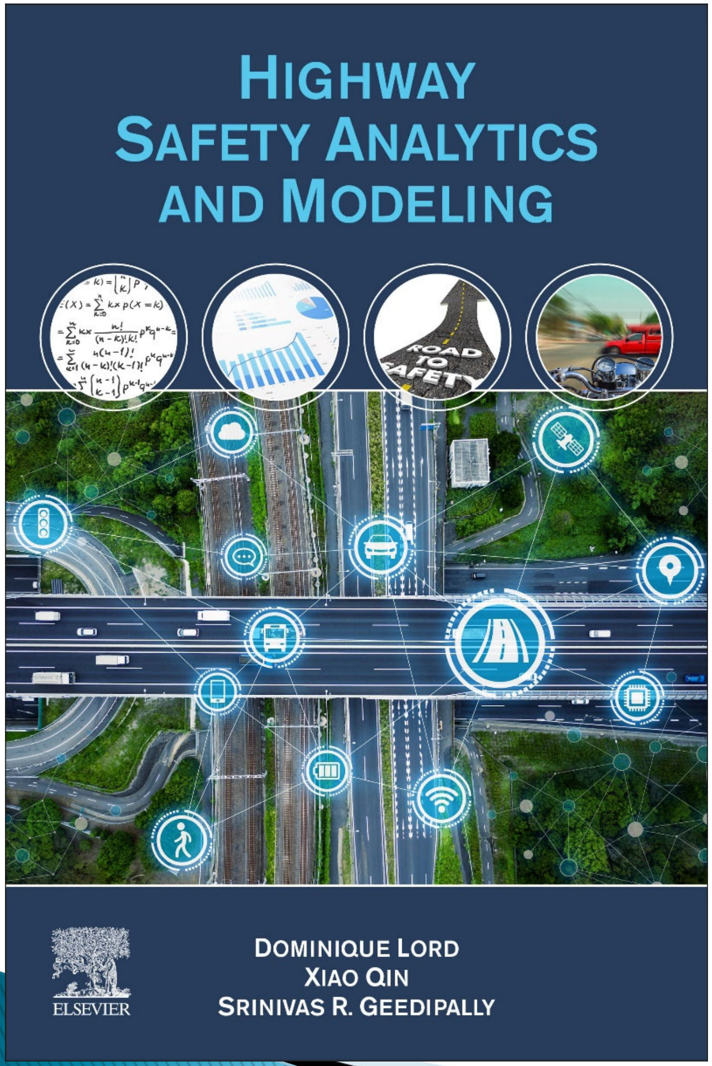
Fundamentals and Data Collection

Part 1

October 24, 2021

Instructor: Dominique Lord
Texas A&M University
d-lord@tamu.edu

Textbook



The material presented in this series of lectures are taken from this textbook and other sources based on lectures given by the authors.

The textbook is available on Amazon and the Elsevier website below among other places.

https://www.elsevier.com/books/highway-safety-analytics-and-modeling/lord/978-0-12-816818-9

Origin of *Highway Safety Analytics and Modeling*

Dominique Lord is a Professor and A.P. and Florence Wiley Faculty Fellow in the Zachry Department of Civil and Environmental Engineering at Texas A&M University. His highway safety research has led to the development of new and innovative methodologies for analyzing crash data and has been used by researchers across the world in medicine, accounting, mathematics, statistics, biology, and engineering..

Xiao Qin is a Professor of Civil and Environmental Engineering, the Director of the Institute for Physical Infrastructure and Transportation, and a licensed professional engineer in Civil Engineering. He chairs the Transportation Research Board (TRB) ACS20(1) Subcommittee on Safety Analytical Methods. He is the Editor of Transportation Research Record and Journal of Transportation Safety & Security, and an Board Member of Accident Analysis & Prevention.

Srinivas R. Geedipally is a Research Engineer in the Center for Transportation Safety at Texas A&M Transportation (TTI) and a registered Professional Engineer in the state of Texas. He received his doctorate from Texas A&M University and has been with TTI since 2005. Dr. Geedipally is an Advisory Board Member of Analytic Methods in Accident Research and has numerous papers published in high-standard international journals and conferences.

Objectives of Textbook

- Provide **information** for practitioners, engineers, scientists, students, and researchers who are interested in analyzing safety data to make engineering-or policy-based decisions.
- Cover all aspects of the **decision-making** process, from collecting and assembling data to making decisions based on the results of the analyses.
- Provide the latest **tools and methods** documented in the literature for analyzing crash data.
- Provide **examples and exercises** to help understand models and methods commonly used for analyzing crash data.

Synopsis of Textbook

Theory and Background

- Describe fundamental and theoretical principles associated with safety data analyses.
- Covers the nature of crash data from the human and statistical/mathematical perspectives, as well as key crash-frequency and crash-severity models.

Highway Safety Analyses

- Describe how the models described in the first area are applied for analyzing safety data.
- Include methods for exploring safety data, conducting cross-sectional and before-after studies, identifying hazardous sites or sites with promise as well as tools for incorporating spatial correlation, and identifying crash risk on a near real-time basis.

Alternative Safety Analyses

- Assemble alternative safety analysis tools.
- Include methods for surrogate measures of safety and data mining techniques for extracting relevant information from datasets, including those categorized as big data.

Outline of Textbook

Chapters
Theory and Background
Chapter 2—Fundamentals and Data Collection
Chapter 3—Crash-Frequency Modeling
Chapter 4—Crash-Severity Modeling
Highway Safety Analyses
Chapter 5—Exploratory Analyses of Safety Data
Chapter 6—Cross-Sectional and Panel Studies in Safety
Chapter 7—Before-After Studies in Safety
Chapter 8—Identification of Hazardous Sites
Chapter 9—Models for Spatial Data
Chapter 10—Capacity, Mobility, and Safety
Alternative Safety Analyses
Chapter 11—Surrogate Safety Measures
Chapter 12—Data Mining and Machine Learning Techniques

What is safety?

Highway Safety

- Significant global issue both in terms of morbidity and economic losses (discussed in greater details below).
- Main goal consists of reducing the number and severity associated with crashes (vehicles, pedestrians and bicyclists). (Civil engineers focus on the design and operation of highways and their effects on safety)
- Was not a hot topic (research wise) until the early to mid-90s. (Few researchers)
- Now, a lot of funding, research and application of methods are devoted to highway safety. (Lots of researchers)
- Vision-Zero/Safe Systems

(Perception of safety)

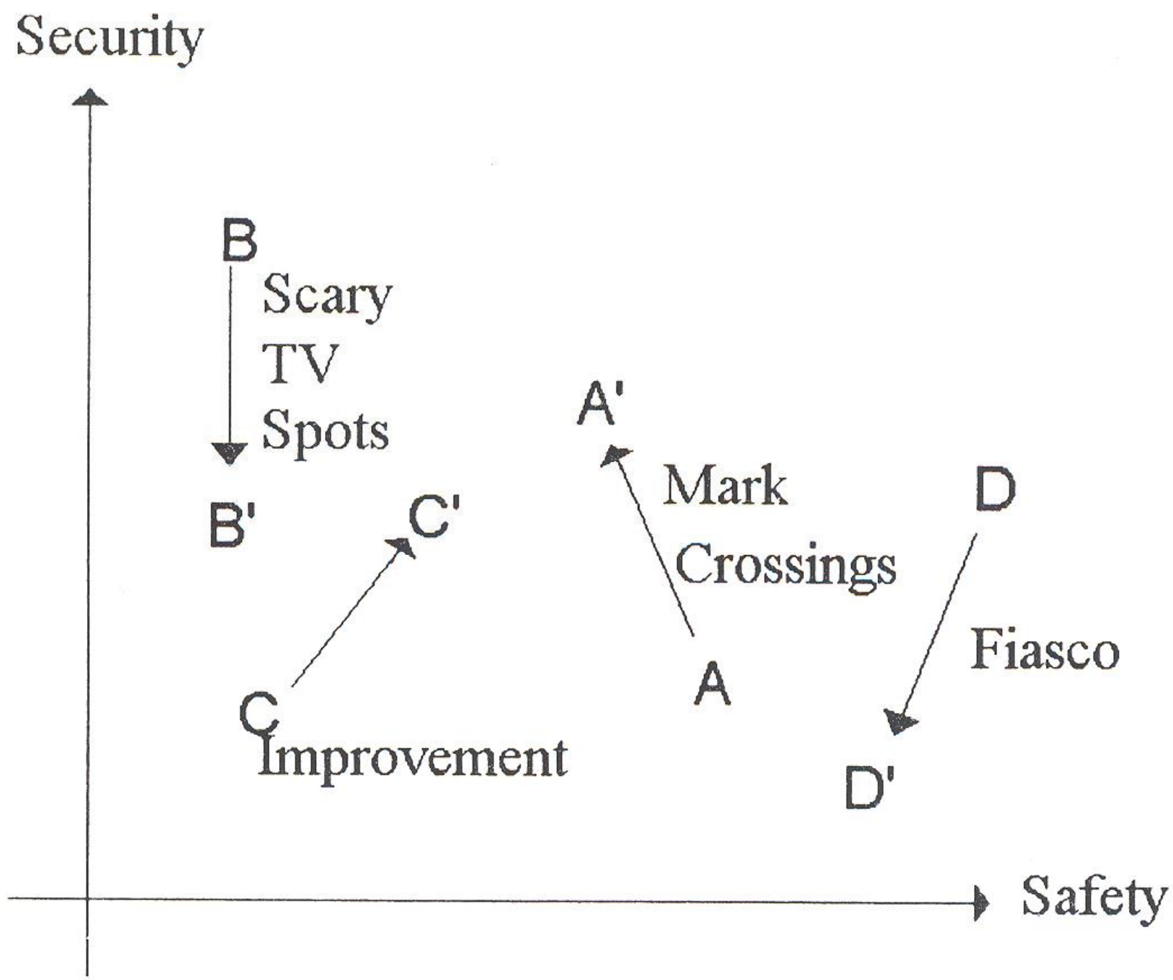


Figure 3.1. Changes in 'safety' and 'security'.

