International Transport Forum

Reporting on Serious Road Traffic Casualties

Combining and using different data sources to improve understanding of non-fatal road traffic crashes

irtad
International Traffic Safety Data and Analysis Group

OECD

Reporting on Serious Road Traffic Casualties

Combining and using different data sources
to improve understanding
of non-fatal road traffic crashes

International Traffic Safety
Data and Analysis Group

FOREWORD

It is with great pleasure that I present this special IRTAD report prepared by a Working Group on *Linking Police and Hospital Data* with a view to identify and assess methodologies for linking different sources of accident data in order to develop better estimates of the real number of road traffic casualties.

International comparisons of road safety necessarily limit themselves to data on fatal crashes and fatalities. These data are of good quality in many countries in the world. Because we use the same 30-day definition (dying within 30 days after a crash occurred) and if countries collect good quality data we can make meaningful international comparisons. But we all know that road crashes can also result in very serious injuries. We know as well that these injuries can have a serious impact on citizen's lives. Moreover, the associated economic costs are considerable. Many OECD/ITF countries have set road targets: for example not more than 1000 deaths in 2020. More and more, countries have also formulated targets for injuries (for example not more than 10 000 serious road crash casualties in 2020) as an expression of a growing interest in dealing with wider issues of road safety. Internationally harmonized and accepted definitions of (serious) injuries are not yet available. This means that it is not possible to compare the wider safety record of countries as long as data definitions and data collection procedures differ. In addition, the problem of underreporting of serious casualties is widely recognised, and many countries work hard to eliminate this problem.

The aim of the IRTAD Group is to co-operate and support its members, and also non-members, to improve road safety data quality in general terms. Our database aims to deliver good quality data. Currently, these data are from fatal crashes only. It is our understanding that data on road traffic injuries are also urgently needed. For this reason we started to work on this topic. The IRTAD Group hopes that this report will make a contribution to the aim that more and more countries will use data on road crash injuries in road safety policies and research, not only IRTAD members but other countries in the world as well. We believe that major steps forward will be made if data on road traffic casualties are widely available, of good quality, and comparable between countries. We trust that this report can contribute to that goal.

The information included in this report is based on:

- Discussions at the Workshop held in London on 30 November – 1 December 2009
- A survey among all IRTAD Members to collect information on:
  - Availability of information sources.
  - Definitions of serious injuries.
  - Methodologies used to link different sources of data.
- Literature reviews of recent research work on linking hospital and police data

This means that our research is mainly based on experiences and knowledge collected in IRTAD-member states and this report reflects the state of the art in these countries, many of which are in a position to work towards better, comparable, data on non-fatal casualties.

It is not easy to answer the question how to apply this way of thinking and working in other countries such as low and middle income countries. For most of these, the priority will continue to be to collect accurate and comprehensive data on fatalities. IRTAD offers its capabilities to support these aims.

We expect the audience for this report to be mainly data specialists in the road safety area, both in transport and health related institutions. Some of the less technical parts may be of wider interest.

Prof. Fred Wegman
Chairman of IRTAD

TABLE OF CONTENTS

CONCLUSIONS AND RECOMMENDATIONS

1. For some purposes it is essential to know the true road casualty total, for example:

   - Understanding trends in collision rates and collision risk, especially when making international comparisons, since reporting practices vary from country to country.
   - Enabling road safety to be highlighted as a priority for action and developing robust arguments for the adoption of interventions (OECD, 2008).
   - Comparing road crashes with other major causes of death (such as other accidental deaths, HIV, malaria).
   - When evaluating the cost of road casualties, in order to estimate medical costs, the burden of road accidents in terms of future disability, and the cost/ benefit ratio for schemes to reduce casualties
   - Efficiently deploying these interventions across areas of higher risk and/or where the greatest potential improvement can be achieved, such as in different road user groups, area types (urban/ rural), or road types etc.

Recommendation 1:

A complete picture of casualty totals from road crashes is needed to fully assess the consequences of road crashes and monitor progress.

2. Road traffic crash fatality data are now well established in IRTAD, and efforts are being made to extend coverage to other casualties, particularly the most serious. Two significant areas need to be addressed to make progress in the collection of comparable data on non-fatal accidents from as many countries as possible:

   - How to define “serious” accidents?
   - How to allow for under-reporting to make the best possible estimate of the number of serious casualties, possibly by linking data from police and hospital sources?

Recommendation 2:

Injury information should complement information on fatal crashes to give a fuller picture of road crashes. Information on injuries should become more important for international comparisons.

3. The police do not collect information on all non-fatal accidents, partly because the legal requirement to report accidents to the police varies from country to country. Making better use of hospital road casualty data should therefore be encouraged. It is clear from research in a number of countries that many road traffic casualties are admitted to hospitals who are not known to the police. On their own, hospital data are not a substitute for police data, as they are likely to include only minimal information on the circumstances of a crash. However, using hospital data in addition to police data can provide valuable information, including:

   - Some basic information on casualties not reported to the police, such as age, gender and vehicle type.
   - Better understanding of the total number of casualties.
   - Better understanding of injury severity.

- In-depth understanding of the medical consequences of particular types of crash, if police and hospital records are linked.

Recommendation 3:

Police data should remain the main source for road crash statistics. However, because of under-reporting problems and possible bias (for example with differing rates of reporting by vehicle type), police data should be complemented by hospital data, which are the next most useful source.

4. There are clear benefits using data on those admitted to hospital following a road traffic crash to supplement police data when studying road casualties. However it is not so easy to control the quality and completeness of data that are collected for medical purposes and hospital administration:

   - Hospital practices change over time, according to the administrative needs of each hospital (for example for billing patients or insurance companies, or claiming costs from elsewhere).
   - Practices also vary between countries, according to data requirements set by central or state governments.
   - Hospital practices may vary from place to place in the same country. In particular, practice may differ on whether casualties should be admitted to hospital for observation, especially if head injuries are suspected.
   - Data may be collected mainly for financial purposes, with little requirement for fields indicating the origin of the crash, making it difficult to identify road casualties.
   - Data systems may differ from hospital to hospital (and even within different hospital departments), and data are rarely gathered into a national system.
   - Inevitably, medical staff does not always treat data entry as high priority.
   - There may be ethical concerns about releasing confidential medical information. Again, different hospitals may have different policies in this area.

   As with police data, some of these issues will be less significant with more serious casualties, especially variation in admission practices.

   In addition to data on those admitted to hospital, some countries may have records on people who attend the "emergency room" only. These systems are less developed, but may have potential in future to provide more information.

Recommendation 4:

The data from hospital emergency departments, available in some countries, should be monitored regularly and researched to determine if they might shed more light on road casualties.

5. In many countries, the police are legally obliged to go to the scene of a crash where there is at least one injured person and one moving vehicle (motorized or not) involved. The police are then responsible for collecting information on the number of casualties, assessing the severity of injuries, and the overall severity of the crash. In most countries, the police define the following levels of injury severity:

   - Fatalities: when someone die within 30 days after a traffic crash.
   - Injured persons, usually split in two categories :
     - Seriously injured.
     - Slightly injured.

The information on crash severity, as reported by the police, is rarely checked later with medical records, except when the injured person dies in hospital. In that case, the information is sent from the hospital to the police who can modify their report, so an injury crash becomes a fatal crash. This process is a source of potential underreporting of fatalities in the police data.

In many countries, hospitals maintain their own database of injuries. But systematic cross reference to police databases is very rare. Hospital data are mainly used for medical purposes and are seldomly used for official statistics on road safety.

Therefore in most countries, official statistics on injury severity (with the exception of fatalities) are only based on the assessment of the police officer at the scene of the crash or on the information transmitted to the police in a short time after the crash.

Recommendation 5:

The assessment of the severity of injuries should preferably be done by medical professionals, and not by the police officer at the scene of the crash.

6. There are no commonly agreed definitions on injury severity. Criteria used in the police record and official statistics to classify the severity of a crash vary from country to country and include:

- The length of hospitalization: in many countries a person seriously injured is a person hospitalized, other than for observation, for more than 24 hours.
- The type of injuries. In some countries, "seriously injured" is based on some specific types of injury.
- The inability to work.
- The length of recovery.
- Long term disability.

In most countries, hospitals do not define levels of injuries as such, but use the International Classification of Diseases (ICD9 or ICD10), which is derived from the medical diagnosis, to describe the injuries.

In some countries, injury severity is defined based on indices such the Abbreviated Injury Scale (AIS), the Maximum Abbreviated Injury Scale (MAIS) and the Injury Severity Score (ISS), which can be derived from the International Classification of Diseases.

Recommendation 6:

Medical staff should be trained in order to systematically classify (road traffic) injuries using ICD International Classification of Diseases and to assess severities with indices such as the Abbreviated Injury Scale (AIS) or the Maximum Abbreviated Injury (MAIS). This information -- without personal information -- should be made easily available for statistical purposes, policymaking and research.

7. Although for all countries police records are the primary source of data on road accidents and many countries also make use hospital data, there are a number of other sources used for road safety throughout the world. There are considerable differences in what data are collected, their quality and the extent to which they are used to evaluate effective countermeasures to improve road safety in the respective countries. The other sources discussed include:

- Mortality registers.
- Forensic agencies.

- Emergency ambulance services.
- Fire services.
- Surveys.
- Insurance related records.
- In depth studies, such as crash investigations.

Police data provide detailed information about crash circumstances, location, and vehicles involved to inform development of road safety policies but they are not complete or perfect sources. It is therefore desirable to use complementary sources to build a balanced and comprehensive picture of the nature and extent of road accidents. Other datasets can be useful both as a check on the quality and completeness of police data and in providing information which is not collected by the police, for example relating to more detailed medical consequences of road accidents.

The other sources mentioned above can provide additional information about the numbers of road crash casualties but generally do not provide the detailed information about the circumstances or causes of the crash available from the police data. Their value is therefore limited in developing preventative measures to improve road safety. In addition there may be issues about quality and availability of data – often there is no national database with information only available at the local level.

Recommendation 7:

Besides police data and hospital data, other data sources are available. These have a limited value on their own, and cannot replace police or hospital data, but can be used to build a more balanced and comprehensive picture, to enrich the main data sources, and as a quality check.

8. Due to the limitations of individual databases it has become increasingly common to link databases from different sources in order to improve road safety research to support road safety policies.

There are three main methods to link two databases such as police and hospital databases: *manual, deterministic and probabilistic. Manual* linkage determines visually the possible matching of each record in a database with all records in another database. Several methods of computer assisted linking have been developed. In a *deterministic* approach the easiest is to link records using one or more unique personal identifiers. This approach is also referred to as “rules based”. It is based on the existence of a unique identifier - or combination of variables – common to both databases being linked. In essence, the *probabilistic* linkage process consists of matching two or more records that come from different data sources and are believed to belong to the same individual. It is based in two probabilities: the probability of matching given that both records belong to the same individual and the probability of matching by chance.

Recommendation 8:

For linking data, the deterministic method is preferred if a unique personal identifier is available; otherwise the probabilistic method is a good alternative.

9. When two or more databases of road casualties exist, such as police and hospital records, it is possible to estimate the number of subjects missed by both registrations, and hence estimate the total number of subjects of interest by using the capture-recapture approach, providing a number of assumptions are satisfied.

The capture-recapture approach is based on six assumptions:

- Closed population.

- Perfect identification of subjects common to both registrations.
- Independence between the registrations.
- Homogeneity of capture by a given registration.
- Same geographical area and same time period.
- Perfect identification of the subjects of interest.

In practice, it is difficult to meet these assumptions fully, but two procedures will help. Firstly, the capture-recapture method relies on a perfect record-linkage between the two or more sources of registration. This is of course difficult to achieve. The capture-recapture estimate is very sensitive to the output of the record-linkage, so there should be formal checking of its quality.

Secondly, it is helpful to stratify on the characteristics that are associated with police under-reporting. If there is no country-specific information on the characteristics associated with police under-reporting, injury severity, road user type, and whether the crash was a single-vehicle or multi-vehicle crash are likely to be important stratification characteristics.

Recommendation 9:

The six assumptions needed to use the capture-recapture method must be considered carefully. Using this method combined with linking police and hospital data may be appropriate to give a fuller picture of road casualties.

10. This report reviews the choice of a medically based definition of a "serious" injury. The objective is to identify a definition of a serious non-fatal injury suitable for monitoring trends within individual countries, and comparing rates and trends between countries.

    Length of stay in hospital is currently the indicator most often used in IRTAD countries. It presents however a number of drawbacks, including the fact that it does not necessarily reflect the severity of injuries. Therefore it is not appropriate for international comparisons given the diversity of hospital admission practices.

Recommendation 10:

Having an internationally agreed definition of "serious" injuries will help the safety research community to better understand the consequences of road crashes and to monitor progress.

Given the existing knowledge and practices, IRTAD proposes to define a 'seriously injured road casualty' as a person with injuries assessed at level 3 or more on the Maximum Abbreviated Injury Scale i.e. "MAIS3+"

CHAPTER 1
Introduction to linking police and hospital road safety data

1.1. Why do we need more information on traffic injuries?

Most countries use data recorded by the police to track trends in road crashes and provide detailed evidence to help reduce the number of casualties. However, it is widely recognised that not all crashes are reported to the police, or recorded by them, and many countries are now using hospital data on road casualties to supplement police data.

For some purposes it is essential to know the true casualty total, for example:

- Understanding trends in collision rates and collision risk, especially when making international comparisons, since reporting practices vary from country to country.
- Enabling road safety to be highlighted as a priority for action and developing robust arguments for the adoption of interventions (ITF/OECD, 2008). Road fatalities should be compared with other major causes of death (such as other accidental deaths, HIV, malaria). WHO estimate that in 2004 road traffic fatalities ranked ninth in terms of leading causes of death. It has been predicted that this will rise to fifth place in 2030 (WHO, 2009).
- When evaluating the cost of road casualties, in order to estimate medical costs, the burden of road crashes in terms of future disability, and the cost/ benefit ratio for schemes to reduce casualties.
- Efficiently deploying road safety interventions across areas of higher risk and/or where the greatest potential improvement can be achieved, such as in different road user groups, area types (urban/ rural), or road types etc.

Gathering more knowledge on serious injuries is a common concern in many countries. The European Commission has committed to launch a global strategy of actions concerning road injuries and first aid in the Communication "Towards a European Road Safety Area: Policy Orientations on Road Safety 2011-2020" Improving crash data systems is also an important priority of the Global Plan for the Decade of Action for Road Safety, adopted by the United Nations, which recommends "to establish and support national and local systems to measure road traffic deaths, injuries and crashes" and to "improve the quality of road safety data collected (UN road safety collaboration, 2011).

In 2009 the World Health Organization's Global Status Report on Road Safety concluded that two main actions were required to improve the quality of data on fatal and non-fatal road traffic injuries:

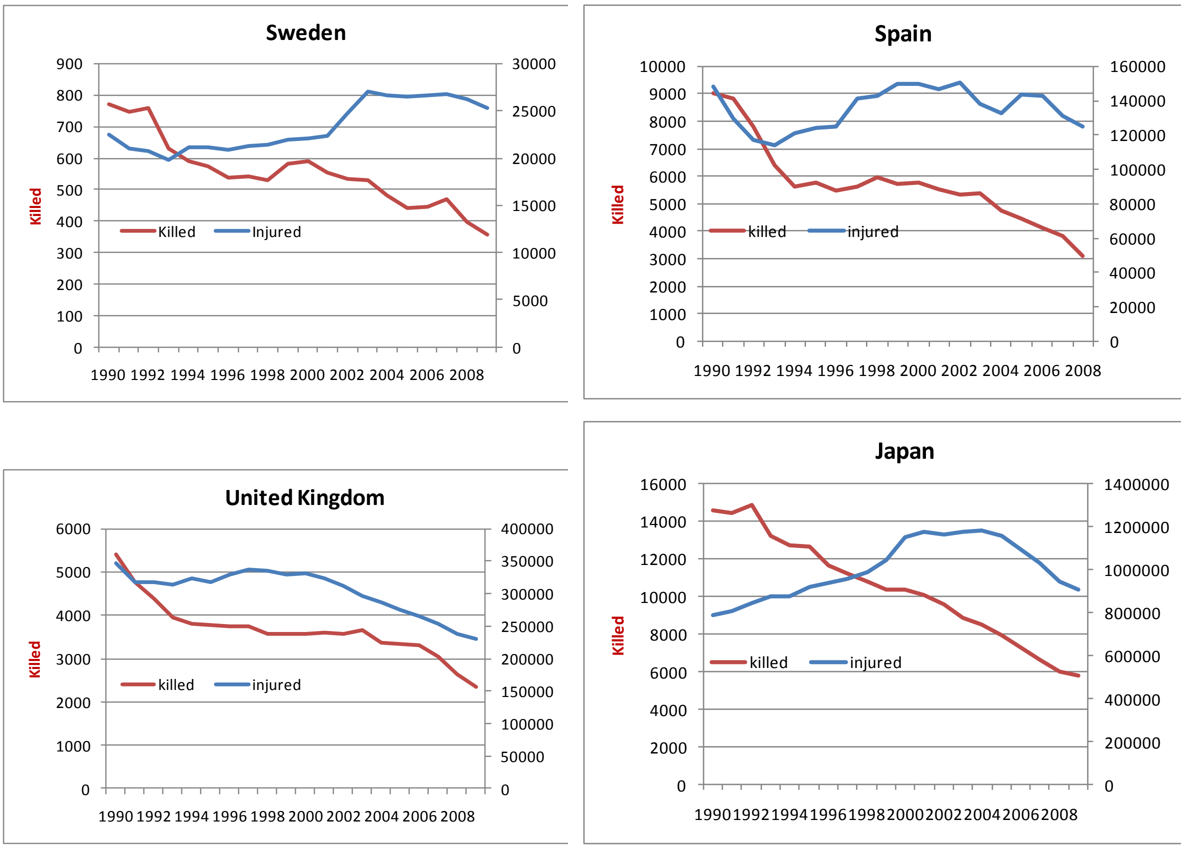
1. Encourage use of the 30-day definition of road traffic death and standardize terminology for classifying the severity of non-fatal injuries.
2. Improve the data linkages between police, transport and health services to address under-reporting.

These issues will be discussed briefly in this chapter, and explored in more depth in later chapters.

Members of the IRTAD group already use the 30 day fatal definition, but we are far from a clear standard definition of what we mean by a "serious" injury. This is becoming more important, as recent declines in fatalities in many developed countries focuses more attention on seriously injured survivors (see figure 1.1).

Figure 1.1. **Trends in the number of *reported* fatalities and injured in a selection of IRTAD countries**

1990-2009



1.2. Underreporting issues

Comparisons with mortality data are useful to show whether police records are complete. Some countries, for example Great Britain, assume that the recording of fatalities (using the agreed international definition of death within 30 days) is almost complete. Other countries, such as the Netherlands, show underreporting of fatalities, which can be taken into account for particular purposes. This issue is not covered in any more detail in this report.

While road traffic crash fatality data are now well established in IRTAD, efforts are being made to extend coverage to other casualties, particularly the most serious. As suggested by the WHO, two significant areas need to be addressed in order for IRTAD to make progress in the collection of comparable data on non-fatal crashes from as many countries as possible:

- How to define “serious” casualties.
- How to allow for underreporting to make the best possible estimate of the number of serious casualties, possibly by linking data from police and hospital sources.

This report aims to address these issues.

1.3. Issues with police data

The police do not collect information on all non-fatal crashes. The legal requirement to report crashes to the police varies from country to country, as reflected in the responses to the survey among IRTAD countries. In some countries the police are not obliged to record details of some minor crashes. More generally, the police may never hear of a crash in a rural area involving a single vehicle with a slightly injured casualty. At the other end of the scale, a crash on a major road involving several vehicles with casualties with serious injuries is very likely to be reported.

In 2007, IRTAD published a special report on the Underreporting of Road Traffic Casualties (Derriks and Mak, 2007). A number of the report's recommendations encouraged better use of hospital data in addition to police data. It is clear from research in a number of countries that many road traffic casualties are admitted to hospital who are not known to the police. On their own, hospital data are not a substitute for police data, as they are likely to include only minimal information on the circumstances of a crash. However, using hospital data in addition to police data can provide valuable information, including:

- Some basic information on casualties not reported to the police, such as age, gender and vehicle type.
- Better understanding of the total number of casualties.
- Better understanding of injury severity.
- In-depth understanding of the medical consequences of particular types of crash, if police and hospital records are linked.

Box 1.1. **Recommendations of the SafetyNet project**

This EC funded project included Task 1.5 which reported in 2008, aiming:

> *to estimate actual numbers of road casualties in Europe from the CARE database by addressing two issues:*
>
> - *the under-reporting in national accident databases and*
> - *the differences between countries of the definitions used to classify injury severity.*

Eight countries took part in the project, but it was possible to match police and hospital records nationally in only two (Austria and the Netherlands). France, Greece, Spain and the UK matched data on a regional basis, and the Czech Republic and Hungary more locally.

It had been hoped that 'conversion factors' derived from these studies could be used more widely to estimate the total numbers of serious injuries, but there were significant differences between countries and over time which made this impossible.

Crash severity

The SafetyNet report also looked at the varying definitions of a 'serious' crash. Given differences in police practice, and the problems non-medically trained police have in determining severity at the scene of a crash, there are advantages in using hospital data for this purpose. In particular, length of stay in hospital, and accident severity measured using the Maximum Injury Severity Score (MAIS) of 3 or more were considered. The conclusion was that:

> *Results based on MAIS are more likely to monitor casualty and severity trends reliably than results based on length of stay.*

1.4. Issues with hospital data

There are clear benefits using data on those admitted to hospital following a road traffic crash to supplement police data when studying road casualties. However it is not so easy to control the quality and completeness of data that are collected for medical purposes and hospital administration:

- Hospital practices change over time, according to the administrative needs of each hospital (for example for billing patients or insurance companies, or claiming costs from elsewhere).
- Practices also vary between countries, according to data requirements set by central or state governments.
- Hospital practices may vary from place to place in the same country. In particular, practice may differ on whether casualties should be admitted to hospital for observation, especially if head injuries are suspected.
- Data may be collected mainly for financial purposes, with little requirement for fields indicating road traffic crashes.

- Data systems may differ from hospital to hospital (and even within different hospital departments), and data may not be gathered into a national system.
- Inevitably, medical staff do not always treat data entry as high priority.
- There may be ethical concerns about releasing confidential medical information. Again, different hospitals may have different policies in this area.

It should be noted that some of these issues also apply to police data, especially changes in practice over time, and variations by country and within countries.

As with police data, some of these issues will be less significant with more serious casualties, especially variation in admission practices.

In addition to data on those admitted to hospital, some countries may have records on people who attend the "emergency room" only. These systems are less developed, but may have potential in future to provide more information.

Box 1.2. **Health Data**

The health sector usually keeps data on most types of injuries, covering the whole spectrum of injury from exposure to death. Data on fatal road traffic injuries may be extracted from "vital registration" data or mortality registers (derived from death certificates completed by medical doctors, which state the cause and underlying cause of death) or where these do not exist from verbal autopsy surveys. Information on non fatal road traffic injuries is kept in hospital in–patient records, trauma registries and may be collected by ambulance services or other emergency services. Some health agencies develop injury surveillance systems for ongoing systematic collection, analysis, interpretation and dissemination of health information on injuries and, it is possible to extract road traffic injury data from these systems. Minor injuries, which usually don't present to hospitals and are the most difficult to quantify.

Source: WHO (2010).

1.5. The need for various sources of information

No single database provides enough information to give a complete picture of road traffic injuries and to fully understand the underlying injury mechanisms (Lujic et al., 2008; Clark, 2004). Traditionally data collected from police have been the main source used to study road traffic injuries. However, although the police usually collect very detailed information on the circumstances of the crash, there is an important degree of underreporting, and lack of reliable data on the nature and severity of injuries. Furthermore, the level of underreporting varies according to the profile of people injured. There is greater underreporting of elderly casualties, urban crashes, slightly injured, users of two-wheeled vehicles and car occupants (Amoros *et al*., 2007; Amoros *et a*l., 2006; Alsop and Langley, 2001, Perez *et al.,* 2006).

On the other hand health care databases, such as hospital admissions records or forensic records, lack information on the circumstances of the crash, but have proved to be very useful to complement police data by capturing missing cases, and also provide detailed information on injury diagnosis (Newcombe, 1959; Winkler, 2010: Cirera et al., 2001; Newcombe, 1988; Jaro 1995). The use of these sources allows international comparisons; estimation of the actual costs to society posed by accidents; and the planning of health care resources.

1.6. Working method of the IRTAD sub group

Workshop on linking hospital and police data

A workshop was held in London in late 2009, attended by representatives from 15 countries. The group agreed the scope of this report, to include police, hospital and other data sources, and that further information was needed on current practices.

IRTAD survey

Following the workshop, a questionnaire was sent to IRTAD member countries to collect a range of information about data collected on non-fatal accidents, in order to establish practical ways to improve the evidence in this area. The areas covered include:

- Sources of data available (police, hospital, insurance etc), and their quality.
- The definition of 'seriously' injured person.
- Severity indices derived from medical data.
- Linking of datasets from different sources (such as police and hospital data).

23 countries responded, and we are grateful for the extensive information they provided. Chapters 2-5 give details of the responses to this survey.

1.7. Audience and scope of the report

This report is based on information provided by 23 countries and reflects the state of the art in these countries regarding the use of several data sources to better understand the real number of traffic casualties. It is mainly focused on high income OECD countries and aims at highlighting current best practice on linking different sources of data. The authors are aware that it may not be possible to implement some of the recommendations in the short term in all countries, but these should certainly be considered in many OECD and some non OECD countries. Where countries do not yet have good quality comprehensive data on fatalities, this should remain the first priority to help improve road safety.

The audience for this report is mainly data specialists in the road safety area, both in transport and health related institutions. Some of the less technical parts may be of wider interest.

1.8 Content and structure of the report

- Chapter 2 reports on current national definitions of a road casualty, and shows the difficulty in establishing a common international definition from the wide range currently in place. In most countries, police at the scene of a crash decide on casualty severity, but with few exceptions do not check later with the hospital which has looked after the casualty.

  Chapter 2 also describes various standard hospital codes and injury severity indices which may be helpful in determining severity. Most countries reported here use the International Classification of Diseases, either Revision 9 or 10 (ICD-10). Some countries also used one of the standard indices, such as the Abbreviated Injury Scale (AIS).

- Chapter 3 looks in more detail at the main sources of information, police and hospital data, (including admissions, and those treated only in the emergency room) and at other data sources. All countries used police data as their main source on fatalities, and the great majority also made use of hospital information on serious casualties. Not all countries collect information on slight casualties. In addition, a few countries used information from emergency ambulances in some areas, but no country had a national source of ambulance data.

- Chapter 4 considers the two main methods of linking of records from more than one source, the deterministic and probabilistic methods. It is not the purpose of this report to provide very detailed technical information on linking, but Chapter 4 is a useful introduction to the methods and issues associated with them, and includes an extensive reference list for further information. It also includes seven case studies of linkage in different countries.

- Chapter 5 covers the simple 'capture- recapture' or dual system method, which uses data linkage to estimate a total population. For example, by knowing numbers of casualties recorded by the police, by hospitals and by both (linked records), it is possible to estimate a lower bound of the number recorded by neither, and therefore to estimate the total number of casualties.

This chapter also includes three case studies and a number of references for readers who would like to explore this technical area in more detail.

- Chapter 6 considers a range of options using hospital data to consider the 'best' option for a practical definition of a serious casualty, in terms of threat to life. The chapter considers various injury severity scales, the length of stay in hospital, and particular injury diagnoses.

References

Derriks H and Mak P. Underreporting of Road Traffic Casualties. Organisation for Economic Co-operation and Development. IRTAD 2007. www.who.int/roadsafety/publications/irtad_underreporting.pdf [acceded 6/4/2010]

ITF/OECD (2008), *Towards Zero: ambitious road safety targets and the safe system approach,* OECD, Paris.

United Nations Road Safety Collaboration (2011), *Global Plan for the Decade of Action for Road Safety 2011-2020,* www.who.int/roadsafety/decade_of_action/plan/english.pdf

World Health Organization (2010). *Data systems: a road safety manual for decision –makers and practitioners,* WHO, Geneva.

CHAPTER 2
Serious road traffic injuries

Abstract

This Chapter reports on current national definitions of a road casualty, and shows the difficulty in establishing a common international definition from the wide range currently in place. In most countries, police at the scene of a crash decide on casualty severity, but with few exceptions do not check later with the hospital which has looked after the casualty.

2.1. Introduction

Until recently most analysis of road crash data has been focused on fatality data. Too little is reported or published on injury accidents and the severity of accidents, which gives biased information on actual safety performance during a period of significant decline in road crash fatalities in many OECD countries.

It is difficult to collect reliable data on injury accidents, and there is no harmonized definition of injury severity. It is desirable to develop an internationally definition of injury severity, and collect more systematic international information on the severity and type of injuries to contribute to further progress on safety research.

This chapter reviews the process for determining the severity of a crash and the current definitions of injuries in member countries used in official statistics, as well as experience of assessing injury severity. It is based on the responses to the survey undertaken for this project, to which 23 countries responded.

2.2. Determining the severity of a crash in official reports

In many countries, the police are obliged to go to the scene of a crash where there is at least one injured person and one moving vehicle (motorized or not) involved. In all 23 countries, except Japan and the Czech Republic (see below) , the police are then responsible for collecting information on the number of casualties, assessing the severity of injuries, and the overall severity of the crash. In most countries, three levels of severity are recorded by the police: fatality; serious injury; slight injury.

The information on crash severity, as reported by the police, is rarely checked later with medical records, except when the injured person dies in hospital. In that case, the information is sent from the hospital to the police who can modify their report, so an injury crash becomes a fatal crash. This process is a source of potential underreporting of fatalities in the police data (Derriks and Mak, 2007).

