DaCoTA

Road Safety Data Collection, Transfer and Analysis

Deliverable No. 5.6: Evaluation Tools

Please refer to this report as follows:
Hautzinger, Pfeiffer, Simon (2012), Evaluation Tools, Deliverable 5.6 of the EC FP7 project DaCoTA.

Grant agreement No TREN / FP7 / TR / 233659 /"DaCoTA"
Theme: Sustainable Surface Transport: Collaborative project
Project Coordinator:
Professor Pete Thomas, Transport Safety Research Centre
Loughborough University, Ashby Road, Loughborough, LE11 3TU, UK

Project Start date: 01/01/2010 **Duration 36 months**

Organisation name of lead contractor for this deliverable:
IVT Research GmbH (IVT Research), Mannheim, Germany

Report Authors:

Heinz Hautzinger, Institut für angewandte Verkehrs- und Tourismusforschung e.V.

Manfred Pfeiffer, IVT Research GmbH

Marie-Christine Simon, LAB

Due date of deliverable	30/09/2012	Submission date:	31/10/2012

Project co-funded by the European Commission within the Seventh Framework Programme	
Dissemination Level (delete as appropriate)	
PU	Public

Transport

Project co-financed by the European Commission, Directorate-General Transport and Energy

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Deliverable 5.6 "Evaluation Tools" represents the essential outcome of Sub-Task 5.3.2 "Methods and Tools" of the DaCoTA project. The intention of the Deliverable is to give an overview on the state of art of evaluation tools and by this providing some kind of reference book for the application of these tools. It provides methodology and examples for the evaluation of (mainly) vehicle safety systems with regard to

- data collection methods
- data analysis methods
- socio-economic methods
- pitfalls and difficulties.

The Deliverable is structured into 3 main parts. The first part (Chapter 2) is dedicated to the assessment of safety benefits of systems in terms of measuring the effects on accident involvement (and injury) risk. After some general remarks on non-statistical considerations crucial for the evaluation of safety systems the following study designs are presented.

- accident involvement survey
- cohort study
- case-control study
- comparative accident study based on the concept of induced exposure
- matched pairs design
    - matched case-control design: one accident involved vehicle (case) and one not involved vehicle (control) are paired
    - matched cohort design: pairing a fitted (e.g. ESP) vehicle (protective factor present) with an unequipped one (protective factor absent).

Subsequently epidemiological data analysis techniques suitable for the assessment of accident involvement risk and injury risk like

- relative risk
- odds ratio
- incidence density
- statistical models for different types of risk measures

are introduced and illustrated by selected empirical examples. Statistical analyses are provided for a cohort study, the induced exposure technique, and the matched pairs design (matched cohort as well as matched case-control design). Up to now, matched pairs designs have relatively seldom been applied in accidentology.

Both the analysis of the cohort study and the matched pairs analysis are based on a Germany study where (among other things) the combined effects of three different eSafety systems fitted in heavy trucks (ESP, ACC, Lane Guard System (LGS)) were investigated (1 250 heavy goods vehicles; n=715 vehicles fitted with the three systems; n=535 vehicles without the systems). It can be said that heavy goods vehicles fitted with ESP, ACC and LGS show a considerably lower accident involvement risk compared to vehicles without these systems. The observed group differences with respect to accident involvement (all accident types) are statistically significant. For the matched pairs design special statistical analysis techniques

(particularly matched odds ratio and conditional logistic regression) are introduced and explained.

The induced exposure example is based on GIDAS data, the sample consists of n=10 270 accident involved passenger cars (study period 1995 to 2011) with and without ESP, i.e. the system to be evaluated is ESP. For the evaluation of ESP the accident characteristic “skidding” has been chosen to distinguish between system-specific accidents (car was skidding) and neutral accidents (no skidding). It appears that the chance of skidding is approx. two-thirds lower for cars with ESP compared to cars without this feature.

At the end of Chapter 2 some additional methodological hints on the evaluation of infrastructure measures are given.

The second part (Chapter 3) is dealing with the evaluation of systems in terms of socio-economic benefits. In this part methods for assessing potential benefits of safety applications are presented. After the description of potential impacts of safety systems with respect to costs, benefits, and factors influencing the effects in different countries, the assessment of the impact of safety systems on the number and severity of accidents is dealt with. Subsequently, a short overview on Efficiency Assessment Tools (Cost Benefit Analysis, Cost Effectiveness Analysis) is given. The main focus of this chapter is on the application of a Cost Benefit Analysis and the provision of standard values for accident costs.

Part 3 (Chapter 4) contains some considerations on the expansion of evaluation results from selected regions or countries to the EU27-level by using the so-called iterative proportional fitting procedure. The method is especially relevant for expanding results on the a priori evaluation of vehicle safety systems by using a simulation tool, for instance. Here, for each case it can be determined whether or not the presence of the system would have avoided or mitigated the accident. Thus, the distributions (regarding e.g. injury severity, lighting conditions, etc.) of both the affected and unaffected accidents are known and can be expanded to a wider accident population. A crucial prerequisite for the application of this method is, of course, that the relevant marginal distributions (accidents, casualties) are available at EU-27 level.

1. INTRODUCTION

The European Road Safety Observatory (ERSO) has been established by the European Commission in order to support safety policy-making in Europe by providing comprehensive data and knowledge on road safety. The main objective of the DaCoTA project is to advance the Observatory by enhancing the existing data and adding new road safety information including the evaluation of intelligent safety systems (Work Package 5).

The basic research question of WP5 is “How does technology contribute to road safety?” The objective is to develop methodologies and approaches that will enable future evaluation of the safety impact of emerging intelligent technologies. This is done by

- identifying and updating the user’s needs in term of accident risk prevention and injury risk prevention
- identifying and updating how current technology can address these needs
- providing methodology on assessing the potential benefits of the relevant safety applications (not only the safety benefits).

The present Deliverable is dealing with the latter issue, i.e. the determination of the relevant methods and tools necessary to perform the evaluation of the benefits of vehicle safety systems. In general, an exact quantification of the safety effect of a given system alone is difficult because accident risk depends on a variety of factors. Statistical methods of such investigations of effectiveness are, therefore, of special significance. It is the purpose of this Deliverable to develop standardised and practically applicable methods for the evaluation of safety systems. Hereby we basically have to differentiate between two types of evaluation:

- The evaluation of safety benefits based on real-world accidents by applying epidemiological methods: As accident involvement is an event occurring in time and space, the general epidemiological concept of disease incidence (incidence = number of new cases of a disease within a specified period of time) applies to studies on accident involvement risk. These approaches are described in Chapter 2.
- The evaluation of systems which are not yet on the market: This is done by the application of simulation software or by case-by-case-analyses of in-depth accident data. A detailed example for the evaluation by means of a simulation tool can be found in Deliverable 5.7 (“Real World and Procedures”) of DaCoTA. In the present deliverable this “a priori evaluation” of safety benefits is dealt with in the framework of socio-economic methods (section 3.3.5), since it is a crucial prerequisite for the application of socio-economic evaluation (Chapter 3). A socio-economic analysis is a decision-support tool. The basis of the analysis is the evaluation of the costs and the economic and social benefits related to the application of a system. This evaluation is then compared to the situation where the system is not applied. Socio-economic analysis allows making statements about social return of an investment. An overview of social costs and benefits can serve as basis for prioritizing separate measures or measure packages.

  In this context methods for expanding evaluation results to EU 27 are of importance in order to be able to make statements for the European Union as a whole. In Chapter 4 methods for expanding evaluation results obtained from selected countries to EU 27-level are presented.

The general intention of the Deliverable is to give an overview on the state of art of evaluation tools and by this providing some kind of reference book for the application of these tools.

2. EVALUATION IN TERMS OF MEASURING ACCIDENT INVOLVEMENT RISK AND INJURY RISK

After some general remarks on non-statistical considerations crucial for the evaluation of safety systems various study designs – including the matched pairs design - are presented in this chapter. Subsequently epidemiological data analysis techniques suitable for the assessment of accident involvement risk (and injury risk) are introduced and illustrated by selected empirical examples. Concretely, statistical analyses are provided for a cohort study, the induced exposure technique, and the matched pairs design (matched cohort as well as matched case-control design).

2.1. Study Designs for Measuring Accident Involvement Risk

2.1.1. Preliminary considerations

The effects of eSafety systems on accident involvement risk of vehicles may be assessed under different study designs offered by the statistical sciences. Practical examples can be found, for instance, under www.iMobility-effects-database.org. Irrespective of the study design actually chosen, however, there are various design elements which cannot be specified on purely statistical grounds. Rather, detailed knowledge of the special features of the topic to be investigated is required. Thus, in our case expertise in automotive engineering and safety technology as well as know-how in accident research is needed.

Broadly speaking, eSafety systems as our study subject are special purpose technical systems designed to assist, inform or alert the driver by

- improving the driver's perceptive faculty (e.g. a night vision system provides to the driver enhanced vision in unlit areas)
- amplifying driver actions (e.g. the emergency brake assist reduces the time necessary to reach ABS regulation)
- correcting a driving mistake (e.g. ESC recovers loss of control)
- helping the driver to come out of a critical situation (e.g. frontal collision warning and lane departure warning systems)
- providing car occupant or external user protection in the case of a crash (e.g. pre-crash systems detecting an imminent crash may move the passenger seat to a better (i.e. less injury-prone) position or retract the seat belts removing excess slack)
- relieving the driver of certain tasks (e.g. Intelligent Speed Adaptation systems can, to a certain extent, replace the driver for speed regulation).

As can be seen, eSafety systems are markedly heterogeneous with respect to function and safety effects. In this situation, for any eSafety system *S* under study the following largely non-statistical considerations are crucial for proper study design:

- Which types of accidents are to be considered as relevant for the study ("system-relevant" or "system-affine" accidents) in the sense that system S is intended or expected to reduce the risk of involvement in such accidents?

- Which analysis level (accident level, vehicle level, vehicle occupant level) is appropriate when assessing a given system S? For instance, pre-crash systems might to be assessed preferably at the vehicle occupant level (occupant injury in crashes) whereas for night vision systems the accident level (e.g. nightly car-pedestrian crashes) could be more appropriate.

All possible study designs are faced with the problem that in addition to the eSafety system to be assessed (say system S) also other eSafety systems (say systems A, B, C, …) may be present in a vehicle. It might even be the case that if system S is to be found in a vehicle one always or nearly always will also find system A in the vehicle. Under these circumstances it might be difficult or even impossible to measure the "pure" effect of system S.

Consequently, in addition to the above considerations one has to answer the following questions:

- What are the most frequent combinations of eSafety systems ("bundles" of systems) fitted in vehicles?
- Are there significant interactions between individual systems in the sense that the efficacy of system S is affected by the presence of system A?

Finally, when speaking of safety system assessment, two approaches may generally be distinguished. The efficacy of a safety system can be measured by the relative risk of involvement in a system-relevant accident (fitted versus non-fitted vehicles). Under a broader public health perspective, however, the societal benefit of a safety system not only depends on its "technical" efficacy but also on the incidence of accidents which can be considered as system-relevant. Thus, when efficacy is to be measured one would restrict risk analysis to involvement in system-relevant accidents. If, on the other hand, the study focuses on societal benefit, involvement in any kind of accident is the "disease" status variable of interest. I.e., in some ways the "success" of a system depends on the point of view of the stakeholders involved (automotive industry, policy,…).

In Chapter 2.1.2 and 2.1.3, alternative designs for eSafety system assessment studies will be presented under a mainly statistical perspective. Chapter 2.1.4 contains the non-statistical (technological and functional) considerations associated with eSafety system assessment.

2.1.2. Overview of study designs for eSafety system assessment

There are several ways in which an empirical study can be designed so as to collect conclusive data on the efficacy of eSafety systems (see, for instance, Woodward 2005). Two basic principles should always be followed: the study should be comparative and we should seek to avoid all potential causes of bias (ibid., p. 24). In the sequel, four different study types suitable for assessing the impact of technologies on the incidence of traffic accidents and casualties are described.

Accident involvement surveys and cohort studies are data collection techniques in the sense that – as a rule - interviews with e.g. vehicle holders have to be conducted. This is not mandatory for case-control studies and the induced exposure concept. Here, the analyses can often be performed solely on the basis of already existing data from different sources (accident data from in-depth investigations or police-recorded accident statistics, data from vehicle registers).

All the approaches described in the following have in common, that they can only be applied to safety systems which can be found in large parts of the vehicle (and accident) population, so that quantitative analyses are possible.

2.1.2.1. Accident involvement survey

Accident involvement surveys (also referred to as retrospective cohort studies) are empirical studies based on a sample of vehicles drawn from a certain target population of vehicles. Typically, national motor vehicle registers may serve as a sampling frame. Clearly, the sample of vehicles should be selected from the frame in accordance with a sampling design that specifies a probability mechanism and a sample size.

In order to obtain data on the characteristics of interest (vehicle fitment, accident involvement of vehicle during a certain study period etc.), the holder of each selected vehicle has to be interviewed. The period for which accident involvement has to be reported is a time period in the past (e.g. last 12 months). Thus, the survey is a retrospective study where accident involvement incidence is investigated at the vehicle-year level.

Information on vehicle fitment (equipment of vehicle with specified eSafety systems) may either be collected in the interview – this, of course, requires that the vehicle holders are correctly informed about the equipment of their vehicles - or by utilizing appropriate external data sources like manufacturer information. The data thus obtained describe the risk factor status[1] of the vehicles in the sample. As accident involvement is only possible for vehicles participating in traffic, data on traffic participation (e.g. annual mileage of vehicle) is also to be collected in the interview. If possible, additional characteristics of the vehicle, its drivers and its use should be reported as these might also affect accident involvement (confounders).

Under the above study design both the absolute and relative risk of accident involvement can be estimated. The effect of vehicle fitment on accident involvement can therefore be assessed.

Normally, not only data on accident involvement as such (involvement yes/no) will be collected but also data on the accident (if involved). This enables the researcher to assess different types of risk (e.g. risk of involvement in accidents with personal injury). Moreover, the effect of eSafety systems on other criterion variables of interest (e.g. casualties) can be determined in this case.

The advantage of this method compared to cohort studies (see next section) is that studies of this type can be conducted within a relatively short period of time. On the other hand the results can be severely biased due to methodological reasons. If the sample of vehicles is drawn from a register, vehicles which have been scrapped due to a (severe) accident might not be included in the register any more. Moreover, if a vehicle was sold after having been involved in an accident (e.g. because the vehicle holder was killed in the accident), the new holder might not be able to provide any information about the accident. The more these conditions apply to unequipped vehicles compared to equipped ones (as it can be expected, if the system under investigation is effective), the more the efficiency of the system – in terms of mitigating severe accidents - will be underestimated.

---

[1] As eSafety systems are expected to reduce accident involvement risk it would be more appropriate to speak of *protective* factors rather than risk factors.