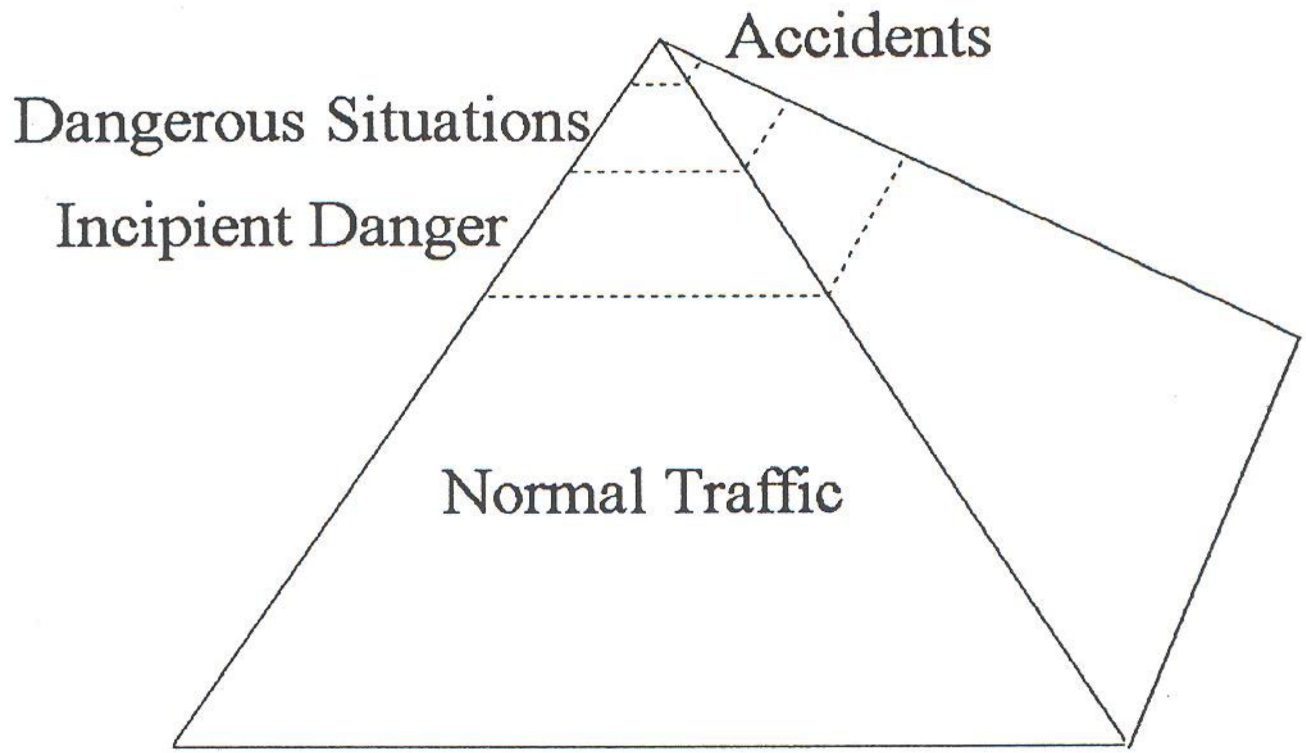
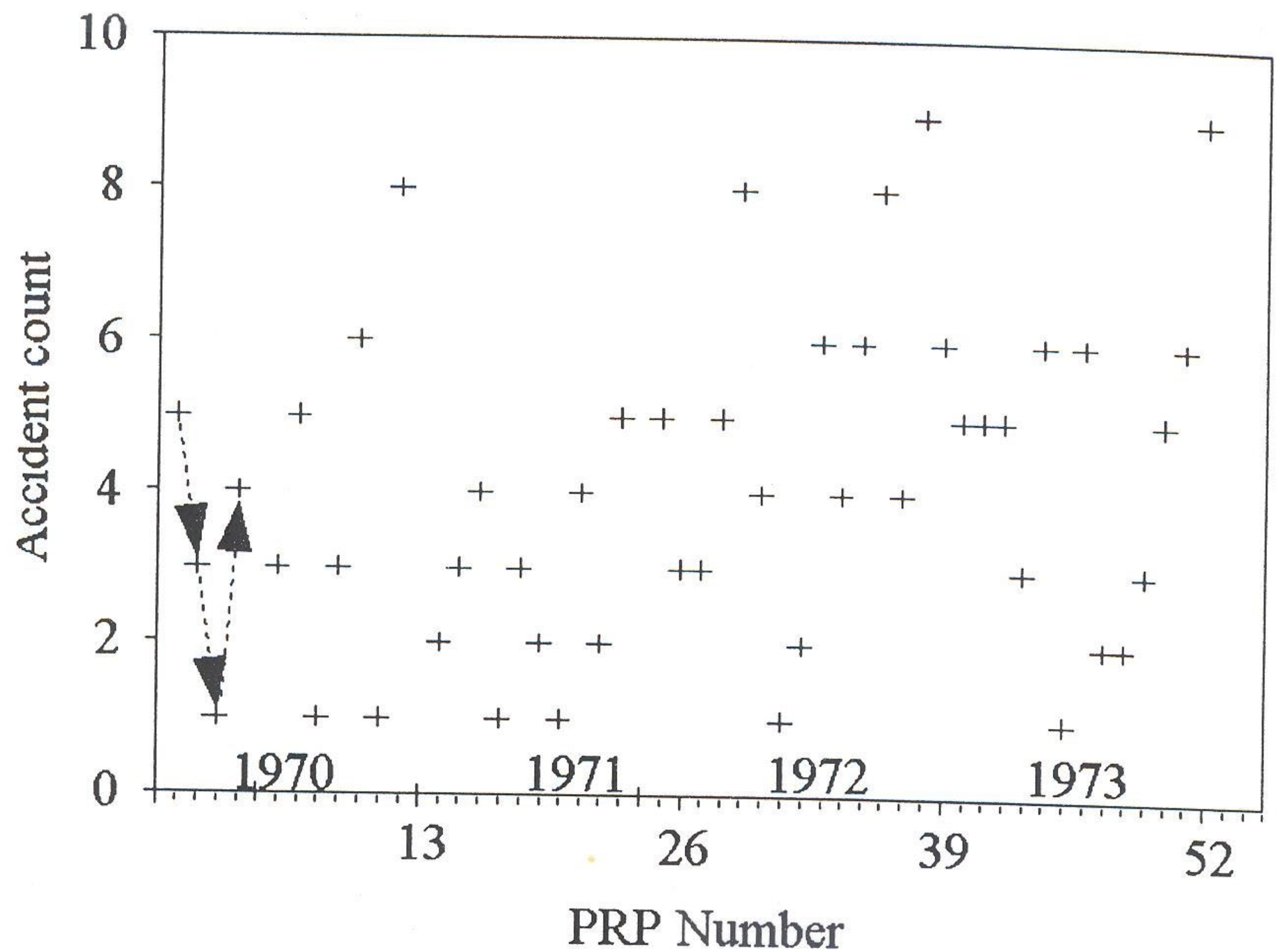
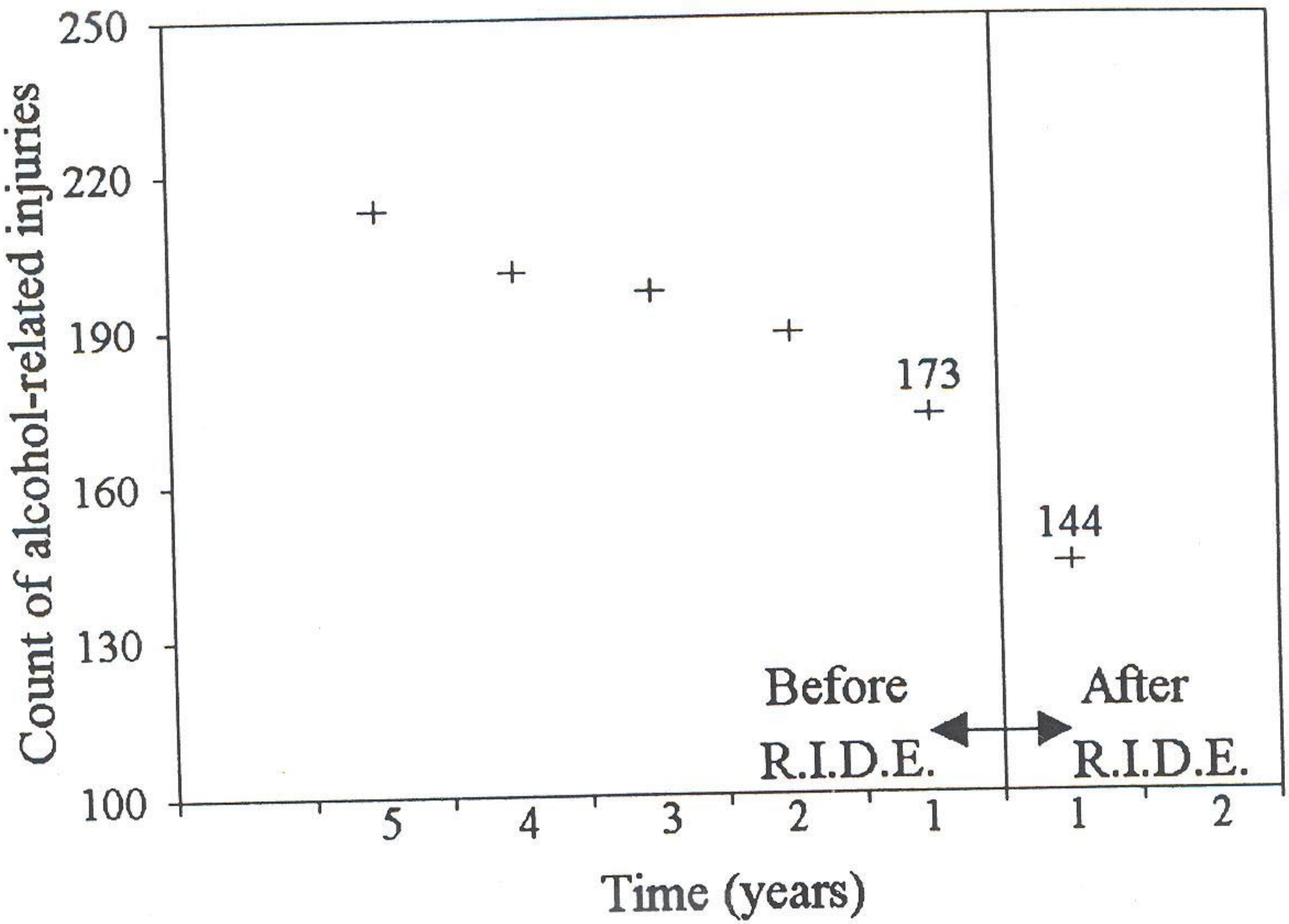


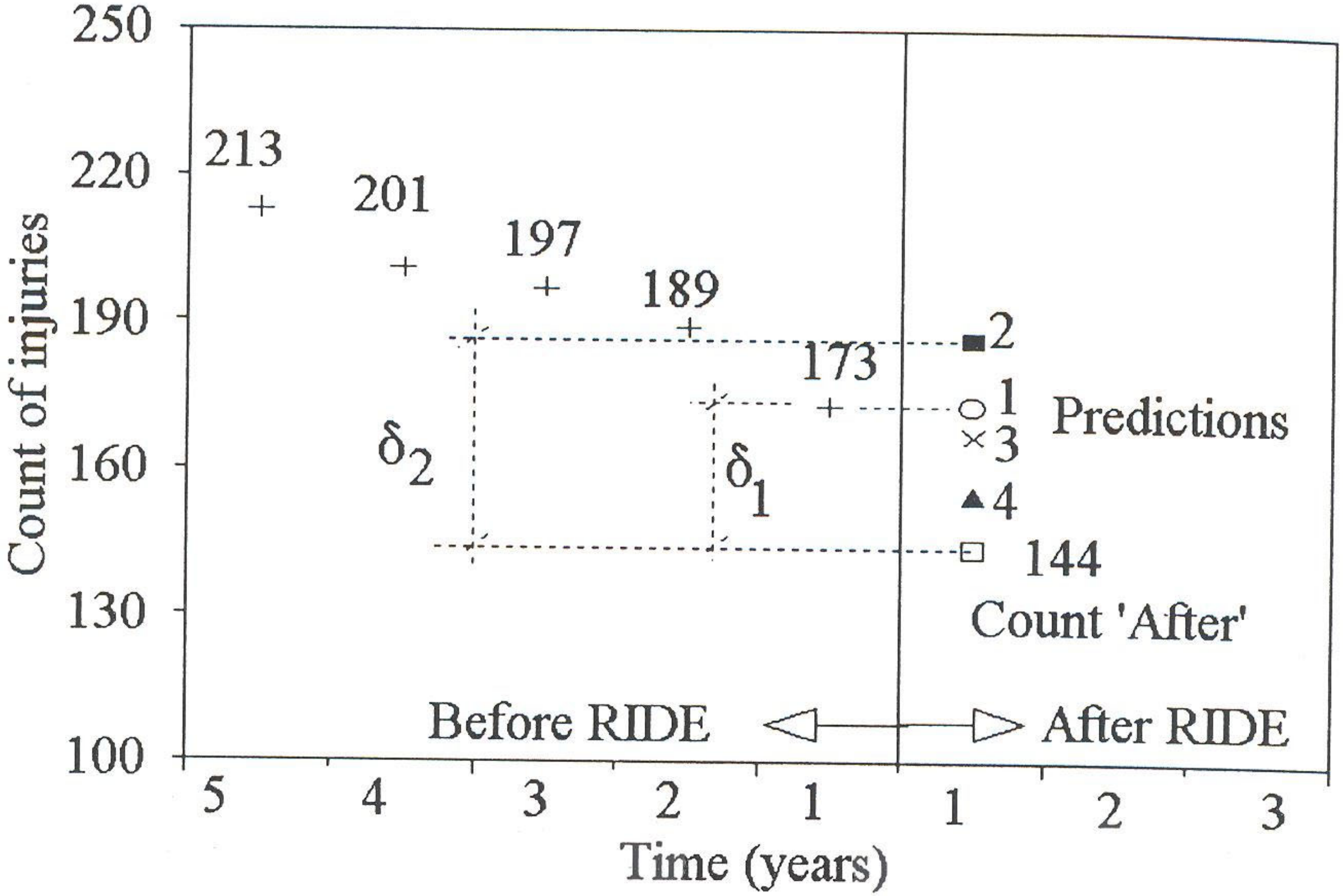
Figure 3.2. The continuum of events.