In many countries, hospitals maintain their own database of injuries. But systematic cross reference to police databases is very rare. Hospital data are mainly used for medical purposes and are not used for official statistics.

Therefore in most countries, official statistics on injury severity (with the exception of fatalities) are only based on the assessment of the police officer at the scene of the crash or on the information transmitted to the police in a short time after the crash. Japan and the Czech Republic are two exceptions where the severity of a crash is determined by the medical doctor at the hospital.

Based on the responses received, Germany is an example of a country with a more comprehensive procedure. The severity of a crash is documented in the notice of accident, compiled by the police. Afterwards the police are obliged to check if the injured person had to stay in hospital for more than 24 hours, which is the definition of a serious injury. For fatalities Germany applies the international definition, i.e. killed = persons who die within 30 days as a result of the crash. When a crash victim dies in hospital the hospital is bound to inform the police. Therefore police have to wait at least 30 days before closing the case if accidents with personal injury took place. The crash category is adjusted by the police based on the death information provided by the hospitals.

Table 2.1. **Who decides on the severity of a crash for recording official statistics?**

	The Police at the scene of the crash	**The hospital**	**Checked later with hospital for injury severity (exc. Fatalities)**	**Comment**
Australia	√		No	The hospital keeps separate records of its own. No data matching except in Western Australia.
Austria	√		No	If the police officer in charge is uncertain of the severity of injury, the box "NEG" (unknown injury level) is ticked.
Czech Republic		√		The severity is determined by the acting doctor. It can be modified later by the hospital (up to 30 days).
Denmark	√			If only slight injury is indicated on the form the severity is regarded as slight. Otherwise it is regarded as serious, and then information is often gathered by asking at the emergency department.
Finland	√	No	No	
France	√	No	Not systematically	Checks can be made for the seriously injured.
Germany	√		√	Referring to fatalities the police have to wait at least 30 days before finishing the report in the case of an injury crash. The crash category is adjusted by the police by contact with the hospital.
Great Britain	√		Not always	Data protection issues may prevent the police obtaining information from medical staff.
Hungary	√	-	√	
Israel	√	-	Not always	
Japan	No	√		The doctor predicts days of recovery.
Lithuania	√	-	No	
Netherlands	√		√	The police officers basically judge severity by the fact that an involved person is transported by ambulance, They ask to which hospital the victim is transported. Later they receive or ask information from/at the hospital whether the person involved is still there, and whether he/she was admitted for in-patient treatment; this also includes a stay overnight for observation. If the person dies this is an unnatural death and it is always reported back to the police, but possibly not always to the police officer that reported the crash.
Poland	√		Yes	
Spain	√	√	No	Police gather on the spot information based on their own judgement. Also, hospital details are taken from ambulance personnel at the scene of the crash.
Sweden	√		Not systematically	With STRADA, there is great potential to compare the police assessment with the hospitals. However, major flaws in the police assessment are found.
Switzerland				
United States	√		No	Specific procedures can vary from state to state. In general the assessment is not checked against hospital or medical records.

Source: Survey among IRTAD countries (2010).

Box 2.1. **Case study: How the police classify the severity of a crash in the Netherlands**

The police use several severity indicators to assess the gravity of a crash. This information is used for road safety analysis.

0	Death on the spot
1	Death later on the same day (before admittance)
2,3,4,5	Death later, within 30 days
6	Hospitalized (in-patient, at least one night)
7	Transported to hospital, but not admitted (treated at Accident & Emergency)
8	Transported to hospital, but admittance unknown
9	Treated on the spot, transport to hospital (by ambulance or other means) not necessary
10.	Treated on the spot, transport to hospital unknown
11	Not injured, only available for drivers (of the vehicle of the victim, or another vehicle involved)

Of course, the severity assigned may not be always correct. In the Dutch linking studies, the police data are matched with the hospital records and compared with the ICD-derived MAIS (see definition in section 4). The estimated true number of serious injuries is now based on the MAIS and no longer on police severity.

It was found that:

Police 1-5:	63% can be matched with a hospital in-patient, of which 90% have a MAIS of 2 or more
Police 6:	67% can be matched with a hospital in-patient, of which 80% have a MAIS2+.
Police 7+8:	10% can be matched with a hospital in-patient, of which 70% have a MAIS2+.
Police 9+10:	5% can be matched with a hospital in-patient, of which 70% have a MAIS2+.
Police 11:	1.2% can be matched with a hospital in-patient, of which 70% have a MAIS2+.

The average numbers of matched cases concerned are:

Police 1-5:	260
Police 6:	7 240
Police 7+8:	1 330
Police 9+10:	900
Police 11:	390

Besides these 10 130 matches it is estimated that there are about 7 000 additional cases, available in the medical file only.

The Netherlands have recently adopted the new indicator of Serious injuries using medical information (MAIS2+, complete estimate).

2.3. Definitions of injuries in official statistics

In most countries, the police define the following levels of injury severity:

- Fatalities
- Injured persons, usually split in two categories:
  - Seriously injured.
  - Slightly injured.

The Netherlands include a third category “very slightly injured”.

2.3.1. Fatalities

In all IRTAD countries a road fatality is a person who died in a traffic crash, within 30 days of the crash. This is the internationally agreed definition, recommended by IRTAD, UNECE, WHO, etc.

2.3.2. Injured person

There are no commonly agreed definitions on injured road casualties. The International Transport Forum – Eurostat - UNECE has developed a classification (see box 2.2.), which is diversely applied in IRTAD countries (see table 2.2). In the United States, the KABCO scale is used [1].

Box 2.2. **Definitions of the ITF/EUROSTAT/UNECE glossary**

Injury crash (or injury accident)

Any accident involving at least one road vehicle in motion on a public road or private road to which the public has right of access, resulting in at least one injured or killed person. A suicide or an attempted suicide is not an accident but an incident caused by a deliberate act to injure oneself fatally. However, if a suicide or an attempted suicide causes injury to another road user, then the incident is regarded as an injury accident. Included are: collisions between road vehicles; between road vehicles and pedestrians; between road vehicles and animals or fixed obstacles and with one road vehicle alone. Included are collisions between road and rail vehicles. Multi-vehicle collisions are counted as only one accident provided that any successive collisions happen within a very short time period. Injury accidents exclude accidents incurring only material damage.

Person injured

Any person who as result of an injury accident was not killed immediately or not dying within 30 days, but sustained an injury, normally needing medical treatment, excluding attempted suicides. Persons with lesser wounds, such as minor cuts and bruises are not normally recorded as injured. An injured person is excluded if the competent authority declares the cause of the injury to be attempted suicide by that person, i.e. a deliberate act to injure oneself resulting in injury, but not in death.

Person seriously injured

Any person injured who was hospitalized for a period of more than 24 hours.

Person slightly injured

Any person injured excluding persons killed or seriously injured. Persons with lesser wounds, such as minor cuts and bruises are not normally recorded as injured.

Source: International Transport Forum, Eurostat, UNECE, *Illustrated Glossary for Transport Statistics, 4th edition, Final version 14/07/2009.*
http://internationaltransportforum.org/Pub/pdf/09GloStat.pdf

The following table presents the current ITF-EUROSTAT-UNECE classification of injured persons[2] and Table 2.3 summarises its application in IRTAD countries.

1. KABCO: *K = Killed; A = Incapacitating Injury; B = Non-Incapacitating Injury; C = Possible Injury; O=No Injury; and U=Injured, severity unknown*

2. Glossary for Transport Statistics. Eurostat – ITF - UNECE

Table 2.2. **Road injuries – ITF-EUROSTAT- UNECE definition and application in IRTAD countries**

	Injured					**Seriously Injured**			**Slightly Injured**
Criteria / Country	**All injured road users (including death >30 days)**	**Normally needing medical treatment**	**lesser wound, such as minor cuts and bruises**	**Attempted suicides**	**Inabiliy to work**	**Hospitali-sation**	**Hospitalisa -tion or serious injuries**	**Inability to work**	**Other than seriously injured**
ITF-Eurostat-UNECE classification	Yes	Yes	No	No	-	> 24 hours	-	-	Yes
Australia	-	-	-	-	-	Admitted to hospital			-
Austria	Yes	Yes	No	No				Inability to work or health problem > 24 days	
Belgium	Yes	Yes	No	Yes		> 24 hours			Yes
Canada	Yes	Yes	No	No		> 24 hours			Yes
Czech Republic	Yes	Yes	Yes	Yes			Yes		Yes
Denmark	Yes	Yes	No	No			Yes		Yes
Finland	Yes		No	Yes		> 1 day			
France	Yes					> 24 hours			Yes
Germany	Yes	Yes	Yes	No		> 24 hours			Yes
Great Britain	Yes	Yes	Yes	Yes			Yes		Yes
Greece	Yes		Yes	No		> 24 hours			
Hungary	Yes	Yes	No	No			Yes		Yes
Iceland	Yes								
Ireland	Yes	Yes	No	No			Yes		Yes
Israel	Yes	Yes	Yes			> 24 hours			
Italy	Yes								
Japan	Yes	Yes	No	No					Yes
Netherlands	Yes	Yes	No	-	-	> 1 night	-	-	Yes
New Zealand	Yes		Yes	Yes			Yes		Yes
Poland	Yes	Yes	Yes	No	Yes	> 7 days			-
Slovenia	Yes	Yes	Yes			> 24 hours			Yes
Spain	Yes	Yes	Yes	Yes		> 24 hours			Yes
Sweden	Yes	Yes	No	No					Yes
Switzerland	Yes	Yes						> 24 hours	Yes
United States	Yes								

Source: annual survey among IRTAD members.

The following table highlights complementary information that was given for the survey conducted by the IRTAD Group for this report (April 2010).

Table 2.3. **Definition of slight and seriously injured – variations from the ITF-EUROSTAT-UNECE classification**

	Injured person
Austria	**Seriously injured:** person suffering an injury which entails an inability to work or personal difficulty for more than 24 days. Personal difficulty: injury or health problem which affects an important organ, or results in health handicap or with an uncertain healing process or leading to the fear of "additional effects". **Slightly injured:** person injured but not seriously injured.
France	**Injured**: A person requiring medical treatment. **Seriously injured**: persons who are hospitalized. **Slightly injured**: persons who are not hospitalized (treated in the emergency departments). Before 2005, the definition was: -seriously injured: injured persons requiring 6 or more days of hospital stay. -slightly injured: injured persons requiring less than 6 days of hospital stay (including outpatients).
Hungary	**Seriously injured:** i.e. bone fracture, dislocation of a joint, injury of a cavity of the body (skull, chest, abdomen) or injury of vessels, tendon or nerves, which take longer than 8 days to heal. **Slightly injured:** any person who suffered slight injury and whose wound, sprain, or bruise recovers within 8 days of the crash.
Netherlands	Someone having sustained injury due to a road traffic crash with at least one moving vehicle involved on a public road in The Netherlands. The vehicle is not necessarily a motor vehicle but can also be a bicycle. For crashes among pedestrians (including in-line skaters and skate boards) there is no information. The injury varies from: **Very slight** (bandage/plaster, treated on the scene by local help/bystanders/first aid assistant. NO transport to hospital/Emergency Unit). **Slight** (treated initially by GP, Ambulance staff, medical team. Transport to Hospital/Emergency unit). **Hospitalized** (stabilised by Ambulance staff, medical team and transported to hospital/Emergency unit or brought to the hospital by other means and admitted for at least one night). **Fatal** (death upon arrival of emergency responders / died during stabilisation / died while on transport to hospital/Emergency unit, died from complications due to crash within 30 days).
Poland	**Serious injury**: serious disability, serious incurable illness or a long term illness actually endangering life, permanent mental illness, complete or a significant loss of ability to work or a permanent disfigurement of the body as well as injuries such as. Fractures, damage of the internal organs, serious cut or irregular wounds. **Slight injury**: loss of health other than described above, disturbing functions of the body organ or having health dysfunctions for a period not longer than 7 days, diagnosed by the doctor.
Sweden	**Seriously injured:** Any person who suffered fracture, crush injury, laceration, severe cuts, concussion or internal injury. In addition, a person is defined as seriously injured if the injuries are expected to lead to hospitalization. **Slightly injured:** Other personal injury. Assessment of whether an injury is serious or slight is conducted by the police on site at the time of the crash.
United Kingdom	Casualties are sub-divided into killed, seriously injured and slightly injured – see below. **Killed** - Human casualties who sustained injuries which caused death less than 30 days (before 1954, about two months) after the crash. Confirmed suicides are excluded. **Serious injury**: An injury for which a person is detained in hospital as an "in patient", or any of the following injuries whether or not they are detained in hospital: fractures, concussion, internal injuries, crushings, burns (excluding friction burns), severe cuts, severe general shock requiring medical treatment and injuries causing death 30 or more days after the crash. **Slight injury**: An injury of a minor character such as a sprain (including neck whiplash injury), bruise or cut which are not judged to be serious, or slight shock requiring roadside attention. This definition includes injuries not requiring medical treatment. An injured casualty is recorded as seriously or slightly injured by the police on the basis of information available within a short time of the crash. This generally will not reflect the results of a medical examination, but may be influenced according to whether the casualty is hospitalised or not. Hospitalisation procedures vary.
United States	State crash reports generally use the KABCO scale (K = Killed; A = Incapacitating Injury; B = Non-Incapacitating Injury; C = Possible Injury; O=No Injury; and U=Injured, severity unknown.

Source: 2010 survey.

2.3.2.1. Seriously injured

Criteria used in police record and official statistics to classify the severity of a crash vary from country to country and include:

- **The length of hospitalization**: in many countries a person seriously injured is a person hospitalized, other than for observation, for more than 24 hours (7 days in Poland)
- **The type of injuries**. In some countries, "seriously injured" is based on some specific types of injury. For example, in Sweden, a person seriously injured is defined as "Any person who obtained fracture, crush injury, laceration, severe cutting injury, concussion or internal injury". In addition, a person is defined as seriously injured if the injuries are expected to lead to hospitalization.
- **The inability to work** (Austria, Poland, Switzerland).
- **The length of recovery** (Hungary, Japan). In Hungary, an injury is serious if it requires treatment for more than 8 days. In Japan, an injury is serious if the predicted length of recovery is more than 30 days.
- **Long term disability** (Poland). This is however rarely mentioned as a criterion for determining the severity of an injury.

2.3.2.2. Slightly injured

In most countries, slightly injured persons include persons injured in a crash but not seriously injured.

Box 2.3. **Classification of road injuries in Switzerland**

In Switzerland, BfU (the Bureau for accident prevention) calculates once a year the whole extent of road traffic crashes based on insurance data. The levels of injury are defined as:

- Fatal injured.
- Serious injury (length of stay in hospital: 7 or more days).
- Medium serious injury (length of stay in hospital : 1 to 6 days).
- Slight injury (ambulant medical treatment).
- One special category is "disabled". This category is not only defined by medical criteria but by some criteria of the loss of working ability.

2.4. Determination of injury levels in hospitals

In most countries, hospitals do not define levels of injuries as such, but use the International Classification of Diseases (ICD9 or ICD 10), which is derived from the medical diagnosis, to describe the injuries.

In some countries, injury severity is defined based on indices such the Abbreviated Injury Scale (AIS), the Maximum Abbreviated Injury Scale (MAIS) and the Injury Severity Score (ISS), which can be derived from the International Classification of Diseases.

In most countries, each hospital has its own database/ registers. Only a few countries have a national database of injury severity.

Box 2.4. **Definition of injury classification and severity indices**

The International Classification of Diseases and related Health Problems (ICD)

The ICD is published by the World Health Organisation[3]. It provides codes to classify diseases as well as signs, symptoms and external causes of injury or disease. Every health condition can be assigned to a unique category and given a code, of up to six characters. In addition to enabling the storage and retrieval of diagnostic information for clinical, epidemiological and quality purposes, these records also provide the basis for the compilation of national mortality and morbidity statistics by WHO Member States. The ICD is revised periodically and is currently in its tenth edition (ICD 10). The 9th edition is still widely used (ICD9).

Causes of accidents are classified. Traffic injuries have a specific code in the section "external cause", as well as codes to describe the injury.

Abbreviated Injury Scale (AIS)

The AIS – published by the Association for the Advancement of Automotive Medicine – is an internationally agreed tool to describe the severity of injury for each of nine regions of the body: 1 Minor, 2 Moderate, 3 Serious, 4 Severe, 5 Critical, 6 Unsurvivable. The regions are 1 Head, 2 Face, 3 Neck, 4 Thorax, 5 Abdomen, 6 Spine, 7 Upper Extremity, 8 Lower Extremity, 9 External and other.

AIS does not reflect the combined effects of multiple injuries. It was initially developed for crash investigation purposes to provide researchers with a simple numerical method for ranking and comparing injuries by severity, and to standardize the terminology used to describe injuries.

It is possible to convert ICD9 or 10 codes into AIS. More details are given below.

Maximum Abbreviated Injury Scale (MAIS)

MAIS is the maximum of the AIS scores for each region of the body. It is used to assess the overall severity of the various injuries.

Injury Severity Score (ISS)

The Injury Severity Scoring is a process by which complex and variable patient data are reduced to a single number. The Injury Severity Score (ISS) is an anatomical scoring system that provides an overall score for patients with multiple injuries. Each injury is assigned an AIS and is allocated to one of six body regions (Head, Face, Chest, Abdomen, Extremities (including Pelvis), External). Only the highest AIS score in each body region is used. The three most severely injured body regions have their score squared and added together to produce the ISS score. ISS was developed to predict mortality.

ICISS is the ICD-derived injury severity score (ICISS)

New Injury Severity Score (NISS)

The new injury severity score (NISS) is computed as the simple sum of squares of the three most severe AIS injuries, regardless of body region.

2.4.1. International Classification of Diseases

The World Health Organisation (WHO) requires all vital registration systems reporting to it to use ICD10 or versions of it. Whereas most OECD countries fill in to the 3rd digit, they do not all do so to the 4th digit.

Some countries highlighted the incompleteness of the external codes of ICD which enable identification of road traffic accidents. As an example, in Switzerland external causes of injuries are not coded any more in the ICD10 system. A systematic coding by a medical doctor of the external causes of

3. More information available on the WHO website: www.who.int/classifications/icd/en/

injuries would be useful to identify traffic injuries more easily. This is currently not always the case. In addition, there are rarely national databases gathering data from all hospitals.

The table below summarises the information provided by IRTAD members regarding the use of ICD in their respective countries to describe traffic injuries as well as experience in converting information from ICD9 system to ICD10.

Table 2.4. **Use of the International Classification of Diseases in IRTAD countries**

	International Classification of Disease – comments
Australia	ICD 10
Austria	ICD 10. Austria has worked on a conversion programme from ICD 9 to ICD 10
Belgium	ICD 9
Czech Rep	ICD 10
Denmark	ICD 10
Finland	ICD 10
France	ICD 9 until 1999 for mortality data. ICD 10 is now used in all hospitals. Incomplete codes relating to external causes.
Germany	ICD 10.
Hungary	ICD10 is the basis for overall injuries - countrywide
Ireland	Project underway.
Israel	ICD9
Japan	ICD 10 used by forensic doctors
Lithuania	ICD 10
Netherlands	ICD 9 ICD 10 since 2010
Poland	ICD 10
Spain	ICD 9. It is planned to shift to ICD 10. (ICD 10 already used in mortality and forensic registers)
Sweden	ICD 10
Switzerland	ICD 10 since 2009. External causes (such as traffic injuries) are not coded anymore.
United Kingdom	ICD 10
United States	ICD 9 used in hospital billing. ICD 10 used by state on death certificates. No immediate plan to convert hospital coding to ICD10

Table 2.5. **External cause codes for road traffic injuries in the International Classification of Diseases – ICD10**

V01-V09	Pedestrian injured in transport accident
V10-V19	Pedal cyclist injured in transport accident
V20-V29	Motorcycle rider injured in transport accident
V30-V39	Occupant of three –wheeled motor vehicle injured in transport accident
V40-V49	Car occupant injured in transport accident
V50-V59	Occupant of pickup truck or van injured in transport accident
V60-V69	Occupant of heavy transport vehicle injured in transport accident
V70-V79	Bus occupant injured in transport accident
V80-V89	Other land transport accidents
V90-V94	Water transport accidents
V95-V97	Air and space transport accidents
V98-V99	Other and unspecified transport accidents

Source: WHO (2010).

2.4.2. Severity indexes: AIS, MAIS, ISS and other indexes

Many countries use the abbreviated Injury Scale (AIS), the Maximum Abbreviated Injury Scale (MAIS) or the Injury Severity Scores.

Determining AIS

There are two ways to determine the AIS:

- AIS may be directly coded by trained medical staff, on the basis of the available medical files regarding the injuries of the patient.
- It can also be derived from the ICD9 or 10 classification. The two main software packages to map AIS from ICD9-10 have been developed by the European Centre for Injury Prevention (University of Navarra, Spain) and the Johns Hopkins University.

Table 2.6. **Methods to determine AIS in IRTAD countries**

Czech Rep	Derived from the diagnosis expressed in ICD-10 classification.
Denmark	Determined by medical doctors.
France	The medical diagnosis is directly coded into the Abbreviated Injury Scale (which includes the AIS severity score). This is done by a trained physician; the diagnosis is the result of all text injury descriptions from all hospital departments the person has attended.
Japan	Determined by medical doctors. The Japan Association for the Surgery of Trauma periodically trained medical doctors and other relevant staff on AIS coding.
Netherlands	AIS is derived from the ICD-9 by use of ICDmap90 (Johns Hopkins 2002).
Spain	Software can convert ICD9-CM codes to AIS: - ICDMAP (Johns Hopkins University). - ICDPIC: (Boston College Department of Economics). This is a STATA module to provide methods for translating ICD9-CM diagnosis codes into standard injury categories and/or scores.
United Kingdom	Mapping from ICD-10 codes using coding developed by University of Navarra (European Centre for Injury Prevention, University of Navarra, Algorithm to transform ICD-10 codes AIS and ISS, version 1 for SPSS. Pamplona, Spain 2006).
United States	AIS derived either from ICD-9 codes provided by hospitals, or, in the case of NASS-CDS, by forensic analysts reading the case file.

Box 2.5. **How to determine MAIS and ISS: real case example of a Woman, Pedestrian, 77 years, knocked down by a truck**

Region injury	**Injury description**	**ICD9-CM**	**AIS**	**Square of top 3**
Chest	Multiple and unspecified intrathoracic organs, without mention of open wound into cavity, Multiple intrathoracic organs	862.8	**6**	36
Abdomen	Multiple pelvic fractures with disruption of pelvic circle	808.43	4	16
Extremity	Fracture of other and unspecified parts of femur, Lower end, unspecified part	821.20	3	9
Extremity	Fracture of tibia and fibula	823.92	2	
Extremity	Open wound of hip and thigh	890.1	1	
Extremity	Other, multiple, and ill-defined fractures of lower limb, open	827.1	2	
Extremity	Crushing injury of lower limb , Knee and lower leg	928.10	2	
Chest	Pneumothorax without mention of open wound into thorax	860.0	3	
	MAIS			**6**
	Injury Severity Score (= some of the square of top 3 or 75 if one of the AIS =6)			**75**

Using AIS, MAIS, ISS and other severity indexes.

MAIS and ISS are directly calculated from the AIS (see box 2.4).

AIS/MAIS are used for research purposes, to assess hospital costs, and to monitor injury severity. The US also use them to assess potential benefits from new safety measures. Only Sweden and the Netherlands use MAIS to produce information in official national statistics on road injuries.

The ISS is a unique figure synthesizing the severity of injuries of the different parts of the body. It is usually used to identify the most critical cases.

Other indices include TRISS (combined trauma score and ISS), the National Advisory Committee for Aeronautics (NACA) index, the Organ Injury Scale (OIS) and the Glasgow Coma Scale[4].

Table 2.7. **Use of AIS, MAIS, ISS and other severity indices in IRTAD countries**

	Use of AIS / MAIS	Use of ISS	Use of other severity indexes
Australia	No	Yes To select high threat to life serious injury cases	
Austria	Yes, in some projects. Not in official statistics.		Yes: Police definitions of severity of injury; these definitions heavily rely on the penal code (e.g. serious injuries are stated as in the Austrian criminal code)
Belgium	No	No	
Canada	No	No	
Czech Rep	Yes MAIS is used for injury monitoring	Yes For the purpose of road crash statistics injuries with ISS > 16 are considered as serious	TRISS (trauma score and ISS) OIS NACA index
Denmark	Yes, in one region only. AIS used for local registration.		
Finland	No	No	
France	Yes, in the Rhône road trauma registry. It is used for research. Serious casualties = MAIS3+ Not used in official statistics. Version AIS 1990	Rhone Registry: Yes Not used in official data ISS and New ISS used for research	Injury Impairment Scale (ISS) Used for research
Germany	Yes Not on a regular or national basis. AIS is registered in the Trauma Registry by DGU and GIDAS (German In-depth Accident Study) Version AIS 1998	Yes, ISS and New ISS	Glasgow Coma Scale NACA index Not on a regular national basis. See AIS MAIS
Hungary	No	Yes in some trauma departments (not all) Only used in some emergency hospital departments for scientific statistical purpose	Trauma score (TS) at the National Ambulance Service but only for multiple casualty accidents (5 or more victims concurrently at the scene)
Ireland	No, project underway		
Israel	No	Yes	Yes, duration of hospitalisation ICU Duration - Person seriously injured: A person who was hospitalized as a result of a crash for a period of 24 hours or more, not for observation only ICU - Stay in ICU (yes/ no)

4. The Glasgow Coma Scale (GCS) is scored between 3 and 15, 3 being the worst, and 15 the best. It is composed of three parameters: Best Eye Response, Best Verbal Response, Best Motor Response See information on www.trauma.org.

Japan	Yes. In some hospitals which have voluntarily registered with the Trauma Data Bank. ITARDA uses it for in-depth depth investigation Version AIS 1998	Yes	TRISS (Trauma Score – Injury Severity Score) Japan Coma Scale Glasgow Coma Scale, etc. Used for Estimation and/or evaluation of individual degree of injury (for head/brain trauma etc
Lithuania	No	Yes	ICF in near future
Netherlands	Yes. Since 2009, MAIS2+ cases are reported to authorities for monitoring measures, policies etc. For the first time, MAIS2+ numbers for 1993 – 2009 were reported to the Parliament in April 2011 Version AIS 1990, plan to migrate to 2005 version	Available but no official use	No
Poland	No	-	
Portugal	No	-	
Spain	Yes Used for injury monitoring and for studying specific types of injuries Version AIS 1990 and 1998	Yes Same software as for AIS	No
Sweden	Yes MAIS used to measure the most serious injuries AIS Version 2005	Yes As a measure of person's total injury	Yes There is a criteria for long term impairment Functional Capacity Index (FCI) was introduced in STRADA in October 2010. The FCI is specified for every injury. Used to monitor national target [5]
Switzerland	No	No	Yes In some cases, use of a proxy indicator coded from the length of stay in hospital: Slight injury: outpatient care Medium severe injury: inpatient 1-6 days Serious injury: inpatient 7 or more days
United Kingdom	Yes, In some regions AIS version 1998	Not currently	
United States	Yes Used to calculate injury severity comparisons, societal cost of injuries, benefits from proposed legislation AIS version 1990-98 update and 2005-08 update	Yes ISS can be calculated if all MAIS by body region are available, and is used at the discretion of the researcher	Yes KABCO is typically used on crash reports Barrel Injury Diagnosis Matrix is used at the discretion of the researcher when ICD-9 codes are available Two important databases, FARS and NASS-GES, only have KABCO but no other severity indexes, like ICD-9 or MAIS. KABCO is used in many publications. "Serious injury" is often defined as K or A injury

5. In May 2009 the Swedish Parliament set itself the objective to cut the number of serious injuries leading to long term impairment by 25%. This dimension is needed for follow up reasons.

Box 2.6. **STRADA – Swedish Traffic Accident Data Acquisition**

STRADA is a national information system collecting data on injuries and crashes in the entire road transport system. STRADA is based on information from the police as well as the hospitals.

STRADA was created in close collaboration with all parties concerned. 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. When hospital data are included there is also a decrease in the number of unrecorded cases, since the police have limited knowledge about some road traffic crashes (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. 18 counties report to STRADA on a complete or partial basis. The remaining three counties are yet to join.

STRADA is a GIS-based system, which means that it largely makes use of a mapping tool both to enter accidents and perform analysis. It is the police and medical staff themselves that enter data into STRADA which means a minimum risk of data distortion. From both sources information is collected regarding where, when, and how the accident happened and who was involved. But only the police collect information on the licence numbers of the vehicles, which is used to get technical data on vehicles involved in accidents. Hospitals provide detailed information on each injury: both the AIS- and ICD10-code as well as the FCI-code.

By accessing STRADA's web-based system for extraction of data or by requesting information from the Swedish Transport Agency, municipalities, researchers, *etc*. can make use of the information. Since 2003 the official statistics of road traffic injuries are based on data extracted from STRADA. 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.

2.5. Conclusion

This chapter shows that there are several distinct ways in which to judge the severity of a crash, as well as to define its seriousness.

Official statistics on road injuries are usually obtained from police records. There is rarely systematic follow-up from hospital data based on the medical diagnoses and longer-term impairment of patients.

Hospitals keep files on the injury severity and causes of trauma of their patients. The WHO requires all hospitals to codify and classify these injuries, through ICD classification. Hospitals also work with severity indexes (AIS, MAIS, ISS), which can be coded either directly, or derived, from the ICD. In most countries, this information is used locally to follow the patient, or for the hospital management~~,~~ or used in research. Except in the Netherlands and Sweden, there is no central database for this type of information. In Sweden and the Netherlands, MAIS data are reported in statistics to provide more information on serious injuries. Sweden set up an indicator of Long Term Impairment, on which targets to reduce road injures are based.