2.1.2.2. Cohort study

As with accident involvement surveys, cohort studies require a random sample to be drawn from the population at risk which in our context is formed by a certain collective of vehicles. After selection, the sample is subdivided into two groups, vehicles with and without the safety system to be evaluated[2]. This, of course, requires a sufficiently high penetration rate of the system in the vehicle population.

In contrast to accident involvement surveys, data on traffic participation and accident involvement are collected prospectively by following up the sampled vehicles ("cohort") through time. In practice, this can be accomplished by periodically conducting short interviews with each vehicle holder asking about accident involvement of his/her vehicle during the most recent period.

If a vehicle changes hands during the whole observation period it should be removed from the sample since the vehicle holder characteristics are not unique.

Generally, cohort studies are considered to be the best type of observational study. For the two groups of fitted and non-fitted vehicles the absolute risk of accident involvement may be estimated under this design (and thus also the relative risk).

As, however, accident involvement is a rare event cohort studies have the disadvantage, that they may require a large sample and a long observation period in order to obtain a reasonable number of accident involvements. In this connection it is important to make arrangements for keeping up the motivation of the interviewees to participate in the study throughout the entire observation period.

2.1.2.3. Case-control study

Accident involvement risk analyses may also be based on two independent random samples of accident-involved ("cases") and not involved ("controls") vehicles that belong to the same general population. The cases, for instance, could be accident-involved vehicles recorded in national traffic accident statistics; the controls, on the other hand, could be vehicles drawn from the national vehicle register. As every accident involvement corresponds to an accidental vehicle trip, the control group may also be formed by non-accidental vehicle trips recorded in a representative mobility survey. Clearly, both for cases and controls the risk factor status (fitted with safety device yes/no) has to be ascertained, which can be a difficult task when using routine data. (Deliverable 5.2 of DaCoTA is dealing with the problem of obtaining information on vehicle fitment in routine data bases.)

Case-control studies are especially useful when the "disease" to be studied is a rare event. By comparing the two groups with regard to the risk factor, the relative risk (not the absolute risks) of accident involvement may be assessed (fitted versus non-fitted vehicles).

Since case-control studies are based on two independent samples which very often will not represent the respective population share (different sampling fractions) the appropriate risk measure for such studies is odds ratio rather than relative risk (cf. section 2.2.2).

2.1.2.4. Induced exposure

In a situation, where only accidental units, i.e. vehicles involved in traffic accidents, have been observed and no information on the population at risk (vehicles exposed to the risk of accident involvement) is available, meaningful accident involvement risk

[2] If vehicles were drawn from a vehicle register containing information on vehicle fitment, one could, of course, draw a stratified random sample.

analyses may be conducted provided that a so-called “induced exposure” method can be applied. Several induced exposure methods have been developed so far. The approaches mainly differ in the way an appropriate “control group” among the accident-involved road users is defined.

Two study types based exclusively on accident data can be distinguished:

- Comparison of responsible and non-responsible drivers involved in two-vehicle collisions with regard to vehicle fitment (with versus without eSafety system). Note: responsible drivers actually correspond to accidental vehicle trips, non-responsible drivers are *assumed* to represent all vehicle trips made in the road system.
- Comparison of fitted and non-fitted vehicles involved in single or multi vehicle crashes with regard to type of accident (system-specific versus neutral). An example on the use of an "induced exposure" methodology to assess specific eSafety systems by estimating the relative risk of accident involvement (comparing fitted to non-fitted vehicles) can be found in section 2.3.2 (based on GIDAS data).

2.1.2.5. Criteria for choosing among alternative study designs

As a rule, special samples drawn from the population at risk or some sort of secondary (routine) accident and exposure data are needed to estimate measures of accident involvement.

The design of an empirical traffic accident involvement risk study will mainly depend on two circumstances:

- feasibility of sampling from the population at risk according to a study design (e.g. survey, cohort study, case-control study) specifically developed for the investigation to be conducted
- possibility of preparing routine traffic accident and mobility behaviour data from external sources in such a way that the structure of the resulting data set corresponds to a certain epidemiological study design.

Normally, special data collection offers the best possibility to assess the determinants of accident involvement. Under the second approach, only already existing databases are used: For instance, accident involvement counts from national traffic accident statistics are related to estimates of involvement risk exposure quantities from representative mobility surveys (total number of trips, total traffic participation time or total distance travelled). As compared to special data collection, the potential of studies using routine data is usually limited. By proper data preparation and adequate application of statistical models and methods, however, the accident researcher can make the most of it.

If exposure data is completely missing, the induced exposure method can sometimes offer a way to estimate relative risks provided that an appropriate control group can be found among the accidental units themselves.

2.1.3. Matched-pairs design

As for all observational studies it might be necessary in accident involvement investigations to adjust for confounding variables. Quite naturally, adjustment can be made at the stage of data analysis using, for example, appropriate regression techniques. One may, however, adjust for confounding factors already at the stage of study design. This leads to so-called “matched designs” which are quite common in various fields of applied statistics.

The matched-pairs design has its origin in the field of experimental research. It is an experimental design structure which matches subjects as closely as possible, and then assigns one member of each pair to the test group and the other member to the control group. Normally, the allocation of the paired subjects to the test group or control group, respectively, should be randomized. Then the differences in outcome are measured. Regarding the evaluation of safety systems, randomisation in the above sense is hardly feasible since the variable “vehicle fitment” can usually not be varied experimentally.

The strength of this design lies in reducing the amount of variation between subjects so that any actual differences due to the experimental conditions are more easily identified.

In the context of evaluating safety systems matching means, that prior to statistical data analysis pairs of vehicles are formed according to a certain matching concept. As a consequence, pairs of vehicles instead of individual vehicles are now the observational units of the study (Hautzinger 2006). Two types of matched-pairs designs can be distinguished:

- In a matched case-control design one accident involved vehicle (case) and one not involved vehicle (control) are paired.
- The matched cohort design consists of pairing a fitted (e.g. ESP) vehicle (protective factor present) with an unequipped one (protective factor absent).

The results from a matched case-control study can be displayed in the format of Table 1:

Accident-involved vehicle (case)	**Not involved vehicle (control)**	
	fitted	**not fitted**
fitted	a	b
not fitted	c	d

Table 1 Data scheme for a matched case-control design

Remark: *The total number of* <u>*pairs*</u> *of vehicles is given by a+b+c+d=n.*

Similarly, the format of Table 2 is appropriate to display the data resulting from a matched cohort study:

Equipped vehicle (test group)	**Reference vehicle (control group)**	
	involved in accident	**not involved**
involved in accident	r	s
not involved	t	u

Table 2 Data scheme for a matched cohort design

Remark: *The total number of* <u>*pairs*</u> *of vehicles is given by r+s+t+u=n.*

As an example for a matched cohort study a large-scale German research project ("FAS[3] project") can be quoted where (among other things) a cohort study on the combined effects of three different eSafety systems fitted in heavy trucks (ESP, ACC, Lane Guard System (LGS)) was conducted. The study population consisted of 1 250 heavy goods vehicles with valid specification of mileage in the reference period (only vehicles with mileage greater than 0). These 1 250 vehicles come from 270 companies participating in the study. The sample of vehicles was split up into two subgroups (cohorts). The test group consisting of n=715 vehicles fitted with the three systems was compared to a reference group of n=535 vehicles without the systems under consideration. Evaluation criterion was the number of accident involvements during the observation period.

In both groups accident involvement has been observed over a time period of approximately 2 years (equipped vehicles: 2.1 years on average; reference vehicles: 1.9 years on average). Each accident involvement (occurring on public roads) had to be reported to the research team by the vehicle holder. Average mileage within the investigation period is 287 000 km for equipped vehicles and 248 000 km for the unequipped ones.

In both subgroups approx. 95 % of the vehicles are semi-trailer tractors, the remaining 5 % are motor vehicles. The main operation of the vehicles area is long-distance hauling (equipped vehicles: 88 %; reference vehicles: 78 %). The vehicles in the test group are markedly younger than the reference vehicles. Whereas all equipped trucks have been registered after 2006, this holds true for approx. half the reference vehicles.

The matching procedure applied to these data can be described as a 1:1 matching by random within each company. For example, if a company has provided 9 vehicles for the study (5 equipped vehicles, 4 not equipped ones), 4 pairs have been formed using random numbers. Since group size differs between the two cohorts the matching procedure resulted in n=527 pairs of vehicles.

By using this kind of matched design, it is possible to control for differences between participating companies which might affect the results (as regards different "safety cultures", different levels of driver training, different accuracy of accident reporting). Analyses of these data can be found in section 2.3.1 and section 2.3.3.

2.1.4. Characteristics of eSafety systems significant for study design

2.1.4.1. Accident and injury types addressed by the different eSafety systems

For each system S to be investigated the following questions should be answered:

- Which types of accidents are to be considered as relevant for the study ("system-relevant" or "system-affine" accidents) in the sense that system S is intended or expected to reduce the risk of involvement in such accidents?
- Which types of accidents are not primarily considered as relevant but may be affected by (side effects of) the system S? This type of accidents is normally not part of the case group but it should also be avoided to include it into the control group.

[3] The acronym FAS is the shortform of "Fahrerassistenzsystem" (driver assistance system).

- Which types of accidents are clearly not affected by the system S (they qualify as a control group).
- Does the selection of relevant accidents depend on specific implementations of the system S? For instance, optical sensors (LIDAR, camera) are affected by adverse weather much more than radar sensors which may result in a very different set of accidents. In this case results will be more meaningful if the different implementation will be analysed separately.
- Which analysis level (accident level, vehicle level, and vehicle occupant level) is appropriate when assessing a given system S?
- What metrics (number of casualties, severe casualties, fatalities, property damage, societal cost, QALY[4], DALY[5],…) are appropriate when assessing a given System S?

2.1.4.2. Assessment problems arising from multiple fitment of vehicles

In addition to the above considerations referring to single eSafety systems (system-by-system discussion), the following questions referring to combinations of systems are to be addressed:

- What are the most frequent combinations of eSafety systems ("bundles" of systems) fitted in vehicles? Experience shows that bundling of systems so far has mainly occurred for technical reasons, e.g. when sensors or actuators can be shared between multiple systems (e.g., forward looking radar for adaptive cruise control as well as collision warning / mitigation). It is, however, conceivable that future systems will also be bundled for marketing purposes. Since this bundling may be very different by make and model it may become necessary to identify the list of systems per individual vehicle (e.g., based on VIN).
- Are there significant interactions between individual systems in the sense that the efficacy of system S is affected by the presence of system A? For instance, a frequent kind of accident occurs if the driver is tired and / or distracted, leaves the carriageway, steers back and loses control. This scenario is addressed by (at least) three different systems: driver alertness monitoring, lane departure warning and stability control. The efficacy of any of these systems depends highly on the presence of the other two.
- What are the consequences of such interactions, e.g. for the definition of the system-relevant accident types. In case of complicated interactions it may turn out that most accidents are somehow affected, i.e. the control group (not affected by any of the systems) may be very small.

In Deliverable 7.4 of the TRACE project a methodology for investigating the combined effects of safety systems has been proposed (Zangmeister et al. 2007).

2.1.4.3. Assessment problems arising from different fitment of vehicles within different driver groups

Assessment of eSafety systems will be further complicated by the fact that vehicle fitment may be associated with type of vehicle driver and category of vehicle use.

[4] Quality Adjusted Life Year
[5] Disability Adjusted Life Year

Thus, for all systems to be evaluated the question has to be answered, if there are significant interactions between individual systems and the respective driver populations in the sense that the system S is more (or less) likely to be found in vehicles driven by certain groups (defined by age, gender, mileage, driving style, etc.).

2.2. Data Analysis

2.2.1. Conceptual framework for studies on accident involvement and injury risk

As accident involvement is an event occurring in time and space, the general epidemiological concept of disease incidence (incidence = number of new cases of a disease within a specified period of time) applies to studies on accident involvement risk. Epidemiological literature offers a variety of risk measures. As a prerequisite for proper application of these measures, a conceptual framework for traffic accident involvement studies has been developed by Hautzinger et al. (2007).

2.2.1.1. Traffic participation and accident involvement

The conceptual framework for accident involvement and injury risk studies proposed by Hautzinger et al. (2007) is tying together methodological concepts of mobility behaviour analysis and traffic safety research. The idea behind this concept is that “accident involvement” is just another word for “accidental trip”: Whenever a person or vehicle participates in traffic, the corresponding trip (person or vehicle trip) may terminate premature and unplanned due to involvement in a traffic accident. If this should actually happen, the corresponding trip can be classified as “accidental”. All other trips can be termed “non-accidental”. From a mobility research point of view, therefore, accident involvement can simply be regarded as a dichotomous trip characteristic (accidental trip yes/no).

2.2.1.2. Population at risk

Under the above perspective, the universe of all trips on the road system is the most natural “population at risk” of a study of traffic accident involvement. The population at risk consists of all trips generated by the members of a given population of trip makers (road users) during a specific study period. It is of fundamental importance to clearly specify the population at risk for any accident involvement study.

In epidemiology, a single element of the population at risk is called “unit at risk”. In our context trips are the units at risk. This wording simply expresses the fact that virtually any trip may end up in an accident. Correspondingly, one can also say that in the course of his or her trip the road user is exposed to the risk of being involved in an accident.

Obviously, as a prerequisite for any accident involvement investigation the accident involvement status of each road user trip sampled from the population at risk has to be specified as either “accidental” or “non-accidental”. Clearly, accident involvement status of trips corresponds to the general epidemiological term “disease status” of persons from the population at risk.

2.2.1.3. Levels of analysis

Trip level analysis

The above approach provides a clear and unified epidemiological framework for the investigation of accident involvement and injury risk at different levels of aggregation. According to the preceding considerations, road user trips are the elementary study units in empirical investigations on accident involvement. Consequently, the trip level is perhaps the most natural level of analysis for accident involvement investigations. Trip level analysis as outlined above is illustrated in Figure 1.