10
8
6
4
2
0
Accident count
1970
1971
1972
1973
13
26
39
52
PRP Number



250
220
190
160
130
100
Count of alcohol-related injuries
173
144
Before
R.I.D.E.
After
R.I.D.E.
5
4
3
2
1
1
2
Time (years)



250
220
190
160
130
100
Count of injuries
213
201
197
189
173
2
1
3
4
144
Predictions
Count 'After'
δ_2
δ_1
Before RIDE
After RIDE
5
4
3
2
1
1
2
3
Time (years)

Highway Safety Definition

For the purpose of this series of seminars/lectures, the relevant definition is to examine the number of crashes per unit of time by severity level.

The goal is to estimate how the design and traffic operation of different elements of the highway network affect the number and severity of crashes. For us, it is from the perspective of civil engineering (the topic of the textbook).

Note that highway safety involves multiple fields, such as law enforcement, psychology/human factors, mechanical, industrial and electrical engineering, economics, sociology, medicine and policy makers among others. It is a multidisciplinary discipline. Specific for this course:

- Making sense of crash and other safety data
- Examine the relationship between design/operational variables and crash risk
- Understanding crash causation factors
- Estimating risk of injury when a crash occur
- Estimating the safety effects of treatments
- Economic impacts of safety decisions

Global Impacts of Traffic-Related Injuries

Overview - Numbers

- According to the World Health Organization (WHO), between 2000 and 2016, roadway-related crashes increased from about 1.15 million to 1.35 million deaths globally.
- On an annual basis, about 80 million nonfatal injuries warranting medical care occur on highway networks.
- Road traffic injuries are ranked eighth as the leading cause of death (2.5%) among people of all ages, right in front of diarrheal diseases and tuberculosis (WHO, 2018).
- Vulnerable road users (i.e., pedestrians and cyclists) represent 26%of road injury deaths, while drivers and passengers of motorized two-wheel and three-wheel vehicles account for another 28% worldwide.
- While a large proportion of high-income countries have observed either a reduction or no change in traffic-related deaths between 2013 and 2016, a significant number of middle- and low-income countries have observed an increase in traffic-related deaths (WHO, 2018), in large part attributed to the rapid motorization observed in developing countries.

Overview – Economic Burden

- In the United States, for instance, highway crashes are estimated to have caused more than US$871 billion in economic loss and societal harm in 2010.
- In Europe, it is estimated that crashes have cost more than US$325 billion (€280 billion) in economic harm in 2015 (this value is considered underestimated), while in Australia the economic burden was estimated to be US$ 23.9 billion (AU$33.2) in 2016 (Litchfield, 2017).
- Globally, it is estimated that 3% of gross domestic product (GDP) is lost to highway crashes (all severities) and can be as high as 5% for middle- and low-income countries.

TABLE 1.2

Change in rank order of DALYs for the 10 leading causes of the global burden of disease

	1990		2020
Rank	Disease or injury	Rank	Disease or injury
1	Lower respiratory infections	1	Ischaemic heart disease
2	Diarrhoeal diseases	2	Unipolar major depression
3	Perinatal conditions	3	Road traffic injuries
4	Unipolar major depression	4	Cerebrovascular disease
5	Ischaemic heart disease	5	Chronic obstructive pulmonary disease
6	Cerebrovascular disease	6	Lower respiratory infections
7	Tuberculosis	7	Tuberculosis
8	Measles	8	War
9	Road traffic injuries	9	Diarrhoeal diseases
10	Congenital abnormalities	10	HIV

DALY: Disability-adjusted life year. A health-gap measure that combines information on the number of years lost from premature death with the loss of health from disability.

Source: reference 2.

TABLE 1.1

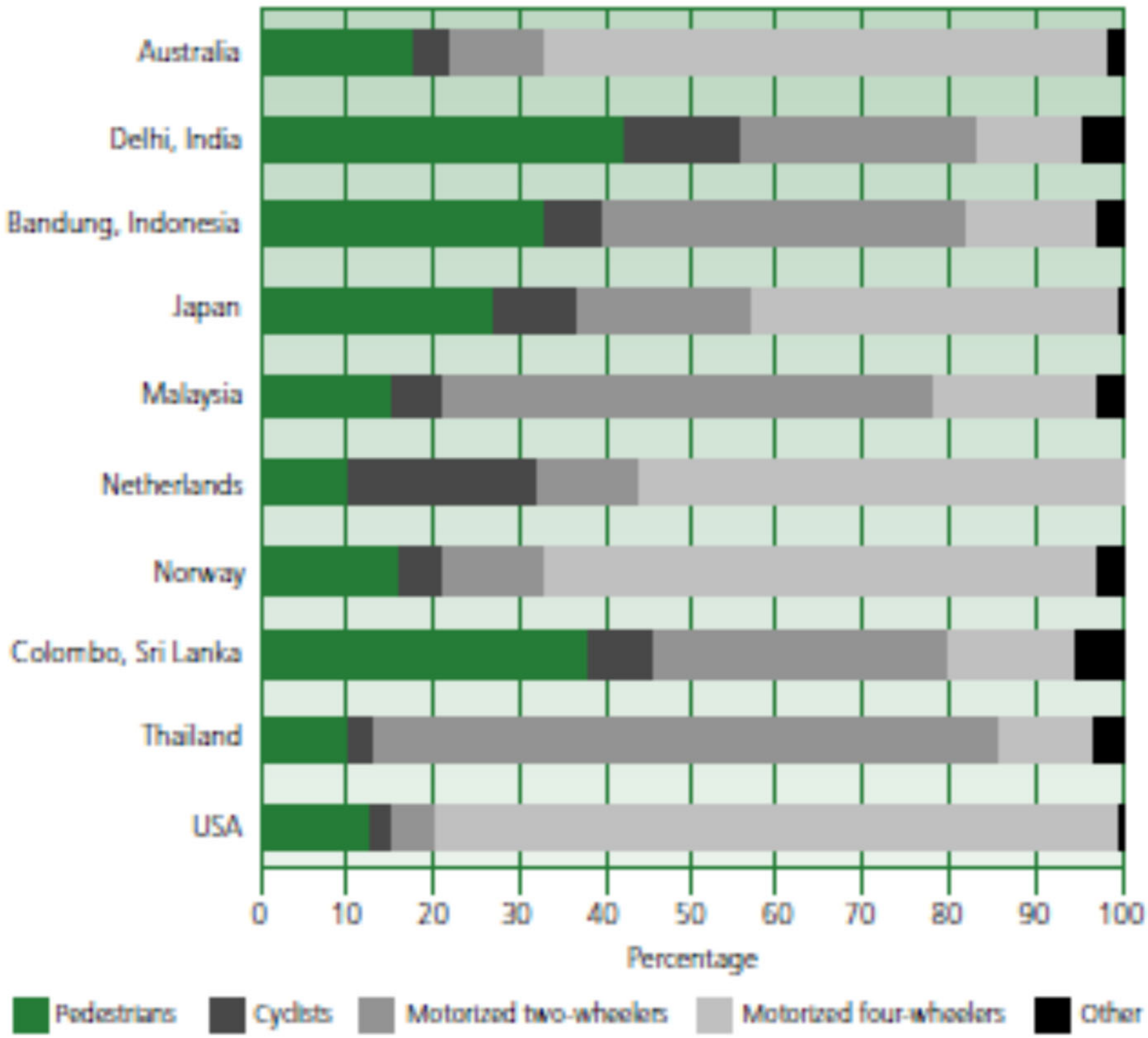
Leading causes of deaths by age group, world, 2002

Rank	0–4 years	5–14 years	15–29 years	30–44 years	45–59 years	≥60 years	All ages
1	Lower respiratory infections 1 890 008	Childhood cluster diseases 219 434	HIV/AIDS 707 277	HIV/AIDS 1 178 856	Ischaemic heart disease 1 043 978	Ischaemic heart disease 5 812 863	Ischaemic heart disease 7 153 056
2	Diarrhoeal diseases 1 577 891	Road traffic injuries 130 835	Road traffic injuries 302 208	Tuberculosis 390 004	Cerebrovascular disease 623 099	Cerebrovascular disease 4 685 722	Cerebrovascular disease 5 489 591
3	Low birth weight 1 149 168	Lower respiratory infections 127 782	Self-inflicted injuries 251 806	Road traffic injuries 285 457	Tuberculosis 400 704	Chronic obstructive pulmonary diseases 2 396 739	Lower respiratory infections 3 764 415
4	Malaria 1 098 446	HIV/AIDS 108 090	Tuberculosis 245 818	Ischaemic heart disease 231 340	HIV/AIDS 390 267	Lower respiratory infections 1 395 611	HIV/AIDS 2 818 762
5	Childhood cluster diseases 1 046 177	Drowning 86 327	Interpersonal violence 216 169	Self-inflicted injuries 230 490	Chronic obstructive pulmonary diseases 309 726	Trachea, bronchus, lung cancers 927 889	Chronic obstructive pulmonary diseases 2 743 509
6	Birth asphyxia and birth trauma 729 066	Malaria 76 257	Lower respiratory infections 92 522	Interpersonal violence 165 796	Trachea, bronchus, lung cancers 261 860	Diabetes mellitus 749 977	Diarrhoeal diseases 1 766 447
7	HIV/AIDS 370 706	Tropical cluster diseases 35 454	Fires 90 845	Cerebrovascular disease 124 417	Cirrhosis of the liver 250 208	Hypertensive heart disease 732 262	Childhood-cluster diseases 1 359 548
8	Congenital heart anomalies 223 569	Fires 33 046	Drowning 87 499	Cirrhosis of the liver 100 101	Road traffic injuries 221 776	Stomach cancer 605 395	Tuberculosis 1 605 063
9	Protein–energy malnutrition 138 197	Tuberculosis 32 762	War 71 680	Lower respiratory infections 98 232	Self-inflicted injuries 189 215	Tuberculosis 495 199	Trachea, bronchus, lung cancers 1 238 417
10	STDs excluding HIV 67 871	Protein–energy malnutrition 30 763	Hypertensive disorders 61 711	Poisonings 81 930	Stomach cancer 185 188	Colon and rectum cancers 476 902	Malaria 1 221 432
11	Meningitis 64 255	Meningitis 30 694	Maternal haemorrhage 56 233	Fires 67 511	Liver cancer 180 117	Nephritis and nephrosis 440 708	Road traffic injuries 1 183 492
12	Drowning 57 287	Leukaemia 21 097	Ischaemic heart disease 53 870	Maternal haemorrhage 63 191	Diabetes mellitus 175 423	Alzheimer and other dementias 382 339	Low birth weight 1 149 172
13	Road traffic injuries 49 736	Falls 20 084	Poisoning 52 956	War 61 018	Lower respiratory infections 160 259	Liver cancer 367 503	Diabetes mellitus 982 175
14	Endocrine disorders 42 619	Violence 18 551	Childhood cluster diseases 48 101	Drowning 56 744	Breast cancer 147 489	Cirrhosis of the liver 366 417	Hypertensive heart disease 903 612
15	Tuberculosis 40 574	Poisonings 18 529	Abortion 43 782	Liver cancer 55 486	Hypertensive heart disease 129 634	Oesophagus cancer 318 112	Self-inflicted injuries 874 955

Source: WHO Global Burden of Disease project, 2002, Version 1 (see Statistical Annex).