Having an internationally agreed definition of "serious" injuries will help the safety research community to better understand the consequences of road crashes and to monitor progress. Further consideration could be given to work on a common definition of "seriously injured", based on the severity indexes determined by medical staff. This issue is discussed further in Chapter 6.

References

Derriks H and Mak P. Underreporting of Road Traffic Casualties. Organisation for Economic Co-operation and Development. IRTAD 2007. www.who.int/roadsafety/publications/irtad_underreporting.pdf [acceded 6/4/2010]

ITF, Eurostat, UNECE (2009), Glossary for Transport Statistics, 4th edition, www.internationaltransportforum.org/Pub/pdf/09GloStat.pdf

WHO (2010), Data systems: road safety manual for decision-makers and practitioners, WHO, Geneva.

CHAPTER 3
Casualty data: use of police, hospital and other sources

Abstract

Chapter 3 looks at the main sources of information, police and hospital data, (including admissions, and those treated only in the emergency room) and at other data sources.

3.1. Introduction

In most countries road traffic collisions are recorded. Such a recording system is vital in the understanding of road "unsafety" and the development of effective countermeasures (Derriks & Mak, 2007), (Rosman, 2001), (Peden *et al*. 2004). Comprehensive collision and road safety performance data collection and analysis are critical requirements for many purposes, as described in Chapter 1.

In the majority, if not all, countries rely on police data as the main source of collision statistics. Some countries supplement this collision information with data collected from hospitals (Ward, Lyons, & Thoreau, 2006). Hospital admission is a larger dataset than fatality (police) data. Police data presents an important but only one aspect of the overall road casualty picture. The use of both hospital and police data would make statistical tests more powerful and enable small effects of road safety measures to be detected (Haque, 2009).

There are a number of other sources used for road safety throughout the world. There are considerable differences in what data are collected, their quality and the extent to which they are used to track trends and evaluate effective countermeasures to improve road safety in the respective countries. Other sources include:

- Mortality registers.
- Forensic agencies.
- Emergency ambulance services.
- Fire services.
- Surveys.
- Insurance related records.
- In depth studies.

Based on the responses to the survey of IRTAD members undertaken for this report, this chapter reviews the details, coverage and quality of each data source.

3.2. Why use several sources of data?

While in nearly all countries, police data are the main source of information for safety statistics, there is growing awareness on the need to also collect and analyse other sources of collision data for road safety analysis. Police data provide detailed information about crash circumstances, location, and vehicles involved to for a number of purposes, but they are not complete or perfect sources.

- Police data suffers from under reporting (Lopez et al., 2000), when collisions are not reported to the police, or when some collisions are reported they may not be recorded.

- Police reports, when the degree of injury is recorded, may include incorrect information (Tsui, et al. 2009).

In the first instance the police may not record collisions due to definitional issues. Police generally report on collisions that occur on a public road, many collisions occur on public place (as distinct from road) and many in private locations. Another issue is the definition of a road fatality, which in most countries, is only included in the data if the person dies within thirty days of the collision.

The definition of what an injury is, and particularly what is a serious injury is another reason why police data is somewhat problematic. The definition of a serious injury, used in many countries, is so broad that it can include very serious injuries resulting in brain damage or spinal cord injury (paralysis) at one end and also relate to minor fractures and sprains (Amoros et al., 2006). It is not within the proficiency of an unqualified police officer to correctly diagnose the medical condition of a casualty, especially when that diagnosis can take place on the side of the road and often in a tense and traumatic situation as at collision scene.

It is therefore desirable to use complementary sources to build a balanced and comprehensive picture of the nature and extent of road accidents. Other datasets can be useful both as a check on the quality and completeness of police data and in providing information which is not collected by the police, for example relating to more detailed medical consequences of road crashes.

However, one must be aware of the limitations (in terms of scope, quality, timeliness, etc.) of each data set.

3.3. How to use different data sources: Comparisons and linking

Additional data sources can be analysed individually and also by combining or linking databases. Comparisons between data sources can provide further information on overall numbers but analysis of the data, for example by age or road user group, enables comparisons of distributions within each dataset, possibly revealing any bias or under-representation of certain groups. However, making comparisons between police and other data sources is not straightforward, as there are often differences in definitions. There may also be changes in data collection and recording practices which can affect comparisons of trends over time.

These other sources of data can provide useful additional information on accidents and their consequences but are not always available at the national level and have their own quality issues. These issues impact on the extent to which the data can be used for linkage, to measure levels of under-reporting or to estimate the "real" number of casualties. Linking records can provide a fuller picture, for example by relating medical consequences to the circumstances of the crash or vehicle. Linking is discussed further in Chapter 4.

Box 3.1. Using several sources of data - Experiences of the United States and the Netherlands

Most linkage of data is between two sources, for example police/hospital data police/mortality data but a few countries bring together a number of datasets to provide a more complete picture of road safety issues, to identify prevention factors and improve road safety decision making.

In the ***United States*** the US Fatality Analysis Reporting System (FARS) contains information on all fatal traffic accidents. The data are compiled from a range of sources including:

- Police reports
- Vehicle registration files
- Driver Licensing files
- Highway Department data
- Vital statistics
- Death certificates
- Coroner/medical examiner reports
- Hospital medical reports
- Emergency Medical Service Reports

The CRASH Outcomes Data Evaluation System (CODES) encourages states to link, using probabilistic methodology, motor vehicle crash data, with other traffic records and injury outcome records (from emergency services, hospitals, death records) and to identify medical and financial consequences.

Further information from: www.nhtsa.gov/NCSA

In ***the Netherlands***, a project is running to improve registration of road traffic accidents through the incorporation of various sources emergency hospital and ambulance services, fire services, forensic services, mortality records, and information from insurance companies.

Formal linkage of police data to other sources is limited. Only two countries use mortality registers, for example, the Netherlands use this to adjust for under-reporting their total number of road crash fatalities derived from police data. A further three countries link with data from forensic agencies, but only one, the United States currently links ambulance service data.

Box 3.2. Use of police and hospital data in Sweden

In Sweden, road safety is defined as the absence of fatalities and serious injuries. Crashes can be accepted but only provided that they do not lead to serious injuries.

The Swedish police are not well equipped to determine the level of injury of a traffic casualty. In-depth traffic safety requires medical data on the type and severity of the injuries.

In 2009, the City of Göteborg developed a road safety action programme for the period 2010-2020. This programme is based on a comprehensive analysis of fatalities and injuries reported by both the health care authorities and the police. Police reporting includes detailed information on where and how the crashes occur. Health data provide a qualified assessment of the degree of injury and the victim´s own description of the crash.

During the eight years 2000-2007 covered by the analysis, 21 580 people were reported as being injured in road crashes in Göteborg. Of these:

- 13 325 sought care at an emergency hospital
- 17 915 were known to police
- 9 660 victims were found in both police and hospital records
- 8 255 were reported by the police but not by the emergency hospitals. It can be assumed that most of these only had slight injuries.
- 6 665 were registered in hospital but not recorded by the police

In addition, the emergency hospitals reported 2 891 pedestrians injured in falls with no vehicle involved. These are by definition not considered as traffic casualties, even if the crash is due to deficiencies or obstacles in the road environment.

Injury severity, as reported by hospitals
Göteborg 2000-2007

	Injured persons		Pedestrians injured in falls	
Degree of injury	Number	%	Number	%
Fatalities	90	1 %	0	0 %
Seriously injured (ISS>8)	948	7 %	583	20 %
Moderately injured (ISS 4–8)	1 941	15 %	1 107	38 %
Slightly injured (ISS 1–3)	10 346	78 %	1 201	42 %
Total	13 325	100 %	2 891	100 %

The analysis of police and health data allowed the City of Göteborg to formulate six strategic areas for road safety.

1. Ensure that road safety is a priority in urban development
2. Give priority to vulnerable road users
3. Give reasons for increased accountability in service
4. Safer vehicles and safer transport
5. Designing streets in order to practically eliminate speeding
6. Safety awareness among children and young people

The objective was set to reduce the number of fatalities and injuries by 75 % over the period 2010-2020.

The success of Göteborg in moving towards its ambitious road safety goal is due to information on crashes being provided by both the police and the hospitals. This has led to an awareness of crashes that were not previously reported as well as to an advanced knowledge of the injuries the victims receive. Thus, the quality of assessments has improved considerably and their conclusions lead to better prioritising and more adequate measures.

3.4. Sources of information and data on fatal crashes casualties – Results of the survey

3.4.1. Police data

As expected all countries surveyed collected their primary data for fatal collisions from the police. The majority collected the data at a national level, whereby larger countries; Australia, Canada and the United States did so at state or regional level.

3.4.2. Hospital data

Of the 23 responding countries, 15 used hospital data as a source for fatal collisions. Australia, Finland, Germany, Hungary, Ireland, Israel, Japan and Norway indicated that they did not use or have access to such data. Of the 15 countries who responded positively to this question 11 indicated they had access to this data at a national level. Canada and the United States use the data at regional or state level. Several countries indicated that people killed on the spot are not registered in the hospital registers.

3.4.3. Emergency Hospital data

Under the heading of "emergency (hospital)", 14 responding countries indicated it was a resource for fatal collision data. Canada, Spain and the United States indicated that such access and use was at local regional or state level. The remaining who responded positively to the question, Denmark, Israel, Japan, Netherlands, Sweden, and Switzerland did so at national level.

3.4.4. Other data sources

As the table below shows mortality registers are also a very common source of information, mentioned by 21 out of 23 countries responding to the survey. Forensic reports provided information in 12 countries, emergency ambulance services also record fatalities in 7 countries but generally only at local levels. Fire-fighters have information on fatal crashes in only 4 out of the 23 responding countries. Insurance organisations are an additional source of information in Finland, Israel, the Netherlands and Switzerland.

Table 3.1. **Sources available for fatal crashes**

Number of countries	Records fatalities (based on 23 responses)
Police data	23
Mortality registers	21
Hospital	15
Emergency hospital	14
Forensic reports	12
Ambulance services	7
Insurance records	5
Other	5
Fire services	4

3.5. Sources of information and data on serious injury and slight Injury - Results of the survey

3.5.1. Police data

As with data collected on fatal collisions, all countries surveyed collected their primary data for serious injury collisions from the police. The majority collected the data at a national level, though larger countries such as Australia, Canada and the United States of America did so at state or regional level.

In the case of data collection for slightly injured, unlike fatal and serious injury data, not all countries collect information. Countries that do not collect data from the police sources are Australia and Lithuania. The majority collected the data at a national level, but once again, larger countries such as Australia, Canada and the United States of America did so at state or regional level.

3.5.2. Hospital admissions data

Of the 23 responding countries, only four did not use hospitals as a data source for serious injury collisions. Ireland, Israel, Japan and Norway indicated that they did not use or have access to such data. Of the 17 countries who responded positively to this question 13 indicated they had access to this data at a national level. Canada and the United States of America use the data at regional or state level; while Germany, who indicated that their data is not exhaustive on a national level and only used in temporary in-depth studies. Denmark indicated that the hospital source for serious injury collision information was “severity only recorded in some systems”.

Of the 23 responding countries, 6 did not use hospitals as a data source for slightly injury collisions. Surprisingly three of these were not the same three countries that did not access hospital data for serious injuries. In the case of serious injury the countries were; Israel, Japan, Ireland and Norway. In the case of slight injuries the countries not using hospital data were Australia, Germany, Netherlands, Norway, Ireland, Spain and Switzerland. Of the 10 countries who responded positively to this question 3 indicated they had access to this data at a national level. Canada and the United States of America use the data at regional or state level; while Denmark indicated that the hospital source for slight injury collision information was “severity only recorded in some systems”.

Box 3.3. **Finland: PRONTO: fire and rescue service data to better understand crash injuries**

Pronto is the data system of the rescue services. It includes information on traffic fatalities as well as seriously and slight injuries.

Rescue services are not called to all crashes, therefore the database is not exhaustive; however it covers quite well the most severe crashes. It provides detailed information on the location.

One advantage of the PRONTO system is that the data are easily and quickly available. Because the data exclude personal information, it is easier to access.

3.5.3. Emergency Hospital data

Under the heading of "emergency (hospital)", 14 responding countries indicated that they used this as a resource for serious injury data. Australia, Canada, Spain, Finland France and the United States of America indicated that such access and use was at local, regional or state level. The remaining who responded positively to the question, Denmark, Israel, Japan, Netherlands, Sweden, and Switzerland did so at national level.

Under the heading of "emergency (hospital)", the same responding countries indicated that they used this as a resource for serious injury data used this resource for slight injuries as well. Australia, Canada, Spain, Finland France and the United States of America indicated that such access and use was at local, regional or state level. The remaining who responded positively to the question, Denmark, Israel, Japan, Netherlands, Sweden, and Switzerland did so at national level.

3.5.4. Other data sources

In all countries responding to the survey, the police are the main source of information for road crash injuries and most countries have some information on injured people from hospitals. As the table above shows the most common other sources of information mentioned by those responding to the survey were emergency services - ambulance (8 countries), fire-fighters (5 countries). Insurance organisations are an additional source of information on injuries in Finland, Israel, the Netherlands, Switzerland and the United Kingdom.

Table 3.2. **Sources available by type of injury (based on 23 responses)**

Number of countries	Records serious injuries	Records slight injuries	Unknown
Police	23	21	
Hospital	19	17	
Emergency - Hospital	14	14	
Mortality registers	2	1	
Forensic reports	5	4	
Ambulance services	8	8	2
Fire services	5	5	
Insurance records	6	6	
Survey	2 (1)	2 (1)	
Other	5	4	

(1) Dutch survey ceased in 2005.

Source: survey among IRTAD Members (2010).

Another aspect of the IRTAD Sub Group on Linking Police and Hospital Data questionnaire was to gather information regarding the quality of collision data collected by the various sources identified. The headings under which the data was collected were:

- Relevant for analysis
- Completeness of the number of serious causalities
- Accuracy
- Compliance with international definitions
- Timeliness

Respondents were asked to rate the quality under each heading as:

- "Good", when data are collected reasonably consistently in all areas, are generally considered accurate (although there may be under-reporting problems), and are a key source of road safety evidence.

- "Sufficient", when data have some issues e.g. not available for all areas, not providing enough information, but still have some useful features and help provide a fuller picture of road safety.

- "Poor", when data are possibly incomplete, inaccurate or not up-to-date, and so are not currently used.

In some cases respondents entered, "unknown", "not sure "or left the response blank. For the purpose of this analysis of the questionnaire all such responses are categorised as "unknown".

The responses are detailed in the following sections, for each source of data.

3.6. Use and limitation of the various data sources

3.6.1 Police data

In almost all countries, police are obliged to go on the scene of a crash where there is at least an injured person. The police officers usually complete a crash form (which can be either digital or on paper) on the scene of the crash, which they complete later when they return to their office, and adjust in some cases, when receiving information from the hospitals on the severity of a victim (see also section 2.2).

Quality of data – results of the survey

A large majority of the respondents (19 of the 23 respondents) indicated that police data were relevant for safety analysis. Police data was found to be "sufficient" by the remaining 4 respondents. When it came to "completeness of the number of serious casualties" only 6 of the 23 respondents indicated police data as good but 15 respondents stated it was sufficient. In the case of accuracy, 20 of 23 of those who responded, found police data either good (10) or sufficient (10). As expected there was a strong indications that police data is in compliance with international definitions with 19 of the 23 respondents indicated that the data was good or sufficient (with only 2 respondents for "sufficient"). Finally respondents were asked about the quality of collision data and the timeliness of it. Once again police data scored highest with 17 of the 21 respondents indicated it was good or sufficient.

General observations

As stated in the introduction, there is a lack of consistency over how police data are collected, what is included and what is excluded. Concerns were expressed by the respondents that in some cases there is substantial data missing regarding pedestrians and cyclists (Finland), the poor quality of original police data that requires validation by other agencies (Ireland), the availability of police data only at state level and not at federal (Australia), the variation of data within countries (United States of America) the exclusion of non public road collisions as well as those caused as a result of suicide

(Germany) and as highlighted (United Kingdom) the fact that collisions can occur but are not reported to the police.

Table 3.3. **Quality of police data – responses of the survey**

POLICE DATA (23 responses)					
	Relevant for analysis	Completeness of the number of serious casualties	Accuracy	Compliance with international definitions	Timeliness
Good	19	6	10	19	15
Sufficient	4	15	10	2	5
Poor		2	2	1	2
Unknown			1	1	1

3.6.2 Hospital admissions data

In most countries each person admitted in a hospital for a disease or following a crash is registered in the records of the hospital, where diagnoses, treatment and other information related to the disease / injuries are reported. In most countries, each hospital has its own database; and there is not often a consolidation of information at regional or national level.

Quality of data – results of the survey

When asked about the relevance for analysis for hospital data, 4 of the 21 respondents (two countries did not respond) stated it was poor while “sufficient” accounted for another 8 and 7 stated hospital data was “good” for analysis. Regarding the completeness of data, 11 of the 19 respondents stated that hospital data was “good” with another 6 respondents indicating it was “sufficient”. As far as accuracy is concerned, 17 of the 19 respondents indicated that hospital data was good or sufficient. Regarding compliance with international definitions, 14 of the 19 responses indicated that hospital data were good, although 4 respondents did indicate the hospital data was only sufficient for this requirement.

General observations

The main issues highlighted by the respondents regarding hospital data is the lack of connectivity between police and hospital data (Czech Republic, Denmark, Ireland), non availability of hospital data (Norway, Ireland) and particularly issues regarding the coding o injury type (Switzerland, Spain, Norway).

Many countries who reported the availability of hospital data indicate that it is based on a regional/local basis (Germany and United States). On the other hand there are positive moves to improve the data provided by this source (Krevin project in the Netherlands on Quality Improvement of the Registration of Road Traffic Accidents in The Netherlands).

Table 3.4. **Quality of the hospital data – responses of the survey**

HOSPITAL DATA (19 responses)					
	Relevant for analysis	Completeness of the number of serious casualties	Accuracy	Compliance with international definitions	Timeliness
Good	7	11	5	11	2
Sufficient	8	6	12	5	10
Poor	4	1	2	1	5
Unknown		1		2	2

3.6.3. Hospital emergency data

In most hospitals, emergency services departments have their own records, which may not be transferred to the main files of the hospital (e.g. when a victim dies in the emergency department). These records can however provide useful information on the type of injuries and its origins.

Only 12 countries commented on the use of data from emergency hospital departments.

Quality of data – results of the survey

For the four criteria (relevant for analysis, completeness, accuracy and timelines), most of the countries which responded were generally satisfied with the quality of the data.

Table 3.5. **Quality of the data from emergency services – responses of the survey**

EMERGENCY (Hospital) (12 responses)				
	Relevant for analysis	Completeness of the number of serious casualties	Accuracy	Timeliness
Good	7	5	4	4
Sufficient	3	5	5	5
Poor	2	2	3	0
Unknown	0	0	0	3

3.6.4. Mortality Registers

As already noted above, mortality registers are a common source of information for fatalities. They provide good information at the national level on overall numbers of road traffic fatalities but with no detail about the circumstances of the crash. The number of deaths is in many countries derived from registrations of death certificates completed by a doctor or coroner. Generally cause of death is coded using International classification of Diseases codes, allowing road transport accidents to be identified.

Table 3.6. **Quality of the data in the mortality registers – results from the survey (based on 21 responses)**

	Relevant for analysis	**Completeness**	**Accuracy**	**Timeliness**
Good	9	8	11	5
Sufficient	8		2	2
Poor	1	3	4	8
Unknown	3	10	4	6

In all 21 of the 23 countries responding to the survey indicated that they currently use information available from mortality registers as a source for fatal collision data. In addition two countries indicated they used this source for serious or slight injuries.

Most countries (19) that had such data indicate that it is available at the national level, and of these 10 also reported they had regional or local level data. In the United States and Canada such data were only available at the state or regional level.

Respondents were again asked to rate the quality of the data. Quality of the data varies from country to country but is generally considered to be good or sufficient in terms of relevance, completeness and accuracy. The definitions are not identical to those used by police for international comparisons, in particular all deaths will be included not just those within the 30 day cut-off. The main limitation is the timeliness of the records, with delays of 2-3 years in some countries.

Only 2 (Israel and Netherlands) countries use mortality registers to link to police data, for example, the Netherlands use this source to adjust for under-reporting of their total number of road crash fatalities derived from police data.

3.6.5. Forensic reports

Eleven countries mentioned use of forensic reports, these are generally based on coroners' or autopsy reports. All responding countries indicated that they used this as a source for fatalities and 5 countries also reported using the data for serious or slight injuries.

Table 3.7. **Quality of the data in the forensic reports (based on 12 responses)**

	Relevant for analysis	Completeness of the number of serious casualties	Accuracy	Timeliness
Good	5	3	4	1
Sufficient	3	4	4	5
Poor	3	1	2	3
Unknown	1	4	2	3

Data quality and coverage between countries varies considerably. In 4 countries the data are only available at local level and other respondents commented that information is not available for all cases or not easily accessible. However, 8 countries considered the data to be "good" or "sufficient" in relation to being relevant for analysis and accuracy although only 6 considered the data to be sufficiently timely.

Only three countries, France, Sweden and the United Kingdom, use the data for linkage. For example, in Great Britain local data on blood alcohol levels of road traffic fatalities is collected from coroners into a central database. It is then linked to police data as part of the process to estimate overall levels of drinking and driving.

3.6.6 Emergency ambulance data

Ambulance operators (which can be either publicly or privately owned) normally record each of their interventions when they are called to take care and pick up an injured person on the road.

Only 9 countries indicated that they currently have information available from Emergency Ambulance services and a further country, Israel, plans to use the data in the future. Two countries, Hungary and Switzerland, do not have data specific to road injuries. 6 responding countries indicated that they used this as a source for fatal, serious and slight injury collision data. In addition Portugal indicated they used this source for data in the case of serious and slight injury collisions but not for fatalities.

Table 3.8. **Quality of the data in the emergency ambulance reports (based on 10 responses)**

	Relevant for analysis	Completeness of the number of serious casualties	Accuracy	Timeliness
Good	1	2	2	
Sufficient	2	1	2	3
Poor	5	2	2	2
Unknown	2	5	4	5

In most countries, these data are for the moment not very useful. Most countries who reported the availability of such data indicate that it is based on a local or regional basis, only Portugal had a national database. The data were generally rated poorly or of unknown quality for relevance for analysis, accuracy and timeliness. However, there are positive moves to improve the data provided by this source in some countries, for example, in the Netherlands as part of a project to improve the registration of traffic accidents. The United States is the only country currently using the data for linkage, although this varies by State, and there is a national initiative under way to standardise the data.

3.6.7. Fire service

Information from Fire services are available in 5 countries; the Netherlands, Portugal, United Kingdom have national data for fatal, serious and slight injury crashes, Sweden have similar data but only at the local level, and France only for serious injuries at regional level.

These data are for the moment not considered very useful and are rated poor for relevance for safety data analysis, completeness and accuracy. The data contain little additional detail on road traffic accidents.

Table 3.9. **Quality of the data in the fire service reports (based on 5 responses)**

	Relevant for analysis	Completeness of the number of serious casualties	Accuracy	Timeliness
Good	0	0	0	0
Sufficient	0	0	0	0
Poor	5	3	4	3
Unknown	0	2	1	2

3.6.8. Insurance Data (vehicle, health insurance)

Data from the insurance industry offers potential to provide a more complete indication of the total number of road collisions than police data, including non-injury collisions, plus details of factors associated with higher propensity to claim (and therefore to be involved in collisions). However, at present there is little experience in using this source. Only 5 countries, Finland, Israel, Netherlands, Switzerland and United Kingdom mentioned this as a source of road safety data. In all these countries except the UK data are available for fatal, serious and slight injuries. The data are rated poorly for use for road safety purposes, only Netherlands and Switzerland rated the data sufficiently relevant for use in analysis. Switzerland also rates the data of being of sufficient accuracy and use the information to produce an annual estimate of road traffic accidents.

Table 3.10. **Quality of the data of insurance (based on 5 responses)**

	Relevant for analysis	Completeness of the number of serious casualties	Accuracy	Timeliness
Good				
Sufficient	2		1	
Poor	1	3	2	3
Unknown	2	2	2	2

There are frequently practical and commercial difficulties in obtaining and collating these data. Different companies may use different systems to hold the data and extracting records consistently is likely to be difficult. Customers can move between companies – data relating to an individual may be held in many places. Only collisions which result in an insurance claim will be recorded. Trends will be affected by the propensity of people to make a claim following a collision, which is believed to be increasing (particularly for bodily injury claims). However there is potential to obtain additional information to help improve road safety for example on vehicle damage, and on drivers (driver licensing and offences) to allow risk factors to be assessed.

3.6.9. Surveys

In general, survey data offers the scope to cover all traffic accidents, including damage only, and allows estimates of total accidents to be made. There is also the opportunity to follow up respondents at a later date, to gather further, more detailed information. However, surveys do not allow the collection of detailed information on crash circumstances and location that is available from the police. The data will be subject to sampling error, as only a proportion of the population is interviewed and, since it is self reported, depends on the ability and willingness of respondents to recall information accurately.

The Netherlands and United Kingdom (Great Britain) both mentioned the use of National Travel Surveys to collect data on injuries (excluding fatalities) in road traffic collisions. Great Britain has used survey data to estimate the total number of casualties, including those not known to police or hospital – the box below provides more detail. Spain undertook a health survey in 2006 which provide estimates on the number of people who were injured in a traffic crash, whether they sought medical advice or were hospitalised. The European Health Survey (EES09) – carried out by Eurostat in 2006-2009 collected households data in 19 countries on state of health, lifestyle and user of health services. One of the questions related to the occurrence of a road traffic crash in the past 12 months[6].

6. See more information : http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/EN/hlth_ehis_esms.htm

Box 3.4. **Great Britain: National Travel Survey: Road accident questions**

Data source and coverage

- Since 2007, questions on road accidents have been asked as part of the National Travel Survey (NTS), a household survey which aims to provide a databank of personal travel information for Great Britain.
- Around 18,000 adults interviewed each year; the response rate to the NTS interview is around 60 per cent.

Strengths and limitations for reporting road casualties

- Includes accidents not resulting in injury (which are not covered by police data).
- Covers private households only. Does not include fatalities, or child casualties.
- Limited detail, which cannot replace in-depth police data used to prevent future casualties.
- Scope to follow up respondents at a later date, to gather further, more detailed information.
- Subject to quite wide sampling error, as only a proportion of the population is interviewed, and few report accidents.
- Self reported results depend on ability and willingness of respondents to recall information accurately. For example, results may be subject to 'telescoping' (tendency to report events which took place before the reference period, particularly if they are memorable), omission (failure to mention events happening during reference period e.g. forgetting minor incidents, reluctance to talk about an incident due to embarrassment or painful memories) or misunderstanding of concepts and questions (e.g. what constitutes an injury).
- Overall it is hard to assess the effect of the above recall issues though further research is being conducted.
- Survey designed to identify long term trends and not suitable for monitoring short term trends.

What it adds

- Provides a broad estimate of the number of casualties not known to police *or* hospital.
- Trends over time not influenced by changes in reporting practices by either the police or hospitals.
- In time, will offer an opportunity to assess to what extent reporting and recording practice may influence police reporting levels and trends at national level.
- Allows linking to extensive personal travel information collected in the NTS (for example, possible to look at accident rates by relative exposure to risk for different groups).
- For further information see: Article 5 in Reported Road casualties Great Britain: 2009 at www.dft.gov.uk/adobepdf/162469/221412/221549/227755/rrcgb2009articles.pdf

Box 3.5. Dutch National Travel Survey

The Dutch National Travel Survey about mobility (annual sample 60,000 households, covering 140,000 persons, response rate 70%) in the years 1996-2005 included a screening question about involvement in a traffic crash. The approximate 7 per cent of those who reported to have had a crash were asked to fill in a further questionnaire. The information was mainly used to estimate slight injuries and damage only crashes. Because of the correlation of accidents and being a member of the same household, it was methodologically difficult to estimate a country number from the data. After 2005 the survey was stopped.

Reference: AVV, Verkeersongevallen in Nederland 2001, chapter 3. (2002), Ministerie van Verkeer en Waterstaat, DG Rijkswaterstaat, Adviesdienst Verkeer en Vervoer (AVV). Rotterdam/Heerlen

http://cardweb.swov.nl/swov/website_uk_advanced_detail.html?Zoek=Zoek&f=TI&pg=q&q=verkeersongevallen+nederland+2001

3.6.10. Other

Finland, France and the United States mentioned the use of other sources for fatal, serious and slight injury accidents. A number of these sources combine a range of information – the Rhone Road Trauma Registry (France) links various sources of data and is discussed in the box below. The United States Fatal Accident Reporting System (FARS) – discussed in the box earlier in the chapter which provide good quality data for analysis. The United States - NASS – GES system is a nationally based system which uses a sample of police accident reports to provide more in-depth analysis of types of accidents and how they occur, for example to answer motor vehicle safety questions.

Box 3.6. **The Rhone Road Trauma Registry (France)**

France: The Rhône road trauma registry is based on the participation of all health departments (public and private) in the Rhône county (and its close surroundings) that may receive road crash casualties. It collects data through notification forms, with the participation of pre-hospital care, hospital departments (from emergency, to surgery, to rehabilitation), and from the forensic department. It also gets electronic data from the fire-fighters. It contains both inpatients (15%) and outpatients (85%).

3.6.11. In depth studies

Although not specifically requested in the survey a number of countries reported that they conduct in-depth crash investigation studies. These sources typically include much more detail than is contained in police, hospital and other records but cover only a small, often non representative sample of crashes and casualties. These studies provide a fuller understanding of the factors that contribute to collisions, their severity and their impacts, including the interaction of vehicle design (primary and secondary safety features), highway and environmental factors and human factors.

Such studies collect detailed and factual (almost microscopic) information from independent investigations of accidents, usually conducted by trained experts using multi-disciplinary approach. The detailed information includes:

- The vehicles.
- The road environment.
- The road-users.
- The interactions of these.