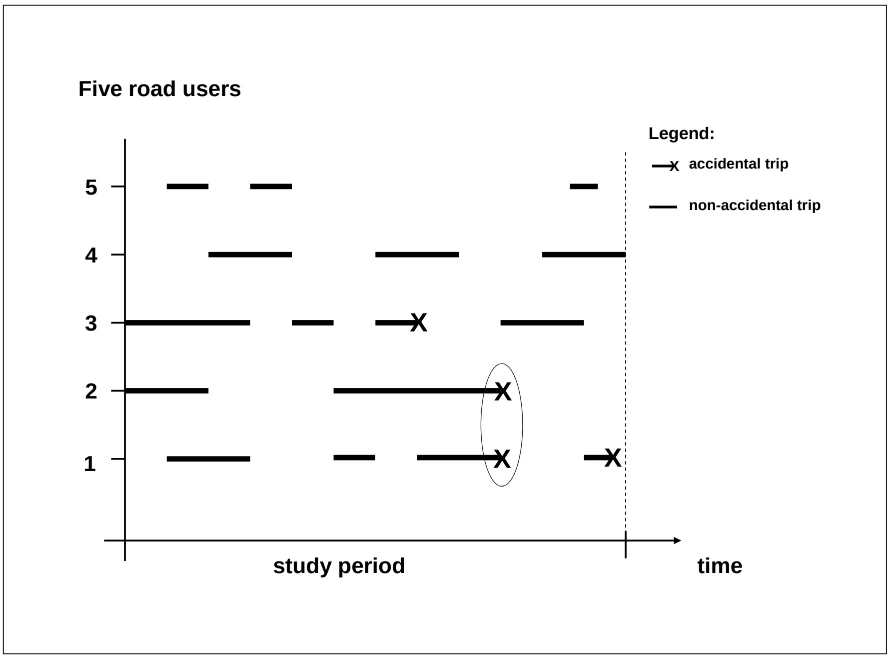

Figure 1 Example of a population of trips exposed to the risk of accident involvement

In the above hypothetical example the population of road users consists of N=5 persons which are observed over a certain study period. The persons participate in traffic from time to time. While participating in traffic the person is termed road user. For example, person 3 appears four times in the traffic system as a road user. In total, M=16 road user trips have been made by the members of the small human population during the study period. The population at risk thus consists of M=16 elements (road user trips). In Figure 1 each road user trip is represented by a horizontal line where the length of the line corresponds to the duration of the trip. In the risk population of all road user trips we find Y=4 accidental trips (marked by "X") and M-Y=12 non-accidental trips.

It is important to note that in the above example the second accident involvement of person 1 as well as the accident involvement of person 3 corresponds to a single road user accident; these two accident involvements occur independently of each other (different accident time and possibly also different location). In contrast to this, the first accident involvement of person 1 and the accident involvement of person 2 are forming a "cluster" (of size two) as these two persons (road users) are involved in the same traffic accident (e.g. a two-car crash). Consequently, in our example we have X=3 accidents (two single road user accidents and one multiple road user accident). Naturally, the clustering in the set of accidental trips (a subset of the population at risk) must be taken into account in statistical analyses of traffic accident involvement incidence. In our example the set of accidental units has 4 elements (accidental road user trips) which belong to 3 different clusters (accidents), one cluster of size two and two clusters of size one.

Person-year level analysis

One may investigate accident involvement incidence also at the person-year level. Depending on the study purpose, a person-year in the above sense corresponds to a person, who is observed over a certain time interval, e.g. a calendar year. In our context one may think of a person-year as a statistical unit corresponding to the set ("cluster") of all trips generated by a certain person during a specific year of study. Clearly, if a person does not participate in traffic during the study period, the corresponding person-year does not belong to the population at risk as he or she is not exposed to the risk of traffic accident involvement. Therefore, only person-years of mobile persons are relevant.

Every person who participates in traffic during the study period as a road user is exposed to the risk of accident involvement. Thus, at the person-year level the binary characteristic "involvement in at least one traffic accident during study period yes/no" or, more specifically, the count variable "number of accident involvements during the study period" describe the accident involvement status of a unit at risk.

In the example shown in Figure 1 the population at risk consists of N=5 elements (person-years). As can be seen, person 1 has two accident involvements (accidental trips) during the study period. Person 2 has exactly one accident involvement just like person 3. Persons 4 and 5 are the two not accident-involved persons in the hypothetical human population considered here. Thus, the subset of accidental person-years (persons involved in at least one accident during the study period) from the population at risk consists of N*=3 elements.

It should be noted here, that the question whether or not a certain individual from the human population under consideration is mobile and thus belongs to the population at risk cannot be answered before the end of the study period.

2.2.1.4. Risk factors for accident involvement

Risk factors as attributes of the units at risk

Basically, accident involvement studies at the trip level are dealing with the probability of a trip to end up in an accident, i.e. to be an accidental trip. Rarely, however, one is interested in the probability that an arbitrary trip from the population at risk is an accidental trip. Rather, one aims at evaluating the chance that a trip which possesses a certain attribute ends up in an accident. The probability of a trip being accidental given that the trip (or the corresponding vehicle) has the particular attribute under consideration is called the risk of accident involvement and the attribute considered is called risk factor.

Those units at risk which have the attribute under consideration are said to be "exposed to the risk factor". Correspondingly, the units at risk which do not have the attribute considered are said to be "not exposed to the risk factor". Frequently, the group without the risk factor will serve as a comparison group leading to the definition of the relative risk as the ratio of the risk of accident involvement for those with the risk factor to the risk of accident involvement for those without the risk factor. If the relative risk is above unity, then the factor under investigation increases risk; if less than unity it reduces risk. A factor which has a relative risk less than unity is referred to as a protective factor. As a rule, safety systems like ESP, for example, can be regarded as protective factors. It is expected that vehicle trips made by vehicles which have this attribute are less prone to be accidental compared to trips made by vehicles without the system.

***Remark**: The meaning of the epidemiological term "factor" (risk factor or protective factor) is different from the meaning of this term in analysis of variance (ANOVA). In the ANOVA terminology a "factor" is a categorical explanatory variable which might affect the distribution of a certain criterion variable. The possible values of a factor are called "levels". Thus, a risk factor in the epidemiological sense is a specific level of a factor according to the ANOVA terminology.*

In this report the term "risk factor" always refers to a specific level of the explanatory variable "risk factor status". In an ANOVA context, one would, for instance, say that the factor "vehicle equipped with ESP" is measured at two levels (yes/ no). In epidemiology, the trips belonging to the first category are said to be those with the protective factor whereas the trips belonging to the second category are said to be those without the protective factor "ESP". In the context of risk factors one may speak of those exposed and those not exposed to the risk factor.

Measuring risk factors

The risk factor status of a unit at risk is frequently measured at only two levels: "exposed" and "not exposed". One may, however, also have a set of possible categorical or ordinal outcomes of the risk factor status. In this situation one would choose one level of the risk factor status to be the base level and compare all other levels to this base. Risk factor status may also be a continuous variable or a discrete variable with a large number of outcomes. In this case the risk factor status can be grouped.

Frequently, the possibilities of assessing risk factors are limited by the fact that the characteristics recorded for accidental trips are not identical with the characteristics recorded for non-accidental trips. For instance, whereas in some in-depth data bases information on the equipment with safety systems can be found (accidental trips), this is usually not the case in mobility surveys (non-accidental trips).

2.2.2. Basic risk measures

2.2.2.1. Overview

According to epidemiological standards, different measures may be used to quantify the "chance" or "relative incidence" of traffic accident involvement:

- risk, relative risk and attributable risk
- odds and odds ratio
- incidence rate and relative incidence rate
- incidence density and relative incidence density.

Measures of chance of traffic accident involvement may be considered at different levels of analysis, especially at the trip level and the person-year level. Clearly, all these measures can only be computed in situations where the number of accident involvement events (accidental trips and accident-involved trip makers, respectively) is known for the study period under consideration.

It will appear that risk and odds can only be found in situations where in addition to the accident involvement count the size of the population at risk is precisely known (like, for instance, in a specifically designed accident involvement incidence investigation at the vehicle-year level). If only a rough estimate of the number at risk is available (e.g. total number of registered vehicles as a surrogate for the number of vehicle trips), one may use the incidence rate instead of risk or odds. Finally, if the total length (duration) of all trips belonging to the population at risk is known, we are able to compute an incidence density measure which in fact is a special type of rate.

The various risk measures presented below may refer either to the population at risk or to a sample drawn from this population. When a risk measure has been found from sample data, one may use it as an estimate for its population equivalent. Different descriptive measures of chance of accident involvement and accidental injury are considered in this section.

2.2.2.2. Risk, relative risk and attributable risk

Risk (cumulative incidence rate CIR)

At the trip level, accident involvement risk is to be understood as the number of accidental trips related to the total number of trips (accidental and non-accidental), i.e. to the size of the population at risk. Thus, accident involvement risk is the proportion of trips ending up in an accident among so many trips at risk. According to this definition, accident involvement risk always refers to a specific population of trips at risk generated by certain universe of "trip makers" during a specified period of time.

The trip-related accident involvement risk is defined as

(1.1) $\mathbf{R_T = Y/M}$ = number of accidental trips (**Y**) / number of trips at risk (**M**).

Obviously, the risk $\mathbf{R_T}$ is simply the proportion of accidental trips among all trips generated by the human or vehicle population under consideration during the study period. In the numerical example presented above we have $R_T = 4/16 = 0.25$. This means that 25 percent of all road user trips during the study period are accidental trips.

In a purely descriptive analysis referring to the complete population at risk or to a concrete sample from this population, the "empirical" risk (1.1) is frequently also called cumulative incidence rate (CIR). The term "risk" may then be reserved for the probability of the event that an arbitrary trip is an accidental trip. In this report the empirical population risk is always denoted by the capital letter R. Sample values that are estimates of their population equivalents are denoted by the lowercase letter r.

At the person-year level we may define accident involvement risk as the ratio of two counts, namely, the number **N*** of accident-involved road users and the total number **N** of all persons exposed to accident involvement risk during the study period of one year:

(1.1a) $\mathbf{R_P = N^*/N}$.

In our numerical example we have $R_P = 3/5 = 0.6$, i.e. 60 percent of the persons observed over the study period are involved in at least one traffic accident.

In real-world situations multiple accident involvement of a specific road user during a study period of standard length (e.g. one calendar year) is an extremely rare event. Therefore, **N*** will normally be only slightly smaller than the number of accidental trips **Y**. On the other hand, the total number **M** of road user trips will normally be considerably larger than the number N of persons (about 1 000 trips per person and year). In practice, therefore, the numerical value of the trip-related accident involvement risk $\mathbf{R_T}$ will be by far smaller (factor 1/1 000) than the accident involvement risk $\mathbf{R_P}$ at the person-year level.

It should be reminded that all empirical risk quantities introduced above are proportions in the sense that the numerator is part of the denominator. Thus, the traffic accident involvement risks $\mathbf{R_T}$ and $\mathbf{R_P}$ always lie between 0 and 1. The other three measures of chance of accident involvement incidence (odds, rate and density) do not have this property.

Relative risk

If in an analysis at the trip level the population of trips at risk is subdivided according to a certain characteristic (e.g. equipment of a vehicle with a certain safety system) into two groups 1 and 2 (e.g. trips made using equipped and unequipped vehicles, respectively) the group-specific risks are defined according to (1.1). Given the two group-specific risks, the relative risk of accident involvement for trips belonging to group 2, compared to those belonging to group 1, is defined as

(1.2) $$\mathbf{\Lambda = R_{T2} / R_{T1}}.$$

If more than two groups are distinguished (risk factor status measured at several levels), one group (e.g. group 1) may be considered as the reference group (also termed base group) and the analyst may relate the risk of the other groups to that of the reference group.

The Greek letter $\mathbf{\Lambda}$ (lambda) represents the population relative risk. Its estimate in a sample will be denoted by the lowercase letter lambda ($\boldsymbol{\lambda}$).

Attributable risk

The relative risk, of course, tells nothing about the overall importance of a certain risk factor. This is because it does not take into account how the units at risk are distributed over the different categories of the risk factor status variable. Let the accident involvement risk for trips made by equipped vehicles be denoted by $\mathbf{R_{T1}}$ and the overall accident involvement risk by $\mathbf{R_T}$. In the hypothetical situation where all unequipped vehicles were substituted by equipped ones, trip-making using an unequipped vehicle would no longer be present and all members of the population at risk would experience the risk of the group with the protection factor (i.e. $\mathbf{R_T} = \mathbf{R_{T1}}$).

Thus, the difference $\mathbf{R_T - R_{T1}}$ may be interpreted as the absolute increase in overall population risk due to the fact that some trip makers use unequipped vehicles instead of equipped ones. Similarly, the ratio $\mathbf{R_{T1} / R_T}$ tells the analyst something about the percentage reduction in population risk if exposure to the risk factor was completely removed. Consequently, the difference $\boldsymbol{\theta} \mathbf{= 1 - R_{T1} / R_T}$ denotes the proportion of the observed overall population risk R_T which can be attributed to the missing of the protection factor “trip-making using an equipped vehicle”. If, for instance, the difference takes on the value $\theta = 0.22$, this would mean in our example that 22 % of the overall risk of accident involvement is attributed to driving an unequipped vehicle instead of an equipped vehicle.

In epidemiology, the quantity

(1.3) $$\boldsymbol{\theta = 1 - R_{T1}/ R_T = (R_T - R_{T1})/ R_T}$$

is termed attributable risk. The attributable risk tends to be large,

- if the risk factor under consideration is rare provided the relative risk is high or
- if the relative risk is low provided the risk factor is common.

"Attributable" does not imply causation. In the above example one could, for instance, conclude that 22 % of the cases of accident involvement would be removed, if all drivers of unequipped vehicles would switch to equipped ones. This, however, would be over-optimistic if there is a third ("confounding") factor involved which determines both vehicle fitment (unequipped versus equipped) and accident involvement. Age of vehicle could, for instance, be such a confounder.

2.2.2.3. Odds and odds ratio

Odds

At the trip level, the chance of accident involvement can also be measured by relating the number of accidental trips to the number of non-accidental trips:

(1.4) $$\boldsymbol{\Omega_T = Y/(M-Y)}.$$

This measure is called the odds of accident involvement and is to be understood here as a population value. In our hypothetical example we have $\boldsymbol{\Omega_T}$ = 4/(16 - 4) = 4/12 = 0.33. This result tells us that the number of accidental trips is just one third of the number of non-accidental trips, i.e. non-accidental trips are three times more frequent than accidental trips.

The odds of accident involvement may, of course, also be defined at the person-year level:

(1.4a) $$\boldsymbol{\Omega_P = N^*/(N - N^*)}.$$

In our example we find $\boldsymbol{\Omega_P}$ = 3/(5 – 3) = 3/2 = 1.5, i.e. the number of accident-involved drivers exceeds the number of not involved (accident-free) drivers by 50 percent.

Odds ratio

If group-specific odds, i.e. odds of accident involvement for units (trips or person-years) belonging to group 1 and group 2, respectively have been determined, the odds ratio for units belonging to group 2, compared to units belonging to group 1, is given by

(1.5) $$\boldsymbol{\Psi = \Omega_2 / \Omega_1}.$$

The Greek letter $\boldsymbol{\Psi}$ (psi) is used to denote the population odds ratio. A sample value that is an estimate of the population odds ratio will be denoted by the lowercase letter $\boldsymbol{\psi}$.

