS) data with LRS reference points.

Improve training

Training police officers complete the data collection forms properly, and in data entry techniques if relevant, can improve data quality. Police officers who are responsible for filling in data at the crash scene should receive training in the following areas:

- The purpose of data collection (i.e. why it is critical for road traffic injury prevention) and the importance of their role as data collectors. The WHO/TRIPP *Road traffic injury prevention training manual* is a helpful resource (14).
- Which forms and/or reports they are responsible for completing, and when.
- Which information is needed for each field on the data collection form.
- Definitions of terms and associated data values, abbreviations and codes (especially important for variables that require subjective judgement, such as injury severity or light conditions).
- Which fields are mandatory to complete at the crash scene.
- Interview techniques for eliciting information from people involved in the crash and for obtaining eyewitness testimony.
- How to take and record required measurements (e.g. crash location, skid marks and point of impact).
- How to sketch a crash diagram.
- Techniques to validate data during the collection process (e.g. taking multiple measurements, clarifying conflicting statements).
- Data collection and entry procedures, and their responsibilities in these processes.

Staff with data entry responsibilities should receive training in the following areas:

- What forms and/or reports they are responsible for completing, and when.
- What information is needed for each field on the data collection form.
- Definitions of terms and associated data values, abbreviations and codes.
- If staff will be entering data that are not pre-coded, special training is needed for proper extraction and coding of data.
- Validation checks and techniques that can be used in the data entry process.
- Data collection and entry procedures, and their responsibilities in these processes.

If the traffic police force is large, it may be more efficient to train a select group of officers who can then train others in their jurisdiction (the 'train-the-trainer' approach).

Appropriate training is an essential part of data quality assurance, but it does not guarantee quality improvements. There are many reasons why a police officer might

not complete a data collection form properly, even when trained to do so. Competing priorities (e.g. pressure to clear the crash site and minimize congestion), time pressures and perceived stake in road safety data collection all influence ability and willingness to collect crash data. Data collection forms and procedures for data collection, entry and reporting should be designed to make the police officer's job as simple, quick and easy as possible. Remember that data collection is just one part of the overall police responsibility for crash investigation and any prosecution procedures.

Quality assurance measures

Quality assurance measures are planned, systematic checks built into data collection and entry procedures to ensure that data captured in the system are accurate and reliable. These checks should be performed on a regular basis, and might include:

- periodic observation of police officers recording data at the crash scene;
- tracking the number of crashes which have been reported to police but do not yet have a crash record in the system (report the number of pending records on a weekly or monthly basis);
- periodic checks of a random sample of electronic records against their matching original source (e.g. hard copy of data collection form, police report) for completeness and accuracy;
- periodic checks of a random sample of records for correct classification of injury severity and crash severity (using detailed police reports or hospital data as reference);
- running statistical tests to determine if certain fields are more prone to being incomplete (see Module 2), so that potential bias can be identified and addressed through changes to data collection instruments or training.

Quality assurance also includes planning for in-depth retrospective evaluations and for assessments of under-reporting (see module 2), which would be done less frequently than the monitoring activities described above.

3.4.2 Strategies to improve data system performance

The situational assessment and in-depth evaluation may indicate that data quality is adequate, but changes may be needed to improve the functioning of the system in which the data are stored and processed. Several strategies that may be used to make improvements in these areas are discussed below:

- Review workflow and user requirements
- Assess database system features
- Possibilities for linkage
- Review (or create) a data management plan
- Implement quality assurance measures (see section 3.4.1)

Box 3.3 contains a checklist of questions to help you identify aspects of the existing data system that can be improved.

BOX 3.3: Checklist for data system improvement strategies

- Which departments contribute data, enter data, or analyze data directly from the existing system, and what changes are proposed?
- From workflow mapping, which parts of the process lead to long time delays, duplication of work, or have a negative impact on data quality?
- Do the features of the database system meet users' key requirements? If not, what features are needed, and can the existing data software platform be altered to meet these requirements?
- Is there a need to change the data software platform used?
- Is linkage to other databases feasible and desirable? What are the possible mechanisms?
- Is there a data management plan? Does it specify procedures for data collection, entry, processing and use? Are roles and responsibilities specified and assigned appropriately? Does it contain adequate provisions for data back-up and security?
- What additional quality assurance measures can be introduced?
- Are there enough staff dedicated to the system and do they have sufficient capacity to operate it?