Such studies can inform assessments of primary and secondary safety measures, encouraging developments in vehicle design and informing the legislation for new safety technologies. A number of countries undertake these type of studies including Belgium, Finland, France, Germany, Great Britain, Greece, Italy, Netherlands, Spain, Sweden, and United States; some examples are shown in the box below.

Box 3.7. **Examples of in-Depth Accident Studies**

In ***Germany*** the GIDAS (German In-Depth Accident Study) www.gidas.org/en collects extensive data to a well defined sampling plan to ensure representativeness compared to the federal statistics. The police, rescue services and fire department report all accidents to a research team who select accidents according to a strict selection process. The detailed information collected at the scene is complemented by further medical information about injuries and treatment and extensive accident reconstruction. Since mid 1999, the GIDAS project collects about 2000 accidents in the areas of Hanover and Dresden per year.

In the ***United Kingdom*** there were until 2009 two studies, 'On the Spot' Study (OTS) and the Co-operative Crash Injury Study (CCIS). Both studies are carried out by specialist civilian teams and provide the opportunity to collect more detailed data than could be expected of the Police. The OTS study, which began in 2000, involves in-depth investigations of a sample of collisions of all severity, including damage only incidents. The study collects perishable information at the scene, detailed injury data and information from those involved in incidents to understand their perspective. The findings are used in the development of countermeasures to reduce risk and mitigate injuries. The OTS database includes data from about 4,500 accidents. Further details are at: www.ukots.org/. The CCIS, which began in 1983, monitored the performance of car structures and secondary safety features in relation to car occupant injury causation and severity. Data is collected from about 1100 cars each year. Further details are at: www.ukccis.org.

The EU funded ***SafetyNet*** project established the European Road Safety Observatory to bring together data and knowledge to support safety policy-making. SafetyNet included work packages on the development of new fatal and in-depth independent accident investigation and causation databases. Further information can be found at: http://ec.europa.eu/transport/road_safety/specialist/toolbox/index_en.htm

The ***DaCota*** project will take this work forward on in-depth safety related accident data by means of the identification of suitable crash investigation teams within Member States and assistance to develop the local infrastructure to gather in-depth accident and injury causation data. See www.dacota-project.eu/

The University of Valencia assessed the various data sources available in the city of Valencia and their usefulness for road safety analysis. Their main conclusions is summarised in the following table, presenting the characteristics, advantages and limitation of each data source.

Table 3.11. **Main police and health traffic crash data and traffic crash victim records City of Valencia (Spain) Summary of the ACTIVA Project**

	CHARACTERISTICS	CONTENT	ADVANTAGES	LIMITATIONS
Police records	Information on crashes collected on the scene of the crash, and sometimes completed later. Main data source for studies in road safety and traffic crash at the international level. Fundamental reference for road safety diagnoses.	Data on crash circumstances (location, time, crash type, environmental factors, road characteristics, possible concurrent factors...). Data on the involved vehicles and their characteristics (vehicle type, registration data, state...) Data on the involved persons (driver and pedestrian offenses, gender, age, safety devices, severity...).	Most immediate source of available information. Standardized procedures. Organized and homogeneous information.	No reliable information on severity of injuries. Underreporting of injuries.
Health : Hospital admissions	Information on in-patients, collected by the hospital itself, from the admission, during the stay and at discharge. Important tool for the health system actors: it gathers useful information for financing, organizing and distributing health resources, for carrying out clinical studies, about resource consumption for pathology...	Information on the hospital casuistry and the characteristics of the assisted morbidity in the hospitals. Data on the main and secondary diagnoses, medical criterion, surgical and obstetric procedures, demographic data on assistance, clinical history, financing, admission and discharge data, E code.	Information with uniform, organized and homogeneous codification that can be compared between the different hospital administrations and countries. Quality information already consolidated and comprehensive (compared with the emergency and ambulance ones). Information on the diagnoses, treatment length, possible after-effects and sanitary. It allows introducing techniques of grouping patients.	Poor information on the crash circumstances. The only available information is the one that comes from the use of the E codes. Deficiencies in the way the E code is classified and filled in.

Health data: **Ambulance**	Information of the communication centres that coordinate the emergency assistance requests. Systematic records of the incidents that are treated by the Emergency Ambulance Services.	General data (person that calls, date, time, type of incident, place, number of victims, mobilized resources, end of the incident, health centre...). Incident data, identifying and demographic data, clinical data (Glasgow scale, traumatological evaluation of the injuries, reviewed trauma score, diagnosis, what has been done to the patient and given medicines).	Information in computer format, although not all. It combines with the police crash reports. It gives information on the time and place of the crash. It gives some information on the crash characteristics. It provides demographic data and gives a first diagnostic assessment and the destination of the victim.	The diagnostic assessment might not be the definitive one. Not for investigation. Different guidelines (for completion) for each community. Deficiencies in the structure of the crash information and in the way the data are filled in.
Emergency hospitals	Administrative files where the data on Emergency Services are collected. The main goal is to follow up internally the activity, the management of the resources and the maintenance of clinical data. It represents the starting link of the hospital care chain.	Identification, demographic and financing data. Assistance data: reason, date and time, unit. Clinical data: anamnesis, complementary exploration, given treatments, diagnoses, treatment when discharged. Discharge data: date and time, destination, doctor and unit.	It gathers plenty of information on injuries. Identification data are more exhaustive than in the police and ambulance records. The clinical information already notes down some after-effects.	Heterogeneity as for the completion procedures, the quantity and quality of information and the computerization. Many data on paper. The person that suffered the crash can go to the Emergency hospital several days after the crash happened.

Source: Chisvert et al. (2005)

3.7. Conclusion

Police, followed by hospital data are the most commonly used sources of data for road safety purposes.

Police data are in all countries the main source of information on traffic crashes. They are considered a reliable source of information and the quality of the information from police records is usually recognised. However, underreporting, especially regarding non fatal crashes, is a common issue. Hospital data are a natural complementary source of information on traffic casualties. In a number of countries, many road traffic casualties are admitted to hospital and are not known to the police. On their own, hospital data are not a substitute for police data, but only an addition, as they are likely to include only minimal information on the circumstances of a crash. However, using hospital data in addition to police data can provide valuable information, including:

- Some basic information on casualties not reported to the police, such as age, gender and vehicle type.
- Better understanding of the total number of casualties.
- Better understanding of injury severity.
- In-depth understanding of the medical consequences of particular types of crash, if police and hospital records are linked.

Therefore it is important for efforts to be made to overcome the barriers to using hospital information as a source in its own right, and also to linking hospital data with police data.

Other sources, with the exception of mortality registers for fatalities, are not widely used and are of variable quality and coverage. However, a number of countries do make use of other sources to add to or link with existing police data and increase their wider understanding of road safety issues. It is suggested that countries may find it useful to evaluate the relevance, coverage, quality and availability of all their sources of data on road accidents. The aim is to identify those with most potential to add value and to consider if improvements are possible. However availability and quality issues are likely to limit the extent to which the data can currently be used for linkage, to measure levels of under-reporting or to estimate the “real” number of casualties.

Differences in legal and administrative systems are likely to affect which sources are of most value in each country. It may also only be possible to make use of such sources at a local or regional level. Mortality registers provide an additional source of information on fatalities but there is no single commonly used other source of injury data.

Sources primarily for fatality data:

- **Mortality Registers** are widely available, generally of reasonable quality and offer countries the opportunity to link or compare to police fatality records to check for completeness. Generally contain little detail but may add information on cause of death.
- **Forensic reports** are a further source for fatalities and linkage with police data but are less widely available. Additional details vary but may add information on cause of death and for example, levels of alcohol.

Sources primarily for injuries:

- **Emergency ambulance** data are often only available locally and may not specifically identify road traffic accidents. However, when available may be able to provide additional medical information on injuries.
- **Fire services** data currently of limited value and not widely available.
- **Insurance data** currently used in only a few countries but has the potential to provide information on accidents not reported to the police and additional details, for example, on injuries, vehicle damage, driver details and risk factors. Experience will differ between countries but commercial considerations mean the data are likely to be difficult to access and interpret.
- **Surveys** are only currently used in a few countries, road safety questions are usually part of a wider survey. Offers the potential to gain more information about participants in accidents and to estimate overall numbers of casualties but will be subject to sampling error.

There are also good examples in a number of countries for example, the United States, France (some examples discussed above) where systems have been developed which bring together a range of sources (police, hospital, emergency services, forensic, death, driver and vehicle records etc) which can provide a more complete picture of road safety issues.

References

Amoros E, Martin JL, Laumon B. (2006) Under-reporting of road crash casualties in France 1. Accid Anal Prev 2006; 38(4):627-635.

Chisvert M, Tormo MT, Martín S, Pont P, Montoro L, Ballestar ML, Marín C. (2005), Estudio epidemiológico de los accidentes de tráfico y su incidencia en la Comunidad Valenciana. Caracterización, factores de riesgo y consecuencias. Proyecto ACTIVA. Volumen I. Análisis a partir de los datos de los Centros de Información y Coordinación de urgencias (CICU). Valencia, Spain: Generalitat Valenciana. Conselleria de Sanitat; 2005.

Chisvert, M.; Lopez-de-Cozar, E.; Ballestar, M.L. (2007). Quality and representativity of the traffic accident data in urban areas: State of the Art. Project: SAU - Urban Accident Analysis Systems. (Dirs.: Sanmartín, J. & Chisvert, M.). Valencia: INTRAS linea editorial. Universidad de Valencia (Spain).

Derriks, Harry M, and Peter M Mak. Underreporting of Road Traffic Casualties. IRTAD International Traffic Safety Data and Analysis Group (IRTAD), OECD, International transport Forum, 2007.

Haque, M Ohidul. Road Safety - Data Collection, Analysis, Monitoring and Countermeasure Evaluations with Cases. University Press of America., 2009.

Lopez, Derrick J, Diana J Rosman, George A Jelinek, Garry J Wilkes, and Peter C Spivulis. “Complementing police road-crash records with trauma registry data - an initial evaluation.” Accident Analysis and Prevention 32 (2000): 771-777.

OECD/ITF (2008). Towards Zero Ambitious Road Safety Targets and the Safe Systems Approach. 2008.

Peden, M et al (2004) . World report on road traffic injury prevention. Geneva: World Health Organisation, 2004.

Rosman, Diana J. “The Western Australia Road Injury Database (1987-1996): ten years of linking police, hospital and death records of road crashes and injuries.” Accident Analysis and Prevention 33 (2001): 81-88.

Tsui, K L, F L So, N N Sze, S C Wong, and T F Leung. “Misclassification of injury severity among road casualties in police reports.” Accident Analysis and Prevention 41 (2009): 84-89.

Ward, Heather, Ronan Lyons, and Roselle Thoreau. Under-reporting of Road Casualties - Part 1. Department of Transport, London, 2006.

CHAPTER 4
Methods to link police with hospital data

Abstract

This chapter describes the two main methods of linking of records from more than one source -- the deterministic and probabilistic methods -- includes an extensive reference list for further information as well as several case studies of linkage in different countries.

4.1. Background and objective

Due to the limitations of individual databases it has become increasingly common to link databases from different sources in order to improve road safety research. The term record linkage comes from the public health area when files of individual patients were brought together using name, date of birth and other information. Record linkage began with the pioneering work of Howard Newcombe in 1959 (Newcombe, 1959), and of Fellegi and Sunter in 1969 (Winkler, 2010). The use of record linkage is increasingly widespread in public health and epidemiology, and allows analyses that would otherwise be very difficult or very costly to implement.

This chapter aims to describe the probabilistic weight-based and the distance-based methods applied to the field of road safety and to provide an overview of published and unpublished experiences on record linkage methods.

4.2. Methods of linkage

There are three main methods to link two databases: manual, deterministic and probabilistic. Manual linkage determines visually the possible matching of each record in a database with all records in another database. This method requires that the linking databases to contain few records and little information. It is also subject to error, which increases in proportion to the volume of data. Manual linkage is the simplest and most valid method when there are few records (Cirera et al., 2001), but is usually impractical due to the large number of records.

4.2.1. Deterministic method

Several methods of computer assisted linking have been developed. The easiest is to link records using one or more unique personal identifiers. This approach is referred to as "rules based" or "deterministic" (Jaro, 1995). It is based on the existence of a unique identifier - or combination of variables – common to both databases being linked. This method links records with an exact match of the identification variables. It is therefore subject to the quality of the identification variables. However, linkage entails the use of databases which have not been designed to be linked, and for which there is rarely a unique personal identifier.

In cases where the database is large; does not have a good quality, unique identifier; or the possible variables available are limited or have little discriminatory power, it is recommended that other methods should be used. In such cases without personal information, and when the number of common variables is limited other methods are used such the probabilistic (Newcombe, 1998; Jaro 1995) or the distance-based (Domingo-Ferrer and Torra, 2006) methods. For categorical data as well as for ordinal and nominal scales, distance-based and probabilistic-based record linkage lead to similar results (Domingo-Ferrer and Torra, 2006).

4.2.2. Probabilistic weight-based method

This method is based on the linkage of records with the highest probability of belonging to the same individual. In essence, the probabilistic linkage process consists of matching two or more records that come from different data sources and are believed to belong to the same individual. It is based in two probabilities: the probability of matching given that both records belong to the same individual and the probability of matching by chance. The less likely is a value of variables, the greater is the weight assigned (Cirera *et al*., 2000; Neutel *et al.,* 1991).

More detailed information on record linkage techniques is available in the proceedings of the United States Federal Committee on Statistical Methodology workshops in 1985 and 1997 (Federal Committee on Statistical Methodology, 1985; 1997).

The matching process involves these phases:

1. Data preparation
2. Selection of linkage variables
3. Evaluation of process feasibility
4. Computation of simple weights
5. Restriction of comparison pairs (blocking)
6. Comparison stage (matching)
7. Simple weights assignment
8. Computation of composite weights
9. Decision stage (linking)
10. Threshold determination
11. Review of dubious pairs

1. Data preparation

As reported by Clark (see www.LinkageWiz.com) the success of record linkage is much more dependent on data quality than on software. Processing the data is frequently the most time consuming and most difficult phase. This requires standardization of formats of all fields for the two databases in order they are comparable; addressing missing values; identifying miscoded values; and removing duplicate records. Names and surnames lead to particularly difficult issues due to variations in spelling and coding. This can be taken into account by generating a phonetic code up to 6 letters for each name, and some software like LinkageWiz (Boufous et al., 2008 and LinkageWig.com) allow this.

2. Selection of linkage variables

In absence of a common identifier variables used for linkage are usually age, gender, and date of the crash. It is assumed that the hospitalisation is on the same day as the crash, or in some cases within two or three days.

3. Evaluation of process feasibility

The feasibility study is based in the distribution of frequencies of each variable.

Common variables available for the linking process are gender, age, date of the crash. A simple way to assess basic feasibility is to multiply the number categories for each variable (Roos and Wajda, 1991). For instance, 2 (gender) * 100 (age) * 365 (days)= 73,000. This number must be greater than the total number of records from both databases (for example 155.000 police records + 40,000 hospital records = 195,000). In this case, this total is clearly insufficient to identify true pairs of records.

A more complex way to assess feasibility of linkage considers not only the distribution of categories of variables, but also the information contained in them (Cook *et al.,* 2001). It is based on calculating the weight needed to achieve a specific probability that two records are a true pair.

4. Computation of simple weights

Prior to linking records it is necessary to compute weights that will become useful later in the linking phase. These weights are based on two probabilities, the probability of matching given that both records belong to the same individual, and the probability of matching by chance. The less probable is a value of the variable, the greater is the weight assigned (Jaro, 1995).

For each category, within each variable there are three possible values, all of them based on the distribution of the variables to be compared between both files and taking into account missing values.

A value (1) will be assigned to the pair if they are identical, a value (2) will be assigned if they are not, and zero if one of the two values are missing.

5. Restriction of comparison pairs (blocking)

Once weights have been computed, it is necessary to compare the information obtained for the variables common to both files. This first step, known as the blocking phase (Jaro, 1995) consists of forming blocks in order to reduce the number of comparisons for computing efficiency. The most common blocks are those hospital records for which the date and time of patient attendance is within three days after the crash occurrence reported in the police files.

6. Comparison stage (Matching)

Within each block, comparisons are made: firstly, the contents of the common variables for both files for each hospital record with each police record are compared.

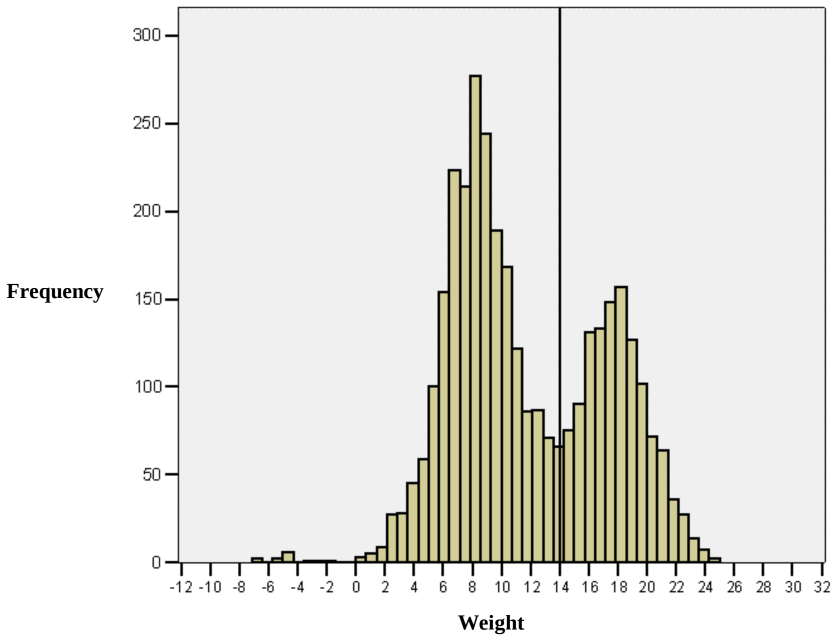
7. Simple weights assignment

Out of every between-variable comparison a weight value is assigned.

8. Computation of composite weights

At a second stage, a composite weight as the sum of the individual weights obtained in between-variable comparisons is generated, allowing the comparison between records. The distribution of the weight is a bimodal distribution. It seems reasonable to think that some characteristic in the registries exists that divides them in two groups, each one with a normal distribution, as can be seen in the figure.

Figure 4.1. **Distribution of the weights**



9. Decision stage (linking)

A hospital record is matched to a police record when it is the record with the highest composite weight after its comparison with all the remaining records in the selected block. If there are two or more records with the same weight for one hospital record, it remains unlinked, because it is impossible to distinguish which record corresponds to the same person. In order to decide which are true pairs, we analyse both the number of variables that agree and the weight of each pair.

10. Threshold determination

Using these weights, two threshold values are defined: the lower-threshold limit, under which all records with such weight value would be considered not corresponding to the same individual, and the upper-threshold limit, above which a record would be considered to belong to the same individual.

11. Manual Review of dubious pairs

For those pairs with a weight value between the two threshold limits, a manual review of the data is recommended using additional information from the records, in order to decide if the linkage should be accepted.

4.2.3. Probabilistic distance-based method

This method is based on the linkage of the two records with the shortest 'distance' between them. In essence, if two records from different sources have distance zero (i.e. they have equal values on a set of key variables, and there is no other record having zero distance as well) then it is believed that both records describe the same individual event. In practice, the values of the key variables are not always equally well defined, or equally accurate, and are sometimes biased in both sets. In the distance-based method these errors and differences are tolerated, weighed and accumulated over the key variables into a distance. Also pairs with a small distance are seen as matches. Which differences are tolerated and how much each weighs in the accumulated distance depends on the variable, and country or database specific situation.

For a variable with only a few possible values and an excellent administration the ‘punishment’ for a difference will be a high distance. For example gender is likely to be a very accurate value in both databases and contains only the values Male and Female. It is very unlikely that a male in one file should describe the same individual as a female in the other database, so the distance will be high. For another variable with many possible values and unclear boundaries the distance will be much lower, not preventing a match if the records are equal on most other key variables. For example, vehicle type: if in one of the databases the definition is not the same as in the other database, or some categories do not exist or are often exchanged, then the distance will be much smaller in case of an unequal value. There may be uncertainty between (light) mopeds and motorcycles, cars and vans, or vans and trucks. The distance assigned can also depend on the prevalence of certain values, e.g. males, drivers, and cars are the most common individuals to be involved in a serious injury crash.

The matching process implies similar phases as explained with the Probabilistic weight-based method above.

1. Data preparation
2. Selection of linkage variables
3. Evaluation of process feasibility
4. Computation of simple weights
5. Restriction of comparison pairs (blocking)
6. Comparison stage (linking)
7. Simple weights assignment
8. Computation of composite weights
9. Decision stage (matching)
10. Threshold determination
11. Review of dubious pairs

Compared to the Probabilistic weight-based method, the following items are different:

1. ***Data preparation:*** It is not necessary to harmonize classifications before continuation to the next steps, the distance functions allow for these differences. Of course the data types should be the same (number, character, date)

4. ***Computation of simple weights****:* Development of distance function with initial weights, see for example SafetyNet (2008)

6,7,8. For each crash record all medical records from the block (defined by an admittance date in a certain period before and after the crash) are ***compared*** and a ***distance is calculated*** as the sum of the distances of each variable.

9. ***Decision stage*****:** For a police casualty record, you choose which is its closest (most similar) medical record. Then you choose from that medical record which is the closest crash record. If this points back to the original crash record then they have each other as closest neighbour and they are now matched. An indicator 'Selectivity' is calculated as the distance to the closest alternative link. Ideally the selectivity is high: there is no alternative pair that could also describe the same person. In the more advanced version of the linking software it is also possible to match with the next best neighbour, in case the closest neighbour is already matched with another record.

10. **Threshold determination:** In absence of a true other record, it is possible to match records at a very high distance. It is also possible to fail to make a clear distinction between two match candidates from the key variables (poor selectivity).

11. **Review of dubious pairs:** If possible cases around the thresholds should be reviewed, in order to refine the initial weights and distance functions (step 3) or the chosen thresholds (step 10).

More detailed information about this method can be obtained from Reurings and Bos (2009).

4.2.4. Validity of record linkage

Although it is not always possible, ideally it is convenient to have a sample of records with personal identifiers to assess the validity of the process. False positive matched records and false negatives must be balanced.

Table 4.1. **True and false positives and negatives according the coincidence with the identifier (ID)**

	Real pair (same ID)	**No real pair (different ID)**
Linked	True Positives (TP)	False Positives (FP)
Non linked	False Negatives (FN)	True Negatives (TN)

The frequency of false positive and false negative pairs can be expressed in the epidemiological terms of sensitivity, specificity, positive predictive value and match rate, and these provide indicators of the quality of the record linkage. From the previous parameters, we can obtain different measures to evaluate the accuracy the record linkage:

- **Sensitivity** (S): S=TP/(TP+FN)

The number of pairs of records true linked pairs divided by the total number of correct pairs of records. It is interpreted as the probability that a concordant pair of records connects in the process.

- **Specificity** (E): E=TN/(FP+TN),

The number of unlinked correctly divided the total number of pairs of incorrect records. It is interpreted as the probability that an incorrect pair are not matched in the process.

- **Positive Predictive Value** (PPV). VPP=TP/(TP+FP)

The number of correct linked records divided by the total number of pairs of linked records. It is interpreted as the probability that a pair linked is a really a pair. The PPV is useful as an indicator accuracy of the linkage process.

- **Match Rate** (MR). MR=(TP+FP)/(TP+FN)

The total number of linked records pairs divided by the total number of pairs of correct records.

4.3. Experiences of police and health data record linkage

Many countries have carried out record linkage procedures of police and health data for a number of years (OECD, 2003; Derriks and Mak, 2007) . Some carry out it periodically, as is the case in some states in the United States (Johnson and Walker, 1996; NHTSA (2009) or in regions of Australia (Ferrante et al., 1993). However, there are great differences in the methodology used. Some countries link the data manually (Cryer *et al*., 2001), while others use deterministic linkage allowing certain level of tolerance in the variables (Amoros et al., 2007;Petridou et al., 2009; Jarvis et al., 2000; Simpson, 1996; Keigan, et al., 1999; Morrison and Ston, 2000, Van et al., 2006, Tercero et al., 2004) (e.g. age + / - 5 years) or matching exactly the identifier variables (Gonzalez et al., 2006; Lai et al., 2006). Other countries used probabilistic record linkage based on weights such the United States , Australia, Spain and Italy (Chini et al., 2008, Cirera et al., 2001' Johnson and Walker, 1996; NHTSA 2009; Ferrante et al., 1993) or distance-based such the Netherlands (SafetyNet, 2008).

There are also differences between countries regarding the number of variables used to link. While some countries have access to initials of the surnames and names (i.e. Australia, New Zealand and the United States), other do not have personal identifiers due to confidentiality and personal data protection (Austria and Spain). Table 4.2 shows the characteristics of the studies that carried out deterministic or probabilistic record linkage in the world.

Table 4.2. **Characteristics of deterministic and probabilistic record linkage between police and hospital records**

First author and year of publication	Country and study period	Type of record linkage	Variables and criteria used for record linkage	Validation
Gonzalez, 2006	7 regions of Alabama, USA. October 2001 – May 2003	Deterministic	Age, gender, date and hour of collision and hospitalization, region of the crash and hospital, and of road user	--
Tecl, 2008 (SafetyNet)	District de Kromeriz, Czech Republic 2003 – 2005	Deterministic with some degree of tolerance	Year of birth (+/-5 years), gender, date of collision and hospitalization (+1 day) and road user	Manual review of the records, but there are no additional quality variables to do that
Merényi, 2008 (SafetyNet)	Hospital Károlyi Sándor en Budapest, Hungary 1 agosto 2004 – 31 enero 2006	Deterministic with some degree of tolerance	Age (+/-1), gender, date and hour of collision (+/-1), place of collision, vehicle, road user, and mechanism of accident	--
Broughton, 2008 (SafetyNet)	Scotland, United Kingdom 1997 –2005	Deterministic with some degree of tolerance	Age (+/-1), gender, injury severity, date and hour of collision (+/-1), date of collision (+/-1), road user (pedestrian and cyclist together), reference area of the hospital and hospital code	--
Jarvis, 2000	Northumbria, United Kingdom April 1990 – March 1995 (Severe injured under 15 years old)	Deterministic with some degree of tolerance	Age (+/-1), date of collision and hospitalization (+1) and hospital code	--
Tercero, 2004	Municipality of e León, Nicaragua 1993	Deterministic with some degree of tolerance	Name, age, gender and date of collision and hospitalization	--
Van, 2006	3 hospitals of the city of Thai Nguyen, Vietnam 2000 – 2004	Deterministic with some degree of tolerance	Name and surname, gender, age (+/-2), date of collision and hospitalization (+/-2), place of residence	--
Amoros, 2008 (SafetyNet)	Region of Rhône, France 1996 – 2001	Deterministic with some degree of tolerance and manual review	Month and year of birth, gender, date (+2) and hour of collision and hospitalization, road user. Address of collision is used for manual review	For year 2001 it was possible to carry out a record linkage through the phonetic code of names and surnames. It gave information on the number of false positive and false negative records
Yannis, 2008 (SafetyNet) Petridou, 2009	Island de Corfu, Greece 1996 – 2003	Deterministic with some degree of tolerance and manual review	Age (+/-5), gender, nationality, date of collision and hospitalization (+2), road user and mode of transport (motorcycle and moped together)	--
Alsop, 2008	New Zealand 1995	Probabilistic, with weights calculation	Date of birth, gender, injury severity, name initials, phonetic code of name, date of collision and hospitalization, place of collision, and hospital	--

First author and year of publication	Country and study period	Type of record linkage	Variables and criteria used for record linkage	Validation
Johnson, 1996 (CODES)	USA, some states	Probabilistic, with weights calculation	States vary. Most of them use date of birth, gender, zip code, date and hour of collision and hospitalization, place of collision, and hospital reference area. Half of states have names and surnames	Manual linkage of a sample of records of people with injuries
Karlson, 1996	Wisconsin, USA 1993	Probabilistic, with weights calculation	Date of birth, gender, zip code, date of collision and hospitalization, place of collision and hospital	Percentage of false positives. Not clear how they derive them.
Ferrante, 1993	Western Australia, Australia October 1987 – December 1988	Probabilistic, with weights calculation	Year of birth, gender, surname, initial of middle name, name, date of collision and hospitalization and road user	Manual review of a sample to derive false positives and false negatives
Lujic, 2008	New South Wales, Australia July 2000 –June 2001	Probabilistic, with weights calculation and manual review of doubtful cases	Date of birth, age, gender, initials, name, phonetic code of name, zip code, date of collision and hospitalization	--
Pérez, 2008 (SafetyNet)	Castilla-León, Spain July 2005 – December 2005	Probabilistic, with weights calculation and manual review of doubtful cases	Age, sex, date of collision and hospitalization, province of collision and hospital	A hospital database was linked with a dummy database containing 70% random hospital records. Calculation of sensitivity and specificity, positive predictive value, and match rate
Novoa, 2010	5 regions, Andalucía, Aragón, Cantabria, Castilla la Mancha, and Galicia, Spain 2000-2007	Probabilistic, with weights calculation and manual review of doubtful cases	Age (+/-5), gender, date of collision and hospitalization (+3), province of residence, vehicle or road user. Anatomical region of the injury has been used in one region	The same as the case before
Jeffrey 2009	West Scotland 1997-2005	A combination of deterministic and probabilistic with some degree of tolerance	Age, post code, gender and date of collision and hospitalization, casualty type and severity. Age (+/-1), gender, injury severity, date and hour of collision (+/-1h), date of collision (+/-1), road user (pedestrian and cyclist together), reference area of the hospital and hospital code	The authors argue that as there is no access to names or unique common person identifiers, it was impossible to estimate the true positive and false negative linkage rates with any degree of accuracy
Hoeglinger, 2008 (SafetyNet)	Austria 2001	Probabilistic distance-based	Age, sex, nationality, date of collision and hospitalization, federal state of the collision and the hospital	1) Analysis of the injury distribution according road user 2) Record linkage with databases of different years

First author and year of publication	Country and study period	Type of record linkage	Variables and criteria used for record linkage	Validation
Bos, 2008 (SafetyNet)	Netherlands 1997 – 2003	Probabilistic distance-based	Date of birth, gender, date and hour of collision and hospitalization, province of collision and hospital	The authors report that it is impossible to validate properly as there is no available information to know whether two records correspond to the same subject. Manual review of mode of transport for pairs of linked records
Reurings (2009)	Netherlands 1993 – 2008	Probabilistic distance-based	Date of birth, gender, date and hour of collision and hospitalization, province of collision and hospital	1) Simulations by linking accident data from one year with medical data from another year resulted in a very limited number of matches 2) Simulation by linking accident data from 3 years with medical data from 3 years in which the before year was shifted 364 days later and in the after year the dates were shifted 364 days earlier, resulted in the fact that 99% of the matches found earlier were reproduced, i.e. not replaced by a coincidentally similar record in another year 3) comparison with data from Accident & Emergency rooms of 12 hospitals (about 15% of NL) confirmed the correctness of missing matches: If it was not possible to match an accident record to one in the medical file, it was in most cases possible to find it in the A&E file, which stated that the person was not admitted, but only treated at A&E and then sent home

4.4. Ethics of record linkage

Clinical records include particularly sensitive information and need high protection. Therefore frequently these data are not accessible. In addition, personal identifiers are frequently not available as these are considered to be confidential information and personal data protection rules are applied. But even without personal identifiers, when two datasets are combined, as pointed out by Clark, the linked database adds new knowledge about individuals that needs to be considered. There may therefore be additional ethical obligations beyond the use of each source alone even if written or oral consent has been obtained for the use of any individual database (Clark, 2004).

