5. Traffic safety data

Introduction

Traditionally, traffic safety relies on police accident records as the primary data source. However, these records may be unreliable, inaccessible, or simply too few or of poor quality to be useful. There are, however, alternative data sources that can complement, and in some cases replace the police data. In particular, the data from the healthcare system should be aimed at as it reveals the true scale and nature of traffic injuries. Other data sources include in-depth accident investigations, insurance data, self-reporting of accidents, traffic conflict and behavioural observations, Safety Performance Indicators and road safety inspections and audits.

Learning outcomes

After completing this module, the students should be able to:

- recall additional data types that can be used to complement police reports
- explain in which way each data source complements the police data
- explain the limitations of each data source
- use several data source to make a holistic assessment of traffic safety at a given location.

Key messages to learners

- *Police records* are the source of the official accident statistics in all countries. The data is usually available long back in time, collected in standardized ways, which allows analysing trends and compare regions or countries. However, police data is known to omit some accidents (under-reporting), contain little detail about accident circumstances, and is not very accurate when it comes to the description of injuries. Historically, the task of the police was to 'find the guilty', which still affects the perception of what is important to record when filling in the forms.

- *Healthcare records* are the most important complement to the police data. When injuries are severe enough, people eventually get in contact with the healthcare system for medical help. Unlike the police data, medical records are not readily available for a safety practitioner and reaching them requires additional efforts. Compared to a police record, a medical report contains much better description of the injuries but is less detailed when it comes to the accident circumstances. By combining police and healthcare records, it is possible to get a rough estimate of how many accidents/injuries have not been captured nether by the police nor by the healthcare system (the 'unknown unknowns').

- *In-depth* accident *investigations* are usually done on a representative sample of accidents, with the focus of recording as many details as possible. Usually, a team of experts from different fields (engineers, psychologist, doctors, etc.) works together to complete the accident report. In-depth investigations are costly but provide very valuable insights into accident contributing factors. Unlike the police, in-depth investigators are not concerned with whom to blame for the accident, but rather how similar accidents can be prevented in the future.

- To be able to settle traffic-related claims, *insurance* companies carry out *damage evaluations* for both property damage and personal injuries. The insurance premium typically takes into account driver and vehicle characteristics, such as accident history and the presence of specific safety features. This information is often much more detailed than what a police officer would normally record. Another advantage of insurance data is that it captures the long-term consequences of

injuries, as victims may continue to receive compensation during rehabilitation, which can last for years—or even a lifetime—after the accident. Although insurance data is generally not accessible to external parties, it is common for companies to conduct their own analyses and present the results publicly in aggregated form.

- *Self-reporting* of accidents typically involves questionnaires in which individuals are asked whether they have been involved in an accident during a recent period (ranging from the past month to the past three years). If the response is positive, follow-up questions are asked about the circumstances of the accident. Alternatively, injured individuals can be approached at medical emergency facilities and asked whether the injury is traffic-related, along with details of the incident. Self-reporting enables capturing unreported accidents—for example, one of the questions might ask whether the police were present at the scene. A drawback of this method is that a large number of people must be contacted to collect a meaningful amount of accident data, which can make the process costly.

- *Traffic conflicts* and other indirect measures of safety rely on observation of events in traffic that do not result in accidents but are safety-critical ('almost accidents'). The advantage of the approach is that safety problems can be identified relatively quickly (after few days of observations), without the ned to wait for years to collect accident records. The weakness of the method is that the conflicts need to be detected (by a human observer or, recently, computer vision tools) which may be labour intensive. Also, the interpretation of conflict counts is not always straightforward, e.g. it is not clear how many accidents one could expect per year if X conflicts were observed during a week.

- *Behavioural observations* are typically used to understand how a particular traffic environment operates in practice. For example, there may be 'unwritten rules' or common behavioural patterns that are not reflected in official traffic regulations and can therefore be difficult to anticipate. Behavioural data can be relatively straightforward to collect; however, it is important to clearly define the purpose of the data collection and have a well-thought-out plan for how the data will be analysed later.

- *Safety Performance Indicators* (*SPIs*) refer to a set of behavioural measures collected in a standardized way and on a relatively large scale. These indicators are typically chosen because they have a clear link to accident risk. Common SPIs include seat belt usage, helmet usage, speeding, and drunk driving. SPIs are useful for setting targets in traffic safety work and are well suited for analysing long-term trends.

- *Naturalistic data* is typically collected by equipping a set of vehicles with various sensors that continuously record information about the driver, the vehicle, and the surrounding environment while the vehicle is in use. Over time, drivers tend to forget that they are being observed and begin to behave more 'naturally'—hence the name. These studies are generally expensive and are primarily conducted for research purposes.

- *Road safety inspections* and *audits* are typically carried out by trained professionals who apply their traffic safety knowledge and experience to identify potential risks. Inspectors are expected to look beyond whether road standards are formally met and consider whether a road that is technically compliant may still pose risks under certain conditions—for example, if drivers are blinded by the sun at specific times of day and fail to see important signs. It is important to have a system in place for documenting identified safety deficiencies and ensuring that they are eventually addressed.

- Particularly in the context of low- and middle-income countries, official sources of traffic safety data are often difficult to access and suffer from poor quality. This has led to the development of some *grassroot data collection methods*—often driven by NGOs—such as analysis of newspaper articles and social media posts to identify and locate traffic accidents, or asking community members to report hazardous locations and witnessed accidents. While these methods play an

important role in the absence of more reliable data, their quality is limited, and they should not be seen as a permanent substitute for more traditional data sources.