Note: The appropriate risk measure for case-control studies is odds ratio rather than relative risk. This can be illustrated by the following hypothetical example:

risk factor status	cases	controls	Total
risk factor absent	30	60	90
risk factor present	70	40	110
Total	n_0=100 accidental trips	n_1=100 non-accidental trips	n_0+n_1=200 trips

Both the sample of cases and the sample of controls consist of n=100 observations. The proportion of observations where the risk factor is present amounts to 70 % for the cases and 40 % for the controls. The respective risk measures are as follows:

Relative Risk: **λ** = (30/90)/(70/110) = 0.5238

Odds Ratio: **ψ** = (30/60)/(70/40) = 0.2857

E.g., the odds ratio value indicates, that the accident involvement risk for trips, where the risk factor is absent, appears to be 71.4 % ((1 - 0.2857) x 100) lower compared to trips where the risk factor is present.

In the table below the conditional distributions of the variable “risk factor status” are exactly the same as in the previous table, however, the number of controls is much larger (n_1=10 000) than the number of cases, which is not unusual for this type of study design:

risk factor status	cases	controls	Total
risk factor absent	30	6.000	6.030
risk factor present	70	4.000	4.070
Total	n_0=100 accidental trips	n_1=10 000 non-accidental trips	n_0+n_1=10 100 trips

Relative Risk: **λ** = (30/6.030)/(70/4.070) = 0.2893

Odds Ratio: **ψ** = (30/6.000)/(70/4.000) = 0.2857

As one can see, the relative risk is different while the odds ratio value remains the same, i.e. the relative risk depends on the – often arbitrarily chosen – sample sizes.

2.2.2.4. Incidence rate and incidence rate ratio (relative rate)

Rate

In quite many analyses at the trip level one knows the number **Y** of accidental trips but not the total number **M** of trips. Similarly, in an analysis at the person-year level one may know rather precisely the number **N*** of trip makers who had an accident but has no information on the number **N** of persons at risk. In both cases we know the numerator but not the denominator required to determine the risk of accident involvement.

If we have at least a rough estimate $\mathbf{M_0}$ of the size of the population at risk or any other quantity to which the number of accident involvement events can be related in a meaningful way, we may use the corresponding quotient to measure the chance of accident involvement. Any quotient of the form

(1.6) $\boldsymbol{\rho = Y/M_0}$

is called accident involvement rate.

Well known examples of accident involvement rates are

- the per-capita accident involvement rate for analyses at the person-year level (**Y** = annual number of accidental person trips; $\mathbf{M_0}$ = mid-year population) and
- the per-vehicle accident involvement rate for analyses at the vehicle-year level (**Y** = annual number of accidental vehicle trips; $\mathbf{M_0}$ = mid-year vehicle stock).

If the study period covers T years, the total number Y of accidental person and vehicle trips should be related to the total number Z of person- or vehicle-years for this period. In the simplest case we may determine the denominator of the rate Y/Z as follows: $Z = M_0(1) + M_0(2) + \ldots + M_0(T)$, where $M_0(t)$ is the mid-year vehicle stock for year t (t=1,…,T).

Relative Rate

To compare two rates one can use the relative rate, group 2 compared to group 1, which is given by

(1.7) $\boldsymbol{\rho_{rel} = \rho_2/\rho_1 = (Y_2/M_{02}) / (Y_1/M_{01})}$.

The relative rate is sometimes also called incidence rate ratio (**IRR**).

2.2.2.5. Incidence density and incidence density ratio (relative density)

Obviously, the characteristic "duration of trip" - which under a different perspective may also be termed "traffic participation time under risk" - varies in the population of trips. Similarly, the "person-time under risk" corresponding to the total duration of all trips made by a given person during the study period varies in the population of persons. The phenomenon of non-constant time under risk (non-constant risk exposure time) for the elements of the population at risk is quite common in epidemiological research and leads to the epidemiological concept of "incidence density".

Accident involvement density

For a given human population which is observed over a certain study period, the time-related accident involvement density is defined as the ratio of incidence of accidental person-trips (**Y**) and total person-time under risk (**T**):

(1.8) $\boldsymbol{\delta_{Time} = Y / T}$.

The denominator **T** is considered as an "aggregate measure of exposure" that can be used to normalise the incidence of accident involvement **Y**. In our context this risk concept is especially reasonable because at every moment in time, while participating in road traffic, the possibility of accident involvement exists.

In our hypothetical example we may assume the mean trip duration to be 30 min per trip. As there are 16 trips in the population at risk, the total person-time under risk is **T** = 16• 30 = 480 min (or 8 hours). Under this assumption we obtain the time-related accident involvement density $\boldsymbol{\delta}_{\mathbf{Time}}$ = 4/8 = 0.5. This measure tells us that on average there are 0.5 accident involvements of persons per person-hour of traffic participation (fortunately, real-world incidence densities are by far lower).

At every spatial point in the road network which is passed by a person while participating in traffic an accident may happen. When person-trips are considered as study units we may, therefore, also use “trip length” or, equivalently, “travel distance under risk” as an appropriate measure of traffic accident involvement risk exposure. Thus, the total person-distance under risk **D** generated by the population of persons during the study period may serve as a standard for comparison.

This concept leads to the distance-related accident involvement density as a relative measure of accident involvement incidence:

(1.9) $\boldsymbol{\delta}_{\mathbf{Distance}} = \mathbf{Y/D}$

Assuming the mean trip length in our hypothetical example to be 15 km per trip, the total person-distance under risk equals **D** = 16 • 15 = 240 km. Hence, the distance-related traffic accident involvement density is $\boldsymbol{\delta}_{\mathbf{Distance}}$ = 4/240 = 0.0167 accident involvements of persons per person-kilometre of traffic participation.

In general, we may denote the accident involvement density by

(1.10) $\boldsymbol{\delta} = \mathbf{Y/X}$,

where **X** represents the population total of a suitable exposure characteristic of the units at risk.

Relative density

To compare two densities one can use the relative density, group 2 compared to group 1, which is given by

(1.11) $\boldsymbol{\delta}_{\mathbf{rel}} = \boldsymbol{\delta}_{\mathbf{2}}/\boldsymbol{\delta}_{\mathbf{1}}$.

In the epidemiological literature the relative density is also called incidence density ratio (**IDR**).

2.2.2.6. A note on the differences between risks, odds, rates and densities

The various measures of chance of accident involvement as introduced above are all deterministic measures referring to a particular well-defined finite population at risk. The numerical values of these measures can be obtained by surveys (accident involvement surveys) or some other types of study. Except for very specific populations at risk, it will not be possible to conduct a complete census yielding the true or exact value of the measure under consideration. Rather, some type of sampling from the population at risk will provide data which allow the measure of chance to be estimated.

For instance, the sample value r of accident involvement risk is to be interpreted as an estimate of the corresponding population risk R. Clearly, both the population and the sample risk are proportions. But whereas R is a fixed (but unknown) quantity, the sample risk r is a random variable. Obviously, odds, rates and densities as defined above are not proportions. Therefore, these quantities are measures of chance but may not be interpreted as “risk” quantities in the above narrow sense. The only exception to this rule is the accident involvement rate ρ which under favourable circumstances might be a good approximation to the population risk R.

Calculation of both the risk and the odds of accident involvement require precise knowledge of the size of the population at risk. It should be noted, that the odds are rarely of interest as the risk is the generally preferred measure of chance. However, in studies on the comparative chance of accident involvement the odds ratio has at least the same importance as the relative risk - either because the odds ratio is all we can estimate (e.g. in case-control studies) or is the more convenient to calculate (e.g. in logistic regression analysis).

As already noted, the accident involvement density (e.g. δ_{Time}) is not a proportion ($\delta_{Time} > 1$ is possible). Rather, it is a ratio of two population characteristics (number of accidental trips related to the total duration of all trips under risk). Therefore, δ_{Time} may not be interpreted as a risk quantity in the above sense: Density is not a measure of accident involvement “risk” but a measure of accident involvement “intensity”. It expresses the incidence of accident involvement per unit of a certain risk exposure quantity, especially the number of accident involvements per hour or per kilometre of traffic participation. As both the duration and length of trips (trips are the elementary units at risk) varies, density measures are appropriate measures of chance of accident involvement.

Formally, there is no difference between the sample values of densities and rates as both measures are quotients of a random variable and a quantity usually assumed to be fixed and known. Conceptually, however, a clear distinction can be made: the denominator T and D, respectively, of the accident involvement density is the population total of a characteristic of the units under risk, whereas the denominator M_0 of the accident involvement rate typically is some crude estimate of the size M of the population at risk.

One conclusion that can be drawn from these considerations is that the data analysis phase should be taken into account before starting the data collection phase.

2.2.2.7. Statistical models for measures of chance of accident involvement

Not surprisingly, different statistical models have to be applied when specific measures of chance of accident involvement are to be estimated from sample data (rates and densities, however, may be estimated and analysed using the same models). Various statistical models suitable for traffic accident involvement risk studies are available enabling the researcher to identify and assess risk factors for accident involvement and thus accident causes (see Hautzinger et al. 2007):

- Models for risk, relative risk and attributable risk
  - binomial model for accident involvement risk
  - normal distribution model for the log of the relative risk
  - model for the attributable risk
  - logistic regression model for involvement risk

- Models for odds and odds ratio
  - normal distribution model for the log of the odds ratio
  - regression model for the log odds of accident involvement

- Models for counts, rates and densities
    - Poisson model for accident involvement counts
    - Poisson model for accident involvement rates and densities
    - Log-linear models for counts, rates and densities

The above risk measures and models are also suitable to assess the risk of being injured in a road traffic accident. Here, a distinction has been made between the unconditional injury risk associated with traffic participation and the road user's risk to receive an injury given that he or she is involved in an accident (conditional injury risk).

2.2.2.8. Measures of chance of accidental injury

The statistical concepts and methods for measuring accident involvement risk as presented above can also be applied to assess the risk of being injured in a road traffic accident. It has, however, to be specified clearly what is meant by "injury risk". In a study on injury risk of road users we may distinguish the following two different types of injury risk:

- risk of being involved and injured in an accident (unconditional road user injury risk)
- risk of being injured given accident involvement (conditional road user injury risk)

The two concepts of injury risk are considered in the sequel.

Unconditional road user injury risk

From a descriptive point of view the unconditional road user injury risk corresponds to the proportion of trips having the following two characteristics:

- trip ends up in an accident ("accidental trip")
- trip maker (road user) is injured in the accident .

Thus, for investigations on unconditional road user injury risk at the trip level, the accident involvement status of the units at risk (road user trips) is to be considered as a characteristic measured at three levels:

- non-accidental trip,
- accidental trip where road user remains uninjured and
- accidental trip where road user is injured.

Consequently, unconditional road user injury risk $\mathbf{R_{inj}}$ is defined as

(1.12) $\mathbf{R_{inj} = Y_{inj}/M}$

where the symbols $\mathbf{Y_{inj}}$ and $\mathbf{M}$ denote the number of accidental trips where the road user is injured and the total number of road user trips (accidental and non-accidental road user trips), respectively.

Similarly, the odds of road user injury is to be defined as

(1.13) $\mathbf{Q_{inj} = Y_{inj}/(M\text{-}Y_{inj})}$.

With these definitions in mind, all risk measures, models and methods proposed here for quantifying accident involvement risk may also be applied to assess unconditional road user injury risk. As can be seen, unconditional road user injury risk may be considered simply as a specific type of accident involvement risk, namely, the risk to be involved in an accident as an injured road user.

Conditional road user injury risk

Quite often, the injury risk of accident-involved road users is of interest. A typical example is crashworthiness assessment of cars where usually the vehicle's level of passive safety is measured by conditional driver injury risk.

Conditional road user injury risk refers to the chance of accidental injury given that the road user is involved in an accident. Consequently, the population at risk is no longer the universe of all trips but rather the subset of accidental trips and, correspondingly, the binary disease status variable is now defined as "trip maker injured yes/no". Clearly, the conditional road user injury risk is by far larger than the unconditional risk of road user injury.

In order to measure conditional road user injury risk at the trip level, the number $\mathbf{Y_{inj}}$ of accidental trips where the road user is injured has to be related to the total number $\mathbf{Y_{acc}}$ of accidental road user trips:

(1.14) $\mathbf{Q_{inj|acc} = Y_{inj} / Y_{acc}}$

Obviously, information on non-accidental trips is not required when assessing conditional road user injury risk. Rather, the use of accident data is sufficient. However, as the following remark shows, standard traffic accident data as provided by national statistics or regional in-depth studies can only be used for calculating the conditional driver injury risk if all traffic accidents (i.e. injury and non-injury accidents) are registered.

Remark: *Normally, injury accidents are completely registered in national traffic accident statistics. Thus, the annual number* $\boldsymbol{Y_{inj}}$ *of injured road users (more precisely: accidental trips where the road user is injured) should usually be known precisely up to a certain number of unreported cases (so-called dark figure). This, however, cannot always be assumed for the total number* $\boldsymbol{Y_{acc}}$ *of accident-involved road users (accidental road user trips).*
As $\boldsymbol{Y_{acc}}$ *equals* $\boldsymbol{Y_{inj}}$ *plus the number of accidental trips where the road user remained uninjured, the denominator of the conditional injury risk measure* $\boldsymbol{Q_{inj}}$ *requires reporting of uninjured accident-involved road users. Thus, for countries where police is reporting injury accidents only, the conditional road user injury risk* $\boldsymbol{Q_{inj}}$ *cannot be calculated. The same holds for countries where accident registration is tied to the fulfilment of certain requirements referring to the type or severity of the accident (e.g. minimum amount of material damage).*

If the target population of an accident survey is only a subset of all traffic accidents, one could, of course, simply disregard the subgroup of accidents that are not covered by the target population. A typical example would be to investigate the conditional injury risk of road users which are involved in accidents where at least one road user is injured (so-called injury accidents). Normally, it will be difficult to interpret such injury risk estimates in a meaningful way. In this context "matched" studies (e.g. analysis of two-car crashes) offer possibilities to overcome the problem; from matched studies, however, only relative injury risks can be estimated. Cummings et al. (2003) and Hautzinger (2006) describe methods for assessing relative injury risks under matched study designs.

2.3. Empirical Examples for the Evaluation of Vehicle Safety Systems

2.3.1. Example 1: Analysing a cohort study

2.3.1.1. Data

In this section the data from the FAS project presented in chapter 2.1.3 are analysed according to the usual cohort design, i.e. without pairing vehicles prior to data analysis (see section 2.3.3 for the matched-pairs analysis). Accident reporting of the companies participating in the FAS project yielded the following empirical results: Among vehicles *with* ACC, ESP and LGS ("cohort 1") 84 out of 715 vehicles were involved in at least one road traffic accident during the investigation period. Correspondingly, among vehicles *without* ACC, ESP and LGS ("cohort 2") 87 out of 535 vehicles were involved.

2.3.1.2. Empirical risk and odds of accident involvement

According to the data presented above the *empirical accident involvement risk* (also termed cumulative incidence rate, CIR) is

11.8 % for equipped vehicles (84/715=0.118)

and

16.3 % for reference vehicles (87/535=0.163).