FIGURE 2.7

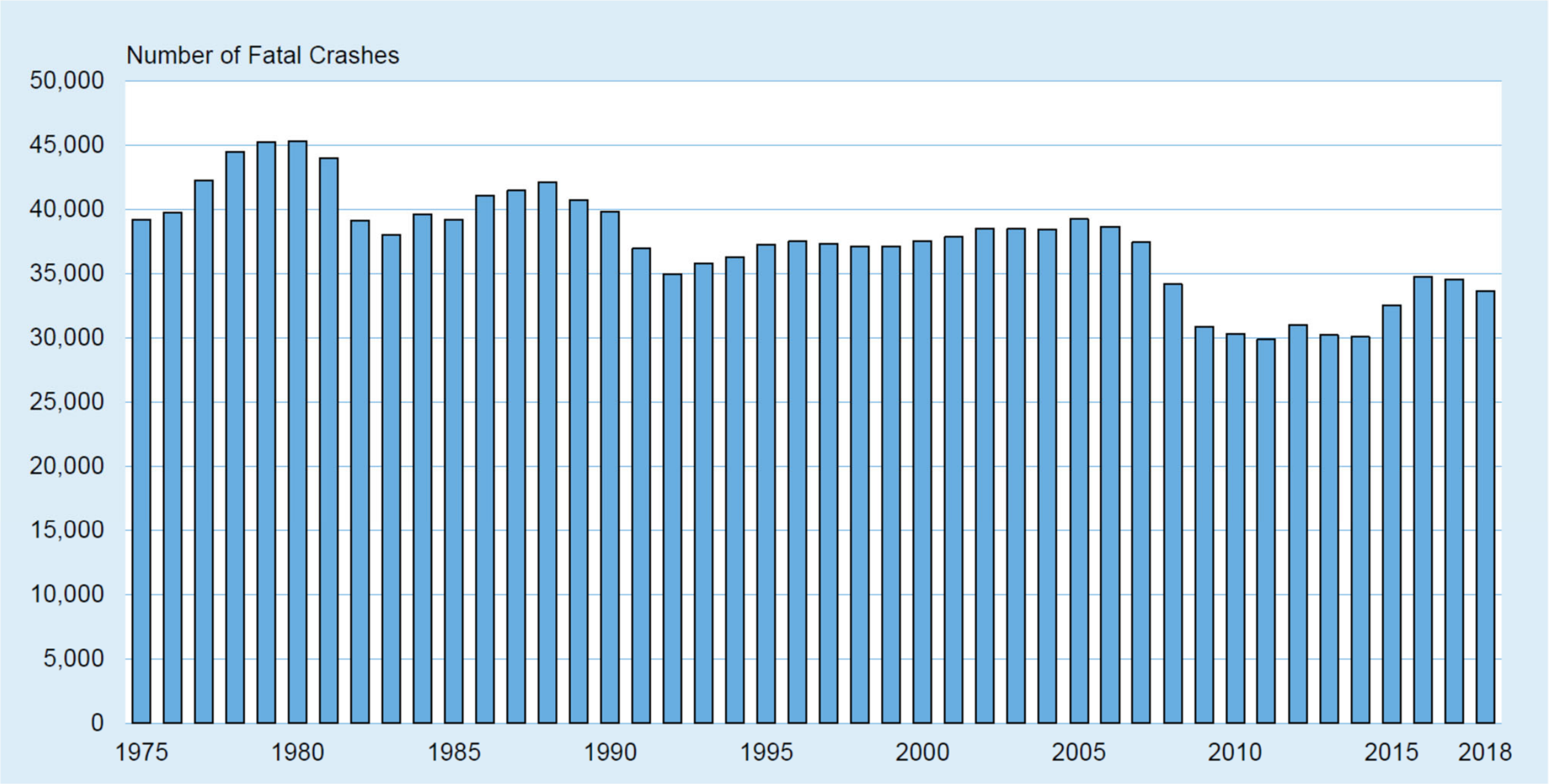
Road users killed in various modes of transport as a proportion of all road traffic deaths



Source: reference *44*.

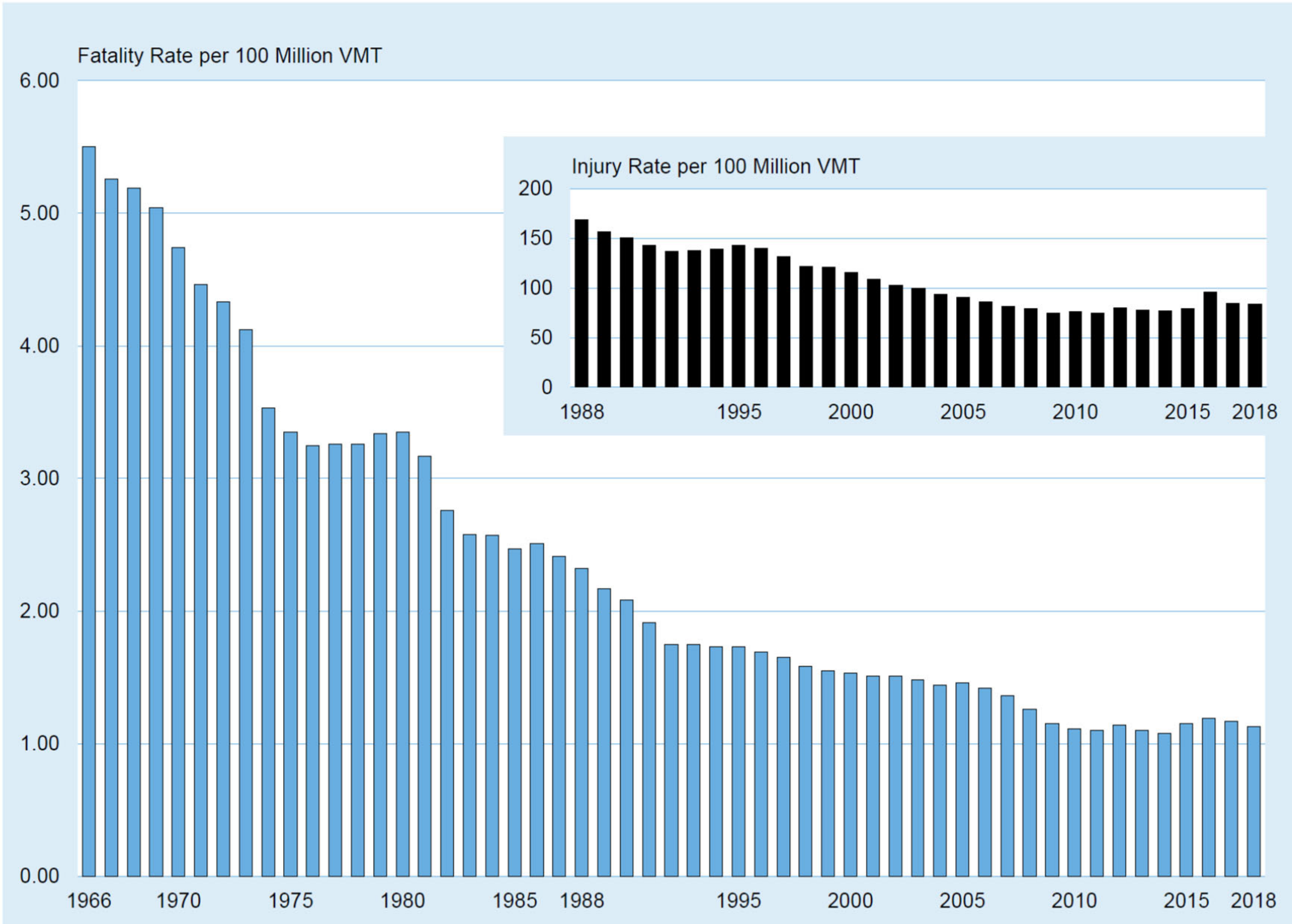
Crash Characteristics in the United States and China

Figure 1. Fatal Crashes, 1975-2018



https://crashstats.nhtsa.dot.gov/#!/DocumentTypeList/12

Figure 2. Motor Vehicle Fatality and Injury Rates per 100 Million Vehicle Miles Traveled, 1966-2018



Source: Vehicle Miles Traveled—FHWA, revised by NHTSA for passenger cars and light trucks

Figure 8. Proportion of People Killed, by Highest Driver Blood Alcohol Concentration in the Crash, 1982-2018