4.5. Record linkage: experience in IRTAD countries

The questionnaire sent to all OECD/ITF countries aimed to collect information on record linkage practices and on methodologies used to link different sources of data. This section describes the results.

Out of 23 countries who answered the questionnaire, 16 reported having **any experience of police and health data record linkage in their country**. (See 0). There is a great variability in the purpose of the linkage, on the setting, and on the methodologies used to carry out the linkage. From now all the description will be based only on those countries who reported any experience of record linkage (n=16 countries).

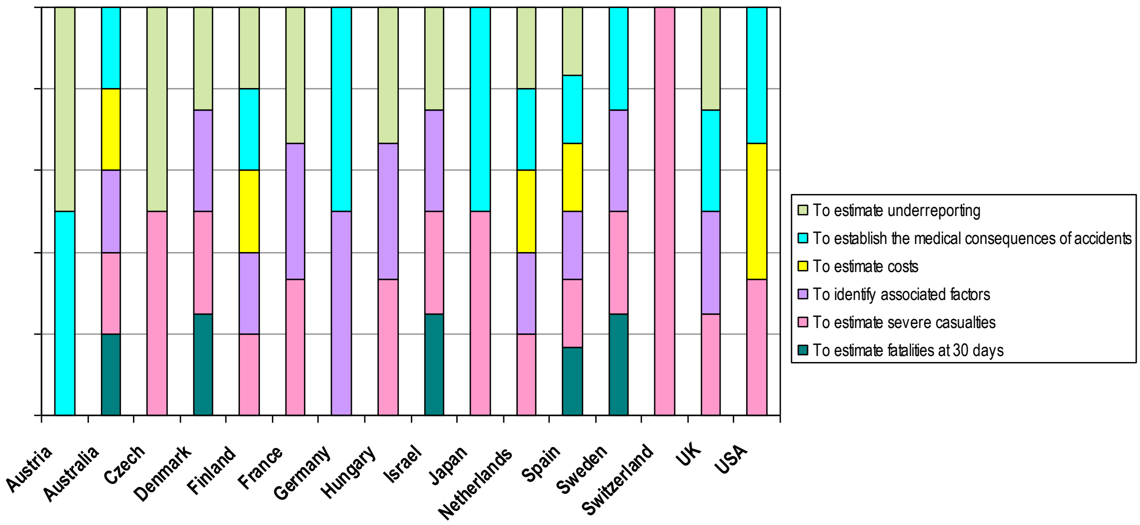
Among those countries who reported any experience of record linkage, the **setting** was national in 9, regional in 7 and local for 9.

Table 4.3. **Experience of police and health data record linkage IRTAD countries 2010**

	Any experience of police and health data record linkage	National	Regional	Local
Austria	yes	Y		
Australia	yes		Y	
Belgium	no			
Canada	no			
Czech Republic	yes			y
Denmark	yes	Y		
Finland	yes	Y		
France	yes		y	
Germany	yes			y
Hungary	yes			y
Ireland	no			
Israel	yes	y	y	y
Japan	yes	y		
Lithuania	no			
Netherlands	yes	y	y	y
Norway	no			
Poland	no			
Portugal	no			
Spain	yes		y	y
Sweden	yes	y	y	y
Switzerland	yes			y
United Kingdom	yes	y	y	y
United States	yes		y	
Total (out of 23)	**16**	**9**	**7**	**9**

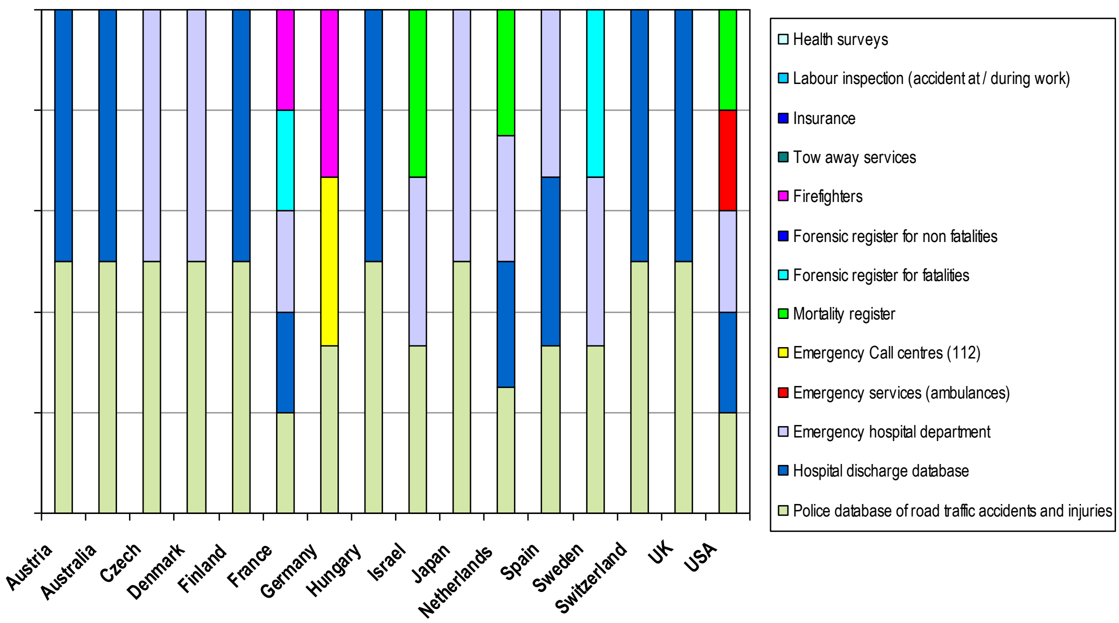
Regarding the **purpose of the linkage**, the most frequent aims were to estimate serious casualties (reported by 14 out of 16 of the countries) and to estimate associated factors (11/16). Other reasons where to estimate underreporting (10/16), to establish the medical consequences of accidents (10/63), to estimate fatalities at 30 days (5/16) or to estimate costs (5/16). (0).

Figure 4.2. **Purpose of the record linkage. IRTAD countries 2010**



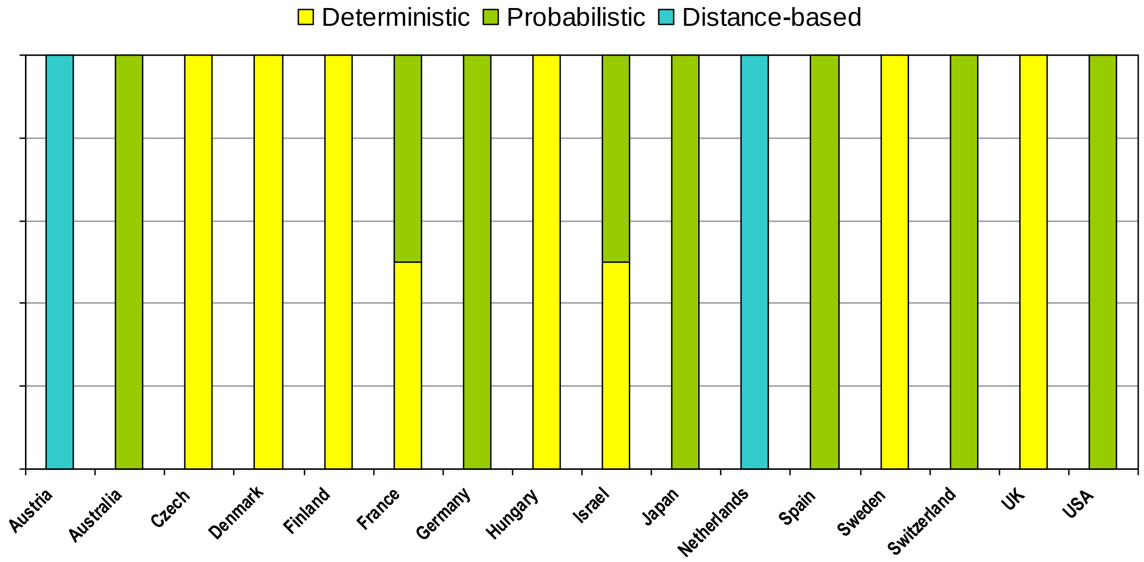
All countries used the police database to link to other **data sources**. The most frequently data source used was the Hospital Discharge Database (10/16) and the Hospital Emergency Records (9/10). There are few experiences of linking to other databases. Three countries linked to the Mortality Register, two to the Forensic Register for fatalities and for non fatalities, one to Emergency Services (Ambulances) and one to Emergency Call Centres. (figure 4.3).

Figure 4.3. **Data sources for record linkage, IRTAD countries 2010**



Deterministic and probabilistic **methods for record linkage** were reported by eight countries each. On the other hand, the distance-based method was reported only by two countries, Austria and the Netherlands. There is a high variability in the reported percentage of records linked, from one fifth to two thirds of police records linked, and from one to two thirds of hospital records linked.

Figure 4.4. **Methods of record linkage. IRTAD countries 2010**



There is also a lot of variability regarding the **variables used to link** the records. Almost all countries include the date of the crash, and half of them include the date of the birth or the age of the person injured. Around one third include the type of vehicle or the place of the crash. Gender was reported by more than one third of the countries (it was not asked explicitly). Personal identifiers such as names, surnames, or initials are hardly used. Five countries have a personal number as identifier (Denmark, Finland, Israel, Sweden and some States in the United States) (figure 4.5).

Figure 4.5. **Variables used to record linkage, IRTAD countries 2010**

	AUT	AUS	CZE	DEN	FIN	FRA	GER	HUN	ISR	JAP	NED	ESP	SWE	SUI	GBR
Gender							X		X	X	X		X		X
Mode of transport (vehicle type) in the medical file / E-code			X			X									X
Localisation (region) of the crash / localisation of the hospital	X	X								X	X	X		X	X
Place of the crash						X	X	X		X	X		X		
Position in the vehicle (driver, passenger, pedestrian)			X			X						X			X
Type of vehicle (car, truck, motorised 2 wheelers, etc.)		X	X			X						X			X
Date of the crash / Date of admittance to health service	X	X	X	X	X	X	X		X	X		X	X	X	X
Date of birth		X				X		X	X		X	X	X	X	
Age of victim	X		X			X			X	X	X	X			X
Names / surnames initials								X				X			
Surnames									X						
Names		X							X						
Personal identification number				X	X				X				x		

*** The table does not include information from the United States, because there is a high variability between states**

Some countries are routinely carrying out police record linkage with data since the 1980s (Denmark, Netherlands), late 1990s (France, Sweden, United Kingdom, United States of America, City of Barcelona) or early 21st century (Japan). Others carried out the record linkage as part of a research project such SafetyNet (Austria, Czech Republic, Hungary) or as a pilot study (Israel, Switzerland, Finland).

4.6. Conclusion

This chapter shows that there is extensive experience of record linkage in the field of road safety in the IRTAD countries. There is a great variability in the methods used for linkage, availability of linkage variables and procedures. There are few experiences at national level, and few are carried out as a routine practice. The most frequent health data sources are Hospital discharge registers and Hospital Emergency Departments. The most frequent methods are the deterministic and weight-based probabilistic, while few applied the distance-based methods. The most frequent variables for linking are date, age and place of the collision, and there are some experiences with ID numbers. It is to be welcomed that some countries are also planning to extend or to start police and health records linkage, and this should be encouraged as a next step for other countries when the quality of the relevant datasets is sufficiently good.

The best method for a given linkage project depends on its purpose. For example, a probabilistic approach may be more efficient to estimate a population parameter in a statistical study that aims to assess the effectiveness of a road safety intervention. For other applications, where the purpose is to make inferences about specific individuals based upon their data, a completely probabilistic approach would be unlikely to give acceptable results. In this case we must be quite sure that records from different sources truly refer to the same person and this might favour a deterministic method, although a probabilistic method with careful clerical review may also be useful.

Box 4.1. **Case study: Austria**

A computer assisted linking process of the Austrian hospital and the police database tried to calculate two matrices in order to enable further calculation of underreporting rates (conversion factors). The design of the calculated matrices fits to the requirements of SafetyNet WP1 Task 5.

“Useable matched” calculated records are likely not always to refer to the same persons. It can be assumed that 30.5% of these matches are wrong. The reason for this is not the linking procedure itself, but the lack of comparable information on both databases. This lack explains the high share of “not useable matched” records due to selectivity, distance (56%) and the “double zero” records (5%).

To calculate underreporting rates for different modes of transport, information about the mode of transport is needed in both databases. The Austrian hospital discharge database does not contain this information. No conversion factors for different modes of transport can be calculated with the result of this report.

As the Austrian hospital discharge database contains only in-patients, information about slightly injured persons, who do not go to the hospital are not included in this report.

Future steps & recommendations:

a.) Results of the linking procedure could be highly improved if e.g. the date of birth is also mentioned in the police database.

b.) More analysis is required to detect why there are so many “unmatchable” records.

c.) More investigation is needed about the recording procedure of persons who are admitted to a hospital, but then transferred to another hospital later.

d.) The new IDB (Injury Database) of DG-Sanco will also contain information about the mode of transport in traffic accidents in near future. In contrast to the Austrian hospital discharge database the IDB is based on a survey carried out in different member states of the EU. More information about the accident will be available with this new database which could be used to re-launch of the calculation of underreporting rates.

Link: https://webgate.ec.europa.eu/idb

Box 4.2. **Case study: Germany**

Objective

The objective of the pilot study was to integrate data on road accidents which is routinely registered by the police and the emergency call centres. The focus of the study was to develop an algorithm for data linking and to answer questions regarding data protection/privacy and administrative issues.

Methodology

Data was gathered for the Lahn-Dill-Kreis for the year 2006. Key variables for the linking procedure were: data of the accident, time of the accident, and place of the accident. It was defined that a match in all three variables had to be given to be a true link. The goodness of a link was specified by weights.

Results

For 549 patients a link could be found. Regarding the initial number of patients treated by the emergency services (N=854) 64% of these cases could be linked with police data. Vice versa 426 accidents (83%) of the initially 514 registered accidents by the police could be liked with emergency data.

Additionally some analyses referring to injury severity etc. independent of the accident situation were conducted.

Conclusion

The linking approach is promising though the procedure itself is very simple. Questions regarding data protection have still to be resolved.

Box 4.3. **Case study: France**

Objective: the record-linkage of the police database and the road trauma registry is one step in the estimation of the total number of road casualties, at the Rhône county level first, and later at the national level.

Methodology: A probabilistic record-linkage has been conducted on indirect identifying variables. The linking variables that we use are: date of the crash, crash location (town/village and possibly street/road), mode of transport (van/ car/ motorised two-wheel, bicycle, pedestrian), age (year and month of birth), and gender of the casualty. The procedure was only partially automated, since a major linking variable, namely crash location, was left as unformatted free text to avoid losing any valuable information. Please note that date and location of the crash are available in the road trauma registry (part of the notification form).

Main results

For the period 1996-2004, Rhône county, there were about 36,300 injured people recorded in the police data, about 83,400 in the road trauma registry, and 26,100 identified in both databases by record-linkage. This yields a total of 93,600 of injured persons (NB: the capture-recapture method applied to these data yields an estimate of 113,800).

References

Amoros E, Martin JL, Laumon B, 2007, Estimating non-fatal road casualties in a large French county, using the capture-recapture method. Accident Analysis and Prevention, 39(3), 483-490

Box 4.4. **Case study: Sweden**

Experience when linking accidents known to the police with injuries known by the hospitals in Sweden

In Sweden, the police reports all accidents with injured people who come to their attention, to the Swedish Transport Agency. In addition about 93 percent of all emergency hospitals are currently reporting to the Swedish Transport Agency injured persons who seek treatment for an accident. This is done on a voluntary basis, and after reimbursement from the Transportation Board. The cost is currently 1 million Euros per year. The data are entered by the local police and hospitals into STRADA, where a match immediately takes place. It is first done with the personal identification number and date. If a personal identification number does not exist, then time and place are used instead with a special algorithm. The requirements are a maximum distance of 1 km and the time difference is less than 24 hours. Each month a manual check takes place of accidents with large distance or time differences.

Our experience is that a very few accidents are matched that should not be. Also, but this is more difficult to verify, we do not miss many accidents that should been matched. In 2009, the police recorded 17 800 accidents with 25 500 injured persons. They are matched with 25 500 persons who sought treatment in hospital. Of these, 10 200 were matched (25%), that is known to both sources. It also means that 40% of those known to the police were also known to the hospitals, also 40% of those registered by the hospitals were registered by the police. One conclusion from this match is thus that the police reporting is far from comprehensive.

One explanation that the police registered injured people who do not appear in hospital data may be that they sought treatment at a Medical Centre and therefore are not become registered in any emergency hospital. It may also have been that their injuries were so mild, that they did not require any type of health care intervention. Some of the casualties that were reported by the hospitals may have occurred in places that are not included by the definition of a road traffic accident, for example, in the country, playground or cemetery. The main reason, however, would be that many injuries recorded by the hospitals are to cyclists who crashed and then failed to call the police but went to the hospital instead. That kind of accident with only a cyclist involved are seldom investigated by the police.

Overall, 15% of those who the police identified as injured, hospitals assessed as no injury. Of those who the police estimated slightly injured, 15 percent had a MAIS of 2 or higher and likewise those who the police estimated seriously injured 47 percent had a MAIS of 1 or lower. The conclusion is that the police often make an incorrect assessment of the injury severity.

From accidents that are matched we get a more complete description of both the circumstances that influenced the accident sequence, as the injuries to those involved. This results in a more accurate information of the real traffic safety problems which in turn facilitates the planning and prioritization of traffic safety measures.

In Sweden 62 out of 67 emergency hospitals currently reports to STRADA. Of the 21 counties in Sweden all emergency hospitals in 18 counties report to STRADA. In addition some hospitals in 1 county report injuries, but no hospitals in the last 2 counties yet report injuries to STRADA. In addition to 25 500 injuries the hospitals reported in 2009, the hospitals also reported some 6 300 pedestrians who slipped and fell down in the traffic environment. They are not included by the definition of a road traffic accident and consequently not matched with the reported injuries from the police. Of course both sources did report fatalities but they are not included in the numbers mentioned above.

Box 4.5. **Case study: Spain**

Experiences of hospital and police record linkage in Spain

Several experiences of police and record linkage have been developed in Spain since early 1998. In the city of Barcelona there is a routine record linkage system that links information from the Urban Police with a database gathered routinely from road injuries hospital emergency casualties, which includes both outpatients and inpatients (DUHAT System). Currently the process is carried out yearly and is used to study related crashes circumstances with their consequence for health, thus being an essential element of support for road safety policies in the city. Some specific studies were to identify factors associated with whiplash or traumatic brain injuries. It has also been used for estimating the real number of casualties and fatalities.

Within the framework of the SafetyNet project, funded by the European Union, hospital and police records were linked for one autonomous region for 6 months. Correction factors were derived to estimate the number of serious casualties. More recently this work has been extended to 6 autonomous regions for the years 2000-2007. The data allow estimation of fatalities within 30 days, and serious casualties at the national level. Main results are described below:

Estimates of road traffic serious injuries and fatalities in Spain, 2005-2007

The aim of this study is to estimate the number of serious traffic injuries and traffic deaths within 30 days in Spain, using a probabilistic record linkage between the records of the police database maintained by the General Traffic Directorate (DGT), and the Minimum Basic Hospital Discharge Data Set (HDR), in 6 autonomous communities.

The study population was traffic injuries and deaths in Spain between 2005 and 2007. We used two databases: 1) the database of traffic accidents and injuries of DGT, and 2) the HDR. Hospital admissions were defined as those traffic injury emergency admissions with a diagnosis of injury (codes 800-959.9 of the ICD-9-CM), an external cause code E810-E819 or E826, and/or an "insurer of an automobile accident "as the funding scheme. Readmissions were excluded. The method of linkage was probabilistic.

The linkage was feasible in 5 out of 6 autonomous communities (it proved to be not feasible for the Community of Madrid). The linkage process assigned a police record from 15.6% to 36.2% of hospital records depending of the region. The proportion of linked records that were recorded as 'slightly injured' in the police database ranged from 30.4% to 45.9%. This result illustrates the misclassification of severity on the police database, as every hospital record is expected to be classified as a "seriously injured" in the police records. The proportions of linked police records ranged from 4.5% to 8.9% for all injuries, and from 16.8% to 27.5% for serious injuries; For unlinked records, the proportion of serious injuries ranged from 10.3% to 21.7%.

The sensitivity of the connection was higher than 98.3% and the specificity higher than 89.7%. Despite the high values of sensitivity and specificity obtained, they must be interpreted with caution due to the limitations of the methods.

The number of serious injuries and fatalities were estimated as a result of the connection. The results show that police data underestimate the number of serious injuries from 19.5% to 25.2%. The number of fatalities within 24 hours, recorded in the police database, underestimates the number of fatalities within 30 days from 17.8% to 20.1%.

The number of serious injuries and traffic deaths should be interpreted cautiously, as the percentage of records matched differs significantly from the expected rate. Therefore, it is expected that the conversion factors obtained are higher than real, since more police and hospital records should have been matched.

Box 4.6. **Case study: United Kingdom**

Main results

Initial results of matching The Health and Social Care Information Centre Hospital Episodes Statistics (HES) and police data (STATS19) on road accidents suggest that the proportion of road accident casualties admitted to hospital that are known to the police has remained relatively constant over recent years.

There is, however, some evidence of an increase in the proportion of casualties admitted to hospital that are recorded as slightly injured in STATS19. This could be due to changes in police recording of severity, or changes in hospital admissions practices, or a combination of both factors.

Methodology

Matching was carried out using the fields common to both sources. These are: age; gender; home postcode; date (of accident in STATS19, of admission to hospital in HES); casualty type; casualty class; local authority (of accident in STATS19, of patient's home in HES).

In the STATS19 database, records corresponding to seriously or slightly injured casualties were selected – around 250 000 records per year. The HES dataset extract contained around 50,000 records per year – one record for each emergency admission episode where the external cause was recorded as a road transport accident. Those patients who subsequently died in hospital were excluded.

To reduce the computational effort required, a geographical variable based on regions was used to split each file into blocks for matching. Candidate matches were then identified within these blocks using gender, age and date. Some tolerance was allowed in matching, specifically:

- Date of admission was allowed to be up to two days after the accident date, to allow for cases where a patient did not arrive at hospital immediately.
- The STATS19 age was allowed to be between one and three years different from the age recorded by the hospital, to allow for the fact that the police sometimes have to estimate this.

The remaining variables – postcode, casualty type and casualty class – were then used to choose the most likely match from among the candidates generated, using pre-defined matching rules. Any cases where it was not possible to distinguish the most likely match were flagged as such and excluded from further analysis. For each pair of matched records, the degree of confidence in the match was assigned based on the degree of agreement between matching variables.

References

More information can be found in the Annual Reports on Road Casualties in Great Britain, available from DfT website (www.dft..gov.uk)::

Box 4.7. **Case study: United States**

The Federal program Crash Outcome Data Evaluation System (CODES) is facilitated by National Highway Traffic Safety Administration (NHTSA) in US DOT. Project sites in states enter into cooperative agreements with NHTSA and receive limited funding assistance but also receive indirect benefits of software, training, and others. Two state sites double as resource centres for the program. The program goal is to facilitate crash and outcome linkage at state level and to enable multi-state studies. Participation is voluntary but requires at least crash data, hospital data, and either emergency department data or emergency medical service data. The incorporation of other data sets is also encouraged. Currently 17 out of 50 states participate.

A customized MS Access-based software application is specifically developed for the linkage. The methodology uses multiple imputations for missing links to achieve population representation of links. Multiple imputations are also applied for missing data, and methods of analysis of multiply-imputed data are applied for analyses, usually using SAS PROC MIANALYZE or related procedures. Each state performs its own linkage and analyses with training and guidance provided through a CODES Technical Resource Centre facilitated by NHTSA. For certain studies data are standardized and contributed by participating states for multi-state studies coordinated through US DOT. A strategic plan is now underway to streamline the program while also expanding it. Linkage has shown success in many states, and added uses at national level are currently being cultivated. For more specifics and examples of usage, please refer to:

***www-nrd.nhtsa.dot.gov/Pubs/811181.pdf*.**

Box 4.8. **Case study: Use of hospital and police data in Israel**

Like other countries, the official statistics of road accidents in Israel is based on police data. A survey conducted by the Central Bureau of Statistics (ICBS) in 2004 supported the claim of under-reporting of seriously injured casualties by almost 50%. As of 2008, police data has been linked on a regular basis with hospital data in order to correctly estimate the real number of seriously injured casualties.

The hospital data on seriously injured casualties is collected by the Gertner Institute – The Center for Injury and Emergency Medicine Research, from 17 hospitals and is sent through a virtual vault to ICBS on a monthly basis. This project is founded by The National Road Safety Authority (RSA).

Upon linking the data, approximately 50% of the seriously injured were recorded by the police as slightly injured or were not recorded at all (in 2009 53%, in 2009 50%). The linked data are published in yearly publications of ICBS and RSA. In addition, ICBS provides special files for research. In the near future, ICBS plans to publish data more frequently. The accumulated data will enable examination of longitudinal data trends. The linked data will also be used to improve infrastructure by pointing to the dangerous road locations.

References

Alsop J, Langley J. *Under-reporting of motor vehicle traffic crash victims in New Zealand*. Accid Anal Prev 2001; 33(3):353-359.

Amoros E, Martin JL, Chiron M, Laumon B. *Road crash casualties: characteristics of police injury severity misclassification.* J Trauma 2007; 62(2):482-490.

Amoros E, Martin JL, Laumon B. *Estimating non-fatal road casualties in a large French county, using the capture-recapture method*. Accid Anal Prev. 2007 May;39(3):483-90.

Amoros E, Martin JL, Laumon B. *Under-reporting of road crash casualties in France* Accid Anal Prev 2006; 38(4):627-635.

Boufous S, Finch C, Hayen A, Williamson A. *Data linkage of Hospital and Police Crash Datasets in NSW.* Sydney: New South Wales Injury Management Research Centre, July 2008.

Chini F, Fortini M, Tuoto R, Farchi S, Giorgi Rossi P, Amato R, Borgia P. *Probabilistic record linkage for the integrated surveillance of road traffic injuries when personal identifiers are lacking.* Roma: instituto Nazionale di Statistica; 2008.

Cirera E, Plasencia A, Ferrando J, Arribas P. *Probabilistic linkage of police and emergency department sources of information on motor-vehicle injury cases: a proposal for improvement*. J Crash Prevention Injury Control 2001;2:229-37.

Cirera E, Plasencia A, Ferrando J, Arribas P. *Probabilistic linkage of police and emergency department sources of information on motor-vehicle injury cases: a proposal for improvement*. J Crash Prevention and Injury Control 2000; 2(3):229-237.

Clark DE. *Practical introduction to record linkage for injury research*. Inj Prev 2004; 10(3):186-191.

Cook LJ, Olson LM, Dean JM. Probabilistic *record linkage: Relationships between file sizes, identifiers, and match weights,* 2.001;40:196-203.

Cryer PC, Westrup S, Cook AC, Ashwell V, Bridger P, Clarke C. *Investigation of bias after data linkage of hospital admissions data to police road traffic crash reports.* Inj Prev. 2001 Sep;7(3):234-41.

Derriks H and Mak P (2007). *Underreporting of Road Traffic Casualties. Organisation for Economic Co-operation and Development*. IRTAD 2007. www.who.int/roadsafety/publications/irtad_underreporting.pdf [acceded 6/4/2010]

Domingo-Ferrer J, Torra V. *Distance-based and probabilistic record linkage for re-identification of records the categorical variables.* SCI JCR journals 2006. http://vneumann.etse.urv.es/publications/nonsci/acia28.pdf [acceded 30/3/2010]

Domingo-Ferrer J, Torra V. *Validating Distance-Based Record Linkage with Probabilistic Record Linkage*. SCI JCR Articles- Crises 2002. http://vneumann.etse.urv.es/webCrises/publications/isijcr/lncs2504.pdf [acceded 30/3/2010]

Federal Committee on Statistical Methodology. *Record linkage techniques—1985: Proceedings of the workshop on exact matching methodologies.* Arlington, Virginia, 1985. Available at: www.fcsm.gov/working-papers/RLT_1985.html [acceded 31/3/2010]

Federal Committee on Statistical Methodology. *Record linkage techniques— 1997: Proceedings of an international workshop and exposition.* Arlington, Virginia, 1997. Available at: www.fcsm.gov/working-papers/RLT_1997.html [acceded 31/3/2010]

Ferrante AM, Rosman DL, Knuiman MW. *The construction of a road injury database.Accid Anal Prev*. 1993 Dec;25(6):659-65.