As an alternative measure of chance of accident involvement one may use the *odds of accident involvement* which amounts to

0.1331 for equipped vehicles ((84/(715-84)=84/631=0.1331))

and

0.1942 for reference vehicles ((87/(535-87)=87/448 =0.1942)).

2.3.1.3. Accident involvement rate

Some vehicles were involved in more than one road accident. Calculations gave the total number of accidents as 104 accidents (accident involvements of vehicles) for cohort 1 and 104 accidents for cohort 2. Thus, the accident involvement rate is

0.145 accidents per equipped vehicle (104/715=0.145)

and

0.194 accidents per reference vehicle (104/535=0.194), respectively.

2.3.1.4. Accident involvement density

An appropriate measure of risk exposure is the sum of kilometres driven in the investigation period. Table 3 summarises the results:

Cohort	number of accident involvements	sum of kilometres driven
equipped vehicles	104	205 211 322
reference vehicles	104	132 557 552
Total	208	337 768 874

Table 3 Number of accident involvements and mileage of heavy goods vehicles in the investigation period

A suitable risk measure for cohort studies with varying observation periods (varying amount of risk exposure) is the *incidence density*

$\boldsymbol{\delta = Y/D}$,

which in our context may also be named *accident involvement density*. Here, **Y** denotes the total number of accident involvements and **D** is the sum of vehicle kilometres in the respective cohort. For details see Böhning 1998, p. 55-58.

The incidence density as a measure of the absolute accident involvement risk amounts to

5.068 accidents per 10 Mio. vehicle-km for the test group (equipped vehicles)

and

7.846 accidents per 10 Mio. vehicle-km for the control group.

2.3.1.5. Relative accident involvement density

In order to compare the accident involvement densities for the two groups, the *relative density* or *incidence density ratio (IDR)* is used, which is given by

$\boldsymbol{\delta_{rel} = \delta_2/\delta_1}$.

For the FAS data, the empirical value of the relative density is

$\delta_{rel} = 5.068/7.846 = 0.646$.

That means that the mileage-related accident involvement risk for heavy goods vehicles with ACC, ESP and LGS is 35.4 % lower compared to not equipped vehicles.

In order to calculate a confidence interval for the population incidence density ratio, the variance of log δ_{rel} has to be estimated by

$\boldsymbol{1/Y_0 + 1/Y_1}$.

($\boldsymbol{Y_0}$ = number of accident involvements in the reference group; $\boldsymbol{Y_1}$ = number of accident involvements in the test group). Applied to the FAS data, the variance of log $\boldsymbol{\delta_{rel}}$ is estimated at 0.01923 (=1/104 + 1/104) which leads to a standard deviation of

√0.01923 = 0.138675. Calculations give the 95 % confidence interval for the population incidence density ratio as

$$\delta_{rel}\cdot\exp[\pm 1.96 \cdot 0.138675] = \delta_{rel}\cdot\exp[\pm 0.271803].$$

Hence, the confidence limits are 0.492 (lower limit) and 0.848 (upper limit), respectively. The 95 % confidence interval for the incidence density ratio does not cover unity. Rather, the upper limit of the interval is lower than 1, thus the accident involvement risk of trucks equipped with ESP, ACC and LGS is significantly lower than the corresponding risk of non-equipped trucks.

2.3.1.6. Relative risk and odds ratio

In addition to using the incidence density ratio as a measure of relative accident involvement risk, one could calculate the relative chance of accident involvement by comparing the group-specific empirical risks and odds of accident involvement, respectively. Under this approach the results of the FAS project are to be displayed in the format of Table 4.

Cohort	**accident involvement of vehicle**		**Total**
	yes	**no**	
test group	84	631	715
control group	87	448	535
Total	171	1 079	1 250

Table 4 Heavy goods vehicles broken down by accident involvement and cohort affiliation

From this table the sample relative risk and odds ratio can easily be calculated:

Relative risk: $\lambda = (84/715) / (87/535) = 0.11748 / 0.16262 = 0.7224$

Odds ratio: $\psi = (84/631) / (87/448) = 0.1331 / 0.1942 = 0.6854$

The results indicate that if group differences in vehicle mileage are ignored the accident involvement risk for heavy goods vehicles with ACC, ESP and LGS appears to be 27.8 % and 31.5 %, respectively, lower compared to not equipped vehicles.
A 95 % confidence interval for the population odds ratio **ψ** can be calculated as follows:

$$\exp[\ln \psi \pm 1.96 \cdot SE_{\ln \psi}]$$

where $SE_{\ln \psi} = \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$ (standard error of ln **ψ**; **a**, **b**, **c**, **d** denote the frequency counts in the cells of the 2X2 table).

The resulting confidence interval for the true population odds ratio **ψ** is exp[ln 0.6854 - 1,96 · 0.16497] = 0.496 (lower limit), and exp[ln 0.6854 + 1,96 · 0. 16497] = 0.947 (upper limit), i.e. we are 95 % sure that the interval (0.496, 0.947) contains the true odds ratio[6]. Since the confidence interval does not cover unity it is statistically significant that the relative chance of being involved in an accident is lower for equipped vehicles compared to non-equipped ones.

2.3.1.7. Adjusted odds ratio

Obviously, by using the above "crude" risk measures (relative risk, odds ratio) one does not consider possible confounders. It can, for instance, be assumed that the probability of being involved in an accident during the reference period increases with the kilometres driven in the respective period. In order to account for the impact of the factor "vehicle mileage" on the probability of accident involvement of heavy goods vehicles a (binary) logistic regression model can be used. In such a model the dependent variable is "accident involvement of vehicle in the investigation period (yes/ no)" and the predictor variables are "vehicle mileage in the investigation period (in 100 000 km)" and "cohort affiliation (test group = vehicle equipped with ESP, ACC and LGS/ control group = non-fitted vehicle)"[7].

The logistic model can be formulated as follows:

$$\mathbf{p_{ij} = \exp(u_{ij})/[1+\exp(u_{ij})] = 1/[1+\exp(-u_{ij})]}$$

where $\mathbf{p_{ij}}$ denotes the probability for accident involvement of vehicle **i** given mileage $\mathbf{X_j}$ and cohort affiliation **j** and $\mathbf{u_{ij}}$ is defined as

$$\mathbf{u_{ij} = \mu + \beta X_j + \alpha_i.}$$

Based on 1 250 observations the regression analysis yields the following results:

Parameter	Estimate	Standard Error	Significance
Intercept	-1.979	0.215	<0.0001
Vehicle mileage	0.134	0.069	0.0527
Cohort affiliation			
test group	-0.430	0.167	0.0102
control group	0	-	-

Table 5 Binary logistic regression model for the accident involvement of heavy goods vehicles

According to the sign of the regression coefficients (0.134 and -0.430, respectively) it can be said that the probability of accident involvement increases with mileage in the investigation period (p-value slightly above a significance level of 5 %) and decreases with a better safety equipment of the vehicle (p-value 0.01, i.e. highly significant). Under the logistic regression model *adjusted odds ratios* (ψ_{mod}) related to the model's independent variables (determinants of accident involvement risk) can be calculated. The results of the corresponding analysis can be interpreted as follows:

- If one compares the odds of accident involvement of vehicles which differ by 1 unit regarding mileage (100 000 km), one obtains ψ_{mod} = exp[0.134] = 1.143.

[6] As the true odds ratio **ψ** is a fixed but unknown quantity the calculated interval, of course, may or may not cover **Ψ**. According to statistical theory we can, however, be "almost sure" (more precisely: "95 % sure") that **Ψ** is contained in the interval.

[7] If the dependent variable is "number of accident involvements" (0, 1, 2,…) i.e. if the determinants of accident involvement density are to be assessed, Poisson regression is an appropriate statistical model.

Thus, the chance of accident involvement for vehicles with a higher mileage is 1.143-fold the corresponding chance of vehicles with a lower mileage.

- The odds of accident involvement of vehicles equipped with ESP, ACC and LGS is 0.651-fold the odds of unequipped vehicles (ψ_{mod} = exp[-0.430] / exp[0] = 0.651). As can be seen, the adjusted odds ratio (0.651) is very close to the relative accident involvement density calculated above (0.6854) which, however, refers to the *number* of accident involvements.

It is, of course, possible (and reasonable) to include additional independent variables in the logistic model like, e.g., vehicle age or variables describing the operation area of the vehicle.

All in all it can be said that heavy goods vehicles fitted with ESP, ACC and LGS show a considerably lower accident involvement risk compared to vehicles without these systems. The observed group differences with respect to accident involvement (all accident types) are statistically significant.

2.3.2. Example 2: “Induced exposure”-analysis

In a situation where no exposure data are available the evaluation of safety systems can take place by using the so-called induced exposure approach provided that an appropriate control group can be found among the accidental units themselves. The idea behind this concept is to compare equipped and not equipped vehicles involved in single or multi vehicle crashes with respect to involvement in “system-specific” and “neutral” accidents. System-specific accidents are those which are supposed to be prevented or mitigated by the system under investigation whereas neutral accidents form the control group and thus should not be affected by the system.

2.3.2.1. Data

The following example is based on GIDAS data (German In-Depth Accident Study) which have been provided by Volkswagen AG. The sample consists of n=10 270 accident involved passenger cars (study period 1995 to 2011) with and without ESP, i.e. the system to be evaluated is ESP.

For the evaluation of ESP the accident characteristic “skidding” has been chosen to distinguish between system-specific accidents (car was skidding) and neutral accidents (no skidding). In the GIDAS codebook the variable “skidding prior to impact” is described as follows: “The vehicle is deemed to have skidded if its longitudinal axis and the direction of motion do not correspond. Normal cornering is not skidding”. The code labels are “skidded, no further details” (1), “no skid” (2), and “unknown” (9).

The hypothesis is that the proportion of skidding accidents (more precise: cars that were skidding prior to the crash) is lower among the vehicles fitted with ESP, since ESP is supposed to prevent cars from skidding.

It should be stressed that the adequate choice of both the system-specific and the neutral accident type is the crucial point when performing an induced exposure analysis[8]. Especially if the accidents regarded as neutral are actually also affected by

[8] However, this is not a statistical question. In order to judge the adequacy of both the system-specific and the neutral accident type, detailed knowledge of the technical functionality of the system to be investigated is required. I.e., expertise in automotive engineering and safety technology as well as know-how in accident research is needed.

the system this would lead to biased results. Very often, however, the accident characteristics necessary for properly distinguishing system-specific and neutral accidents are not available in routine data bases (like police recorded accident data e.g.).

2.3.2.2. Relative risk and odds ratio

In the subsequent analyses the two variables

- tyre age (<3 years; 3-5 years; 6 years and older), and
- road condition (dry; other)

are considered as potential confounders. First, however, the data scheme for "crude" analysis (without potential confounders) is shown in Table 6:

ESP	**skidding**		**Total**
	no	**yes**	
yes	2 792	102	2 894
no	6 602	774	7 376
Total	9 394	876	10 270

Table 6 Accident-involved passenger cars by skidding and presence/absence of ESP (Source: GIDAS data 1995 to 2011)

It has to be noted that this table only contains cases where all variables (including the above mentioned potential confounders) have non-missing values. Especially the values of the variable tyre age are unknown for many accident involved cars (which therefore have been ruled out). However, if the table above is produced under inclusion of all available cases, it turns out that the conditional distributions of the variable "skidding" are very similar to those depicted in the table. Therefore, it can be assumed that no substantial bias is introduced by the fact that many cases have missing values[9] for tyre age (and road condition).

The respective risk measures calculated from the table are as follows:

Odds Ratio: $\psi = (102/2\,792)/(774/6\,602) = 0.3116$

Relative Risk: $\lambda = (102/2\,894)/(774/7\,376) = 0.3359$

Thus, the chance of skidding is approx. 68 % lower for cars with ESP compared to cars without this feature.

2.3.2.3. Adjusted odds ratio

If the potential confounders mentioned above are to be included in the analysis, a statistical model must be used. Since the dependent variable "skidding" has two

[9] Almost any database - including accident and exposure data - does have incomplete information on some units. Per default, most software products simply delete all units with missing values relevant for the analysis at hand (so-called complete-case approach). However, there are many situations for which such an approach leads to a substantial loss of information and thus reduces the possibility for accurate estimates, short confidence intervals and statistically significant results. A compendium of the most useful methods for handling missing data in accident research can be found in Grömping et al. 2007.

categories (yes/ no), a binary logit model is the appropriate statistical tool. Here, the probability of "skidding=yes" is modelled. The model estimation results are shown in the following tables:

Predictor variable	df	Wald Chi²	Significance
ESP	1	95.96	<0.0001
Road condition	1	297.86	<0.0001
Tyre age	2	6.37	0.0413

Table 7 Binary logit model for skidding of accident-involved passenger cars

The above Chi²-values indicate that the variable road condition has the strongest impact on skidding followed by the presence/ absence of ESP. The variable tyre age plays a comparatively minor role but is nevertheless significant at a 5 % level.

Parameter	Estimate	Standard Error	Significance
Intercept	-1.2657	0.0769	<0.0001
ESP			
yes	-1.0755	0.1098	<0.0001
no	0	-	-
Road condition			
dry	-1.2515	0.0725	<0.0001
other	0	-	-
Tyre age			
<3 years	-0.2295	0.0912	0.0118
3-5 years	-0.1476	0.0927	0.1115
>5 years	0	-	-

Table 8 Parameter estimates of the binary logit model for skidding of accident-involved passenger cars

The parameter estimates tell us that the probability for skidding is significantly

- lower for cars with ESP compared to those without ESP
- lower on dry road surface compared to "not dry"
- lower for cars with a tyre age of <3 years compared to those where the tyres are at least 6 years old.

Without any doubt, the results are plausible. From the parameter estimates adjusted odds ratios can be calculated, the one for ESP vs. no ESP amounts to

$$\psi_{mod} = \exp[-1.0755] / \exp[0] = 0.3411,$$

i.e. the efficiency of ESP appears to be somewhat lower when additional predictor variables are considered in the analysis (crude odds ratio 0.3188). A possible explanation for this phenomenon is that there is a correlation between presence/ absence of ESP and tyre age. As tyre age tends to be lower for ESP-equipped cars, the efficiency of ESP is slightly overestimated when disregarding the variable tyre age in the analysis.

2.3.3. Matched-pairs analysis: Methodology and example

As already mentioned a matched-pairs study requires a matched-pairs analysis, which can be more complex both to understand and compute (Woodward, 2005, p.300ff), but also more effective than approaches lacking appropriate methodological rigor (detailed information on matched-pairs methodology can, e.g., be found in Cummings, McKnight, and Greenland (2003); Cummings, McKnight, and Weiss(2003); Cummings and McKnight (2004)).