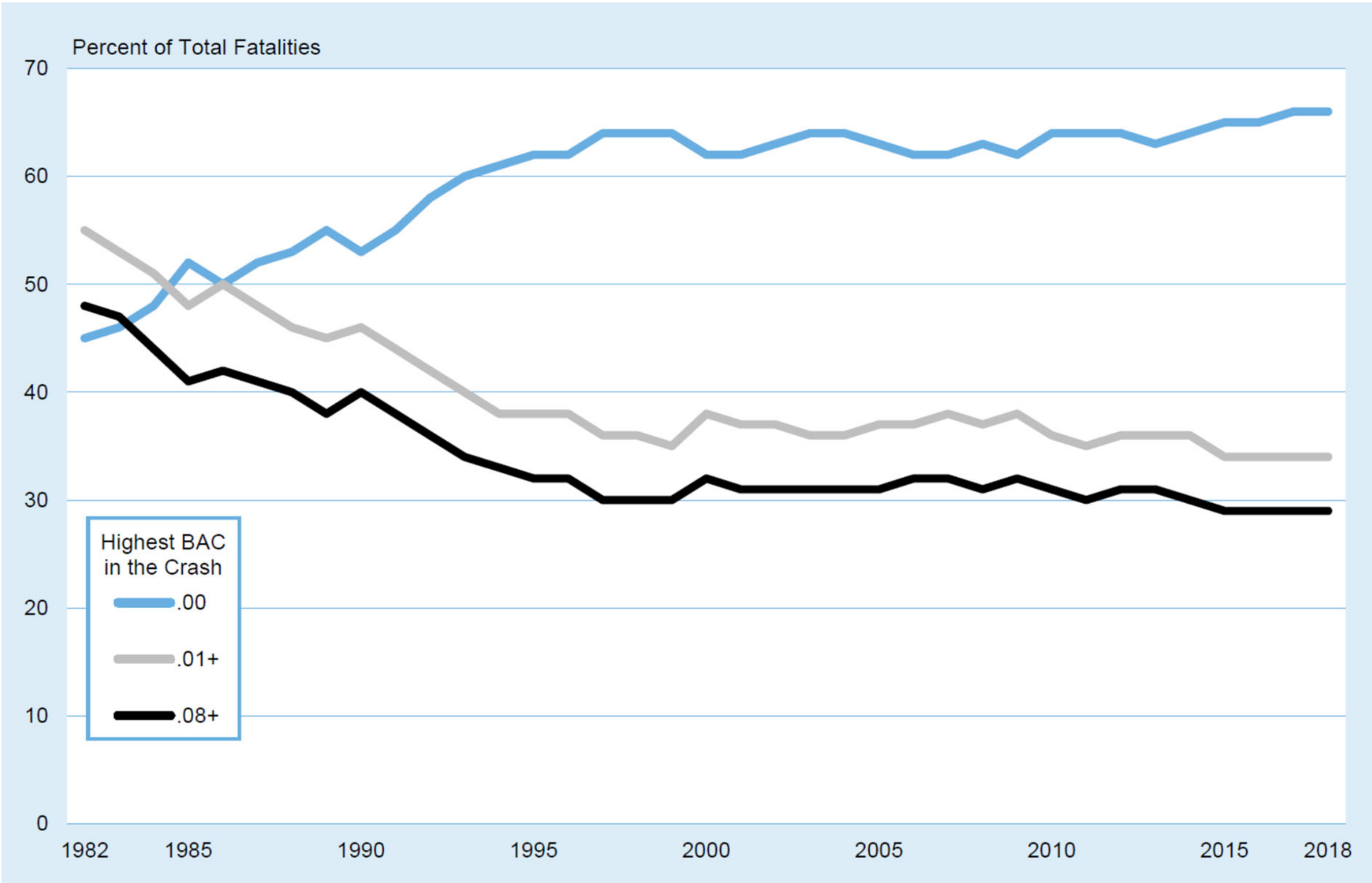
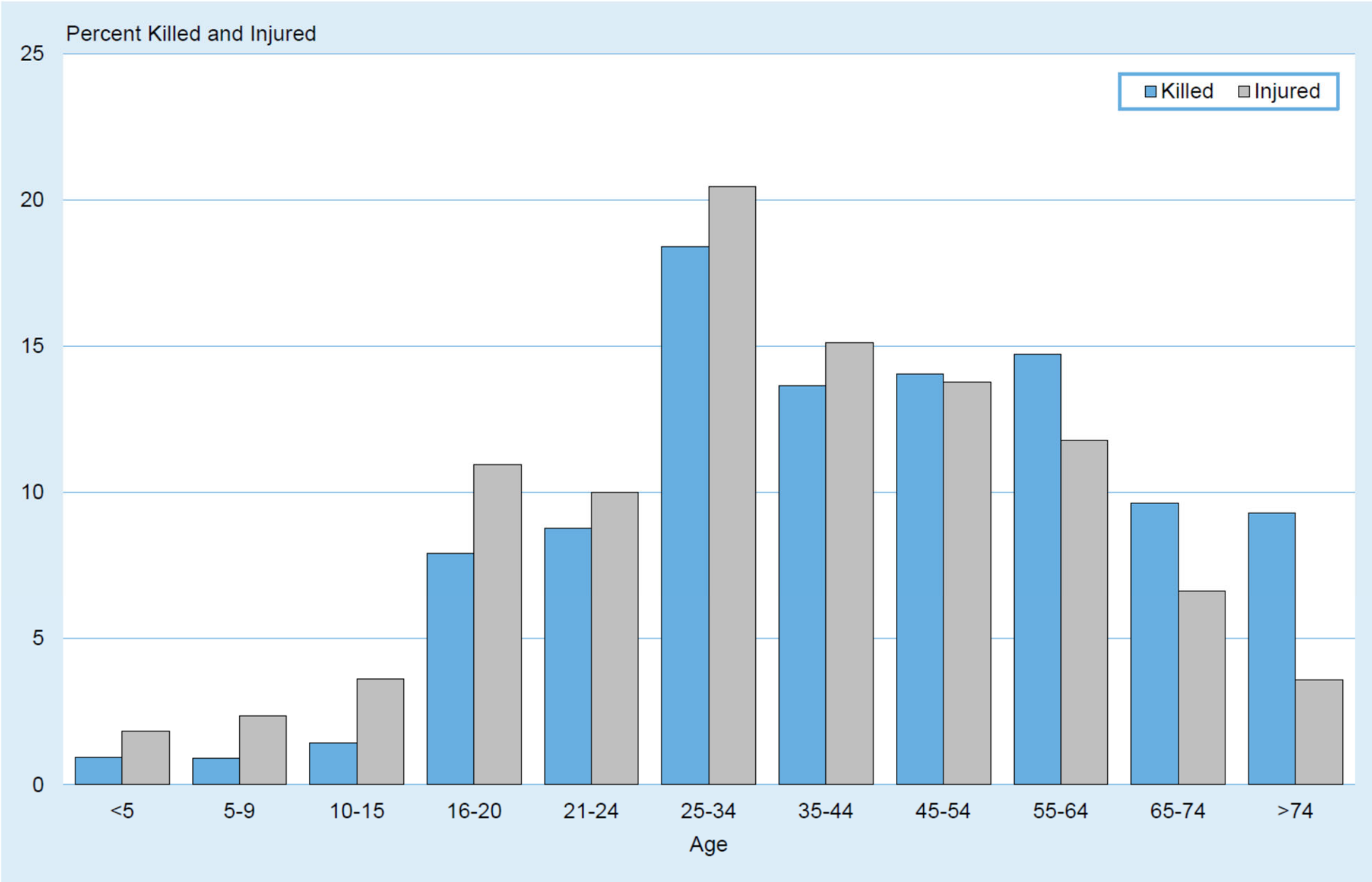
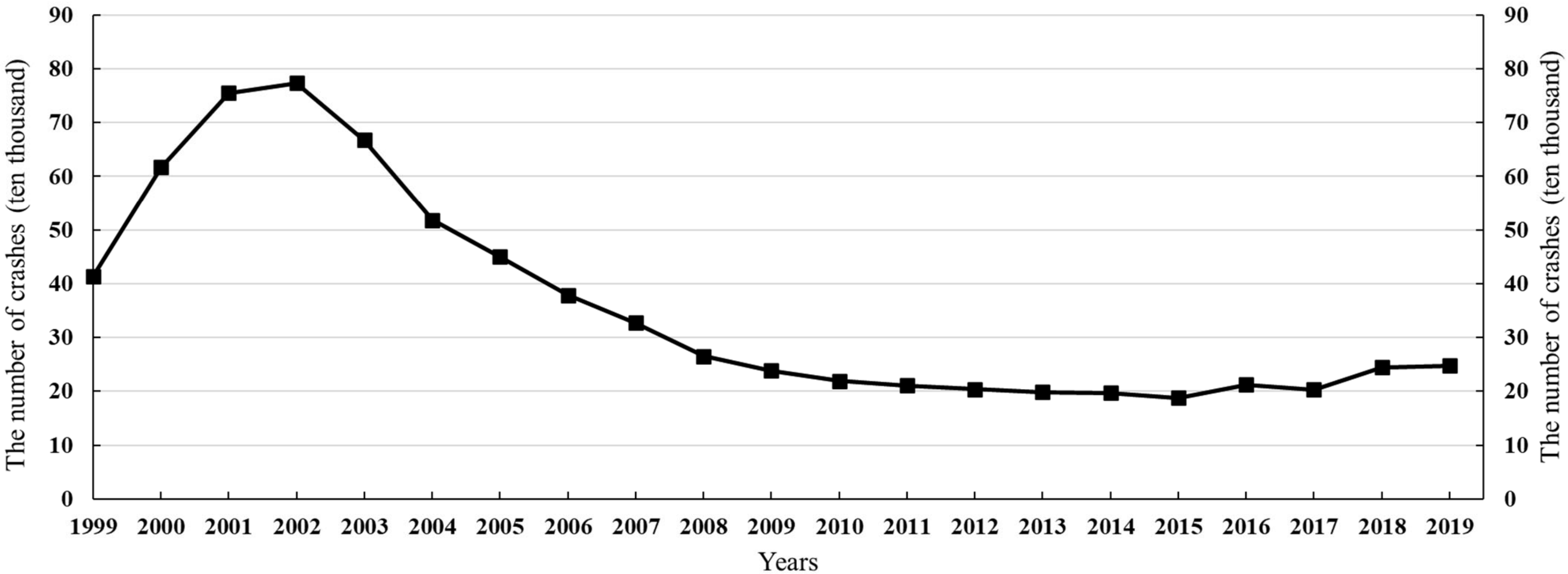


Figure 17. Percentage of People Killed and Injured, by Age



1.3 Development status of road traffic accident

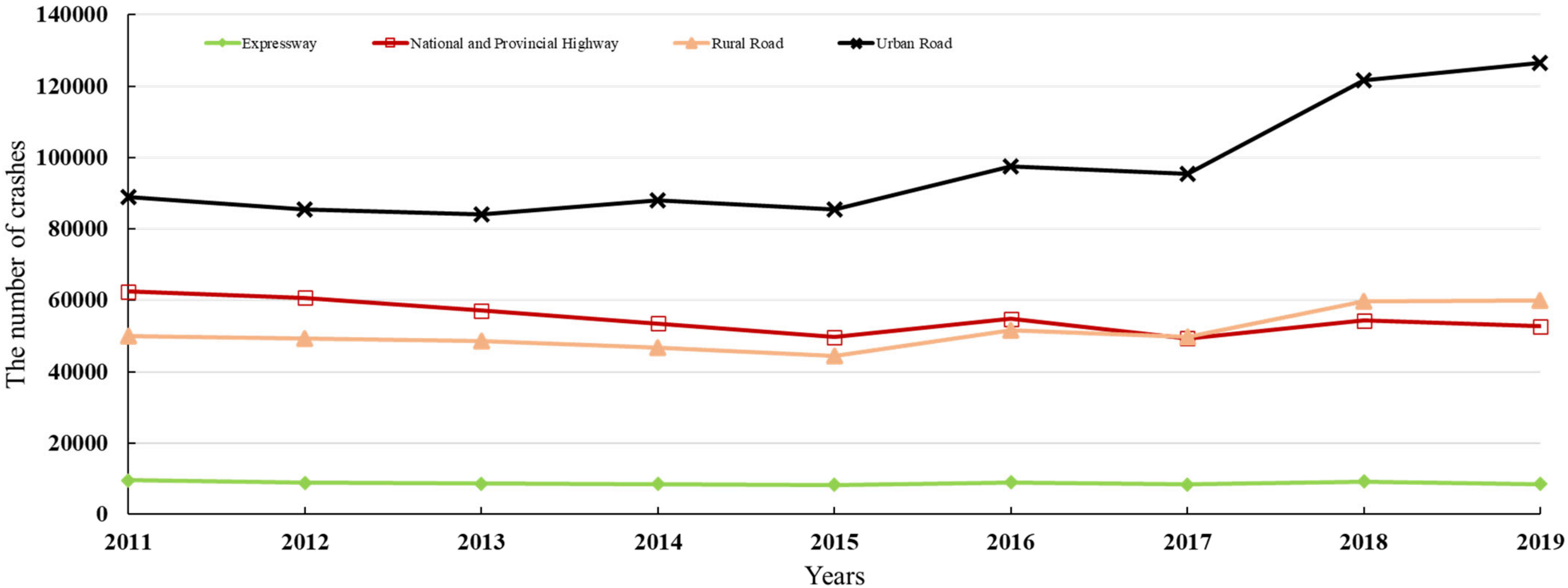


The number of crashes in China (2011-2019)

- From 1999 to 2002, it showed a rapid growth trend, with an average annual growth rate of **24.7%**.
- From 2002 to 2010, it showed a reduction trend, with an average annual decrease rate of **17.2%**.
- From 2017 to 2019, it showed a rapid growth trend, with an average annual growth rate of **10.9%**.

Data sources: National Bureau of Statistics.

1.3 Development status of road traffic accident

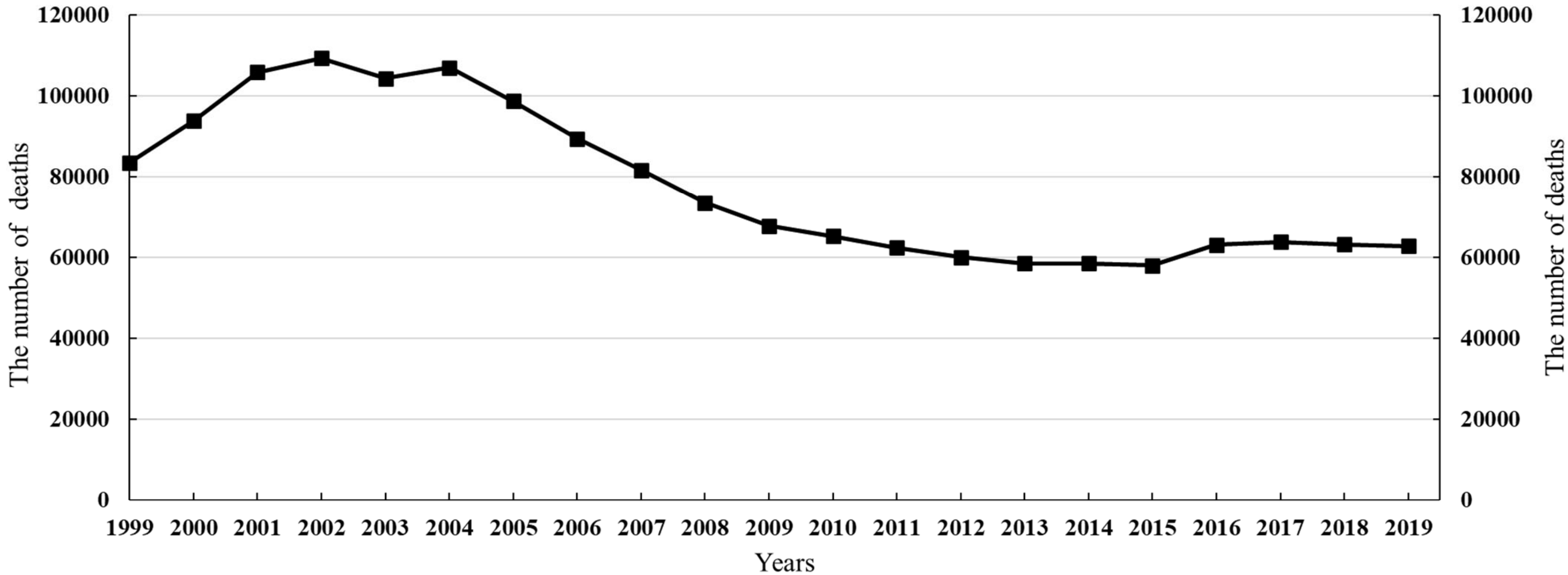


The number of crashes for four road types in China (2011-2019)

- The number of Urban road crashes is the highest, with an average every year is about **9700**.
- The number of Expressway crashes shows a downward trend as a whole, with an average annual decrease of **1.31%**.

Data sources: National Bureau of Statistics.

1.3 Development status of road traffic accident



The number of deaths in China (2011-2019)

- From 1999 to 2002, it showed a rapid growth trend, with an average annual growth rate of **9.5%.**
- From 2004 to 2014, it showed a reduction trend, with an average annual decrease rate of **5.4%**.
- From 2015 to 2019, it showed a trend of growth first and then reduction.