Gonzalez RP, Cummings G, Mulekar M, Rodning CB. *Increased mortality in rural vehicular trauma: identifying contributing factors through data linkage*. J Trauma. 2006 Aug;61(2):404-9.

Gutoskie, P. (2003), *IRTAD Special Report: The availability of hospitalised road user data in OECD member countries.* OECD; 2003.

Jaro MA. *Matching and Record Linkage, in Business Survey Methods*, eds. B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, M. J. Colledge, and P. S. Kott, New York: Wiley, 1995 pp. 355–384.

Jaro MA. *Probabilistic linkage of large public health data files*. Stat Med 1995;14:491–8.

Jarvis SN, Lowe PJ, Avery A, Levene S, Cormack RM. *Children are not goldfish--mark/recapture techniques and their application to injury data.Inj Prev.* 2000 Mar;6(1):46-50.

Jeffrey S, Stone DH, Blamey A, Clark D, Cooper C, Dickson K et al. *An evaluation of police reporting of road casualties.* Injury Prevention 2009; 15(1):13-18.

Johnson SW, Walker J. The *crash outcome data evaluation system (CODES).* Report No. DOT HS 808 338. Washington: National Highway Traffic Safety Administration; 1996.

Keigan M, Broughton J, Tunbridge RJ. *Linkage of STATS19 and Scottish hospital in-patient data* — analyses for 1980–1995. TRL Report 420. Crowthorne: Transport Research Laboratory; 1999.

Lai CH, Huang WS, Chang KK, Jeng MC, Doong JL. *Using data linkage to generate 30-day crash-fatality adjustment factors for Taiwan*. Accid Anal Prev. 2006 Jul;38(4):696-702.

LinkageWiz *Data Matching Software*. LinkageWiz Inc. Adelaide. www.linkagewiz.com [acceded 31/3/2010]

Lujic S, Finch C, Boufous S, Hayen A, Dunsmuir W. *How comparable are road traffic crash cases in hospital admissions data and police records?* An examination of data linkage rates. Aust N Z J Public Health 2008; 32(1):28-33.

Morrison A, Stone DH. *Capture-recapture: a useful methodological tool for counting traffic related injuries?* Inj Prev. 2000 Dec;6(4):299-304.

Neutel CI, Johansen HL, Walop W.*'New data from old': epidemiology and record-linkage*.Prog Food Nutr Sci. 1991;15(3):85-116. Review.

Newcombe HB, Kennedy JM, Axford SJ, et al. *Automatic linkage of vital records*. Science 1959;130:954–9.

Newcombe HB. *Handbook of record linkage: methods for health an statistical studies*. Administration and business. Oxford: Oxford University Press, 1988.

Newcombe HB. *Handbook of record linkage: methods for health an statistical studies.* Administration and business. Oxford: Oxford University Press, 1988.

NHTSA (2009). *The Crash Outcome Data Evaluation System (CODES) and applications to improve traffic safety decision making.* www-nrd.nhtsa.dot.gov/Pubs/811181.PDF [acceded 6/4/2010]

Novoa AM, Pérez K, Santamariña-Rubio E. *Dirección General de Trafico*. Ministerio del Interior. Diciembre 2009. [In Spanish]. www.dgt.es/was6/portal/contenidos/documentos/seguridad_vial/estudios_informes/Estimacion_heridos_graves_y_muertos_final.pdf.

Pérez K, Cirera E, Borrell E, Plasencia A, on behalf of the work group of the Spanish Society of Epidemiology on the Measuring of the Impact on health of road traffic accidents in Spain. *Motor vehicle crash fatalities at 30 days in Spain*. Gac Sanit. 2006;20(2):108-15.

Petridou E, Yannis G, Terzidis A, Dessypris N, Germeni E, Evgenikos P, Tselenti N, Chaziris A, Skalkidis I. *Linking emergency medical department and road traffic police casualty data: a tool in assessing the burden of injuries in less resourced countries*.Traffic Inj Prev. 2009 Mar;10(1):37-43.

Reurings, M.C.B., Bos, N.M. (Seriously injured road crash casualties in the Netherlands in the period 1993-2008; The real number of in-patients with a minimum MAIS of 2) (Dutch Language Only) SWOV report R-2009-12. 2009. www.swov.nl/rapport/R-2009-12.pdf. [acceded 8/10/2010]

Roos LL, Wajda A. *Record linkage strategies.* Methods of information in Medicine, 1.991;30:117-123

SafetyNet. *Building the European Road Safety Observatory.* Deliverable 1.15. Final report on task 1.5. SafetyNet; 2008. http://ec.europa.eu/transport/wcm/road_safety/erso/safetynet/fixed/WP1/D1.15_Estimation_real_number_of_road_accident_casualties_final%20report_3.pdf [acceded 6/4/2010]

Simpson HF. *Comparison of hospital and police casualty data: a national study*. TRL Report 173. Crowthorne: Transport Research Laboratory; 1996.

Tercero F, Andersson R. *Measuring transport injuries in a developing country: an application of the capture-recapture method.* Accid Anal Prev. 2004 Jan;36(1):13-20.

Van HT, Singhasivanon P, Kaewkungwal J, Suriyawongpaisal P, Khai LH. *Estimation of non-fatal road traffic injuries in Thai Nguyen, Vietnam using capture-recapture method.* Southeast Asian J Trop Med Public Health. 2006 Mar;37(2):405-11.

Ward and Lyons (2006), *Underreporting of road casualties, Phase 1,* Dft, London.

Winkler WE (2010). *The state of record linkage and current research problems*. Available at http://eprints.kfupm.edu.sa/71379/1/71379.pdf [acceded 31/3/2010

CHAPTER 5
Estimating the true number of serious road casualties

Abstract

Chapter 5 covers the simple "capture- recapture" or dual system method, which uses data linkage to estimate a total population. This chapter also includes three case studies, and a number of references for readers who would like to explore this technical area in more detail.

5.1. Introduction

When two or more registrations of road casualties exist, such as police and hospital records, it is possible to estimate the number of subjects missed by both registrations, and hence estimate the total number of subjects of interest by using the capture-recapture approach, providing a number of assumptions are satisfied.

Two papers provide a very interesting literature review of the capture-recapture approach, with application to health events in humans (Hook and Regal, 1995; IWGDMF, 1995a); the health event of interest to us is being injured or killed in a road crash.

5.2. History

The capture–recapture approach came from animal zoology, in order to estimate a fish or bird population. For instance, if one wants to estimate the number of fish in a pond, the idea is to make a first catch (capture), to mark / tag the captured fish, release them in the pond, then make another independent catch and count how many are captured for the first time and how many for the second time (recapture). From the frequencies of the three sub-groups (i.e. captured at the 1st catch only, captured at the second catch only, captured at both catches), it is possible to estimate the number of fish that were never caught, and hence the total number of fish in the pond. The method uses basic probability formulae, and is based on a number of assumptions.

So, Similar methods have been used in a number of fields, including demography and health. Within injury epidemiology it has been applied in road traffic injury, usually focusing on some subgroups, such as children (Jarvis et al., 2000; Roberts and Scragg, 1994), teenagers and young adults (Morrison and Stone, 2000), pedestrians and cyclists (Dhillon et al., 2001), truck drivers (Meuleners et al., 2006), on specific injuries (Chiu et al., 1993) or on some specific areas: towns (Razzak and Luby, 1998: Tercero and Andersson, 2004; Van et al., 2006) or islands (Aptel et al., 1999). These studies are summarised in table 5.1.

There are other names for "capture-recapture" approach: "mark-recapture" mainly in animal ecology, "dual system method", "dual-record systems" or "multiple records systems" mainly in demography. In epidemiology "ascertainment corrected rates" may imply that they are based on the capture-recapture method.

5.3. Current views of the capture-recapture methodology

The use of the capture-recapture method in epidemiology, i.e. concerning human health events is subject to some extreme points of view. Some researchers are very enthusiastic about it and argue that it is better to apply it than not, and that every registry wishing to evaluate its coverage should use it. On the contrary, some authors are very reluctant to use it. The issue is quite controversial; in a way, the method suffers from its simplicity of computation: it is sometimes used very quickly without really worrying about whether the underlying assumptions are met or not. Other authors have been trying to make it less controversial, tackling all aspects of the method one by one. A list of seventeen recommendations has been written (Hook and Regal, 1999; Hook and Regal, 2000b) dealing with the implementation of the method, checking that the underlying conditions are met, and if not, or if only partially, what are the consequences; how one should present the analysis itself, and how one should display the results. In the field of injury epidemiology, a few authors are very sceptical (Jarvis et al., 2000; Morrison and Stone, 2000); however, their discussion of the underlying conditions of the method is not detailed.

Table 5.1. **Studies in traffic safety which have used capture-recapture to estimate the real number of casualties, by year of publication**

Author, year published	**Country, state, year**	**Restriction**	**Police records** or other **(frequency)**	**Medical records (frequency, type)**	**Overlap (frequency)**	**Aggregate**	**Capture-recapture estimate***	**Method**	**Strata or covariates**	**Comments**
(Zavareh et al. 2008)	Iran, west Azarbaijan (2004-05)	Fatalities	N=669 Death register	N=665 Forensic medicine register	N=437	N=897	N=1018	Simple 2-list method (NUE), by strata	Road user type OR age, OR sex	Reporting rate depends on road user type
(Amoros et al. 2007)	France, Rhone, 2001	Non-fatal casualties	N=4135	N=10636 Road trauma registry	N=2813 to 3112 (3 matching standards)	N=11,958 to 11,659	N=16,129 to 14,471	2-list method With stratification	Injury severity, road user type, AND third party	3 record-linkage scenarios
(Amoros 2007)	France, Rhône, 1996-2001	Non-fatal casualties	N=36,333	N=83,448 Road trauma registry	N=26,064	N=93,708	N=113,823	Multinomial logit, with covariates	Injury severity, road user type, Third party, road type, AND police type	
Reurings M, To be published	*Netherlands 1993-2008*	*MAIS2+*	*N=1,182,128 persons injured or killed in a traffic accident (N=681,635) or involved as a driver in an injury accident (N=500,493)*	*N=1.297.629 persons admitted to hospital for external cause E800-848 + 880-899 + 928 + 929 + 958 + 988 (of which N=313.178 in E810-829)*	*N=162.135 of which N=126.419 with MAIS=2 or more)*		*NA*		*MAIS OR Mode (E_code) OR Region*	

(Van HT et al. 2006)	Vietnam, Thai Nguyen (city) 2000-2004	Non-fatal casualties	N=1373	N=6069	N=254 to 748	N=7188 to N=6694	N=32,806 to N=11,140	2-list method		4 matching standards
(Meuleners et al. 2006)	Western Australia, July 1999-2000	Heavy vehicle crashes	N=3 deaths	N=2	N=1	N=4	N=5	Simple 2-list method	Fatal/ seriously injured separately	Very small sample sizes
Ibid.			N= 24 seriously injured	N=35 Hospital records	N=14	N=45	N=59	Simple 2-list method	Fatal/ seriously injured separately	small sample sizes
(Dhillon et al. 2001)	USA, California, 1992-1995	Childhood pedestrians & cyclists	N=1015	N=474 ED logs, hospital admission	N=379	N=1110	N=1269	Simple 2-list method	Pedestrians/ cyclists separately	
(Tercero and Andersson 2004)	Nicaragua, Leon, 1993	Fatalities	N=32	N=13	N=7	N=38	N=57	Simple 2-list method	Fatal / non-fatal/ inpatients separately	
Ibid.		Inpatients and outpatients	N=183	N=1334	N=34 non-fatal	N=1483	N=7017	Simple 2-list method	Fatal / non-fatal/ inpatients separately	
Ibid.		Inpatients only	N=183	N=106	N=16	N=1483	N=1157	Simple 2-list method	Fatal / non-fatal/ inpatients separately	
(Razzak and Luby 1998)	Pakistan, Karachi, 1994	Fatalities	N=544 deaths	N=972 deaths	N=68 to 220 deaths (4 matching standards)	N=819 to 667	N=3116 to 963	Simple 2-list method		4 record-linkage scenarios

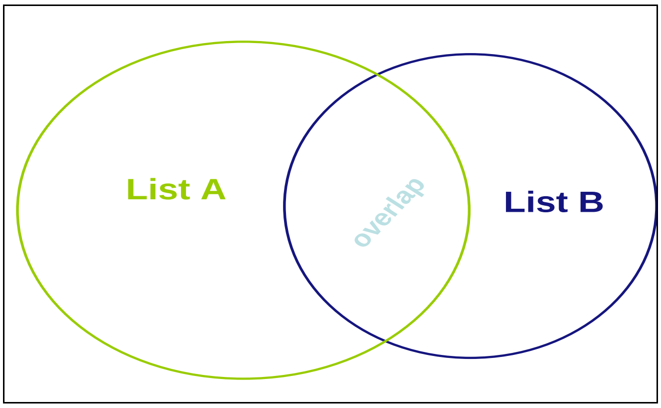
Ibid.	Pakistan, Karachi, 1994	Injured	N=793 injured	N=18936 Non-gov ambulance service	N=39 to 98 injured (4 matching standards)	N=2802 to 2743	N=47,140 to 18,760	Simple 2-list method		4 record-linkag scenarios
(Jarvis et al. 2000)	UK, Northumbria, 1990-1995	Children<15	N=1009	N=836	N=357 or N=488 (2 matching standards)	N=1488 or 1357	N=2743 or 2067	Simple 2-list method, with stratification (on age and road user)	Road user type, sex and age	Hetero-geneity of capture on age and road user type (not on sex)
(Morrison and Stone 2000)	UK, Scotland, 1995	Fatalities, 15-24 y.o.	N=99 deaths	N=97 deaths	N=69 to 92 (4 matching standards)	N=127 to 104	N=139 to 104	2-list method		(4 matching standards)
Ibid.		Seriously injured, 15-24 y.o.	N=1290	N=1458 hospital discharges	N=616 to 955 (4 matching standards)	N=2132 to 1793	N=3052 to 1969	2-list method		(4 matching standards)
(Roberts and Scragg 1994)	New Zealand, 1992-1994	Pedestrian children <15	N=206 (police? Or other?)	N=238	N=184	N=260	N=266	Simple 2-list method,		
(Aptel et al. 1999)	France, la Réunion (island), 1993-94		N=87	N=137 Injury survey	N=34	N=190	N=346	2-list method (NUE)		
(Chiu et al. 1993)	USA, (1 month)	Fatalities	N=253	N=260 (mortality statistics)	N=241	N=299	N=307	Simple 2-list method		

*NB: the method used in the Netherlands is not capture-recapture; it is detailed at the end of this chapter.

5.4. Method

The capture recapture method is based on at least two registration sources (or lists) of subjects of interest. We present the method for two lists, but it can be expanded to three or more lists. It can be displayed in the following way:

Figure 5.1. **Schematic display of two registration sources (two 'lists')**



However, with this display, it is less obvious that subjects of interest may be missing from both lists. In the display below, i.e. a cross-tabulation between list A and list B, it becomes obvious that some subjects of interest may be missing from both lists:

Table 5.2. **Frequency distribution according to being present or absent in the two registrations**

		List B		
		yes	no	
List A	yes	n_{AB}	$n_{A\bar{B}}$	n_A
	no	$n_{\bar{A}B}$	$n_{\overline{AB}}$	
		n_B		n

Here, if one assumes that the probability of being registered in list A is independent of the probability of being registered in list B, this translates[7] into: $n_A/n = n_{AB}/n_B$ and from this we obtain the intuitive Petersen estimate : $\hat{n} = \dfrac{n_A \times n_B}{n_{AB}}$. It also happens to be the maximum likelihood estimator (MLE).

The maximum likelihood estimator is only asymptotically unbiased (i.e. unbiased for large sample sizes). Another estimator has hence been developed : the Nearly Unbiased Estimator (Wittes, 1972):

7. In probability theory, if two events A and B are independent, then P(A)=P(A|B), where P(A)= probability of occurrence of event A, and P(A|B) is the probability of A conditional on B. Besides, by definition P(A|B)= P(A and B) / P(B), so that we get : P(A)= P(A and B)/P(B). A probability can be estimated by the corresponding observed proportion, so that P(A) is estimated by n_A/n and P(A and B) /P(B) is estimated by (n_{AB}/n) / (n_B/n) which simplifies into n_{AB}/n_B. We hence get $n_A/n = n_{AB}/n_B$.

$$\hat{n} = \frac{(n_A + 1) \times (n_B + 1)}{(n_{AB} + 1)} - 1$$

These are the first "simple" estimators; they apply to the case where there are only two registration sources (also called the "2-list method"), and a simple pocket calculator is enough.

The method can be elaborated, for instance by including some stratification (see further on, under the condition of homogeneity of capture).

More recently, explicit modelling has been developed. This enables covariates associated with capture probability to be accounted for. It also makes it possible to allow for interdependency between lists by estimating interaction terms, but this is only possible when there are at least three registrations. The most widespread modelling in animal ecology is based on a Poisson distribution, together with a log-linear model (Chao et al., 2001; Cormack, 1989; IWGDMF, 1995a; IWGDMF, 1995b); this has been developed in the framework of the analysis of contingency tables. Another modelling approach is based on the multinomial distribution and the logistic model (Alho, 1990; Tiling and Sterne, 1999). There also exist some non-parametric development, as well as some Bayesian approach (Hook and Regal, 2000a; Madigan and York, 1997).

5.5. Assumptions

The capture-recapture approach is based on four key assumptions, and also on two implicit ones.

The four key assumptions are (Hook and Regal, 1995; IWGDMF, 1995a):

1) Closed population.
2) Perfect identification of subjects common to both registrations.
3) Independence between the registrations.
4) Homogeneity of capture by a given registration.

The two implicit assumptions are (Gallay *et al*., 2002):

5) Same geographical area and same time period.
6) Perfect identification of the subjects of interest.

We discuss each of these in detail below, starting with the implicit ones.

Assumption n°5: same geographical area and same time period covered by the two (or more) registrations

This is indeed obvious, but it is worth taking a close look at it.

When we deal with police and hospital data, especially at a sub-national level, it is likely that the police area is based on administrative boundaries whereas hospitals have "catchment" areas. These are likely to be different, so one should try to define and use common areas.

For the time period, one has to allow for a small lag time between the date of crash and the date of health service attendance; this is usually allowed for in probabilistic record-linkage, but one should not forget about it for crashes occurring at the very beginning / end of the time period under study.

Assumption n°6: perfect identification of the subjects of interest; in other words no classification error

The criteria for defining a subject of interest should be very precise, and should be the same for the two (or more) registrations.

For road traffic casualties, the criteria would be at least: "injured or killed" and "in a road traffic accident". In some countries, there may be the supplementary criteria such as "at least a motor vehicle involved" (whereas in other countries, it is merely "at least a vehicle involved").

For instance, in the French Rhone trauma registry, the inclusion criteria are much broader than for the police: people injured on their own while roller-skating are included; they were hence excluded from the registry data before applying capture-recapture (in fact, before conducting the record-linkage).

If the inclusion criteria are not exactly the same between the two registrations, this should be kept in mind in the interpretation of the results.

In the Netherlands this assumption is not met. It is not possible to identify the serious road injuries (with a Maximum AIS[8] severity score (MAIS) of at least 2) in the police registry because the police do not register the MAIS; in the hospital registry it is not possible to identify all casualties of road crashes, because not all these casualties have the correct E-code in the registry and are therefore not recognizable as road casualties. Therefore SWOV developed a different method to estimate the number of serious road injuries in the Netherlands (Reurings and Stipdonk, 2010, see also Box 5.3).

Assumption n°1: closed population

This means that there should be no entry or loss between the two registrations. In animal ecology this means no death or birth between the two captures (which in this framework occur at different points of time).

As regards road traffic casualties, this means that the road casualty should not "escape" the possibility of being registered (by any of the two or more sources). It is however likely to happen for those people who are slightly injured, in a crash at some distance from their place of residence; they might wish to attend the hospital closest to their home rather than closest to the place of crash. Similarly, some people may wish to avoid the police: typically when they know they have violated road traffic legislation at the time of the crash (crossing at red light, drinking and driving…). These losses lead to an under-estimation of the number of casualties.

Assumption n°2: there should be perfect identification of subjects of interest common to both sources

This depends on the reliability of the record-linkage. This in turns depends on the quality and the discriminative efficiency of the linking variables. The quality of the record-linkage is difficult to assess without the availability of a gold standard. It can however be evaluated to some extent. For instance, it is possible to evaluate error rates of the linking variables and compare them with usual error rates; it is also possible to estimate the probability and frequency of false positives and of false negatives links.

It may also be worth conducting some sensitivity analysis, that is to say applying capture-recapture on different linking standards, using less and less restrictive matching criteria. This has been done in a number of the studies on road traffic casualties (see table 5.1). For instance, in estimating the number of road traffic casualties in Scotland among young people aged 14-25 (Morrison and Stone, 2000), standard A was the result of the record-linkage requiring exact matching on age, sex, day of crash, month of crash, year of crash and crash location; standard B allowed for the day of crash in police data to be one day before the day in the hospital file; standard C only allowed for ages in the two files to differ by one year; standard D allowed for both of these discordances. The table below provides another example of applying capture-recapture on three matching standards.

8. Abbreviated Injury Scale.

Table 5.3. **Registrations frequencies and estimates of stratified capture-recapture approach, for three linking standards, Rhône, France, 2001**

Record linkages	police data	medical registry	overlap	altogether	capture-recapture estimate
standard 1	4 135	10 636	2 813	11 958	16 130
standard 2	4 135	10 636	2 956	11 815	15 294
standard 3	4 135	10 636	3 112	11 659	14 471

Standard 2= standard 1 + additional matching using names

Standard 3= standard 1 + correction for false negatives and false positives

Assumption n°3: There should be independence between the registrations

The subjects' probability of being registered in one source should be independent on the probability of being registered by the other source. This is the basic underlying assumption for establishing the Petersen estimator (as previously mentioned).

In practice, it is often the case that the two (or more) registrations are not independent. It has been established (IWGDMF, 1995a) that, when there is a negative dependency, the estimator of the total number of subjects of interest is biased upwards; when there is a positive dependency, which is the most common situation, the estimator is biased downwards.

For road casualties, police and hospital registrations are likely to be positively dependent. This is especially the case for seriously injured people. Typically, if the police are first on the scene, they are highly likely to alert the pre-hospital health services. The reverse, that is to say the pre-hospital service alerting the police, may be less frequent. This positive dependency implies that the obtained estimate of the number of road traffic casualties is likely a lower bound.

When there at least three registrations, it is possible to take account of two-by-two dependencies between registrations using a log-linear model and including corresponding interaction terms. However the order of the interaction terms must be lower than the number of different registration sources; otherwise the model is non-identifiable. In particular, it is not possible to estimate an interaction term when there are only two registrations.

Assumption n°4: There should be homogeneity of capture by a given source/registration

This means that all subjects of interest should have the same probability of being registered by a given source; the probability of being registered by the other source could be different (and usually is), but still has to be equal between all the subjects of interest.

This is usually not the case. Typically, the degree of severity plays a role in the probability of being registered: the higher the severity, the higher the probability of being registered. This is seen for road traffic casualties: the most seriously injured casualties have a higher probability of being reported than the slightly injured. But other characteristics usually influence the reporting probability: the number of vehicles involved in the crash, the road user type (Derriks and Mak, 2007; Elvik and Mysen, 1999; Hvoslef, 1994). Indeed, multi-vehicle crashes are more likely to be reported than single-vehicle crashes; injured cyclists are less likely to be reported than injured car occupants (all other things being equal).

In such cases, the homogeneity of capture is only valid within sub-groups (for example within cyclist, within car occupants, etc). It is however possible to account for lack of homogeneity. The first way to do it is to stratify on these sub-groups, i.e. to stratify on the variable which is associated with the probability of registration, and which defines the sub-groups. More precisely, one should estimate the number of subjects of interest in each stratum, and then one should sum up the estimates obtained over the strata to get the total number of subjects of interest.

As an example of implicit stratification, most studies of road traffic casualties estimate fatal and non-fatal casualties separately. The registration probability is indeed much higher for fatalities than for non-fatal casualties.

This stratification approach can however in practice only deal with 2 or 3 variables associated with probability of registration. The strata are indeed defined by the combination of the 2 or 3 variables, and hence the cells frequencies get smaller and smaller with the number of stratification variables, up to a point where the estimation would not be valid. The table below gives an example of stratified capture-recapture.

Table 5.4. **Capture-recapture stratified on road user type, third party involvement and injury severity (NISS), linking standard n° 1, Rhône, France, non –fatal casualties, 2001**

Road user type and third party Strata	*Injury severity*	*Police data*	*medical registry*	*overlap*	*altogether*	*capture-recapture estimate*
pedestrians	NISS 1-3	286	589	187	687	897
	NISS 4-8	138	192	101	230	264
	NISS 9+	125	138	99	164	174
cyclists	NISS 1-3	62	836	47	850	1090
	NISS 4-8	21	322	15	328	448
	NISS 9+	18	75	14	79	96
M2W users	NISS 1-3	299	682	225	756	905
With third party	NISS 4-8	191	226	117	300	368
	NISS 9+	149	177	125	201	211
Without 3rd party	NISS 1-3	27	699	14	712	1310
	NISS 4-8	34	266	16	285	555
	NISS 9+	25	110	22	113	123
Car occupants	NISS 1-3	1576	3644	1019	4201	5636
With third party	NISS 4-8	325	440	228	538	629
	NISS 9+	158	179	116	222	245
Without third party	NISS 1-3	322	1108	211	1219	1691
	NISS 4-8	130	244	89	285	357
	NISS 9+	82	149	71	161	174
others	NISS 1-3	134	414	71	477	776
	NISS 4-8	21	98	15	105	137
	NISS 9+	13	44	13	44	44
total						**16130**

NB : cyclists were not split into the two strata with / without third party because cell frequencies were too small.

A better way to deal with heterogeneity of capture is in fact to use explicit modelling and to include as covariates the variables that influence registration probability. The number of covariates one can take into account is hence higher.

As mentioned already, the most widespread modelling (at least in animal ecology) is based on a Poisson distribution, together with a log-linear model; this has been developed in the framework of the analysis of contingency tables. Another modelling approach is based on the Multinomial distribution and the logistic model; this seems more common in epidemiology.

In France, the estimation using the multinomial logit model was based on a paper by Tilling and Sterne describing how to include covariates (Tilling and Sterne, 1999). The French model included five major variables influencing the registration probability: injury severity, road user type, whether a third party was involved in the crash, road type and police type.

5.6. Conclusion

It can be very easy and quick to calculate a capture-recapture estimate (especially using the two-list method) and it is hence very tempting. However a correct implementation of the method is not so quick and easy. One must really consider the six assumptions one by one and see how much they are met or not, and whether some departure away for the assumptions can be taken into account.

In the case of road traffic casualties (killed and injured), two recommendations can be made. The first recommendation is at least to stratify on the characteristics that are associated with police under-reporting. If there is no country-specific information on the characteristics associated with police under-reporting, one can consider that injury severity, road user type, and whether the crash was a single-vehicle or multi-vehicle crash are important stratification characteristics (Elvik and Mysen, 1999).

A second recommendation deals with record-linkage. Indeed the capture-recapture method relies on a perfect record-linkage between the two or more sources of registration. This is of course difficult to achieve. Hence, unless there is formal checking that the quality of the record-linkage is good, applying capture-recapture should be based on different linking standards, as had been done in the literature (see table 5.1). The capture-recapture estimate is indeed very sensitive to the output of the record-linkage.

Box 5.1. **The experience in Spain**

Estimated number of road traffic fatalities at 30 days in the city of Barcelona: assessment of the method capture-recapture

Authors: Marc Marí-Dell'Olmo, Catherine Pérez, Isabel Ricart and Carme Borrell

Background

Fatality in a traffic crash is defined as a person "who dies at the scene of the collision or in the 30 days following the crash". According to this definition a follow up within 30 days should be done for all people injured in a traffic crash. Nevertheless, in practice it is not feasible to follow up all collisions because of resource constraints.

In Catalunya the "*Institut de Medicina Legal de Catalunya*" (IMLC) -Forensic Institute- collaborates with "*Servei Català de Trànsit*" -Autonomic Transport Authority- notifying, in an active way, all the traffic deaths. In the rest of Spain, crash-fatality adjustment factors are used to estimate these fatalities, though it is known that these factors involve important biases.

Objective

The aim of the present study was to assess the capture-recapture method, by using police and hospital records, to see if this is an effective tool for estimating traffic deaths occurring within 30 days.

Setting, subjects and methods

A cross-sectional design study was carried out. The population of the study was the traffic deaths, in the place of the crash or in the 30 following days in Barcelona city during the year 2004. The information sources used were the reports of traffic accidents of the "Guardia Urbana de Barcelona" (GUB) -police data-, the records of the emergency services of the most important Barcelona hospitals (DUHAT), the "Conjunto Mínimo y Básico de Datos de Altas Hospitalarias" (CMBDAH) -data on hospitalized patients- and the autopsy reports of IMLC. All the sources were matched by a probabilistic linkage and a later manual review. A descriptive analysis of the linkage results was carried out and three methods were used to estimate the deaths occurring within 30 days of the injury event.