2.3.3.1. Assessment of risk factors without adjustment for vehicle-specific variables

In this case the data to be analysed can be displayed in a two-dimensional contingency table. When both accident involvement status and vehicle equipment are binary variables a 2x2 table will arise under a matched cohort study design (cross-tabulation of pairs of vehicles by accident involvement status of equipped vehicle and accident involvement status of reference vehicle; see Table 2 in Section 2.1.3).

For matched studies the *matched odds ratio* is the appropriate measure of comparative chance of accident involvement. The matched odds ratio can be estimated from the 2x2 table; it appears that only the off-diagonal elements of the table are relevant for point estimation. Confidence intervals for the population value of the matched odds ratio can be computed using the F-distribution.

For 2x2 tables the null hypothesis of no association between the risk factor vehicle fitment and the criterion variable accident involvement status (corresponding to no effect of vehicle fitment, i.e. equal accident involvement risk of equipped and not equipped vehicles) can be tested using a so-called symmetry test (*McNemar's test*).

When there are more than two levels of accident involvement (e.g. not involved in an accident/ involved in an accident with only material damage/ involved in an accident with personal injury) the empirical frequency data will be displayed in r x r tables. For r x r tables *Bouwker's test* is the appropriate method for testing the hypothesis of symmetry (Hautzinger 2006). However, as for r x r tables symmetry and no association are no longer equivalent, specific tests of "marginal homogeneity" (corresponding to no effect of vehicle fitment, i.e. equipped and not equipped vehicles do not differ with respect to distribution of accident involvement status) are required to assess the effect of vehicle fitment on accident involvement status. Among other approaches the Stuart-Maxwell test and the Bhapkar test can be used. For statistical details see, for instance, Agresti (2002).

In order to demonstrate application of matched-pairs (matched cohort) analysis, data from the German FAS project introduced in Section 2.1.3 are used. Here, 527 pairs of heavy goods vehicles have been formed. The data on accident involvement of the pair members is depicted in the following table:

Equipped vehicle (test group)	**Reference vehicle (control group)**		**Total number of pairs**
	involved in accident	**not involved in accident**	
involved in accident	19	43	62
not involved	67	398	465
Total number of pairs	86	441	527

Table 9 Pairs of heavy goods vehicles by accident involvement of the pair members

In order to calculate the matched odds ratio only those pairs are considered where accident involvement status differs between the respective pair members (43 + 67 = 110 “discordant” pairs of vehicles). The pairs where both the equipped and the reference vehicle have or have not been involved in an accident contain no relevant information for the assessment of relative risk. Thus, the *matched odds ratio* is calculated by

$$\psi_m = 43/67 = 0.6418.$$

If we compare the result to the unmatched odds ratio which has been estimated at ψ=0.6854 (see Section 2.3.1), the matching effect turns out to be relatively small in our example. As usual, however, the matched odds ratio proves to be numerically smaller than the unmatched odds ratio.

As already mentioned, for 2x2 tables the null hypothesis of “no association” or “marginal homogeneity” (i.e. equal accident involvement risk of equipped and not equipped vehicles) can be tested by using McNemar’s test. As for the matched odds ratio point estimator, the test only uses the discordant pairs and is computed as follows:

$$\chi^2 = (43-67)^2/(43+67) = 5.236$$

The value obtained is to be compared to a chi-square distribution with 1 degree of freedom, which yields a p-value of 0.0221. Thus, the null hypothesis of no effect of vehicle fitment on accident involvement can be rejected at the 5 % and even the 2.5 % significance level.

McNemar’s test is actually a two-sided test:

H_0 : Matched Odds Ratio = 1 against H_1 : Matched Odds Ratio ≠ 1.

Frequently, however, it will be of interest whether or not there is a positive effect of vehicle fitment. In such a situation we want to test

H_0 : Matched Odds Ratio = 1 against H_1 : Matched Odds Ratio < 1

corresponding to

$H_0 : \pi = 1/2$ against $H_1 : \pi < 1/2$

where $\boldsymbol{\pi}$ denotes the conditional probability Pr {“equipped vehicle involved, not equipped vehicle not involved” I “discordant pair”}. For the one-sided test we may use the usual Gauss test:

$$Z = [43/110 – 1/2] / \sqrt{[(1/2)\cdot(1/2)/110]} = -2.288.$$

From a table of the standard normal distribution we find the p-value 0.011. Thus, the positive safety effect of fitting heavy goods vehicles with ACC, ESP and LGS proves to be highly significant (note that $Z = -\sqrt{5.236} = -2.288$). For details see Woodward (2005), p. 303.

Additional analyses reveal that equipped and unequipped vehicles differ as regards vehicle age and distance driven in the study period – on average reference vehicles are older and have a slightly lower mileage. Therefore, it is reasonable to control for confounders by using an appropriate statistical model.

2.3.3.2. Assessment of risk factors with adjustment for vehicle-specific variables

When the adjusted odds ratio for the risk factor (better: protection factor) "vehicle fitment" is of interest (adjustment for confounding vehicle-specific variables), the statistical analysis of two-dimensional contingency tables is no longer sufficient. Rather, specific regression models for accident involvement status are needed which in addition to the risk factor also contain confounding factors as explanatory variables. As accident involvement status of the vehicles clustered in the same pair cannot be regarded as two independent observations, the cluster or multilevel structure of the data (level 1: pairs; level 2: vehicles) must be taken into account.

Among several alternative statistical models the fixed effects logit model appears to be most suitable for the analysis of matched-pairs data, especially when both theoretical and practical considerations play an important role. In order to obtain empirical estimates of the regression parameters and the corresponding (adjusted) odds ratios one can transform the fixed effects logit model in a specific way ("conditioning out pair-specific fixed effects") leading to the so-called *conditional logistic regression model for matched-pairs data*. This model can be estimated using standard logistic regression software. Here, the assumption has to be made, that for each pair the accident involvement of vehicle 1 (equipped vehicle) is independent from the accident involvement of vehicle 2 (reference vehicle)[10].

Very briefly, the method for estimating the parameters can be described as follows:

- Eliminate all vehicles where accident involvement status of the two vehicles does not differ.
- Create difference scores for all vehicle-specific covariates (variable value for vehicle 1 minus variable value for vehicle 2).
- Use maximum likelihood to estimate the logistic regression predicting accident involvement status of vehicle 1 with the difference scores as predictor variables.

Instead of calculating difference scores (where each observation in the data set represents one vehicle *pair*) it is also possible to store each vehicle in a data line and to link the paired vehicles by a variable which indicates pair membership (e.g. "pair number"). In SAS, e.g., the conditional logistic regression model can be estimated with the LOGISTIC procedure where the variable denoting pair membership is referred in the STRATA statement. It should be noted that from matched-pairs data we cannot estimate the absolute risk of accident involvement but only the comparative chance of accident involvement.

The conditional logistic regression model for the FAS data is based on the discordant pairs[11] (as in the "crude analysis" only discordant pairs are relevant). The dependent variable in this model is "accident involvement of vehicle in the investigation period" which is either "yes" for the equipped vehicle and "no" for the unequipped one, or "no" for the equipped vehicle and "yes" for the reference vehicle. Besides the variable "cohort affiliation" (equipped/ not equipped) there are 4 additional (vehicle-specific) predictor variables in the model:

- Distance driven in the study period (mileage – in 100 000 km)
- Year of manufacture
- Truck manufacturer (2 categories named A, B[12])
- Area of operation (local or regional transport / long distance transport)

---

[10] In applications where this assumption is not realistic one can use the random effects probit model.
[11] Due to missing values only 106 out of 110 discordant pairs can actually be used in the model.
[12] A is the manufacturer with the highest frequency in the sample, B is the residual category.

The model estimation results are shown in the following table:

Predictor variable	df	Wald Chi²	Significance
cohort affiliation	1	6.5426	0.0105
mileage (in 100 000 km)	1	7.1693	0.0074
year of manufacture	1	0.3853	0.5348
truck manufacturer	1	0.5601	0.4542
area of operation	1	0.0126	0.9107

Table 10 Conditional logistic regression model for the accident involvement of heavy goods vehicles (matched cohort design)

The results show that only 2 out of 5 predictor variables prove to be significant at the 5 %-level. The variables "year of manufacture", "truck manufacturer", and "operation area" are far from being statistically significant whereas mileage (i.e. level of risk exposure) and cohort membership (i.e. vehicle fitment) are relevant factors for accident involvement risk.

The parameter estimates are shown in Table 11. In the conditional logistic regression analysis the probability of "accident involvement = yes" is modelled.

Parameter	Estimate	Standard Error	Significance
cohort affiliation			
test group (equipped)	-0.8841	0.3457	0.0105
control group (unequipped)	0	-	-
mileage (in 100.000 km)	0.5509	0.2057	0.0074
year of manufacture	0.0723	0.1165	0.5348
truck manufacturer			
manufacturer A	0.3896	0.5206	0.4542
manufacturer B	0	-	-
area of operation			
long distance transport	-0.0707	0.6301	0.9107
local and regional transport	0	-	-

Table 11 Parameter estimates of the conditional logistic regression model for the accident involvement of heavy goods vehicles (matched cohort design)

As usual, adjusted odds ratios can be calculated using the respective parameter estimates. One obtains the result that the accident involvement risk of vehicles equipped with ESP, ACC and LGS is 0.41-fold the risk of unequipped vehicles:

$$\psi_{m;mod} = \exp[-0.8841] / \exp[0] = 0.413$$

Thus, the adjusted matched odds ratio resulting from the conditional logistic regression model is by far lower than the "crude" matched odds ratio ($\psi_m = 0.6418$; see above). This is due to the inclusion of the variable vehicle mileage in the model. The probability of accident involvement in the study period tends to increase with the number of kilometres driven in this period. Since on average the equipped vehicles show a higher mileage, the actual safety effect of the three driver assistance systems considered here is partially lessened (or "hidden") in the crude analysis. By controlling for the confounder "vehicle mileage" in a multivariate analysis (here: conditional logistic regression) the "true" effect of the safety systems comes to light.

2.3.3.3. Matched case-control design

Finally, the application of a matched case-control analysis shall be demonstrated (1:1 matching). The analyses are again based on data from the German FAS project. In order to generate an appropriate data set, controls (accident-free vehicles) were matched to cases (accident-involved vehicles) by company affiliation, i.e. the two members of each pair are vehicles belonging to the same company. For each case the corresponding control was chosen randomly from the subset of the company's accident-free vehicles.

As a result, the above matching procedure yielded 151 pairs of vehicles[13]. The table below shows the vehicle fitment status of the pair members:

Case (accident-involved vehicle) is equipped with ACC, ESP and LGS?	Control (accident-free vehicle) is equipped with ACC, ESP and LGS?		Total number of pairs
	yes	no	
yes	35	35	70
no	55	26	81
Total number of pairs	90	61	151

Table 12 Case-control pairs of heavy goods vehicles by vehicle fitment status of the pair members

Considering again only discordant pairs (n=35+55=90) the matched odds ratio calculated from these data amounts to

$$\psi_m = 35/55 = 0.636,$$

which is very close to the value resulting from the matched cohort analysis (0.642). McNemar's test of symmetry indicates a statistically significant result (p-value 0.035).

If one applies the conditional logistic regression model which has been used for the matched cohort analysis to the data arranged according to the case-control design, the following results are obtained:

Predictor variable	df	Wald Chi²	Significance
vehicle fitment	1	6.5354	0.0106
mileage (in 100 000 km)	1	2.3538	0.1250
year of manufacture	1	0.4869	0.4853
truck manufacturer	1	0.8043	0.3698
area of operation	1	0.0850	0.7706

Table 13 Conditional logistic regression model for the accident involvement of heavy goods vehicles (matched case-control design)

In contrast to the matched cohort analysis, the factor vehicle mileage is not significant in the case-control model. The parameter estimates are shown in Table

[13] Although 171 accident-involved vehicles are contained in the original FAS data set it was only possible to build 151 case-control pairs. Among other things this is due to the fact that several companies had provided only a single (equipped) vehicle for participation in the study.

14. As for the matched cohort analysis the conditional logistic regression analysis models the probability of “accident involvement = yes”.

Parameter	Estimate	Standard Error	Significance
vehicle fitment			
equipped	-0.9519	0.3724	0.0106
unequipped	0	-	-
mileage (in 100.000 km)	0.3461	0.2265	0.1250
year of manufacture	0.0901	0.1292	0.4853
truck manufacturer			
manufacturer A	0.5265	0.5871	0.3698
manufacturer B	0	-	-
area of operation			
long distance transport	0.2236	0.7667	0.7706
local and regional transport	0	-	-

Table 14 Parameter estimates of the conditional logistic regression model for the accident involvement of heavy goods vehicles (matched case-control design)

Using the respective parameter estimates it appears that for vehicles equipped with ESP, ACC and LGS the risk of accident involvement is 0.39-fold the risk of unequipped vehicles,

$$\psi_{m;mod} = \exp[-0.9519] / \exp[0] = 0.386.$$

Hence, regarding the efficiency of the three vehicle safety systems considered here the adjusted odds ratio under the case-control design is very close to the value resulting from the matched cohort design (0.413).

2.4. Note on the Evaluation of Infrastructure Measures

2.4.1. Research design

One important component for the evaluation of infrastructure measures (like e.g. dynamic traffic management systems or local danger warnings) is the selection of research design. Basically, one distinguishes between experimental and non-experimental designs (for example, retrospective observational studies). For a sound evaluation of infrastructure measures it is recommended to apply quasi-experimental designs. Basically this means that an area or region where the measure is implemented (test area) is compared to a similar area but without the measure (control area). Hereby, one can distinguish between the following basic variants (see, for example, Cook and Campbell 1979):

- with/without comparison (1 test or area as well as 1 control area at one point in time)
- pre-/post-comparison with control group(s) (1 test area before and after the introduction of the measure; 1 or more control area(s) at the same points in time, but without introduction of the measure)

- pre-/post - with/without comparison (1 test area before and after the introduction of the measure; 1 control area at the same points in time; measure is implemented in the pre-time frame and will be withdrawn in the post-time frame)

In addition to the above-mentioned basic variants, many other designs are possible (for example, time series analysis with and without control group; investigations with multiple and repetitious application of the measure). As a minimal requirement of an evaluation of infrastructure measures one should consider a pre-/post-investigation with control group. For methodological reasons, however, a pre-/post - with/without design would be preferable, because here, sources of danger to the validity of the results can be better controlled than in a pre-/post-comparison with control group. Furthermore, in a pre-post - with/without design it is possible to gain information about the permanence of a measure, i.e. whether the effects of a measure survive over time or disappear after its removal. The disadvantage of this procedure lies in the fact that it requires a longer period of investigation because the measure must already be introduced for a sufficiently long time in the pre-time frame of the control area.

2.4.2. Selection of areas

Against the background of quasi-experimental research designs it is obvious that an important criterion for the proper evaluation of infrastructure measures is the selection of appropriate test and control areas (or test and control sections), and linked to this, consideration of the accident volume (problem of required sample size).