Data sources: National Bureau of Statistics.

Crash Process: Drivers, Roadways and Vehicles

Human-Environment-Vehicle System

Provide a conceptual framework to analyze motor vehicle collisions.

Example: A 20-year-old man with little driving experience, is taking an unfamiliar road on his way to an interview. His vehicle is not properly maintained. In fact, the tires are in poor shape. At some point on the trip, the rain starts to fall. Shortly thereafter he enters a horizontal curve with a radius below minimum standards, loses control of the vehicle and run off the road into a tree located within a few feet of the traveled way.

Question: What are the contributing factors that have lead to this crash?

Human-Environment-Vehicle System

System Component	Event	Circumstance
Human	Trip	Young, inexperienced, stressed
Vehicle	Choice of vehicle	Tires in poor condition
Environment	Rain	Wet and slippery surface
Environment	Curve	Below standard
Human	Steering maneuver	Oversteering
Human	Loss of control	Unstabilized shoulder
Environment	Roadside conditions	Tree too close traveled way
Outcome	Impact (crash)	

Crash Contributing Factors

TABLE 2.1 Critical reasons for crash occurrences (NHTSA, 2018).

Crash causation	Percentage (standard error)
Drivers	94% (2.2%)
Vehicles	2% (0.7%)
Roadways	2% (1.3%)
Unknown	2% (1.4%)
Total	100%

Crash Contributing Factors

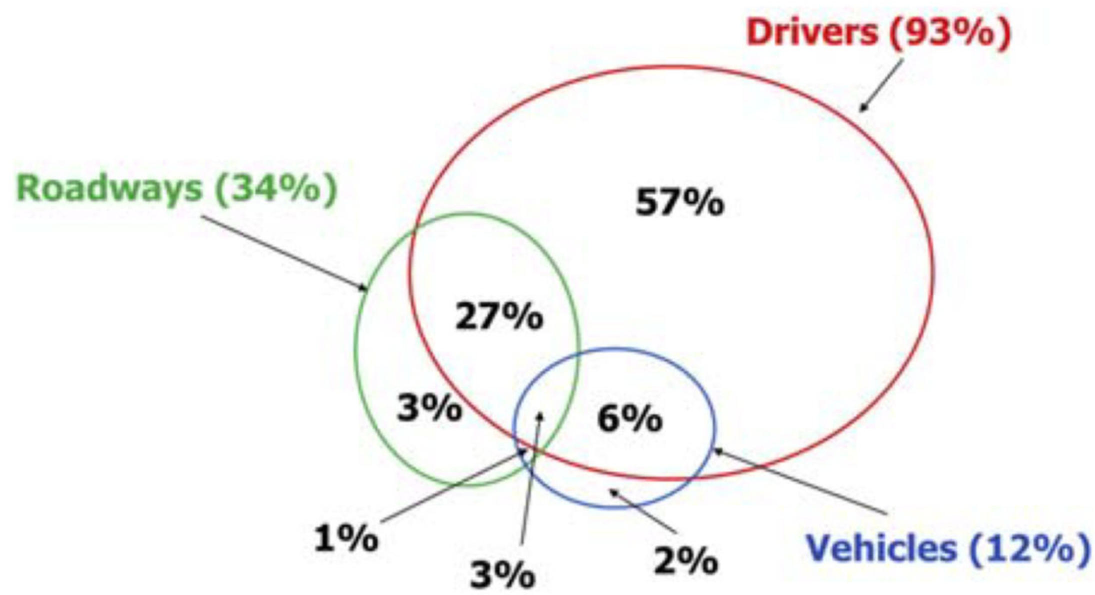


FIGURE 2.1 Precrash causation factors for roadways, drivers, and vehicles (Rumar, 1985).

http://cyberlaw.stanford.edu/blog/2013/12/human-error-cause-vehicle-crashes
https://www.fhwa.dot.gov/publications/publicroads/95winter/p95wi14.cfm

Haddon Matrix

	Agent	Host	Physical Environment	Social Environment
Pre-event				
During the event				
Post-event				

William H. Haddon, Jr., came up with a Matrix to systematically analyze car crashes in a 1972 paper.

Using the Haddon Matrix

TABLE 2.2 Haddon matrix for urban crashes (Herbel et al., 2010).

Event	Driver	Vehicle	Roadway	Social environment
Precrash	Poor vision, speeding	Failed brakes, worn out tire	Poorly timed traffic lights	Speeding culture, red light running
Crash	Failure to use seatbelt	Air bag failure	Poorly designed brake-away pole	Lack of vehicle regulations
Postcrash	Age (to sustain injury), alcohol	Poorly design fuel tank	Poor emergency communication	Lack of support for EMS trauma systems

Sources of Data & Crash Data Management

Sources of Data

- Traditional Data
  - Crash Data (Police reports, self-reporting, etc.)
  - Traffic Conflicts/Surrogate Measures of Safety
  - Hospital
  - Crash Reconstruction
  - Insurance (Private)
- Naturalistic Data
  - Instrumented Vehicles
  - Vehicle Blackboxes
  - Safety Systems (Onstar, etc.)
- Social Media/Crowd Sourcing Data
  - Twitter, Facebook, Instagram, Waze
- Disruptive Technological Data (enhance or new technology)
  - Cellphones
  - Video Recording/Processing
  - Roadside Sensors (Toll Roads, etc.)

Traditional Data

- General Characteristics
  - Despite the flaws, crash data provide the ultimate sources of information about the safety performance of roadway elements. (others: traffic conflicts, driving task analysis)
  - Unfortunately, people get hurt or killed and properties get damaged before finding this performance.
  - Highly dependent on severity of the crash
- Crash data used for:
  - Identification of hazardous sites
  - Benefit-costs analysis
  - Safety relationships (statistical models)
  - Highway design process
  - Policy development

Table 2.4 Important Variables Collected from Crash Data

Variable	Description
Identification Number	Each crash report should have its own identification number. This ensures that each crash is unique and can be easily traceable.
Location	The location can be identified using a linear system, such as control-section mile point on a predefined maps maintained by the transportation agency. More recently, most agencies are now reliably coding crash data using the geographic information system (GIS) technology.
Date and Time	These two variables can be used to assign crashes for different seasons and whether the crash occurred during nighttime, dusk, dawn or daytime conditions.
Severity	This is used to characterize the most severe injuries among all the occupants or vulnerable road users (pedestrians or bicyclists). For example, if a crash has three injures, one incapacitated (Type A) and two possible injuries (Type C), the crash will be classified as incapacitating injury (Type A).
Collision Type or Manner of Collision	This variable describes the characteristic of the crash, such as right-angle, side-swipe or left-turn/through collision.
Direction of Travel	This variable explains the direction or trajectory of each vehicle or road user involved in the crash.
Alcohol or Drugs	This variable explains if any of the drivers or vulnerable road users was under the influence of alcohol or drugs. This variable will be often be updated in the report after the crash to account for the time needed to get laboratory results back.
Vehicle Occupants	This variable describes the gender and age of each vehicle occupant or road user. It may include the legal driving and insurance statuses of drivers.
Vehicles Involved	This one describes the characteristics of each vehicle. This variable defines the crash as being a single-vehicle or multivehicle event.
Narratives	This section of the report is usually not coded electronically. However, the narrative is very important since it provides information about the crash process (based on the testimony of witnesses and the visual assessment of the officer). It is usually accompanied by one or more figures or sketches which helps explain what happened. Based on the authors' experience, many research projects involved the review of these narratives for validating the electronic databases. This is a very time consuming and costly process.

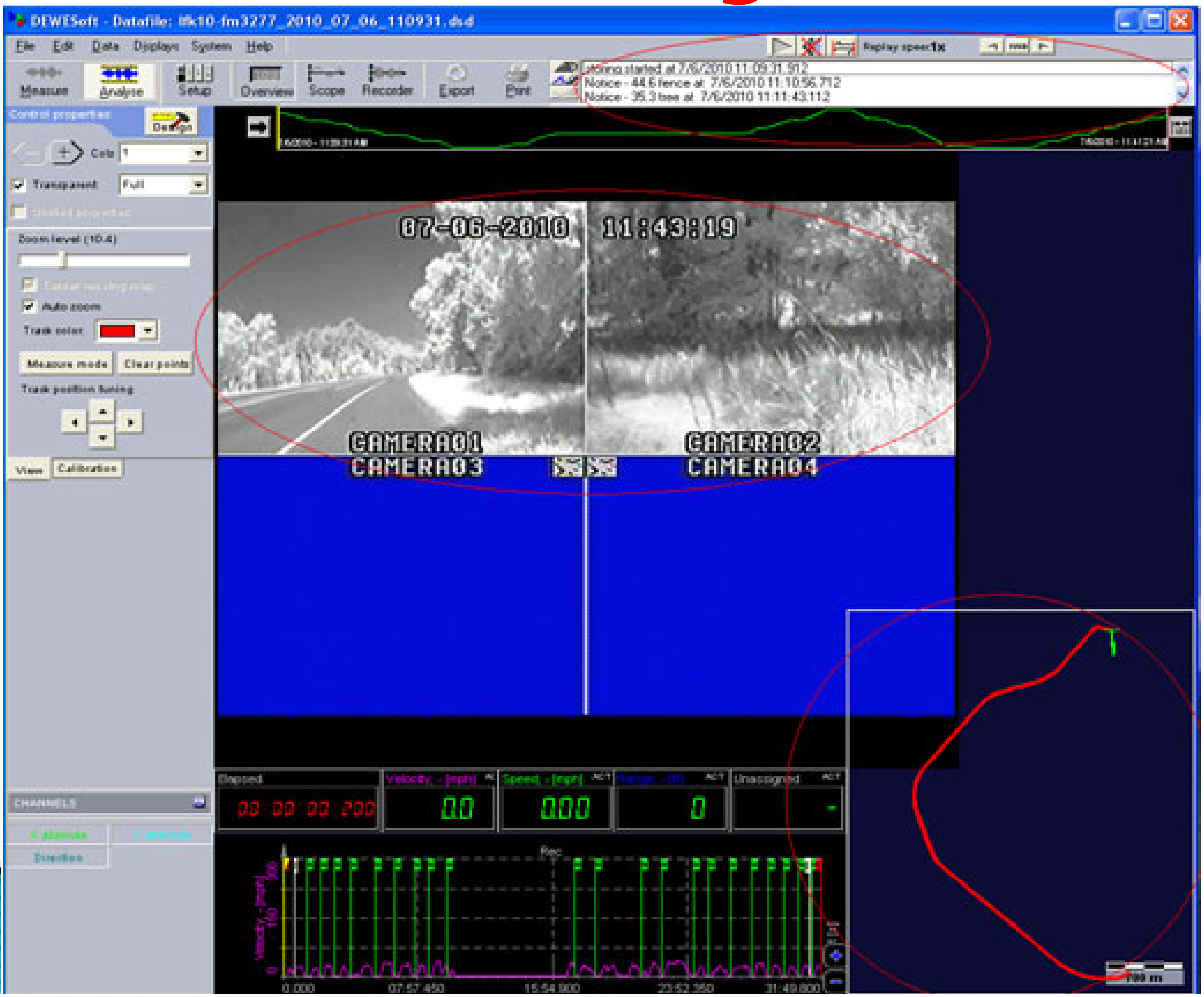
Table 2.5 Common Variables Found in Roadway Data

Location (control-section mile point or GIS) Segment Length Type of Pavement Traffic Control at Intersections Speed Limit Road Alignment (tangent, curve) Road Surface Condition Right-of-Way Width Parking	Highway Classification (Freeway, Arterials, etc.) Type of Lane and Width Type of Shoulder and Width Number of lanes Divided/Undivided Lighting

Table 2.6 Traffic Flow Data

Location (control-section mile point or GIS)
Annual Average Daily Traffic/AADT (vehicles/day)
Average Daily Traffic/ADT (vehicles/day)
Traffic Mix (heavy vehicles, motorcycles, passenger cars, etc.)
Speed Distribution
Short Counts (hourly volumes, 15-minute values, etc.)
Vehicle Occupancy (on instrumented urban freeway segments)
Traffic Density (on instrumented urban freeway segments)
Turning Movements at Intersections