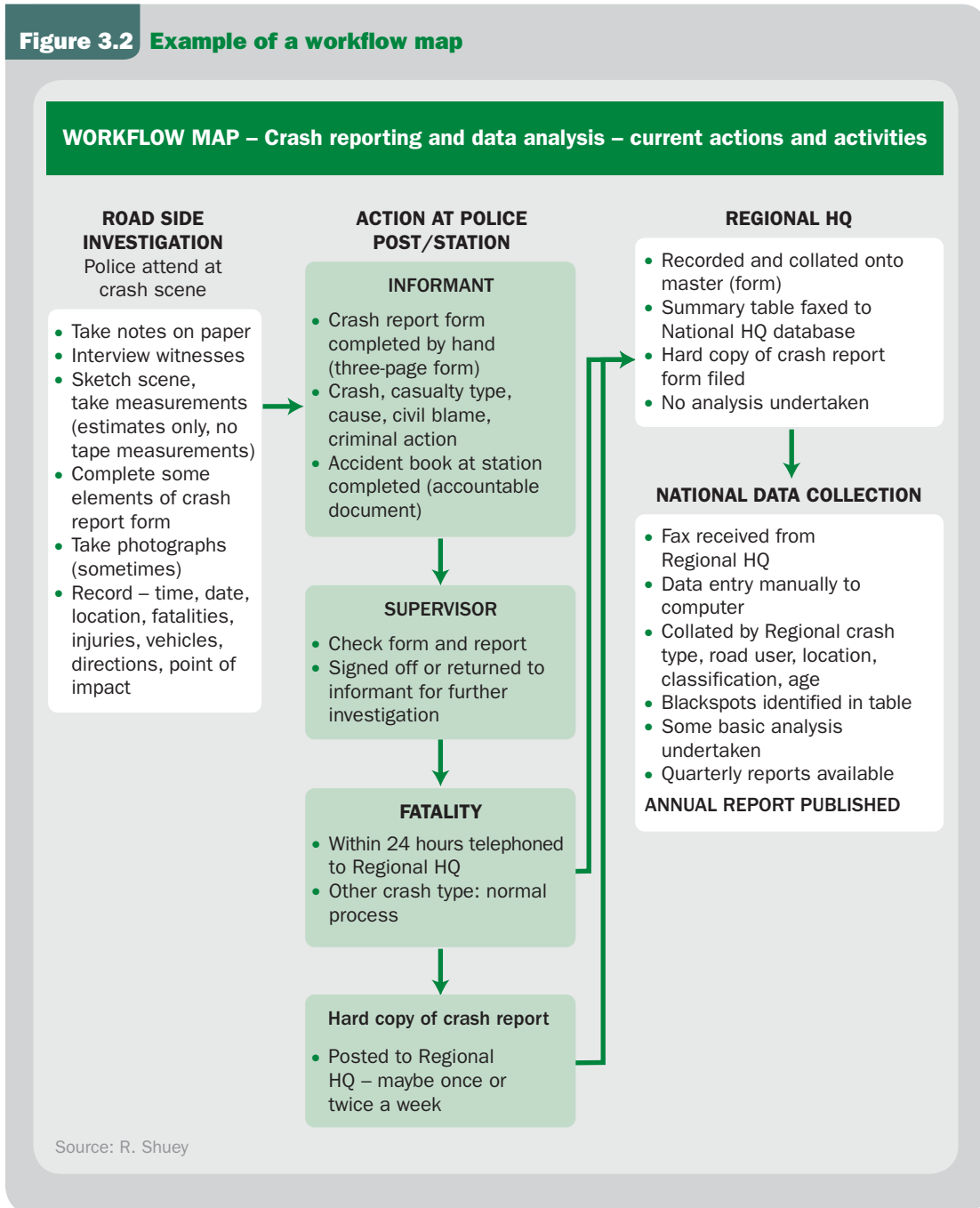
Review workflow and user requirements

Workflow mapping shows how crash data move through the system, from collection at crash scene to analysis and dissemination. This visual representation can help identify procedural problems that have a negative impact on the overall data system. A workflow map should have been created during the situational assessment or in-depth evaluation. If not, it should be done now. The purpose is to identify the processes by which data move through the system and who is responsible at each step. Figure 3.2 shows a workflow map for a hypothetical road crash data system where data are collected by the police, and data entry and analysis are centralized at the National Road Safety Council.

The workflow map should document how things actually happen, not what the process is supposed to be. For example, if the hard copies of crash data records should be sent monthly to the central processing agency for data entry, but in reality they are sent only twice a year, this should be captured in the workflow map. Further guidance on workflow mapping can be found in (15).

Once the workflow mapping is complete, validate it with the working group and use it to identify areas where improvements can be made, either by changing procedures or practices, or changing the data software platform. Mapping the ideal workflow may assist this.

Figure 3.2 Example of a workflow map



Centralized data entry, where data collection forms or crash reports completed by police officers are forwarded to a single location for coding and entry to the crash database, can be an effective way to improve data quality and system efficiency if the central body is adequately staffed with well-trained people (see Case study 3.4).



CASE STUDY 3.4: Road Crash and Victim Information System, Cambodia