The basic criterion for the selection of the test and control area is its comparability with respect to traffic, population and accident structure. If the measure refers to a single section or segments of a section (for example, in the case of automatic speed surveillance on a particular road), one also has to pay attention to comparability with regard to traffic and road type of the test and control sections. Test and control areas should be clearly separated from each other in their geographic dimensions because the measures could also affect areas bordering on the test area. Numbers of accidents in the selected areas[14] should also not differ greatly.

In selecting a test area or test section, one has to pay attention that no areas with remarkably high numbers of accidents are chosen. If one takes, for example, different segments of a section into consideration, one can compute the mean value of accident frequency. Individual segments will deviate more or less from the mean. The more the number of accidents exceeds the mean of the pre-time interval, the larger the probability is that it will carry a smaller value in the post-time interval under otherwise similar conditions. From this follows that in areas with an especially high frequency of accidents in the pre-time interval, on average there will be a reduced number of accidents in the post-time interval, even in the absence of any measures (the so-called “regression-to-the-mean-effect”). A quantitative evaluation of the effectiveness of a measure in the test areas that were chosen because of their significant number of accidents leads, therefore, to biased results (see Hauer 1980).

[14] Strictly speaking and according to the principles of experimental design, one should require that the decision as to which of the two areas will serve as the test area and which as the control area should be made randomly (so-called randomisation).

In reference to the time intervals of the investigation (as well as the size of the areas), they should be oriented with respect to the number of accidents that can be expected. Granted, road safety measures often allow only for relatively small effects on accidents.

2.4.3. Evaluation criteria

Absolute accident numbers (frequency count), degree of accident risks (for example, accidents per 1 million vehicle kilometres) as well as ratio and proportion numbers can be applied as evaluation criteria. With respect to the evaluation of time- and space-limited measures, it seems sufficient to apply the absolute accident frequency, but this also depends on the underlying research design. Thus, for a simple with/ without comparison, for example, only the ratio or proportion numbers can be reasonably interpreted. These can also be derived from the absolute numbers. Basically, it is advisable to only take into consideration accidents with personal injuries because in case of accidents without personal injuries, the number of unreported cases can falsify the results.

2.4.4. Data analysis

If one uses a pre-/post-design with control group for the evaluation of an infrastructure measure, the data structure results as represented in a two-by-two table:

	pre	**post**	**Total**
test area	n_{11}	n_{12}	$n_{1\cdot}$
control area	n_{21}	n_{22}	$n_{2\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	n

The case presented here shows absolute frequencies in the cells of the table; for example, number of accidents that are labelled n_{ij} (i,j=1,2). Change of given indicators of effectiveness over time can be directly read from the table. The question whether possibly noticed differences are random or can be linked to the measure can, as usually, be answered by testing levels of significance (null hypothesis: no connection between the area and the time period of investigation).

In the case of absolute accident frequency, the likelihood-ratio-test is an appropriate and relatively simple procedure for the evaluation of significance, requiring no special statistics-software. The logic of the test consists of a comparison of the actual, empirically collected accident frequency in the four cells of the table with the expected theoretical frequency under the assumption of independence of the two criteria (area and time of investigation). In the case of independence, expected theoretical cell values are computed from the empirical marginal frequencies of the table. The test statistic t of this test looks like:

$$t = 2\sum_i \sum_j n_{ij}\left(\ln n_{ij} - \ln\left[n_{i.} n_{.j} / n\right]\right)$$

Here, **ln** is the symbol for “natural logarithm”. The t-value is compared to a chi-square distribution on 1 degree of freedom. As usual, the level of significance must be determined before the data analysis.

3. EVALUATION IN TERMS OF SOCIO-ECONOMIC BENEFITS

After the description of potential impacts of safety systems with respect to costs, benefits, and factors influencing the effects in different countries the assessment of the impact of safety systems on the number and severity of accidents is presented in this chapter. Subsequently, a short overview on Efficiency Assessment Tools (Cost Benefit Analysis, Cost Effectiveness Analysis) is given. The main focus of this chapter is on the application of a Cost Benefit Analysis and the provision of standard values for accident costs.

3.1. The Context of Socio-economic Analysis

In order to optimize the improvement of road safety, it is necessary to dedicate public money to it. It is also necessary to choose the best and less expensive programme or system which could reach the predefined targets. In this connection the question arises which methods could be applied to identify the needs and to assess the effectiveness.

One could answer that the best system is the system which can avoid the biggest number of accidents or the biggest number of fatalities. But one can also say that the best system is the one which costs less money to the society. Literature shows that for avoiding accidents and save money we should consider both approaches. This common approach takes into account several components such as the system specification, the penetration rate of the system (consumption), the related stakes (accidents, injuries, production lost, harm consequences, pollution, time, traffic, …), and the costs of the impacts (direct, indirect or side) to identify the potential benefit we can obtain by applying the “best of the best” system.

Obviously, comparison of programmes or systems requires common indicators related to safety, impacts, and costs and thus common methods. The common methods involve an evaluation of the effectiveness of a programme or a system, the related benefits, and the cost savings. In the literature a great number of socio-economic methods can be found. In this chapter it is tried to give the state of the art on socio-economic assessment of road safety measures.

A socio-economic analysis is a decision-support tool. The basis of the analysis is the evaluation of the costs and the economic and social benefits related to the application of a program (project, system) for the society. This evaluation is then compared to the situation where the program is not applied. The analysis leads to the direct and indirect impacts and to the distribution of these impacts on the different actors of society. Socio-economic analysis allows making statements about social return of an investment. An overview of social costs and benefits can serve as basis for prioritizing separate measures or measure packages.

However, not all road safety measures lend themselves equally well to socio-economic analysis. According to the Handbook of Road Safety Measures (Elvik & Vaa 2004) general purpose policy instruments (e.g. motor vehicle taxation, regulation of commercial transport, urban and regional planning, access to medical services) don’t lend themselves very well to socio-economic analysis. On the other hand, measures referring to

- Road design and road furniture
- Road maintenance

- Traffic control
- Vehicle design and personal protection
- Vehicle inspection
- Public education and information campaigns
- Police enforcement

are suited for such an analysis. For example, in the eIMPACT project the following Intelligent Vehicle Safety Systems (IVSS) have been recommended for in-depth socio-economic assessment (eIMPACT 2008):

- Electronic Stability Control, ESC
- Full Speed Range ACC, FSR
- Emergency Braking, EBR
- Pre-Crash Protection of Vulnerable Road Users, PCV
- Lane Change Assistant (Warning), LCA
- Lane Keeping Support, LKS
- NightVisionWarn, NIW
- Driver Drowsiness Monitoring and Warning, DDM
- eCall (one-way communication), ECA
- Intersection Safety, INS
- Wireless Local Danger Warning, WLD
- SpeedAlert, SPE

3.2. Potential Impacts of Safety Systems

3.2.1. Types of impacts

According to the literature there are different ways to illustrate the relationship between road safety and the application of safety systems. The SEiSS project (Abele et al. 2005) proposed a generic framework with common steps and elements:

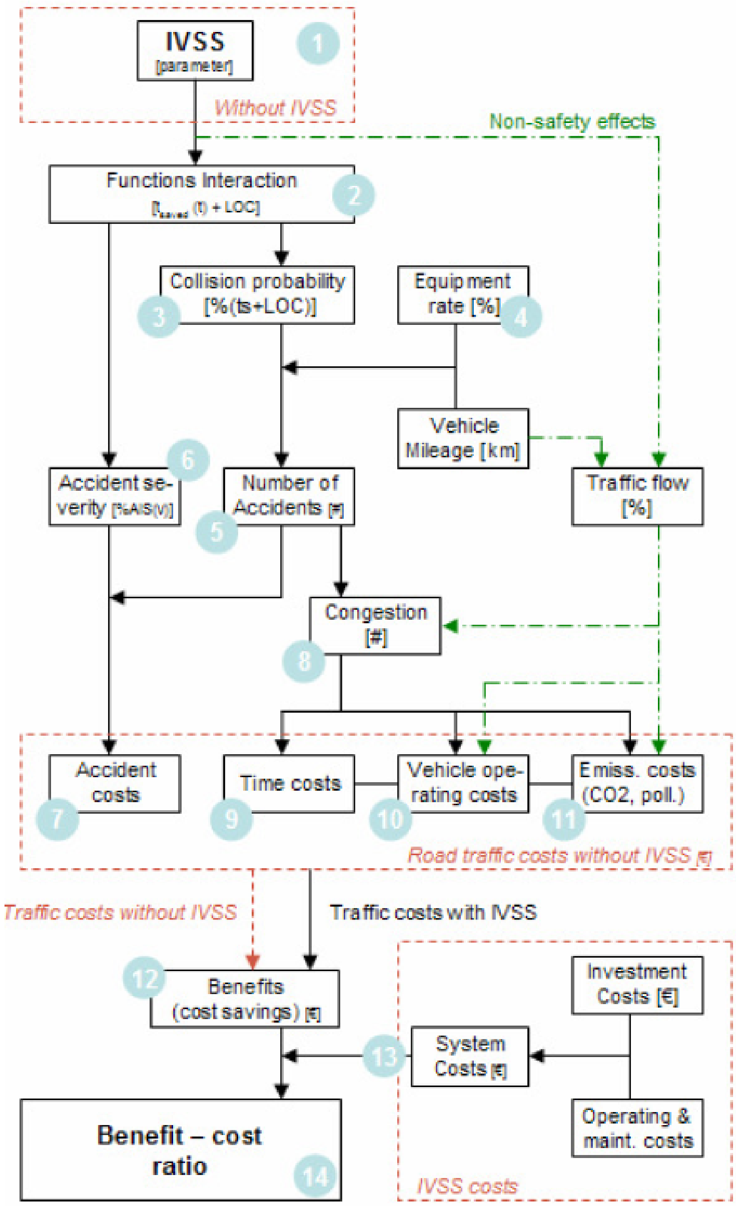

Figure 2 Synopsis of the impacts of IVSS on accident occurrence and severity (Source: Abele et al. 2005, p 3)

This synopsis describes the impacts of IVSS on accident occurrence and severity and consequently on direct (accident costs) or indirect impacts (time, pollution…), costs related to the absence or the presence of IVSS and consequently the benefits of the application of IVSS.

According to the ARCOS project (Action de Recherche pour une COnduite Sécurisée, 2004) the socio-economic costs can be classified into three categories*:*

- direct market value
- indirect market value
- indirect non-market value.

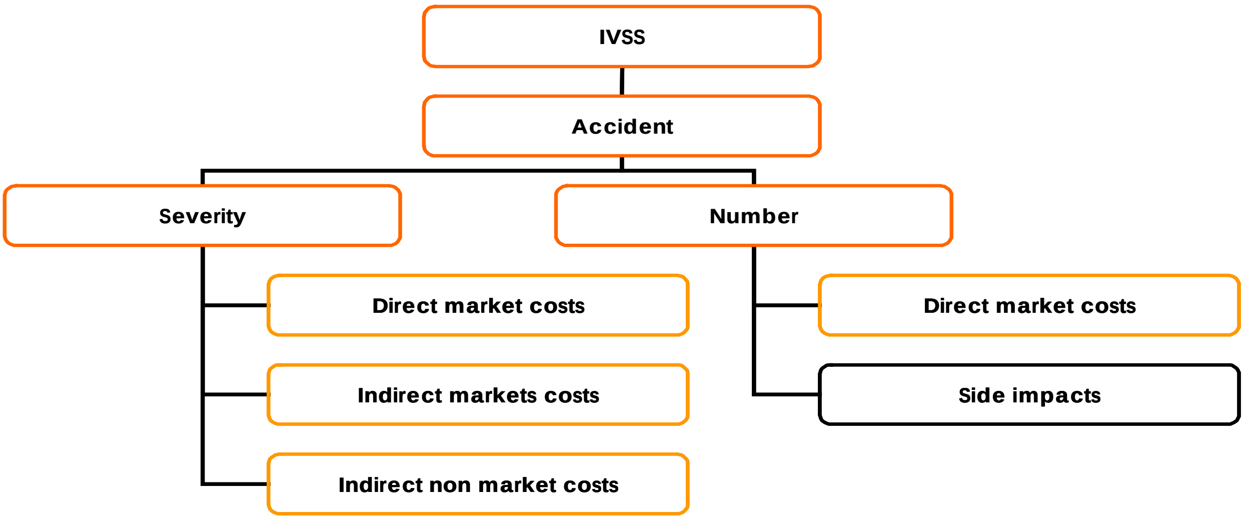

Figure 3 Classification of accident costs

Direct market value concerns:

- Medical and social costs:
    - Sanitary transport
    - First emergency
    - Medical care
    - Medicine
    - Convalescence
    - Funeral
    - Re-education
    - Rehabilitation
    - Home care
    - …
- Property costs:
    - Vehicles
    - Street furniture
    - Property
    - Objects
    - Travel, repair related to the accidents
    - Environment
    - Fuel consumption related to the congestion of the traffic.
    - …
- General costs:
    - Fire services
    - Police
    - Expertise
    - Justice
    - Insurances
    - …

Indirect market costs concerns:

- Loss of future production of the fatality, the injured or the jailed.
- Loss of potential production of the descendants, of the unemployed.
- …

Indirect non-market value concerns:

- In fatality case:
    - Harm
    - Pretium mortis (estimation, cost, value of the death)
    - Transfer of the pretium mortis to the heir.
    - …

- In case of injuries:
    - Pretium doloris (estimation, price, value of the pain, value of the sorrow)
    - Esthetical harm
    - Pleasure harm
    - Sexual harm
    - Third party harm
    - …

Other authors (e.g. Elvik 2001) tend to include side impacts as relevant in a socio-economic analysis. Elvik noted that although a large number of possible impacts of road investment projects can be valued in monetary terms, there is still a substantial number of impacts that are not included in socio-economic analyses. Inclusion of these impacts could make a major difference for the results of a socio-economic analysis.

Besides the direct and indirect impacts of a road safety measure, one can identify other marginal impacts that could be taken into account in a socio-economic analysis such as:

- Infrastructure (impact of infrastructure measures on the number and the severity of accidents according to the level of investment discriminated (ARCOS))
- Adaptation behaviour (risk compensation, risk homeostasis)
- Travel demand
- Other effects.

With respect to **adaptation behaviour**, it can be said that measures in general might change the habits of the road users in a negative (and unintended) way. If people perceive relevant changes in their environment they adapt their behaviour to meet the new challenges or to benefit from new chances. They may try to act more cautious if changes are perceived as having dangerous impacts or they may try to capitalize on new possibilities to meet own aims more effectively (Schlag, 2008). Thus, countermeasures against risk, e.g. driver support systems to reduce risk when driving, may be counterbalanced by behavioural adaptations using the improvement to act more risky, e.g. to drive faster. Whether the net outcome is positive or negative depends on the amount of unintended behavioural adaptations.