Video Logs



DEWESoft - Datafile: tfk10-fm3277_2010_07_06_110931.dsd
File Edit Data Displays System Help
Measure Analyse Setup Overview Scope Recorder Export Print
Notice - 44.6 fence at 7/6/2010 11:10:56.712
Notice - 35.3 tree at 7/6/2010 11:11:43.112
07-06-2010 11:43:19
CAMERA01
CAMERA02
CAMERA03
CAMERA04
Transparent
Auto zoom
Measure mode
Clear points
Track position tuning
View
Calibration
CHANNELS
Elapsed
Unassigned

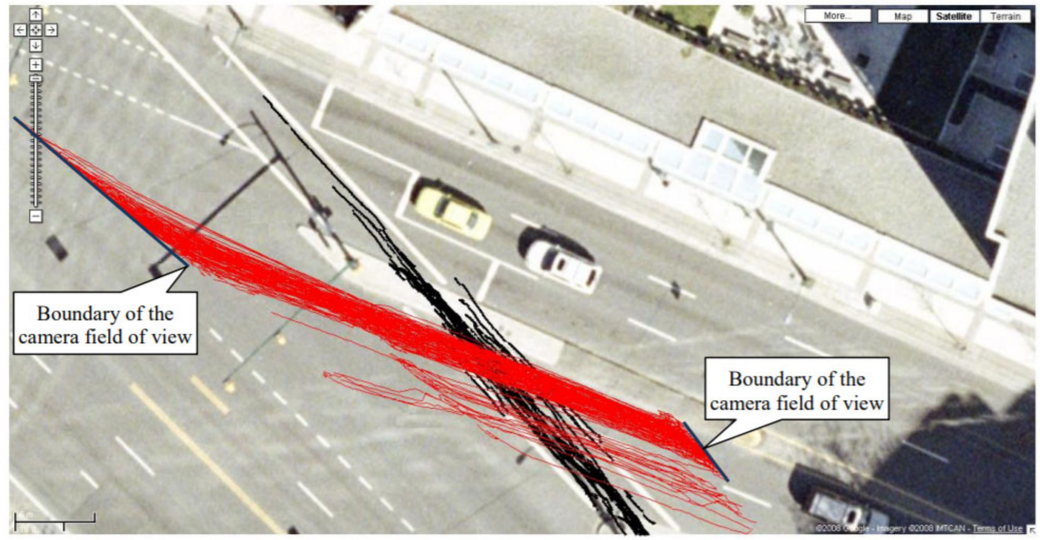
Google Earth

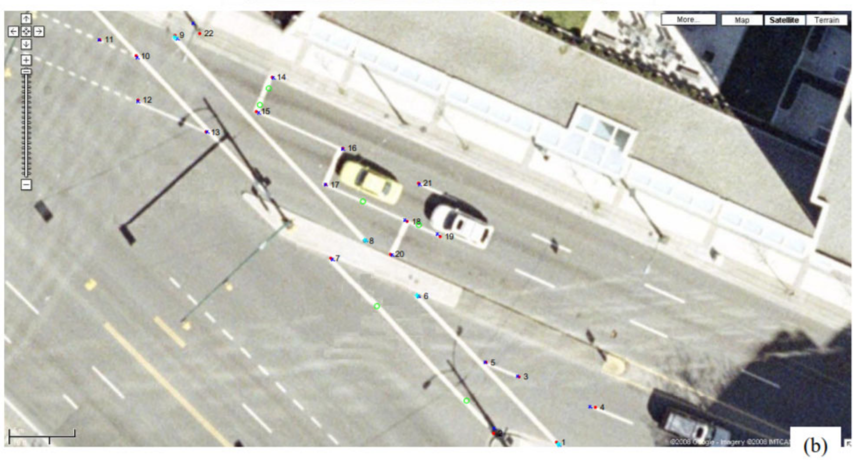
StreetView

StreetView

Video Processing

Video Processing





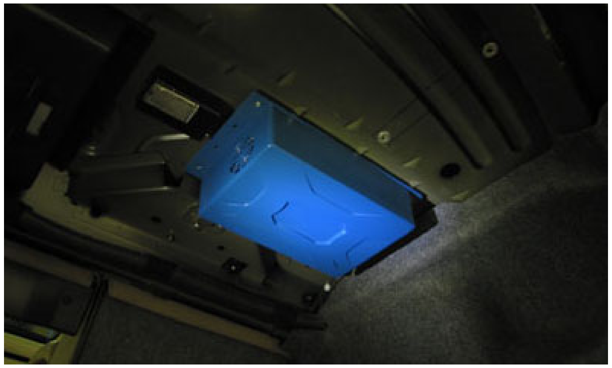
Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data. Ismail et al. (2009) https://doi.org/10.3141/2140-05

Naturalistic Data

Data obtained from instrumented vehicles. The installed instrumentation or data acquisition system collect data using several sensors and video cameras whenever the vehicle is running. Next slide shows different naturalistic studies/data availability from around the world.

Below, the pictures show some of the instrumentations.

Head Unit



Main Unit

Radar

Sources of Data – ND/DTD

UDRIVE (Europe): https://results.udrive.eu/

UMTRI (Michigan): http://www.umtri.umich.edu/our-focus/naturalistic-driving-data

Australian Naturalistic Study: http://www.ands.unsw.edu.au/

Streetlight Data: https://www.streetlightdata.com/ (exposure)

Safe2save Data: https://safe2save.org/ (exposure)

Data Users

- Traffic Engineers: elimination of hazardous sites; highway design, etc.
- Police force: enforcement location, etc.
- Researchers: understanding the crash process; safety relationships, etc.
- Decision-makers: alcohol measures, speed limit
- Prosecutors: transportation-related criminal lawsuits, eye witness statements, etc.
- Insurance companies: set premiums, types of vehicles, age of drivers,
- Vehicle manufacturers: research for safer vehicles

Crash Report

- Quality of data highly dependent on the officer at the scene of the crash
- Most important data are collected for potential criminal prosecution; other data less important
- Sometimes officers fill out the report on the scene while others do it at the end of the day
- Important to have an open line of communication between the engineering and police departments
- Now, crash reports and process for collecting all electronic

Critical Information

- Geographical location (intersections, mile-point, GIS)
- Date (year, day of week, time of day, etc.)
- Type of involvement (vehicle, driver, occupants, etc.)
- Outcome (severity)
- Environmental conditions (weather, lighting, road surface conditions)
- Characteristics of collision (direction of road users, errors, collision type)

Critical Information

Table 4-1 Accident data – Recommended information

ACCIDENT IDENTIFICATION NUMBER	LOCATION
- A unique number, which prevents accident data from being entered twice (should be combined with police station reference number).	- The exact road location where the accident occurred. It can be described by: - narratives; - X, Y coordinates of a uniform coordinate system; - highway number and kilometer post rounded off, for example, to the nearest 100 m; - distance from node; - distance from known point. ***Accident location methods***
DATE AND TIME	COLLISION TYPE
- The exact date (preferably with four digits for the year, two digits for the month and two digits for the day). - Although the day of the week can be computed from the date, it may be useful to have a separate data item to store this piece of information. This is especially true in the case of a non-computerized database.	- By one or some combinations of the followings: - narrative description; - sketch; - code.
VEHICLE	VEHICLE MANOEUVRES AND SKETCH
- For each involved vehicle: - type; - make; - year of manufacture; - vehicle identification number (VIN). - The VIN is a multi-character identification number placed either on the body of the vehicle or included on the vehicle registration certificate. Through the VIN, a number of variables, such as type, model, year of manufacture, body type, engine size, restraint system, etc., can be easily accessed in the vehicle registration file.	- The description codes of the manoeuvre of each vehicle/ participant or an accident sketch. - To be useful, this sketch must include the following information: - each vehicle's direction and manoeuvre; - vehicle identification (veh1, veh2, etc.); - reference points; - important measures; - scale.

Taken from the 2003 version of the PIARC's Road Safety Manual

Critical Information

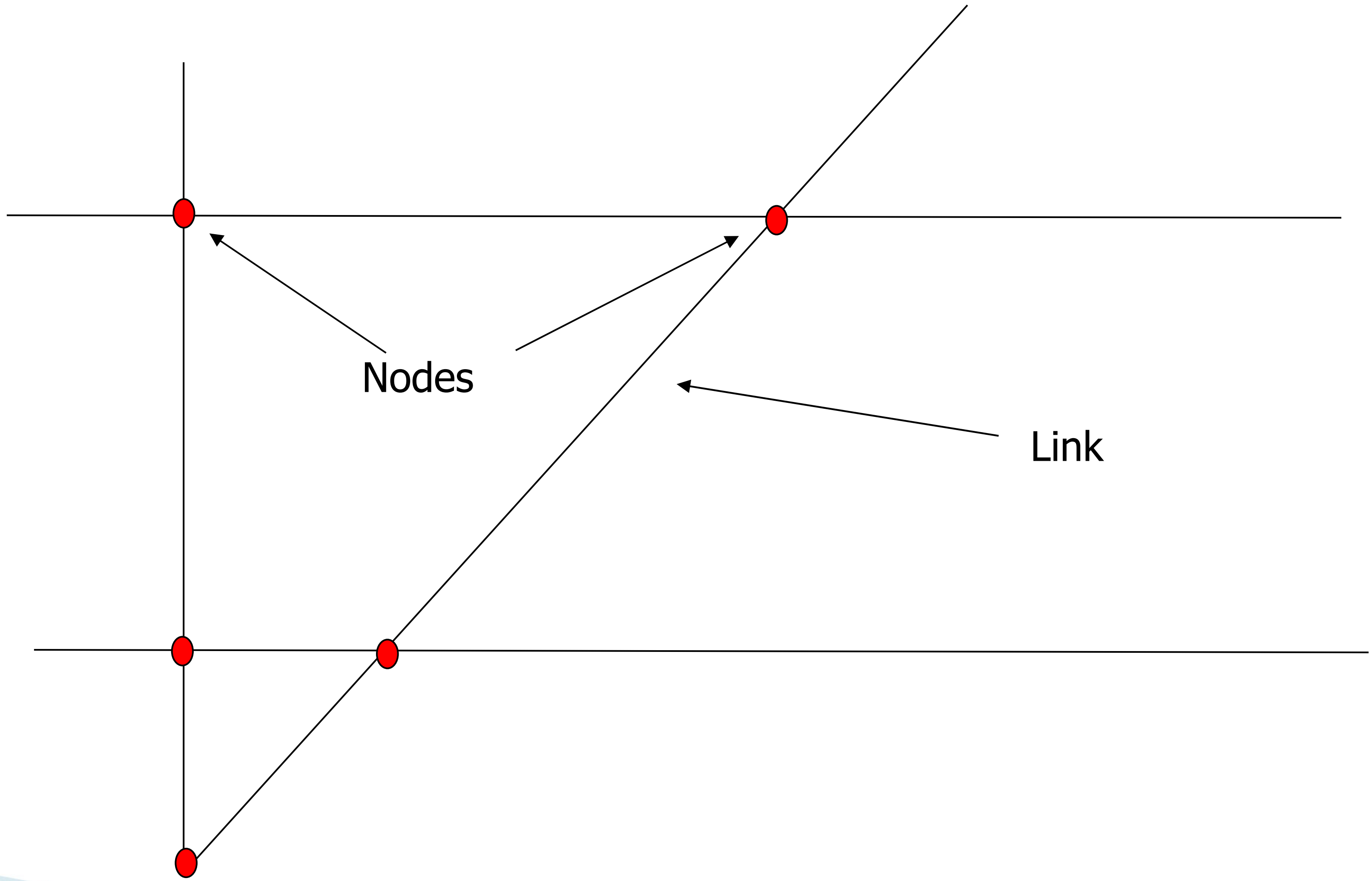
DIRECTION OF TRAVEL	CONTRIBUTING CIRCUMSTANCES
- The direction of each participant, for example: - from A to B; - the direction of increasing/decreasing mileage.	- What were the environmental conditions such as: - whether (rainy, windy, foggy, etc.); - road surface conditions (icy, wet, debrid on the road, etc.); - other contributing factor.
CASUALITIES	**DRIVER / PASSENGER**
- By severity of injury for each casualty, for example: - fatal; - serious injury; - light injury. - The injury severity, determined by a police officer, is often largely subjective and does not always reflect the injury scale used by a hospital. A more accurate injury severity number can be retrieved from hospital files.	- name, sex, age; - location in the vehicle; - driving licence number; - driving experience.
RESTRAINT USE	**ALCOHOL / DRUGS**
- seat belt, helmet, children safety seat; - whether air bag (deployed or non-deployed).	- By the result of the alcohol test: - on the spot; - in the police station or hospital.
SPEED	**EXTENT OF PROPERTY DAMAGE**
- driver and eyewitness statement; - length of skid marks on the road - tachograph of the truck.	- By estimation
NARRATIVES	
- Narratives are key elements of a report form. They often complete the picture about the accident with some useful additional information that cannot be coded. It is also the most time consuming part of the process, therefore tends to be lacking. Enough space should be provided on the report form to accommodate all information that the police officer finds important.	