The volume of motorized road traffic in Cambodia has risen rapidly since 1995. At the same time, weak traffic regulations, insufficient enforcement, speed increases because of improved roads and a lack of road safety education have led to rapidly rising numbers of road crashes and casualties. Inadequate public health care provision for people with road traffic injuries, and poor access to health services generally have made the problem worse.

Until recently, road crash data were collected by three different ministries (Public Work and Transport, Interior, and Health). Although the databases developed by these ministries provided relevant indicators on Cambodia's road safety situation, under-reporting levels were high and the databases were limited in scope, incompatible and inaccurate.

Recognizing the key role of reliable data for effective road safety prevention, the three ministries began to develop a new system in 2004, based on standardized and more detailed data collection forms. The project was led by Handicap International Belgium (HIB), with support from the French Development Cooperation, Belgian Technical Cooperation and the World Health Organization. The system has been developed in accordance with the requirements of ASEAN and the United Nations, and aligned to Action 2 (Road Accident Data Systems) of the National Road Safety Action Plan of the Royal Government of Cambodia. The system has been extended to cover all

provinces in Cambodia and to equip traffic police with global positioning system (GPS) devices.

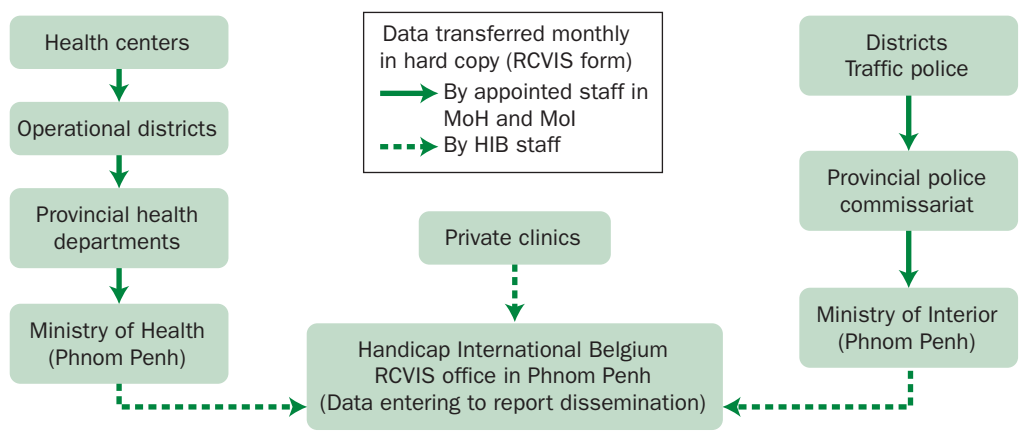
The objective of the Road Crash and Victim Information System (RCVIS) is to provide accurate, continuous and comprehensive information on road crashes and victims. This in turn will increase understanding of the current road safety situation, help plan appropriate responses and policy, and help evaluate the impact of current and future initiatives.