First of all, crash-fatality adjustment factors, obtained by "Dirección General de Tráfico" -Traffic department- for the year 2000 were used. Secondly, the records of all the sources were aggregated but counting only once the overlapping records among them. Thirdly, the capture-recapture method was applied, on the one hand comparing the two sources and applying Lincoln-Petersen and Chapman estimators, and on the other hand analyzing the different subsets of all the three sources, obtained from the different linkages by log-linear models. The obtained results were contrasted with the deaths gathered by the IMLC for comparison.

Results

From the IMLC, between 47 and 71 traffic deaths were registered. The uncertainty is because some of the crashes may not have been in the city. The capture-recapture method overestimated significantly the number of traffic deaths within 30 days. This method was applied to two information sources (89 deaths, Confidence Interval - CI - 95 % = 75-103) and to three information sources by log-linear models (103 deaths CI 95 % = 94-112). Estimations were also obtained using crash-fatality adjustment factors (57 deaths) and when the deaths registered by the GUB or the DUHAT were aggregated (67 deaths).

Conclusions

It is important for public health studies to produce unbiased estimations of the number of traffic fatalities, because they are useful to plan necessary health resources. They also allow more reliable determination of health costs.

The results obtained in this study do not validate the use of the capture-recapture method to estimate the number of traffic fatalities at 30 days in Barcelona. That is due to the violation of all the assumptions that must be checked to use the capture-recapture method. We suggest as an alternative estimating the total number of fatalities by aggregating the number of fatalities identified by a police information source with those identified by a surveillance system based on hospital emergency records.

Reference

M. Marí-Dell'Olmo, K. Pérez, I. Ricart, C. Borrell. Estimación del número de fallecidos a 30 días por lesiones de tráfico en la ciudad de Barcelona: valoración del método de captura-recaptura. XII Congreso de la Sociedad Española de Salud Pública y Administración Saniraria (SESPAS). Barcelona 2007. Gac Sanit. 21 (Extraordinario 2): 14. (Oral Presentation).

Box 5.2. **Estimating the number of serious road casualties in Great Britain using the capture – recapture method**

(contribution from Matthew Tranter)

The capture-recapture method can be applied to estimate the number of serious road casualties in Great Britain using the results of linking data from police records (STATS19) and hospital admissions (HES) data for England.

Overall, the estimated number of serious casualties in Great Britain is around 105 thousand (based on the 1999 to 2007 average). This compares with around 33 thousand serious casualties recorded in the police data for the same period. The key points suggested by these estimates are that:

- Around a third of serious road casualties are recorded in the police STATS19 dataset, with variations by type of road user [chart 1].
- The police data appears to provide a more accurate reflection of the trend than the hospital admissions data [chart 2].

Chart 1. **Proportion of total serious casualties known to police or admitted to hospital, based on capture-recapture estimate for England 1999-2007**

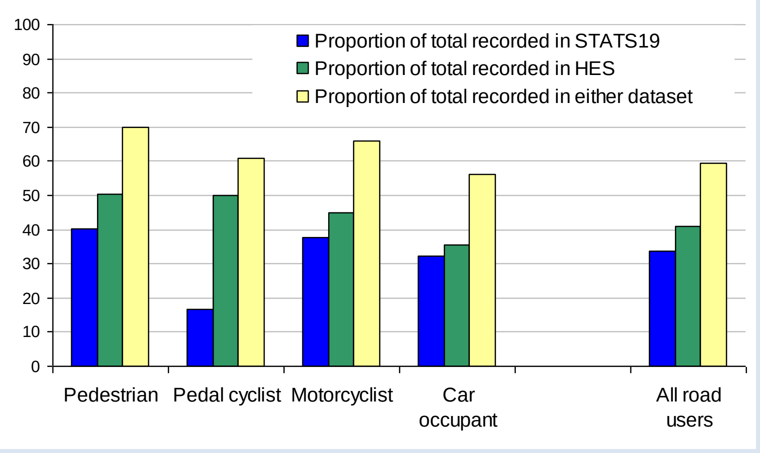
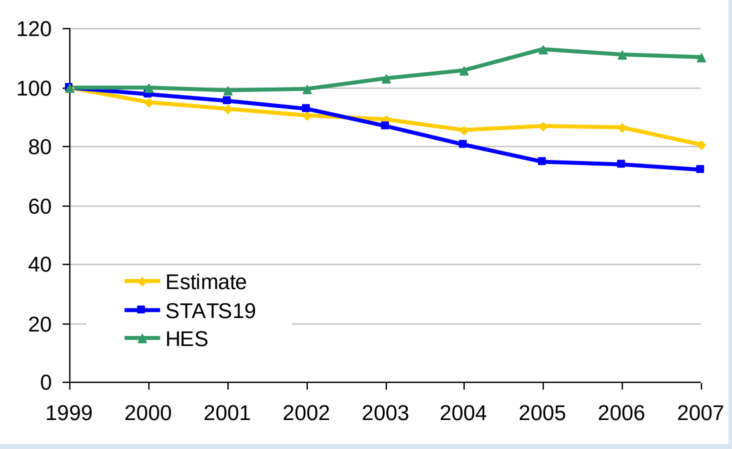


Chart 2. **Trend in serious casualties – estimated, STATS19 and HES (indexed 1999 = 100)**



These results are broadly in line with the findings of previous research into levels of reporting of road accidents in Great Britain. As some of the key assumptions of capture-recapture do not hold when the method is used to estimate road traffic injuries, these results should be interpreted with caution. But despite the limitations, this work shows the potential value of such estimation both in illustrating the incompleteness of police data, and estimating trends which take account of data from different data sources.

Box 5.3. **The experience in the Netherlands – 'Probabilities of registration"**

In the Netherlands, like other countries, the first step in estimating the number of serious road injuries is linking the police registry to the hospital registry, to identify the serious road injuries that are registered in both. Generally the next step would be to apply capture-recapture techniques to the results of the linking procedure in order to estimate the number of serious road injuries not registered at all. This is however not possible in the Netherlands, because Condition 6 (perfect identification of the subjects of interest) is not met.

In the police registry the injury severity expressed in MAIS is not recorded, but the Dutch definition of 'serious road injury' is based on MAIS. Therefore, it is by definition not possible to identify the serious road injuries in the police registries. Besides this, the injury severity that is recorded by the police, i.e. whether or not a road casualty is hospitalized, is not necessarily the truth. Research has shown that only 67% of the casualties that were hospitalized according to the police was indeed hospitalized. On the other hand, there are also serious road injuries in the hospital file, matched to a police record, who were according to the police not hospitalized, or not even injured at all.

The hospital registry does contain information on the injury severity expressed in MAIS.

Another advantage of the hospital registry is that is a complete registry (or at least it is assumed to be), meaning that all in-patients of Dutch hospitals are recorded. Hence also all serious road injuries are also included in the hospital registry. The problem is however that they are not all recognizable as such. It turns out that about 15% of the serious road injuries has received an incorrect E-code in the hospital registry.

The completeness of the hospital file and the 'perfect' identification of subjects common to both registrations implicate that the remainder police file is empty. Therefore, instead of using capture-recapture, SWOV developed a new method to estimate the number of serious road injuries. This method is based on 'probabilities of registration'. By using the results of the linking procedure the following probabilities are estimated:

- The probability of a serious road injury to be registered in the police registry (not necessarily as hospitalized).
- The probability of a serious road injury to be recognizable as such in the hospital registry.

where the effects of mode of transport, region and injury severity on these probabilities are taken into account. Stratification to these variables, it done to meet condition 4 (homogeneity of capture).

Further details of this method are to be published in *Reurings, M.C.B. Stipdonk, H.L. (2011), Estimating the number of road injuries in the Netherlands, Annals of Epidemiology 2011; 21: 648-653*

For results, see also Box 2.1 and the SWOV website.

www.swov.nl/cognos/cgi-bin/ppdscgi.exe?DC=Q&E=/English/Accidents/Real%20number%20of%20seriously%20injured%20people

References

Abeni, D. D., Brancato, G. and Perucci, C. A., 1994. *Capture-recapture to estimate the size of the population with human immunodeficiency virus type 1 infection*. Epidemiology 5 (4), 410-414.

Alho, J., 1990. *Logistic regression in capture-recapture models*. Biometrics 46, 623-635.

Amoros, E., 2007. *Les blessés par accidents de la route : estimation de leur nombre et de leur gravité lésionnelle, France*, 1996-2004 ; *modélisation à partir d'un registre médical (Rhône) et des données policières (France).* Université Lyon 1.

Amoros, E., Martin, J. and Laumon, B., 2007. *Estimating non-fatal road casualties in a large French county, using the capture-recapture method.* Accident Analysis and Prevention 39 (3), 483-490.

Amoros, E., Martin, J. and Laumon, B., 2007. *Estimating non-fatal road casualties in a large French county, using the capture-recapture method.* Accident Analysis and Prevention 39 (3), 483-490.

Aptel, I., Salmi, L. R., Masson, F., Bourdé, A., Henrion, G. and Erny, P., 1999. *Road accident statistics: discrepancies between police and hospital data in a French island.* Accident Analysis and Prevention 31 (1-2), 101-108.

Chao, A., Tsay, P. K., Lin, S. H., Shau, W. Y. and Chao, D. Y., 2001. *The applications of capture-recapture models to epidemiological data.* Statistics in Medecine 20 (20), 3123-3157.

Chiu, W. T., Dearwater, S. R., McCarty, D. J., Songer, T. J. and LaPorte, R. E., 1993. *Establishment of accurate incidence rates for head and spinal cord injuries in developing and developed countries: a capture-recapture approach*. Journal of Trauma 35 (2), 206-211.

Cormack, R. M., 1989. *Loglinear models for capture-recapture.* Biometrics 45, 395-413.

Derriks, H. and Mak, P. (2007), *"Underreporting of road traffic casualties,"* IRTAD special report, 39.

Dhillon, P. K., Lightstone, A. S., Peek-Asa, C. and Kraus, J. F., 2001. *Assessment of hospital and police ascertainment of automobile versus childhood pedestrian and bicyclist collisions*. Accident Analysis and Prevention 33 (4), 529-537.

Elvik, R. and Mysen, A. B., 1999. *Incomplete accident reporting; meta-analysis of studies made in 13 countries*. Transportation Research Record 1665, 133-140.

Fienberg, S. E., 1992. *Bibliography on capture-recapture modelling with applications to census undercount adjustments*. Survey methodology 18, 143-154.

Fisher, N., Turner, S. W., Pugh, R. and Taylor, C., 1994. *Estimating numbers of homeless and homeless mentally ill people in north east Westminster by using capture-recapture analysis.* Bristish Medical Journal 308 (6920), 27-30.

Gallay, A., Nardone, A., Vaillant, V. and Desenclos, J. C., 2002. La *méthode capture-recapture appliquée à l'épidémiologie: principes, limites et applications. [The capture-recapture method applied to epidemiology: principles, limitations and application].* Revue d'Epidémiologie et de Santé Publique 50 (2), 219-232.

Garton, M. J., Abdalla, M. I., Reid, D. M. and Russell, I. T., 1996*. Estimating the point accuracy of population registers using capture-recapture methods in Scotland.* Journal of Epidemiology and Community Health 50 (1), 99-103.

Gemmell, I., Millar, T. and Hay, G., 2004. *Capture-recapture estimates of problem drug use and the use of simulation based confidence intervals in a stratified analysis.* Journal of Epidemiology and Community Health 58 (9), 758-765.

Hook, E. B. and Regal, R. R., 1995. *Capture-recapture methods in epidemiology: methods and limitations.* Epidemiologic Reviews 17 (2), 243-264.

Hook, E. B. and Regal, R. R., 1999. *Recommendations for presentation and evaluation of capture-recapture estimates in epidemiology.* Journal of Clinical Epidemiology 52 (10), 917-926; discussion 929-933.

Hook, E. B. and Regal, R. R., 2000a. *Accuracy of alternative approaches to capture-recapture estimates of disease frequency: internal validity analysis of data from five sources*. American Journal of Epidemiology 152 (8), 771-779.

Hook, E. B. and Regal, R. R., 2000b*. On the need for a 16th and 17th recommendations for capture-recapture analysis.* Journal of Clinical Epidemiology 53 (12), 1275-1277.

Hvoslef, H. (1994), *"Under-reporting of road traffic accidents recorded by the police at the international level,"* 126.

IWGDMF, 1995a. International Working Group for Disease Monitoring and Forecasting. *Capture-recapture and multiple-record systems estimation I: History and theoretical development*. American Journal of Epidemiology 142 (10), 1047-1058.

IWGDMF, 1995b. International Working Group for Disease Monitoring and Forecasting. *Capture-recapture and multiple-record systems estimation II: Applications in human diseases.* American Journal of Epidemiology 142 (10), 1059-1068.

Jarvis, S. N., Lowe, P. J., Avery, A., Levene, S. and Cormack, R. M., 2000. *Children are not goldfish - mark/recapture techniques and their application to injury data.* Injury Prevention 6, 46-50.

Johnson, R. L., Gabella, B. A., Gerhart, K. A., McCray, J., Menconi, J. C. and Whiteneck, G. G., 1997. *Evaluating sources of traumatic spinal cord injury surveillance data in Colorado.* American Journal of Epidemiology 146 (3), 266-272.

LaPorte, R. E., Dearwater, S. R., Chang, Y. F., Songer, T. J., Aaron, D. J., Anderson, R. L. and Olsen, T., 1995. *Efficiency and accuracy of disease monitoring systems: application of capture-recapture methods to injury monitoring.* American Journal of Epidemiology 142 (10), 1069-1077.

Madigan, D. and York, J., 1997. *Bayesian methods for estimation of the size of a closed population*. Biometrika 84 (1), 19-31.

Meuleners, L. B., Lee, A. H., Cercarelli, L. R. and Legge, M., 2006. *Estimating crashes involving heavy vehicles in Western Australia, 1999-2000: A capture-recapture method.* Accident Analysis and Prevention 38 (1), 170-174.

Morrison, A. and Stone, D. H., 2000. *Capture-recapture: a useful methodological tool for counting traffic related injuries?* Injury Prevention 6 (4), 299-304.

Neugebauer, R. and Wittes, J., 1994. *Voluntary and involuntary capture-recapture samples--problems in the estimation of hidden and elusive populations*. American Journal of Public Health 84 (7), 1068-1069.

Razzak, J. A. and Luby, S. P., 1998. *Estimating deaths and injuries due to road traffic accidents in Karachi, Pakistan, through the capture-recapture method.* International Journal of Epidemiology 27, 866-870.

Razzak, J. A. and Luby, S. P., 1998. *Estimating deaths and injuries due to road traffic accidents in Karachi, Pakistan, through the capture-recapture method*. International Journal of Epidemiology 27, 866-870.

Reurings, M.C.B. (2010). *Ernstig verkeersgewonden in Nederland in 1993-2008: in het ziekenhuis opgenomen verkeersslachtoffers met een MAIS-score van ten minste 2.* Beschrijving en verantwoording van de schattingsmethode. (Serious road injuries in the Netherlands in 1993-2008: road casualties admitted to hospital with a MAIS score of at least 2; Description and justification of the estimation method) R-2010-15, SWOV, Leidschendam. www.swov.nl/rapport/R-2010-15.pdf

Reurings, M.C.B., Bos, N.M. (2009). *Ernstig gewonde verkeersslachtoffers in Nederland in 1993-2008*. Het werkelijke aantal in ziekenhuizen opgenomen verkeersslachtoffers met een MAIS van ten minste 2. R-2009-12, SWOV, Leidschendam. www.swov.nl/rapport/R-2009-12.pdf

Roberts, I. and Scragg, R., 1994. Application *of capture-recapture methodology to estimate the completeness of child injury surveillance.* Journal of Paediatrics and Child Health 30 (6), 513-514.

Rosenman, K. D., Kalush, A., Reilly, M. J., Gardiner, J. C., Reeves, M. and Luo, Z., *2006. How much work-related injury and illness is missed by the current national surveillance system?* Journal of Occupational and Environmental Medicine 48 (4), 357-365.

Tercero, F. and Andersson, R., 2004. *Measuring transport injuries in a developing country: an application of the capture-recapture method.* Accident Analysis and Prevention 36 (1), 13-20.

Tilling, K. and Sterne, J. A., 1999. *Capture-recapture models including covariate effects*. American Journal of Epidemiology 149 (4), 392-400.

Van HT, Singhasivanon P, Kaewkungwal J, Suriyawongpaisal P and LH., K., 2006. *Estimation of non-fatal road traffic injuries in Thai Nguyen, Vietnam using capture-recapture method.* Southeast Asian J Trop Med Public Health 37 (2), 405-411.

Wittes, J., 1972. *On the bias and estimated variance of Chapman's two-sample capture-recapture population estimate.* Biometrics 28, 592-597.

Zavareh, D. K., Mohammadi, R., Laflamme, L., Naghavi, M., Zarei, A. and Haglund, B. J. A., 2008. *Estimating road traffic mortality more accurately: Use of the capture-recapture method in the West Azarbaijan Province of Iran.* International Journal of Injury Control and Safety Promotion 15 (1), 9 - 17.

CHAPTER 6
Towards an international definition of a serious road injury

Abstract

Chapter 6 considers a range of options using hospital data to consider the 'best' option for a practical definition of a serious casualty, in terms of threat to life. The chapter considers various injury severity scales, the length of stay in hospital, and particular injury diagnoses.

6.1. Background

6.1.1. Aim

The aim of this chapter is to make practical recommendations on the best definition of a 'serious' injury, for the purpose of international monitoring of trends in non-fatal road casualties. That is, we want to identify a case definition of serious injury suitable for a) monitoring trends within individual countries, b) comparing rates between countries, and c) comparing trends between countries (i.e.. when using and comparing selected indicators of injury incidence within and between countries).

Again, it should be stressed that this discussion is most appropriate for countries which already have reliable systems for monitoring fatalities. However, it may also be useful to countries that are beginning to consider what data may be useful in future.

6.1.2. Constraints and Issues

Earlier chapters indicated a) the limitations of police data for identifying serious injury, and b) the availability of hospital inpatient / discharge data (called here hospital discharge data [HDD]) in most countries. In HDD, diagnosis is usually captured using the WHO International Classification of Diseases (ICD) 9th or 10th revision. This chapter is limited to a consideration of case definitions suitable for use with HDD for the international monitoring of trends in non-fatal road casualties. The chapter is also limited to a consideration of serious injury as measured by threat to life (see the next subsection).

Some potential issues / problems when using hospital discharge data for these sorts of comparisons are as follows (Langley et al., 2002; Langley and Brenner, 2004):

1. There are very different health care delivery systems in each country.
2. There may be variations in what is defined as a hospital inpatient and/or discharge.
3. Electronic hospital records represent different things in different countries (e.g. records may relate to an admission to hospital, or to a specialty area within a hospital).
4. If based on a sample of hospitals, there are differing proportions of hospitals sampled country to country, with varying degrees of representativeness of the sample, and using differing methods to extrapolate to the whole population.
5. Some types of hospitals may not be consistently included in all national level HDD systems (e.g. military, psychological, rehabilitation/long term care).

Each of these needs to be considered when interpreting the results of international comparisons. It is also crucial that hospital staff should be well trained in the various coding systems discussed below.

Additionally, in order to describe incident cases, readmissions and transfers need (as far as is possible) to be removed (or linked to the first admission if considering length of stay in hospital). This pre-processing of the hospital discharge data will be taken as a given.

In Australia, Harrison and Steenkamp (2002) suggested a number of steps to improve indicator reporting including omitting 'same day' cases from HDD (i.e. cases admitted and discharged on the same day), since they found great variability in the proportions of these cases between States and Territories. This same message is applicable to international comparisons. The validity of this step needs to be investigated in constructing specifications of the indicators for the particular international comparison using hospital discharge data considered in this document.

In order to facilitate the achievement of the above, the indicator implied in the aims should be fully specified to allow calculation to be consistent at any place, any time and across any population.

6.1.3. Measuring severity – why threat to life?

Severity can be measured on a number of dimensions: e.g. threat to life; (threat of) disability, quality of life, burden, or cost. Before discussing instruments /methods for identifying cases of serious injury, there is a theoretical question of what one is interested in measuring. Is the interest in, for example, injuries that are serious as measured by threat to life, or in disabling injury? If the interest is in disabling, high cost, or high burden injury, or injury that has a significant impact on quality of life, then it would be necessary to consider not only hospital discharge data but other data sources (e.g. emergency data), too. Hospital discharge data do not capture a high proportion of all disabling, high cost or high burden injuries (Cryer, 2007). For example, casualties with back injuries are often not admitted to hospital; however, back injury can be very disabling - resulting in significant periods of time off work. The same is not true for threat to life injury. Injury with high threat to life typically gets admitted to hospital - irrespective of country or degree of health spending (Cryer et al., 2011).

6.1.4. Threat to life severity measures

Previous chapters have mentioned a number of severity measures including: AIS and its derivatives (which include ISS), ICISS, and length of stay (LoS) in hospital. The AIS derivatives and ICISS have been validated as threat to life measures (Baker et al., 1974; Stephenson et al., 2002; Stephenson et al., 2004). LoS is often proposed as a proxy measure for severity of injury - although it is not necessarily a proxy for threat to life severity of injury. Nevertheless, it will be considered below - since it has been suggested as a potential severity measure in previous chapters and previous documents, and it can clearly be derived from hospital discharge data. A further option, not considered in earlier chapters, is to define "serious injury" in terms of a set of sentinel diagnoses. This too will be considered in this chapter. For the remainder of this chapter, the term "injury severity" will designate 'threat to life injury severity".

6.1.5. Criteria for choosing a serious injury case definition

The bulk of this chapter discusses the options for a serious non-fatal injury case definition. The aim is to determine whether it should be based on AIS or its derivatives (in particular on maximum AIS, i.e. MAIS), ICISS, LoS, a set of sentinel diagnoses, or some other approach. As a basis for making a sensible decision, the next section describes ideal characteristics for a serious non-fatal injury indicator for international monitoring of trends in non-fatal road casualties. This will be followed by a consideration of each method of defining a case of serious non-fatal injury against these ideal characteristics. Finally, the chapter will make recommendations regarding the most appropriate approach to defining serious non-fatal injury for the purposes described above.

6.2. Criteria for judging the 'severity' case definition

This section proposes a set of ideal characteristics of a serious non-fatal injury case definition for international monitoring of trends in non-fatal road casualties.

When using hospital discharge data as the basis for indicators, one is interested in restricting consideration to serious injuries for a number of reasons including:

- An interest in the prevention of "important" injuries.
- To control bias, i.e. the probability of a case being ascertained should be independent of extraneous factors - of social, economic and demographic factors, as well as service supply and access factors.

"Important" injuries are those that result in death, significant disability, or loss of quality of life, carry a significant threat to life or threat of disability, or result in significant cost to the individual or society. In this context, "important" is synonymous with significant threat to life.

In respect to controlling bias, when using hospital discharge data as the primary data source for ascertaining cases, factors other than underlying injury incidence can affect admission and so, potentially, case ascertainment (Lyons et al., 2005). These include health service supply (e.g. availability of consultants and beds) and access factors (e.g. distance from hospital). These will bias time trends and international comparisons if not controlled (see below). However, in most developed countries, serious injury cases (using a high threat-to-life threshold to define "serious") usually get admitted. Hence, a focus on serious injury will help control these biasing influences.

Indicators should be validated before use, to establish whether any potential bias has been controlled (Cryer et al., 2002; .Cryer and Langley, 2006). Within each country, the magnitude and direction of biases can change over time. So if bias is not controlled it can affect both within country trends, and between country comparisons. Assessing potential indicators against criteria for an ideal indicator is a form of face validity (i.e. face validity is an expert judgement that the measure represents what you want to measure "on the face of it").

Several criteria have been identified for judging indicators of injury incidence; a number of these are general criteria, i.e. they do not apply specifically to the question of the severity of injury measure. These include: availability of reliable data; a measurement that is practicable; an indicator that reflects a well-defined information objective; is timely; and is derived from data that are inclusive or representative of the target population that the indicator aims to reflect. For indicators of risk or rate, the denominator should reflect the exposure of the population to relevant injury hazards.

A fundamental criterion for developing serious road traffic injury indicators is the availability of data that is accurate enough to allow the reliable identification of cases of serious road traffic injury. Accurate diagnostic and external cause (mechanism) of injury coding is required to apply case definitions and conduct appropriate case ascertainment. Audit information is required in order that a quality assessment can be made.

Moving now to the criteria that have been previously identified that relate to the measurement of serious injury. These include:

1. The indicator should be readily comprehensible.
2. The case definition should be based on diagnosis – on anatomical or physiological damage.
3. The case definition should ascertain the same injury diagnoses over time and by place.
4. The cases included should be all of those that the indicator aims to reflect or a well defined sample of them.
5. The indicator should be robust to potential or known changes/differences in coding frames or coding practice between places or over time.
6. There should be unbiased case ascertainment: The probability of a case being ascertained should be independent of extraneous factors - of social, economic and demographic factors, as well as service supply and access factors.

Careful definition and selection of serious non-fatal injury cases can mitigate many of the effects of variations in health service provision and health service data collections (Cryer and Langley, 2006; Cryer and Langley, 2004; Du et al., 2008). One key means of addressing this is to define "serious" such that it encompasses diagnoses with a very high probability of admission (e.g. fractured femur [Boufous et al., 2007]). Comparisons of injuries, using hospital discharge data, with consistently high probability of admission will be a comparison of serious injury incidence, rather than of service effects. This can be achieved by explicitly specifying sentinel diagnoses that have a high probability of admission, or by using a high severity threshold for the case definition. In choosing the threshold for the serious injury definition based on hospital discharge data, the goal is to capture as many serious injuries as possible, without compromising the requirement that they have a very high probability of admission to hospital.

6.3. Suitability of AIS and derivatives

The relationship between diagnosis and severity has been used successfully in the development of severity scales such as the Abbreviated Injury Scale (AIS) and the ICD-based Injury Severity Score (ICISS). [Association for the Advancement of Automotive Medicine, 1990; Hobbs et al., 1979].

Diagnoses that carry an AIS score of 3 or above are designated "serious" in the AIS manual (Association for the Advancement of Automotive Medicine, 1990). Such injuries have been found to be associated with a high likelihood of admission to hospital (Hobbs et al., 1979; Cryer et al., 1999). Consequently, it has been argued that the use of hospital discharge data to estimate trends in the frequency and rates of injuries with AIS of 3 or more would be minimally influenced by health service effects (Cryer and Langley, 2007).

In earlier work, it has been proposed that the ICD-10 diagnoses associated with a maximum AIS (MAIS) score of 3 or more be classified as a case of serious threat to life injury (Thomas et al., 2009). It would be feasible for any country with ICD-9 and ICD-10 coded HDD to apply this definition.

In the assessment below, it will be assumed that defining serious injury as an MAIS of 3 or greater, with AIS derived from ICD codes, will capture injured people with diagnoses that have a high probability of admission across all countries considered in the international comparison. It is recommended that this assumption be tested.

1. **Readily comprehensible**

   In the sphere of road traffic injury research, prevention and policy making, MAIS is a familiar method of measuring severity of injury. Consequently, a serious injury indicator for the purposes outlined in the aims, with serious defined as MAIS≥3 will be comprehensible to the target audience.

2. **The case definition should be based on diagnosis – on anatomical or physiological damage**

   In the AIS, severity scores are assigned to diagnosis descriptions. Consequently, using MAIS≥3 to define serious injury is equivalent to defining a case based on those diagnoses with a maximum AIS severity score of 3 or above. Consequently, this criterion is satisfied.

3. **The case definition should ascertain the same injury diagnoses over time and between places**

   Within a country, assuming the same translation between ICD and AIS score is the used over time, this criterion will be satisfied. Whether it is satisfied between places, i.e. between countries, will depend on whether the same translator between ICD and AIS severity score is used. If it is, then the criterion will be satisfied. If it is not, then there will be uncertainty regarding the equivalence of the cases ascertained between countries.

4. **The cases included should be all of those that the indicator aims to reflect or a well defined sample of them**

   MAIS≥3 defines serious injury unequivocally and completely. Such a definition includes all cases that the indicator aims to measure.

5. **Robust to potential or known changes/differences in coding frames or coding practice between places or over time**

   There is a translation between ICD-9 and ICD-10 diagnoses. It is far from perfect, however. For example, there are instances of a 1-to-many mapping in the translation, as well as many-to-1 mapping. The structure of ICD-10 was a radical change from ICD-9. Comparisons over time that span periods where both ICD-9 and ICD-10 are used, are likely to have some discontinuity at the change point. Similarly, if countries that code to ICD-9 are compared to those using ICD-10, there will be some bias introduced in the comparison of serious injury rates and trends. This problem will be compounded by the use, inevitably, of differing translators from ICD to AIS severity score for countries coding diagnosis to ICD-9 compared to countries coding to ICD-10. Consequently, this criterion is not satisfied.

6. **Unbiased case ascertainment**

   Use of MAIS≥3 to define serious injury should result in the capture of injured people with diagnoses that have a high probability of admission. In those circumstances, the probability of a case being ascertained should be independent of extraneous factors - of social, economic and demographic factors, as well as service supply and access factors.

Other than the concern about bias introduced when comparing indicators based on ICD-9 coded diagnostic data with ICD-10 coded data, a definition of serious injury based on MAIS$\geq$3 satisfies all of the above criterion. This is the case if it can be assumed that MAIS$\geq$3 captures diagnoses that have a high probability of admission to hospital.

6.4. Suitability of ICISS

The ICD-based Injury Severity Score (ICISS) is a threat to life severity score that can be used to define a serious injury. [8,9,19,23-26] For example, in New Zealand, their national Injury Prevention Strategy (NZIPS) serious non-fatal injury indicators are based on the frequency of occurrence of injury that exceeds a pre-specified ICISS threshold (Cryer and Langley, 2006; Cryer and Langley, 2004). In that work, the threshold was selected such that the injuries captured are associated with an in-hospital mortality rate of 5.9% or worse. The goal in setting the threshold was to identify cases with injury diagnoses associated with a high probability of admission, in order to minimise service effects from the resulting indicator trends.