Taken from the 2003 version of the PIARC's Road Safety Manual

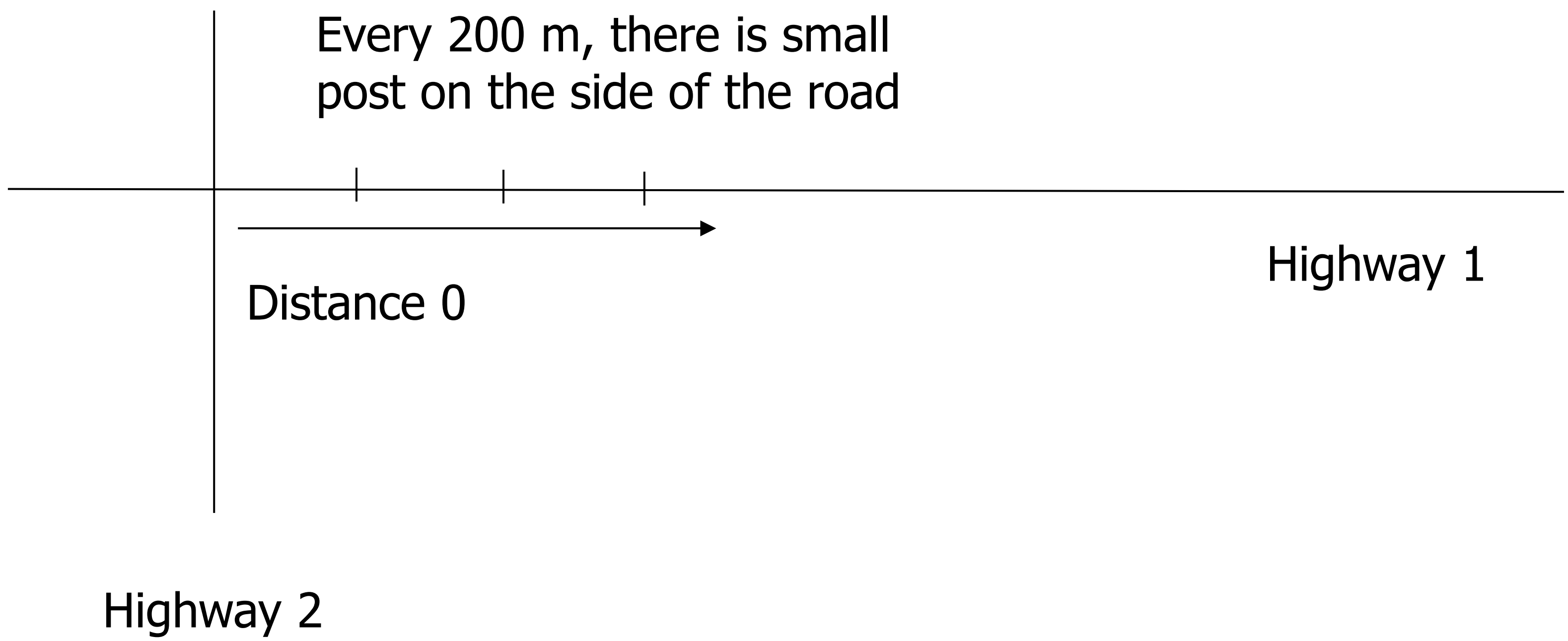
Location Methods

- Location of the crash is the most important aspect of data collection:
  - Estimate hazardous sites that experience more crashes than what would be expected
  - Provide a usually way to link different databases
- Three methods: Link-Node, Route-Km Post (aka Control-Section), GPS coordinate

Link-Node



Route-Km Post



GPS System

- Provides X and Y coordinates of a given geographical coordinate system
- Advantages: human error free; many software tools include GPS coordinate; can estimate in the Z-coordinate (depth)
- Disadvantages: Accuracy used to be an issue, but now the location pinpoint is very accurate (e.g., look at the location app on your smartphone).

Road System Inventory

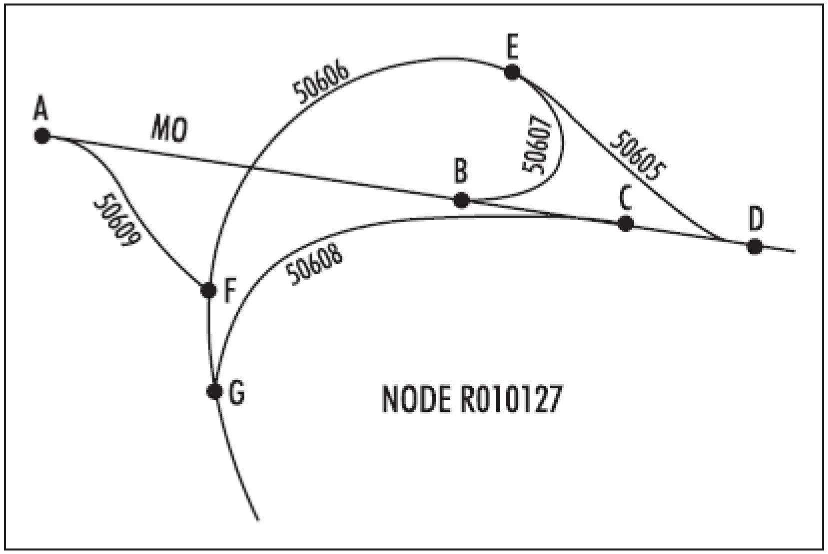
Reference map

Location accuracy depends not only on the officer's description but also on the level of detail and accuracy of the maps the officer uses to identify the location. Detailed maps are crucial when an accident occurs at an intersection with many legs or on a freeway ramp. A unique reference number may be assigned to entrance and exit ramp junction points.

The prerequisites of a successful accident location method (link-node or route-km post) are in general (Figure 4-2):

- easily recognizable reference landmarks along the road;
- a sufficiently detailed map accurately reflecting the road inventory file.

Figure 4-2 Link and node map at interchange



Links between each pair of nodes are assigned a unique number. Such maps help police officers at the scene of an accident identify the location quickly and accurately.

When roadworks have changed the length of a road and the km posts have not been altered to reflect the real distances, the posts no longer indicate the true km point of a location. Calculation algorithms must then be used to convert the reported km posts into true values.

Crash Data Storage

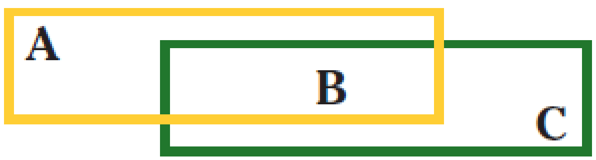
- Old way: paper trails
- All crashes are coded electronically (e.g., CRIS in Texas; FHWA's HSIS)
- Now, commercial programs available that displays visually simple crash statistics
- Internet-based visual tools (using GIS/GPS capabilities: ARCView, etc.)
- DBF, SAS, TXT files

Crash Data Limitations

- Not all crashes are reportable
  - Personal injury (definition)
  - PDO (limit, changes over time)
- Not all crashes are reported
  - Ignorance of the law
  - Victim's unawareness of injury at the time of the collision
  - Desire to avoid bureaucracy
  - Desire to avoid insurance company penalties
  - Type of collision (single vehicle versus multivehicle)
  - Type of users (bicyclists and pedestrians)
  - Police force (report level varies by jurisdiction)
- Reported crashes may contain errors
  - On report, fraudulent claims

Crash Data Limitations

Figure 4-6 Comparison of accident files (police and hospital)



A = accidents recorded by the police only
B = accidents recorded by the police and the hospital
C = accidents recorded by the hospital only

Definition 1 = $\dfrac{B}{B+C}$ This definition covers those hospital file cases for which a police report is found.

Definition 2 = $\dfrac{A+B}{B+C}$ This definition compares the total number of records found in the police file with the total number of cases in the hospital file.

Definition 3 = $\dfrac{A+B}{A+B+C}$ This definition, theoretically the most correct, compares the proportion of the total police records with the total number of cases in both the police and the hospital files.

Taken from the 2003 version of the PIARC's Road Safety Manual

Crash Data Limitations

Figure 4-7 Accident report – The golden mean

An increase in the number of questions on a report form may lead to a sharp drop in the number of reported accidents. Some questions may not be answered by the police officer because of oversight rather than any intention not to answer.

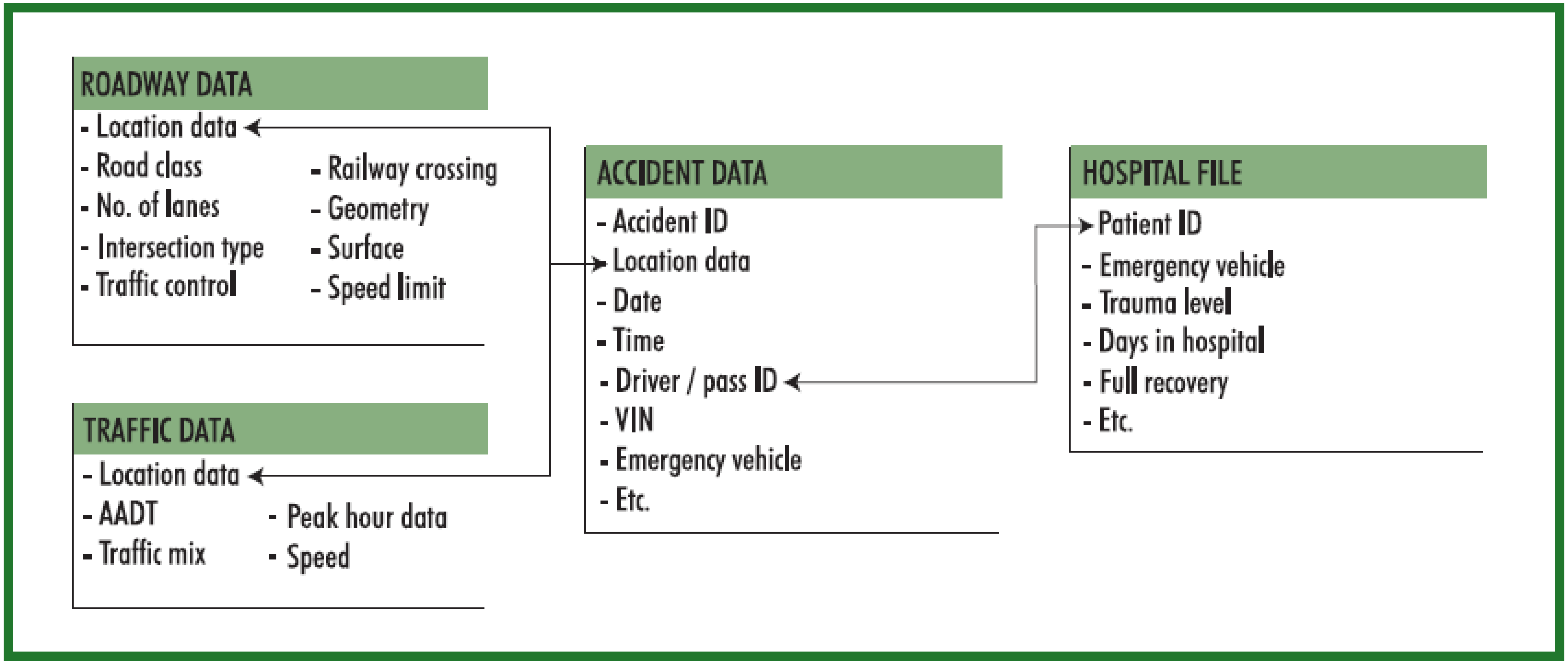
A short, one page report form may lead to higher reporting rates, but provides less detailed information. This may result in insufficient data to carry out comprehensive studies.

Taken from the 2003 version of the PIARC's Road Safety Manual

Integrated Data Files

Linking Files

Figure 4-9 Linked file



Taken from the 2003 version of the PIARC's Road Safety Manual