Data sources

To minimize under-reporting, the RCVIS draws on combined information collected from traffic police and hospitals. To ensure high quality data collection, Handicap International Belgium, in collaboration with the Ministry of Interior and Ministry of Health, ran training sessions on the use of the new standard data collection forms, inviting traffic police from each district, and technical staff from health centres and hospitals located on national roads in 24 provinces.

The following figure shows how data are transferred from district to ministry level. Both ministries are responsible for the collection of data from their provincial officers, while HIB is responsible for centralizing the data from both ministries and private clinics, and analysing the data and publishing it.

Data collection flow



Continues . . .

Continued from previous page

System life cycle

The information cycle of the RCVIS can be described as:

1. **Data collection:** There are two different RCVIS forms. The first, used by traffic police, concentrates on the type and causes of crashes, while the second, filled in by hospital and health centre staff, focuses on the type and severity of injury. In general, traffic police are not present at all crash scenes and do not report all crashes they witness. Hospital data are therefore needed to complement traffic police data. Once complete, the forms are sent to the relevant national levels (Ministry of Health and Ministry of Interior in Phnom Penh) at the end of each month.
2. **Data verification and follow up:** Each month HIB collects the forms from the national levels (ministry offices) and private clinics. Next, data verification ensures the forms are complete and contain accurate information. Additional contact may be made directly to the provincial staff who filled in the forms to get more detailed information.
3. **Data entry and storage:** After verification, the forms are entered into the database through two applications – one for police, and one for hospitals.
The applications are designed to prevent confusion and mistakes during data entry (e.g. if the victim is a motorbike rider, seatbelt information is not needed).
4. **Data check and analysis:** Data checks are made to spot duplicated data entry by health facilities and traffic police. If a casualty is reported by a health facility as well as by the traffic police, it will be entered only once, as hospital data. To eliminate double entries, common core variables such as name of victim, accident date, time, type of road user, type of transport, location of crash, severity of injury, and hospital discharge are checked.
The checking process is complicated and is therefore done manually. After the checking process, all data are centralized into the RCVIS database, which will be analysed and used to produce reports. Data are exported for further analysis using software such as SPSS Statistics and Microsoft Excel.
5. **Monthly/annual report production:** A report is produced to compare the evolution of trends from month to month, or year to year. Unusual trends can be spotted – for example drink-driving during the Khmer New Year, or the number of

victims wearing a helmet after a period of helmet enforcement. These can be reviewed again in the database. If causes and remedies are identified, they will also be included in the report.

6. **Monthly/annual report dissemination:** The reports are disseminated regularly in soft and hard copy to more than 400 end users, including the National Road Safety Committee (NRSC), the ministries of Public Work and Transport, Interior, Health, and Information, the National Assembly, news media and local and international non-governmental organizations.
7. **End user and feedback:** At the end of the annual reports, a feedback form is attached that end users can fill in and return to HIB via e-mail or hard copy.

Impact of the system

- **Increase of political will:** Through data dissemination via media, government officials have contacted HIB for more detailed data to help develop policies, strategies and an action plan to reduce road crashes (e.g. the Cambodian prime minister referred to RCVIS data to call for more commitment and action on road safety).
- **Improvement of blackspots:** The Ministry of Public Work and Transport, in collaboration with the Japan International Cooperation Agency (JICA), is now using blackspot data to plan appropriate remedial actions at dangerous locations along the national road network. TICO, an emergency response organization, uses the data to locate ambulances close to spots with frequent crashes.
- **Reference to develop plans and proposals:** RCVIS data provide the reference for all road safety stakeholders (NRSC, Ministry of Health, WHO, GRSP) to develop strategies, proposals, and documents for the Cambodian road safety sector.
- **Evaluation:** RCVIS data have been used as an evaluation tool to measure the effectiveness and impact of project implementation, such as helmet promotion and community-based education projects.
- **Extension to another system:** Based on the experience of RCVIS, the Ministry of Health has decided to extend the system into a broader Injury Surveillance System (ISS), which will include data on other causes of injury, such as falls, domestic violence, and drowning.