Learning activities

Exercise 1

Preparations (teacher only):

- Identify a location (not far away from the class venue) where you expect some traffic safety problems. It can be an intersection, a pedestrian crossing, or a road section with much activity on the sides that create 'friction' (e.g. parked vehicles stopping and leaving, local traders and customers in on or very close to the road, etc.). Aim for a medium-size location.
- Approach the local police to receive accident reports for the last five years. If this is absolutely not possible, ask for an empty accident reporting form and make your own 'reports' (for exercise purpose only).
- Approach the local traffic engineers to receive any background information about the location—drawings, traffic counts, traffic light scheme, etc.

 With the students:

- Examine the location drawings, GoogleMaps satellite and StreetView images. What potential traffic safety problems could you expect here?
- Examine the police accident reports. Can you see any patterns? Are they in line with what you expected?
- Do a web search on social media. Could you find any accident reports by witnesses from the same location? Are these the same accidents that were registered by the police?
- Visit the location. Spend 20–30 minutes observing how the traffic operates. Can you see any 'disturbances' or unconventional behaviours that were not so obvious when just looking at the drawings? Is the general speed level appropriate for this type of environment? Are the formal traffic rules followed? If not, does it create safety risks? Who yields to whom? Do vulnerable road users have dedicated space and are they respected by motor vehicle users? Could you see any safety-critical situations (near-misses)? Are the observed near-misses follow the same pattern as the police accident reports?

Assessment quiz

The assessment quiz can be used as a part of the examination, or as another form of learning activities.

Question 1

What is true about police accident records:

a. Police records are available over long time periods with relatively consistent methods and definitions used in the data collection. (correct)

b. Police data is particularly good for PDO (property damage only) accidents. (incorrect)

   Comment (shown after the answer has been given): Nope. Large (often largest) share of PDO accidents is not known to the police. The reasons vary—people may mistrust the police and prefer to settle things on their own, the police may refuse to come if the accident is minor, etc. From practical perspective, the police data on PDOs is a highly unreliable source to base decisions on.

c. Police report is based on police officer's and witnesses' observations (often documented directly on the accident spot). (correct)

d. Policemen are good at estimating the types and severity of the injuries caused by the accident. (incorrect)

Comment (shown after the answer has been given): Comparisons between severity estimates done by policemen on the spot and medical professionals later at healthcare facilities show that police reports tend to overestimate the severity. Generally, a policeman is seldom qualified to make a proper medical diagnosis.

e. Historically, the task of the police investigation was to find who is guilty, not how a similar accident can be prevented in the future. This often results in a 'skewed' reports on accident contributing factors with road users being blamed in most cases. (correct)

Question 2

What is true about healthcare records?

a. There are many types of medical facilities in which an accident victim may be treated. It creates problems in data collection since the records are spread over many organizations that often do not 'talk' to each other. (correct)

b. Accident location and circumstances are well documented in a healthcare report. (incorrect)

c. Medical personal feel often that collection of the accident information is an extra burden for them and not a priority given other urgent tasks (like taking care of the injured people). (correct)

d. Matching of the police and healthcare records allows to estimate the scale of under-reporting by each data source, as well as the number of accidents not registered in any of them. (correct)

e. When compared to the healthcare records, it becomes evident that the police under-reports to a greater extent accidents involving vulnerable road users (cyclists, pedestrians, etc.). (correct)

f. It is common that medical personal visits the accident site to collect additional data about the accident in case the patient could not provide all the relevant information. (incorrect)

g. One accident may generate several healthcare reports. (correct)

Comment (shown after the answer has been given): More than one person can get injured in the same accident. To make things even more complicated, they might not be treated at the same hospital but rather sent to different medical facilities depending on what special equipment and expertise is needed and is available. Finding and re-connecting these healthcare records to the original accident may become a difficult quest.

h. Healthcare records has solved the problem of accident under-reporting. (incorrect)

Comment (shown after the answer has been given): While healthcare data complements police reports, accidents/injuries that are not known to any of them still go under-reported. While it is possible to roughly estimate the number of 'unknown accidents', the procedure is not very accurate.

Question 3

What is true about traffic conflicts and behavioural observations?

a. Traffic conflict and behavioural observations are performed on a large scale and there is a special database for that at each city with data going many years back in time. (incorrect)

b. Traffic conflicts are used to assess traffic safety by observing safety-critical events in real traffic (without waiting for real accidents to happen). (correct)

c. It is simple and straightforward to estimate the expected number of accidents knowing how many traffic conflicts occur per day. (incorrect)

d. Behavioural studies are often custom designed depending on the type of problem that needs investigation. (correct)

e. Safety Performance Indicators (SPIs) describe certain types of behaviour known to be strongly related to safety. They are collected repeatedly on a national scale, allowing to study trends over type. (correct)

Recommended reading and resources for students

- Laureshyn, A., Agerholm, N., & Sucha, M. (Eds.). (2025). *Traffic safety data: the way to better knowledge and better traffic safety*. IET. https://doi.org/10.1049/PBTR028E.
- Polders, E., & Brijs, T. (Eds.). (2018). *How to analyse accident causation? A handbook with focus on vulnerable road users*. Deliverable 6.3. Horizon 2020, InDeV project. https://www.bast.de/InDeV/EN/Handbook/Handbook-InDeV.pdf.

Recommended (additional) reading for teacher

- Laureshyn, A., & Varhelyi, A. (2018). The Swedish Traffic Conflict Technique: observer's manual. https://www.ictct.net/wp-content/uploads/SMoS_Library/LIB_Laureshyn_Varhelyi_2020.pdf

Prepared by expert

In case you have specific questions, need a discussion partner, or just want feedback on your lecture materials, you may reach out the author(s) of this module. Please, put 'AfroSAFE curriculum' in the email subject.

Aliaksei Laureshyn
Lund University
Sweden

aliaksei.laureshyn@tft.lth.se