ICISS is constructed from diagnosis-specific survival probabilities (originally called SRRs; now branded DSPs), which are calculated from a training set of data. The training set used in the above example was New Zealand hospital discharges for the period from the introduction of the latest revision of the ICD (i.e. ICD-10 introduced during 1999) to the end of 2001. The threshold was set to capture cases with injury diagnoses associated with a high probability of admission. At the time (2004), this threshold was based on empirical evidence as well as face validity arguments (Cryer and Langley, 2004). Given recent work, the process of setting the threshold can also be informed by empirical estimates of diagnosis-specific probabilities of admission (see later in this chapter) (Cryer et al., 2011).

The ICISS approach to deriving anatomical severity has been empirically validated as a threat to life severity scale in a number of settings. [Stephenson et al., 2002; Stephenson et al., 2004; Osler et al., 1996; Sacco et al., 1999; Meredith et al., 2002; Clark e and Winchell, 2004; Kim et al., 2004). The original research work was based on patients treated in specialist facilities (e.g. trauma centres) and as such are atypical of all seriously injured persons. The New Zealand / Australian work was based on all injury patients admitted to hospital (Stephenson et al., 2002; Stephenson et al., 2004).

ICISS has the potential advantage over AIS in the following ways:

- ICISS scores are empirically derived.
- Recent studies have found the ICISS to perform better than its rivals (ISS, NISS, APS derived from the current version of ICDMAP) (Stephenson et al., 2002; Fingerhut, 2004).

The ICISS threat to life severity score has shown good concordance and calibration when validated against death as an outcome. [9] However, it does have limitations and some disadvantages compared to AIS and its derivatives.

One of the strengths of ICISS, listed above, is, currently, an important limitation in the context of the development of indicators for international comparison. ICISS depends for its calculation on empirically derived DSPs, and these vary from country to country. One country's DSPs are unlikely to be identical to those for another country (e.g. due to differences in the effectiveness of pre-hospital care, transport infrastructure, and different standards of hospital care affecting the likelihood of survival following injury). One way round this problem is to estimate DSPs based on a pooled set of data across several countries. This would result in aggregate estimates of DSPs. The same aggregate DSPs could be applied in each country in the calculation of a patient's ICISS score. A diagnosis-specific DSP would not now be equal to a country's own survival probability for that diagnosis, but rather an approximation to it. This approach is a parallel to age-standardisation of rates where a standard population is used to artificially estimate rates for a particular country. An ICE project to estimate international ICD-10 based DSPs has been carried out and a report is currently being prepared by Rolf Gedeborg and colleagues. .

The assessment below is based on the assumption that acceptable international standard DSPs exist.

1. Readily comprehensible

Although the ICISS is essentially a measure of survival probability, its precise meaning is dependent on how the DSPs are constructed. Typically, the DSPs are derived from hospital discharge data or from trauma centre data alone – and hence ICISS is a measure of survival to discharge, given admission. If international standard DSPs are used, they would provide artificial measures of survival probabilities. The precise meaning of a given ICISS severity threshold, in these circumstances, may not be obvious to a policy maker.

2. **The case definition should be based on diagnosis – on anatomical or physiological damage**

   Strictly speaking, ICISS is a severity measure based on diagnosis. However, where there are multiple injuries, it is constructed from the DSPs relating to each of those injuries. So like ISS, the occurrence of a number of moderately severe injuries could result in a severity score that is classified as 'serious'. A modification of the ICISS approach is to use the worst injury DSP (instead of ICISS) [Kilgo et al., 2003], in a similar way to which MAIS is used rather than ISS. Based on this approach, a set of injury diagnoses that represent serious injury could be identified.

3. **The case definition should ascertain the same injury diagnoses over time and between places**

   This condition would be satisfied if common DSPs are used over time and between places. This is the proposal outlined above.

4. **The cases included should be all of those that the indicator aims to reflect or a well defined sample of them**

   Provided the ICISS (or DSP) threshold, used to define a case of serious injury, captures only those diagnoses that have a high probability of admission, this criterion should be satisfied. The aims (stated at the start of this chapter), in conjunction with this criterion, call for the vast majority of serious injury cases (however they are defined) to be identified from the hospital discharge data. Again, provided the ICISS / DSP threshold that is used to define 'serious' is set appropriately (i.e.. low enough to capture diagnoses with a high probability of admission; high enough not to exclude relevant serious injury cases), then a large majority of relevant cases should be captured.

5. **Robust to potential or known changes/differences in coding frames or coding practice between places or over time**

   The translation between ICD-9 and ICD-10 diagnoses is far from perfect. The New Zealand experience suggests that the change from ICD-9 to ICD-10 resulted in some discontinuity in time trends when monitoring serious non-fatal injury, where the definition of 'serious' was based on ICISS (Gulliver et al., 2010). The NZ experience suggests that the threshold used for defining 'serious' should be ICD revision specific. In spite of the above, with judicious choice of threshold, discontinuities in trends in road traffic injuries may be minimised.

6. **Unbiased case ascertainment**

   Provided the threshold used to define 'serious' is chosen such that only cases with a high probability of admission are captured, then case ascertainment will be unbiased. The probability of a case being ascertained, in these circumstances, will be largely independent of extraneous factors - of social, economic and demographic factors, as well as service supply and access factors.

Although the above assessments are mostly positive, the DSPs on which the proposed approach is based are undergoing peer review.

6.5. Suitability of Length of stay

A number of agencies have used length of stay in hospital (LoS) as a proxy for severity in their national non-fatal indicators of injury incidence (e.g. New Zealand Ministry of Transport, English Department of Health). [Minister of Transport, 2010;)]. The question arises: is LoS a stable proxy measure for severity? Work has suggested that it is not.

Langley, Cryer and colleagues have carried out a number of pieces of work that have shown contradictory trends between the indicators where severity is defined on the basis of service utilisation (e.g. admission to hospital, or days stay in hospital of at least 4 days) and those based on anatomical severity definitions (e.g. diagnosis-based definitions, ICISS, MAIS). [Cryer et al., 2002; Langley et al., 2003] For example, the New Zealand national trends found for the motor vehicle traffic crash (MVTC) serious injury indicators based on an ICISS threshold showed little change in rates over time, whereas MVTC indicators based on admissions showed a significant decline. [Langley et al., 2003; Cryer, 2005] The concern was that the latter trends were driven by service utilisation factors.

Cryer and colleagues found that over the 10 years that were included in their New Zealand study, lengths of stay in hospital for admissions for all diagnoses combined (injury and non-injury), as well as for all injury diagnoses, reduced. Reductions were shown using all of the LoS metrics investigated: geometric mean, interpolated median, percentage of incident injuries with LoS≥3 days, ≥4 days, ≥7 days and ≥14 days. [Cryer et al., 2010]

If LoS was a good proxy for severity and was stable over time, then for a particular group of injuries of constant severity it would be expected that these LoS statistics would remain constant over time. For example, for the group of injuries that have an AIS severity score of 3, it would be expected that the geometric mean or interpolated median LoS would be constant over time. This was found not to be the case. [Cryer et al., 2010]

Furthermore, Cryer and colleagues considered 11 sentinel ICD-9-CM diagnoses for which cases classified to a particular diagnosis mapped to a single AIS severity code, using the high severity and low severity options, in ICDMAP-90 (and so in that sense each diagnosis code was homogenous in severity – ICD-HS diagnoses). They demonstrated that for every ICD-HS diagnoses (e.g. closed fracture of the pubis, closed fractured neck of humerus, closed fracture of the neck of femur) that almost all of the LoS statistics considered[9] reduced (statistically significantly) over time. [Cryer et al., 2010]

The results for the LoS metrics considered by Cryer and colleagues, and for each of the AIS and ICISS severity strata, along with the results for the ICD diagnoses homogeneous in severity considered in their work, suggested that LoS is not a <u>stable</u> proxy for severity. (Cryer et al., 2010)

The above results are consistent with those in a study of selected OECD countries. It was found that there was a decline in LoS for specific diagnoses, including for fractured neck of femur, across all of the 7 OECD countries considered. The decline was likely to be associated with changes in treatment and care services. For example, the UK government has regarded reducing LoS in the National Health Service as a component of increased efficiency, thus resulting in pressures to reduce diagnosis-specific LoSs. In that Clarke and Rosen (2001) review , the authors' state:

"In many health systems, there are managerial and financial incentives to reduce LOS".

They indicate that from a healthcare provider's viewpoint, reasons to reduce LoS include: to tailor care to the individual and their preferences (often home care, if it can be provided safely); and to reduce resources spent on an individual so that the health budget can stretch further.

The above studies show that length of stay is not a stable proxy for severity. Reductions in LoS measures over time, for the groups considered as well as the sentinel diagnoses, were due to factors other than reductions in the incidence of serious injury; e.g. changes in service delivery over time. A LoS threshold should <u>not</u> be used as a proxy for severity of injury if the goal is to monitor time trends in injury incidence. The concern is that using a LoS threshold to identify serious injuries will results in misleading trends; and this applies to all countries (since once can expect diagnosis-specific changes in LoS over time in every country), to any time period, and to any coding frame, be it ICD-9, ICD-10, or other. (Cryer et al., 2002; Cryer et al., 2010)

Despite the problems identified, the following question is still considered: Does the use of length of stay as a measure of severity of injury satisfy the criteria identified for a severity measure?

1. Readily comprehensible

A LoS threshold, when used to define serious injury, appears a fairly straightforward concept and thus could be regarded as readily comprehensible to all. However, given that a LoS threshold will capture particular injuries with differing likelihoods over time and between place (or country), in these terms it is very difficult to understand what it is actually measuring, other than service use.

9. Geometric mean, interpolated median, percentage with length of stay ≥3 days, ≥4 days, ≥7 days and ≥14 days.

2. **The case definition should be based on diagnosis – on anatomical or physiological damage**
   Given that a LoS threshold will capture particular injuries with differing likelihoods over time and between place, a case definition of 'serious' based on LoS cannot be regarded as being based on diagnosis – on anatomical or physiological damage.

3. **The case definition should ascertain the same injury diagnoses over time and between places**
   The discussion above, and the response to criterion 1 indicates that criterion 3 is not be satisfied.

4. **The cases included should be all of those that the indicator aims to reflect or a well defined sample of them**
   The cases that the indicator aims to reflect are all those that can be regarded as serious. However, the work described above shows that, when using a LoS threshold, varying proportions of particular serious injury will be captured over time, and between countries. Hence, this criterion is not satisfied.

5. **Robust to potential or known changes/differences in coding frames or coding practice between places or over time**
   This criterion is satisfied. The changes in coding frames or coding practices will not affect the number of cases identified using a LoS threshold.

6. **Unbiased case ascertainment**
   A LoS threshold will capture differing proportions of particular serious injury diagnoses over time, and between countries. As a consequence, case ascertainment, when using a LoS threshold to define 'serious' is not unbiased.

Most of the above criteria for a good severity measure are not satisfied for LoS.

6.6. Suitability of Sentinel serious injury diagnoses

International comparisons of non-fatal injury, based on hospital discharge data, are often contaminated by differential health service effects between countries (e.g. varying access to, and provision of, inpatient hospital care). One way to control these contamination effects is to make comparisons solely using a case definition based on diagnoses that have a high probability of inpatient admission. That way, health service effects will be largely removed.

One approach is to focus on probability of admission directly; i.e. to focus on injury diagnoses that have a high probability of admission in all countries as the basis for international comparison - as a means of controlling the extraneous factors when making those comparisons (Cryer et al., 2011).

Several authors have found that fractured neck of femur cases have a high probability of admission. Cryer and colleagues found that femoral fracture more generally have high probabilities of admission in their 6 country study (Cryer et al., 2011). That work also identified a list of diagnoses (based on an analysis of the ICD-9-CM and ICD-10 4-digit codes) that were consistent with a high probability of admission across all the countries considered. Those diagnoses included the following:

Brain injury

- Fractured base of skull with brain injury
- Cerebral laceration and contusion – other and unspecified
- Subarachnoid haemorrhage
- Focal brain injury

Spinal cord lesion

- Fractured cervical vertebra with spinal cord lesion
- Fractured dorsal [thoracic] vertebra with spinal cord lesion
- Fractured lumbar vertebra with spinal cord lesion

Open fractures

- Mandible
- Humerus - shaft or other
- Radius and ulna, lower end
- Patella
- Tibia and fibula – unspecified
- Ankle – medial malleolus, lateral malleolus, bimalleolar, trimalleolar
- Heel

Other (open or closed) fractures (see below)

- Femur
- Pelvis – acetabulum

Internal organ injury

- Traumatic haemopneumothorax
- Diaphragm
- Gastrointestinal tract
- Spleen
- Liver or gall bladder

There could be other open fractures that satisfy the condition of high probability of admission for one of the countries but that were not identified due to coding frame limitations, i.e. because the relevant 4-digit ICD-10 codes include a mix of open and closed fractures. These include:

- Lower end of humerus
- Upper end of radius and ulna
- Upper end or shaft of tibia and fibula

They found that the results for USA and for Barcelona, which code to their data to ICD-9-CM, indicated that these can be added to the list of open fractures.

Additionally, Cryer and colleagues found good evidence to include traumatic haemothorax, pneumothorax, and other injury to lung also in the above list.

At recent ICE meeting, it was agreed that the work of Cryer and colleagues (described above), along with information from the 5 country comparison carried out by Lyons and colleagues (Macey, 2010), the work of Gedeborg and colleagues to calculate standard DSPs from several countries data, as well as other identified international comparison work, should form the basis of an agreed list of sentinel diagnoses for international comparisons.

So, coming to the specific question: Does the use of a case definition of severity of injury based on sentinel diagnoses satisfy the criteria identified for a severity measure?

1. **Readily comprehensible**
   For a policy maker to be confronted with a fairly long list of diagnoses is perhaps a little cumbersome; nevertheless an indicator based on a list of diagnoses is completely transparent, and hence comprehensible.

2. **The case definition should be based on diagnosis – on anatomical or physiological damage**
   This definition is clearly based on diagnosis – i.e. on the appropriate ICD-9 or ICD-10 diagnosis codes.

3. **The case definition should ascertain the same injury diagnoses over time and between places**
   Given that the case definition is a list of diagnoses, subject to common coding frames and to accurate coding across time, and between countries, a case definition based on sentinel diagnoses will be stable across time and place.

4. **The cases included should be all of those that the indicator aims to reflect or a well defined sample of them**
   The interest is in serious non-fatal injury. A case definition based on set of sentinel diagnoses will not ascertain all serious non-fatal injury. However, it is a clearly defined subset of the universe of serious injuries. It is recommended that, once the sentinel diagnoses are identified, it is estimated what proportion of serious threat to life injuries are captured by the sentinel diagnoses, and the main other serious injury diagnoses that are not included in the list of sentinel diagnoses. At that point, an assessment should be made regarding whether the sentinel diagnoses represent an acceptable subset.

5. **Robust to potential or known changes/differences in coding frames or coding practice between places or over time**
   The lookup between ICD-9 and ICD-10 diagnoses is far from perfect. Consequently, a case definition based on sentinel diagnoses will not be totally robust to a change from ICD-9 to ICD-10 coding frame. This is also true of all the other viable options for a cases definition, however.

6. **Unbiased case ascertainment**
   Given that the case definition is based on a list of diagnoses, and the diagnoses are chosen since they have a high probability of admission, then, subject to accurate diagnostic coding, the probability of a case being ascertained from the hospital discharge data should be independent of extraneous factors - of social, economic and demographic factors, as well as service supply and access factors.

6.7. Selection of a suitable injury measure to identify serious road casualties

Table 6.1 summarises the analysis provided in the previous sections to identify a severity of injury measure when a) monitoring trends within individual countries, and comparing rates and trends between countries.

Table 6.1. **Summary of the analysis of the candidate severity measures**

	Criteria[a]					
	1	**2**	**3**	**4**	**5**	**6**
MAIS	√	√	√	√	X	√
ICISS	X	√	√	√	? √	√
LoS	? √	X	X	X	√	X
Sentinel	√	√	√	?	?X	√

a) Criteria:
1. Readily comprehensible
2. The case definition should be based on diagnosis – on anatomical or physiological damage.
3. The case definition should ascertain the same injury diagnoses over time and between places.
4. The cases included should be all of those that the indicator aims to reflect or a well defined sample of them.
5. The indicator should be robust to potential or known changes/differences in coding frames or coding practice between places or over time.
6. Unbiased case ascertainment: The probability of a case being ascertained should be independent of extraneous factors - of social, economic and demographic factors, as well as service supply and access factors.

Serious injury based on MAIS3+, with AIS derived from ICD codes, is a strong candidate for a definition of serious injury. Other than the concern about bias introduced when comparing indicators based on ICD-9 coded diagnostic data with ICD-10 coded data, a definition of serious injury based on MAIS3+ satisfies all of the above criterion. In that assessment, it was assumed that defining serious injury as MAIS3+, with AIS derived from ICD codes, will capture injured people with diagnoses that have a high probability of admission across all countries considered in the international comparison. It is recommended that this assumption be tested before a final decision is made.

It is clear that there is still an important problem to overcome before an ICISS-based measure of serious injury, for international comparison, can be proposed. ICISS cannot be considered as a contender for the measurement of severity, in the short term.

Many of the above criteria for a good severity of injury measure for the purposes stated in the aims are not satisfied for LoS. LoS is not, therefore, a contender for the measurement of severity in this context. It is concluded, that it should not be considered further for the purposes described above.

It is recommended that work be carried out to agree a set of sentinel diagnoses that can be used to address the aims described above. Once the sentinel diagnoses are identified, it is recommended to estimate what proportion of serious threat to life injuries are captured by the sentinel diagnoses, and the main other serious injury diagnoses that are not included in the list of sentinel diagnoses.

In summary, none of the options satisfy all of the criteria. Amongst the approaches considered, MAIS, ICISS and the sentinel diagnoses provide the best options. Each requires some work prior to adoption. For MAIS, it needs to be established whether the diagnoses captured using the MAIS3+ threshold have high probabilities of admission. The definition which is closest to being ready to use is MAIS3+.

References

Association for the Advancement of Automotive Medicine. *The Abbreviated Injury Scale*, 1990 revision. Des Plaines, IL: AAAM, 1990.

Baker S, O'Neill B, Haddon W, Long W. *The Injury Severity Score: a method for describing patients with multiple injuries and evaluation of emergency care.* J Trauma 1974;47:187-96.

Boufous S, Finch C, Close J, Day L, Lord S. *Hospital admissions following presentations to emergency departments for a fracture in older people*. Injury Prevention 2007;13:211-14.

Clark DE, Winchell RJ. *Risk Adjustment for injured Patients using Administrative data. The Journal of Trauma: Injury, Infection and Critical Care* 2004;57:130-40.

Clarke A, Rosen R. Length of stay. *How short should hospital care be? European Journal of Public Health* 2001;11(2):166-70.

Connor J, Langley J, Cryer C. *International comparisons of injury. A compilation of reports to the New Zealand Injury Prevention Strategy Secretariat*. *Injury Prevention Research Unit Occasional Report OR 063.* Dunedin: Injury Prevention Research Unit, University of Otago, New Zealand, 2007.

Cryer C, Gulliver P, Davie G, de Graaf B, Langley J, et al. *NZIPS serious injury case definition review. Empirical validation of the New Zealand Injury Prevention Strategy indicators: The identification of ICD diagnoses associated with a high probability of inpatient hospital admission. Injury Prevention Research Unit (University of Otago) OR 086.* Dunedin: New Zealand: Injury Prevention Research Unit, University of Otago, 2011.

Cryer C, Gulliver P, Langley J, Davie G. *To what degree is length of stay in hospital a stable proxy for injury severity?* Injury Prevention 2010;16:254-60.

Cryer C, Langley J. *Developing indicators of injury incidence that can be used to monitor global, regional and local trends. Injury Prevention Research Unit Occasional Report OR 070.* Dunedin: Injury Prevention Research Unit, University of Otago, 2007.
Cryer C, Langley J, S. S. *Developing Valid Injury Outcome Indicators: A report for the New Zealand Injury*

Prevention Strategy. Injury Prevention Research Unit Occasional Report OR 049. Dunedin: Injury Prevention Research Unit, University of Otago, New Zealand., 2004.

Cryer C, Langley JD, Jarvis SN, MacKenzie SG, Stephenson SCR, Heywood P. *Injury outcome indicators: the development of a validation tool.* Injury Prevention 2005;11:53-57.

Cryer C, Langley JD, Stephenson SCR, Jarvis SN, Edwards P. *Measure for measure: The quest for valid indicators of non-fatal injury incidence.* Public Health 2002;116:257-62.

Cryer C, Langley JD. *Developing valid indicators of injury incidence for "all injury".* Injury Prevention 2006;12:202-07.

Cryer C. *Injury outcome indicators - validation matters.* International Journal of Injury Control and Safety Promotion 2005;12(4):219-24.

Cryer PC, Jarvis SN, Edwards P, Langley JD. How *can we reliably measure the occurrence of non-fatal injury?* International Journal of Consumer & Product Safety 1999;6(4):183-91.

Du W, Hayen A, Finch C, Hatfield J. *Comparison of methods to correct the miscounting of multiple episodes of care when estimating incidence of hospitalised injury in child motor vehicle passengers*. Acc Anal Prev 2008;40:1563-68.

Fingerhut LA (Ed). *Discussion document on injury severity measurement in administrative datasets*. Washington DC: National Center for Health Statistics (unpublished), 2004.

Gulliver P, Cryer C, Davie G. Chartbook of the New Zealand Injury Prevention Strategy Serious Injury Outcome: 1994 – 2009. . *Injury Prevention Research Unit (University of Otago) OR 089*. Wellington: New Zealand Injury Prevention Strategy Secretariat, ACC. www.nzips.govt.nz/resources/publications.php, 2010.

Harrison JE, Steenkamp *M. Technical review and documentation of current NHPA injury indicators and data sources. Injury Research and Statistics Series Number 14.* Adelaide: Australian Institute of Health and Welfare, 2002.

Hobbs CA, Grattan E, Hobbs JA. *Classification of injury severity by length of stay in hospital. TRRL Laboratory Report 871*. Crowthorne: Transport and Road Research Laboratory, 1979.

Johansen A, Evans R, Bartlett C, Stone M. *Trauma admissions in the elderly: how does a patient's age affect the likelihood of their being admitted to hospital after fracture?* Injury 1998;29(10):779-84.

Kilgo PD, Turner MO, Meredith W. *The worst injury predicts mortality outcome the best: rethinking the role of multiple injuries in trauma outcome scoring. Journal of Trauma Injury, Infection and Critical Care 55: 599-606*. Journal of Trauma Injury, Infection and Critical Care 2003;55:599-606.

Kim Y, Jung KY, Kim CY, Kim YI, Shin Y. *Validation of the International Classification of Diseases 10th Edition - based Injury Severity Score (ICISS).* The Journal of Trauma: Injury, Infection and Critical Care 2000;4(2):280-85.

Langley J, Stephenson S, Cryer C, Borman B. *Traps for the Unwary in estimating person-based injury incidence.* Injury Prevention 2002;8:332-7.

Langley JD, Brenner R. *What is an injury?* Injury Prevention 2004;10(2):69-71.

Langley JD, Stephenson S, Cryer C. *Measuring Road Traffic Safety Performance: Monitoring Trends in Nonfatal Injury.* Traffic Injury Prevention 2003;4(4):291-96.

Lawrence BA, Miller TR, Weiss HB, Spicer RS. *Issues in using state hospital discharge data in injury control research and surveillance.* Acc Anal Prev 2007;39:319-25.

Lyons RA, Brophy S, Pockett R, John G. *Purpose, development and use of injury indicators. International Journal of Injury Control and Safety Promotion* 2005;12:207-11.

Meredith JW, Evans G, Kilgo PD, Mackenzie E, Osler T, McGwin G, et al. *Comparison of the Abilities of Nine Scoring Aligorithms in Predicting Mortality.* The Journal of Trauma, Injury, Infection and Critical Care 2002;53(4):621-29.

Minister of Transport. *Road Safety to 2010.* Wellington, NZ, 2003:33.

Osler T, Rutledge R, Deis J, Bedrick E. *ICISS: An International Classification of Disease-9 Based Injury Severity Score*. The Journal of Trauma: Injury, Infection and Critical Care 1996;41(3):380-88.

Pasco JA, Sanders KM, Hoekstra FM, Henry MJ, Nicholson GC, Kotowicz MA. *The human cost of fracture.* Osteoporosis Int 2005;16(12):2046-52.

Sacco WJ, MacKenzie EJ, Champion HR, Davis EG, Buckman RF. *Comparison of alternative methods for assessing injury severity based on anatomic descriptors.* J Trauma 1999;47(3):441-6; discussion 46-7.

Secretary of State for Health. *Saving Lives: Our Healthier Natio*n. London: The Stationary Office, 1999.

Secretary of State for Health. The new NHS. Modern and Dependable. London: NHS Executive, Department of Health., 1997.

Stephenson S, Henley G, Harrison JE, Langley JD*. Diagnosis based injury severity scaling: investigation of a method using Australian and New Zealand hospitalisations.* Injury Prevention 2004;10:379-83.

Stephenson SCR, Langley JD, Civil I. *Comparing Measures of Injury Severity for Use with Large Databases.* The Journal of Trauma, Injury, Infection and Critical Care 2002;53:326-32.

Thomas P, Chambers-Smith D, Morris A, Yannis G, et al. *Building the European Road Safety Observatory. SafetyNet Final Activity Report:* http://ec.europa.eu/transport/wcm/road_safety/erso/safetynet/fixed/SafetyNet%20Final%20Activity%20Report_V3_Final.pdf 2009.

MEMBERS OF THE IRTAD WORKING GROUP ON LINKING

Austria	**Mr. Christian BRANDSTÄTTER** *KfV*
	Mr. Christian STEFAN *KfV*
Belgium	**Ms Nina NUYTTENS** IBSR
Czech Republic	**Mr. Josef MIKULIK** *CDV*
	Mr. Josef TESARIK *Traffic Police*
Denmark	**Mr. Stig HEMDORFF** *Danish Road Directorate*
France	**Ms. Emmanuelle AMOROS** *IFSTTAR*
Germany	**Ms Kerstin AUERBACH** *BASt*
	Ms Susanne SCHOENEBECK *BASt*
	Mr. Eike SCHMIDT *BASt*
Ireland	**Mr. Michael C. BROSNAN** *Road Safety Authority*
Israel	**Ms Sarit LEVI** *National Road Safety Authority*
	Mrs. Orit YALON-SHUQRUN *Central Bureau of Statistics*
Malaysia	**Mrs. Sharifa ALLYANA BT SYED MOHAMED RAHIM** *MIROS*
	Mr. Hizal HANIS BIN HASHIM *MIROS*
Netherlands	**Mr. Fred WEGMAN** *SWOV*
	Mr. Niels BOS *SWOV*
Norway	**Ms Marianne ROSTOFT** *Road Administration*

Poland	**Ms Anna ZIELINSKA** *ITS*
	Ms Justyna WACOSWSKA *ITS*
Spain	**Mr. Rafael HEREDERO** *European Center for Injury Prevention*
	Ms Catherine PEREZ *Agència de Salut Pública de Barcelona*
	Ms Maria SEGUI *European Center for Injury Prevention*
	Ms Pilar ZORI *Dirección General de Tráfico*
Sweden	**Mr. Jan IFVER** *Swedish Transport Agency*
	Mr. Thomas LEKANDER *Swedish Transport Administration*
	Ms Kristina MATTSSON *Swedish Transport Agency*
United Kingdom	**Ms Liz BRUTUS** *Department for Transport*
	Ms Pat KILBEY *Department for Transport*
	Ms Barbara NOBLE *Ex Department for Transport*
	Mr. Matthew TRANTER *Department for Transport*
ITF/OECD /IRTAD	**Mrs. Véronique FEYPELL – de LA BEAUMELLE**
	Ms. Susanne REICHWEIN

MAIN AUTHORS OF THE REPORT

France	**Emmanuelle Amoros** *IFSTTAR*
Ireland	**Michael C. Brosnan** *National Road Safety Authority*
Netherlands	**Fred Wegman** *SWOV*
	Niels Bos *SWOV*
Spain	**Catherine Perez** *Agència de Salut Pública de Barcelona*
	Maria Segui *European Center for Injury Prevention*
	Rafael Heredero *European Center for Injury Prevention*
United Kingdom	**Barbara Noble** *ex Department for Transport*
	Pat Kilbey *Department for Transport*
ITF/OECD	**Véronique Feypell**
Consultant	**Colin Cryer** *University of Otago, New Zealand*

EDITOR

The final report was edited by Barbara Noble (ex UK Department for Transport).

PEER REVIEWERS

The final report was reviewed by Margie Peden and Dinesh Sehti from the World Health Organization. The IRTAD Group is very grateful for the comments they provided.

OTHER CONTRIBUTORS

Professors John Langley and Hank Weiss, Injury Prevention Research Unit, University of Otago, New Zealand, for their helpful comments on Chapter 6.

Reporting on Serious Road Traffic Casualties

Combining and using different data sources to improve understanding of non-fatal road traffic crashes

To improve further road safety, it is necessary to have a better understanding of the real number of road traffic casualties, including serious injuries. This is made possible by linking different sources of accident data, including police and hospital data. This report, prepared by the IRTAD Group, reviews how serious injuries are defined in IRTAD countries and identifies and assesses methodologies for linking different sources of crash data.

photo credit: Konstantin Sutyagin

International Transport Forum
2 rue André Pascal
75775 Paris Cedex 16
France
T +33 (0)1 45 24 97 10
F +33 (0)1 45 24 13 22
Email : itf.contact@oecd.org
Web: www.internationaltransportforum.org