HIB will continue supporting the implementation of the RCVIS while the system management is transferred to the ministries of Health and of Information, and the General Secretariat of the National Road Safety Committee.

Database system features

Compare stakeholder data needs, and the related user requirements, with the system features documented in the assessment/evaluation. Changes should be considered for areas where these do not match. For example, a key stakeholder group might place high priority on the ability to generate reports that differ from the system's pre-defined reports. Some of these discrepancies may be possible to address by changing the database architecture or access to the system, while others may require introduction of a different software platform.



A database is a collection of related data organized for storage, search and retrieval. Databases of paper records are organized, searched and retrieved by hand. Electronic databases use computer software platforms to organize, store, search and retrieve the records, as directed by an administrator or user. Electronic databases can be structured according to various models (e.g. hierarchical, relational). The structure, or architecture, of the database directly affects users' ability to search and retrieve records quickly, and the types of analyses that can be performed.

Seek assistance from someone with extensive experience in database creation and management, including up-to-date knowledge of database models, software platforms and developments in technology. This expert should be able to help identify how the existing database system can be altered to better address user requirements, and whether the existing software platform is capable of supporting these changes. They should also be committed to ongoing support (see Box 3.4).

Research on 11 'good practice' crash database systems in Asia, Europe and North America identified several useful features (16):

- 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.
- Option for customized, 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.

BOX 3.4: Working with consultants and commercial suppliers

Road safety consulting and the development of commercial products related to road safety data constitute a growing business. Consultants and suppliers provide an important service and represent significant expertise. The knowledge, skills and time required to design, implement and modify road crash data systems should not be underestimated.

However, consultants and suppliers must be chosen with care to ensure appropriate service and sustainability. Choice of the wrong consultants or products can lead to frustration, wasted resources and failure of the project. This is equally true whether contracting consultants to assist with modifications to an existing road crash data system, or designing a completely new one.

To maximize success of the project:

- choose consultants and suppliers with specialist knowledge, experience working in countries with road safety situations and data systems similar yours, and the capacity to provide ongoing technical support;
- ask other clients about their experience and satisfaction;
- ask about the contractor's staffing and delivery capabilities;
- include follow-up activities in the contract and clarify availability of and mechanisms for support after completion of the project;
- pay contractors on a fixed-fee basis, where they are paid for meeting deliverables, rather than on a time basis, where contractors are paid per hour of work;
- choose commercial products (e.g. software platforms) that have been tested and proven, and that will be adequately supported, (for implementation and long term) by the supplier or other consultants;
- choose commercial products that your staff can be trained to use and support, so you do not have to depend solely on supplier support after implementation.

- Crash density maps.
- Site ranking based on crash rates, numbers, costs.
- Route assessments.
- Ability to monitor sites of interest, i.e. before and after treatments.
- Details of search criteria included on outputs.
- Web-based access for data entry and analysis.
- Public-access version of database.

Integration of hospital data on injury severity and outcomes was also identified as an important feature, though few road crash database systems have achieved this. It is worth noting that crash database systems may be able to perform at an adequate level without these features. However, addition of these features would improve the accuracy, efficiency and utility of data provision through the system (16).

Possibilities for linkage

Linking police data and other data sources is often proposed as a way to improve data quality, but it may not be the best starting place for database system improvement.

Establishing successful linkages among existing databases can be extremely complicated and difficult to do. Resources may be better invested in other strategies.

As a first step, a subset of the multisectoral data working group could meet regularly (weekly, monthly, or quarterly depending on the volume of severe and fatal crashes) to review and compare data from different sources, and discuss possibilities for formal linkage mechanisms. If it is not possible to establish linkages among databases, it may still be possible to include data from other sources by using centralized data entry (see Case studies 3.5 and 3.8).

Under-reporting levels and accuracy of injury severity classification can be assessed using periodic studies (see Module 2) where systematic linkage of databases is not feasible. Further information on data linkage is provided in section 3.5.



CASE STUDY 3.5: **Fatality Analysis Reporting System, USA**

The USA's Fatality Analysis Reporting System (FARS) was conceived, designed, and developed in 1975 by the National Center for Statistics and Analysis (NCSA) – part of the National Highway Traffic Safety Administration (NHTSA). It provides the traffic safety community with the means to identify traffic safety problems, develop appropriate solutions and provide an objective basis for evaluating the effectiveness of motor vehicle safety standards and highway safety programmes.

FARS contains data derived from a census of fatal traffic crashes within the 50 states, the District of Columbia and Puerto Rico. Crashes included in the database involved at least one motor vehicle travelling on a traffic way open to the public, and resulted in the death of a person (vehicle occupant or non-motorist) within 720 hours (30 days) of the crash.

All FARS data on fatal motor vehicle traffic crashes are gathered from each state's own source documents, and are coded on standard FARS forms. Analysts obtain the documents needed to complete the FARS forms, which generally include some or all of the following: police accident reports, state vehicle registration files, state driver licensing files, state highway department data, vital statistics, death certificates, coroner/medical examiner reports, hospital medical reports, and emergency medical service reports. Each case has more than 125 coded data elements that describe the crash, vehicles, and people involved.

FARS data released to the public do not include any personal identifying information such as names, addresses, or social security numbers, and vehicle identification numbers are abbreviated in public access files on the Internet. Thus any data kept in FARS files and available to the public fully conform with Privacy Act laws.

FARS data are used extensively within NHTSA, the national lead agency for road safety. National and state-level analyses are performed. NHTSA publishes FARS data in various formats including fact sheets, an annual statistical compendium, and reports on special themes. FARS data can be queried by the public and can also be made available on CD-Rom and computer tape. NHTSA regularly receives requests for FARS data from state and local governments, research organizations, private citizens, the automobile and insurance industries, Congress, and the media.

For more information about FARS, or to access FARS data, see www.nhtsa.dot.gov/portal/site/nhtsa/menuitem.0efe59a360fbaad24ec86e10dba046a0/

Review (or create) a data management plan

The data management plan should document the intended workflow (or Standard Operating Procedures) of data collection, entry, processing and analysis, specifying roles and responsibilities of the people and agencies involved. Putting this plan in writing provides a 'road map' of how the system should function, and acts as a monitoring tool.

The plan should specify:

- the agency and title of staff member in charge of the overall road crash data system – this is the agency that 'owns' the database and is usually responsible for data processing and analysis;
- summary of key user requirements for database system;
- software platform;
- facilities and equipment requirements;
- description of data to be collected (e.g. index of data elements);
- agency and staff responsible for data collection (may differ by data element);
- data collection instruments and procedures (may differ by data element);
- agency and staff responsible for data entry (may differ by data element);
- data entry procedures;
- data cleaning and processing procedures;
- quality assurance measures (both built-in and manual);
- agency, title and duties of database administrator(s), usually responsible for back-up, security, system performance and availability, development and testing support for hardware and software;
- back-up equipment, software and procedures;
- IT-specific security mechanisms, software and procedures;
- provisions for protecting confidentiality;
- linkage mechanisms (if applicable);
- system access for analysis and reporting;
- dissemination – output formats, frequency, target audience, agency and the person responsible for producing this information.



NOTE Protect your data! Review back-up procedures and security mechanisms with an IT expert to make sure data are as safe as possible from accidental or malicious loss (e.g. through computer hackers). Avoid carrying data on a laptop computer or storage device that may be lost or stolen.

In addition to a data management plan, it is important to ensure that sufficiently trained staff are available to allow the data management system to function properly. A system environment evaluation can reveal if this is the case (see 4). If not, additional staff and training may be required.