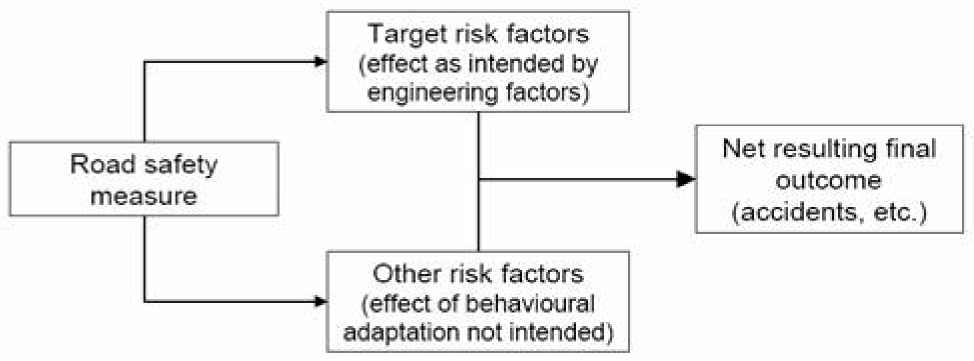

Figure 4 Behavioural adaptation: Net resulting outcome in safety
(Source: Schlag 2008)

One could argue that behavioural adaptation implicates that sole engineering measures cannot fully exploit the potential reduction of accidents. In fact there are publications supporting this assumption. Early evidence was found for technical measures such as Antilock Braking System, ABS and also for educational measures as slippery road training courses, more recently for Adaptive Cruise Control, ACC, and other in-car systems reducing workload and at least partially leading to adaptations such as speeding or fulfilling secondary tasks like using mobile phones more often when driving.

In conclusion, OECD expert group related that "The potential for behavioural adaptation affecting a safety measure should be considered in estimating the costs and benefits of safety programs. Programmes with minimal adaptation may be more effective, in the long run than those which produce large initial safety gains, but also produce adaptations that eliminate the gain" (OECD 1990).

Concerning the impact of the measures on **travel demand**, the amount of travel depends on the generalised costs of travel. The generalised costs of travel are subjective and will vary from one individual to another. One of the problems of using the generalised costs of travel to estimate the benefits of induced travel is the fact that some of the costs that go into the generalised costs of travel may be misperceived by road users. For example, many car drivers tend to consider only fuel costs when they estimate vehicle operating costs. But vehicle operating costs include several other items, of which depreciation (the decline in the value of the car) is not the least important.

Other effects: Most road safety measures have an effect on other policy objectives in addition to safety. This means that socio-economic analysis ought to include these effects in addition to effects on safety. The effects that are likely to be most difficult to include in socio-economic analysis at the current state of knowledge are:

All components of air pollution: There seems to be a tendency for the estimated costs of air pollution to become higher as more types of pollution and more effects of it are included in the estimates. The costs of air pollution are relevant for socio-economic analysis of a number of important road safety measures. All measures that affect speed, for example, also affect air pollution, because vehicle emissions depend on speed. All measures that require the additional use of energy, like daytime running lights, may increase vehicle pollution emission.

The visual intrusion caused by road systems: The visual intrusion of a road system (or its "ugliness") is a factor for which hardly any estimates of the costs to society can be found. Road user insecurity: It is obvious that change in the level of road user insecurity is an important impact of very many road safety measures. Still, too little is known about these impacts to include them in socio-economic analyses.

3.2.2. Variability of the impacts

The effects of safety measures can vary from one country to another. Although evaluation research aims for general knowledge the effects of a measure on safety are likely to vary from one place to another, depending on, for example, design features of the measure, road user behaviour, and a lot of other factors.

Differences in road accident costs are attributable to the differences in market prices and differences in individual preferences for safety. Preferences for road safety are manifested in the amounts people are willing to pay in order to improve road safety. Since road safety is a normal good, willingness to pay for it depends on income. The higher your income, the more you are likely to be willing to pay to have road safety improved. The efficiency criterion of welfare economics implies that the provision of road safety should match exactly the demand for it. At least officially, no country differentiates the provision of road safety based on individual income. If an average value of safety is used for countries with different levels of cost, safety may be overprovided in low-cost (low-income) countries and underprovided in high-cost (high-income) countries. Averaging the value of safety across countries only makes sense if the values being averaged include the same elements and have been estimated the same way in all countries.

In the context of implementation of measures, especially with respect to previously introduced measures, the effects of a certain road safety measure on the number of accidents may depend on whether it is introduced as a stand-alone measure or as part of a package of measures (PROMISING project). The effects may also depend on whether it is introduced into an environment where few measures have been implemented or into an environment where a lot of measures have been implemented. In general, the effect of measure **i**, denoted by $\mathbf{E_i}$, on the number of accidents, is expressed as a percentage change (in most cases a percentage reduction). Then $\mathbf{R_i}$ is the residual number of accidents still expected to occur when measure i has been implemented:

$$\mathbf{R_i = 1 - E_i}$$

If, for example, a measure affects 100 accidents and has a 20% effect on those accidents, the residual when the measure has been implemented is 0.8 (1 – 0.2). A simple model to estimate the combined effect of two measures on the number of accidents, when one of them reduces the number of accidents by 20% and the other by 30%, is:

$$\text{Combined effect} = 1 - (0.8 \times 0.7) = 1 - 0.56 = 0.44$$

That is, the combined effect of the measures is an accident reduction of 44%, not 50%, as the sum of their individual effects would seem to imply. This simple model of estimating the combined effects of several safety measures *that affect the same target accidents* assumes that the percentage effects of the measures remains unaffected when the measures are combined in a package.

An example of this is the proposal of the European Commission to mandate Automated Emergency Braking Systems (AEBS). The systems have an automated braking function, the benefits of which can be predicted using existing accident data. However, it is anticipated that the car manufacturers will include functions such as Adaptive Cruise Control, Forward Collision Warning and pre-impact adaptive restraint systems, which will not be mandatory. These types of systems are already fitted to some vehicles and will be fitted to more vehicles than AEBS. Therefore the fleet penetration of such systems will be ahead of AEBS, thus reducing the benefits of the AEBS function itself; a factor not accounted for in the benefits study undertaken. There are also potential effects on completely separate systems such as antiwhiplash seats, because AEBS will influence the frequency/severity of rear impacts (Robinson & Knight 2009).

3.3. Evaluating the Impacts of Safety Systems on the Number and Severity of Accidents

3.3.1. Evaluation concepts for IVSS

Primary safety measures are designed to help avoid accidents or, if this is not possible, to stabilize respectively reduce the dynamics of the vehicle to such an extent that the secondary safety measures are able to act as good as possible. The efficiency of a primary safety measure is a criterion for the effectiveness, with which a system of primary safety succeeds in avoiding or mitigating the severity of accidents within its range of operation and in interaction with the driver and the vehicle.

The aim of safety evaluation in road safety is to investigate the impact of advanced safety functions on reducing several types of accidents or mitigating accident consequences. The evaluation can be performed from two different perspectives:

- Assessment of the potential proportion of accidents that could be avoided and of the potential proportion of accidents whose severity could be reduced for safety functions that are not yet on the market (so-called ***a priori effectiveness***).
- Assessment of the actual proportion of accidents that could be avoided and of the actual proportion of accidents whose severity could be reduced for safety functions that are already in the market (so-called ***a posteriori effectiveness***) once the vehicles are equipped with existing functions.

After a system is introduced, it takes several years for it to penetrate the market. Only then it is possible to gain information on its efficiency based on real world accident statistics. Many of these systems take more than a decade of years to achieve a sufficient penetration rate. For the optimization of the development process it is therefore essential to have statistically reliable predictions for the expected efficiency available continuously from the choice of a promising idea for the design of a new safety measure to the starting point of its development and throughout the whole process.

So it becomes possible

- to focus on those primary safety measures that address most efficiently relevant accidents and conflict situations resulting from human errors,
- to configure an efficient set of optimal balanced sensors, actuators and algorithms,
- to optimize the efficiency of the function by preliminary design using simulation methods,
- to obtain reliable information that the customer can expect from the system as benefit.

Efficiency analysis or evaluation is the key technology to achieve such an improved development process.

The eIMPACT project ("Socio-economic Impact Assessment of Stand-alone and Co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe") assessed the socio-economic effects of Intelligent Vehicle Safety Systems (IVSS) and their impact on traffic and safety. It addressed policy options and the views and roles of the different stakeholders involved.

The impact assessment dealt with:

- The system specifications (functional descriptions of how, where and when the systems have an effect and interactions) were also the basis for the traffic and safety impact analyses.
- Collision probability
- The estimation of the penetration rates for passenger cars, goods vehicles and buses based on the specification of the selected systems.
- Prediction for number of accidents for specific IVSS set-ups.
- Prediction for accident severity for specific IVSS set-ups.

3.3.2. System specifications - technology and functions interaction matrix (IVSS)

The starting point for a socio-economic analysis is a clear and common definition of the technologies and functions of the IVSS (specifications) to be evaluated. In this context IVSS is generally described as a technology that has a direct influence on safety. Therefore IVSS should be interpreted as functions. However, different functions may influence the same safety problem, making it necessary to define the areas of interaction comprehensively. The availability of technologies and functions of IVSS enables to predict market introduction from a technical point of view. Functions have to be described in parameters to define their effectiveness. Due to performance differences between systems from different suppliers, the definitions have to be based on average parameters (generic systems). In addition, it is assumed that a function keeps constant over time in its effectiveness but decreases in price.

The safety functions of IVSS may depend on each other or affect one another with respect to safety mechanism, effectiveness and accordingly safety impacts. These possible function interactions have to be taken into consideration and analysed to ensure that there is no interference between the systems. The assessment of functions interaction is based on the time correlation approach for IVSS which reverts to the physics of accidents. This approach subdivides an accident into various time phases. The phases are:

- Prior to driving (planning and preparation of a trip),
- Driving (support of the driver by the vehicle in normal vehicle operation),
- Warning (the vehicle technology expects a dangerous situation, the driver is informed),
- Assistance (support of the driver by vehicle systems),
- Pre-crash (time directly before an unavoidable crash),
- Crash (with passive safety systems in operation),
- Post crash (after the crash).

Each phase includes a different level of danger in terms of collision probability as well as a different support of the driver by a particular vehicle system.

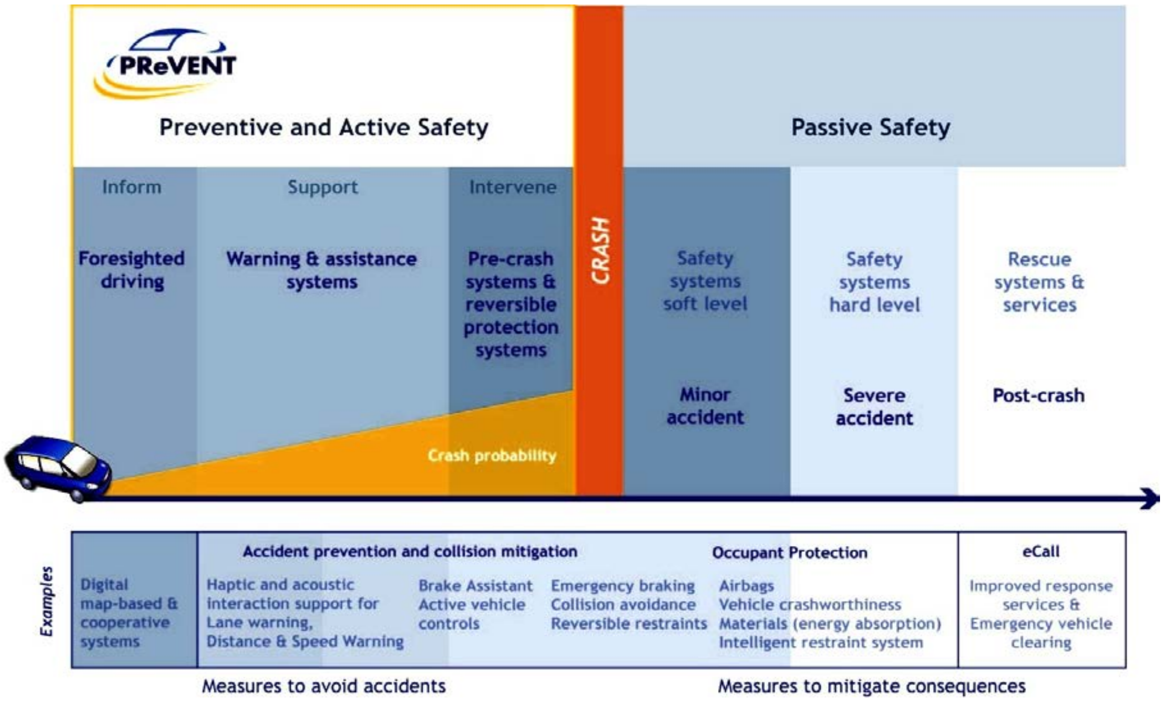

Figure 5 Accident Phases and General IVSS Functions (PReVENT 2006)

As illustrated in Figure 5, IVSSs operate in different phases of an accident.

For the evaluation of IVSS, it is most important to eliminate any possible interactions between the different systems that could arise over time due to technological dependencies. Hence, separate evaluation results for each safety system are ensured. However, since technologically reasonable combinations of IVSS exist, a system package may also be assessed.

3.3.3. Collision probability for IVSS

The performance of IVSS leads to an earlier warning information to the driver, better car stability, faster breaking or fewer driving faults. Hence, IVSS can be represented by time gains or contrary, time losses. In some cases it is rather difficult to correlate IVSS to time. This remains the case for ABS (ensuring steer ability while braking), ESP/ESC (preventing loss of traction through braking of specific wheels) and Safe (Adaptive) Speed (which limits the maximum speed to the physical limits of the vehicle combined with the specific road conditions). However, taking the physics of accidents into account, these overall time savings or losses will lead to a change in the number and severity of accidents related to the IVSS. Since each IVSS can moreover be related to different accident types (e.g. loss of control) correlation tables can be determined for different accident types and speeds for each safety system. These tables can be used to calculate the specific accident probability for each IVSS. Time gains achievable by the use of an IVSS will lead to a reduction in collision probability, no time savings translates into normal accident patterns.

Such correlations should become the result of accident causation analyses providing a standardised basis for further calculation.

3.3.4. Equipment rate for IVSS

The main goal of integrating the market perspective into the proposed model is to find a way of forecasting the penetration of IVSS within the vehicle fleet of the countries considered expressed as equipment rate. Within the socio-economic impact assessment the equipment rate has a threefold influence on the evaluation:

- The first is that vehicles which are equipped with IVSS as well as vehicles or other road users which are involved in crashes with those vehicles benefit from the advantages of the crash avoidance or crash mitigation effects of the IVSS. Therefore, only those vehicles which are equipped with IVSS influence the overall socio-economic impact.
- Secondly, some IVSS may need a certain equipment rate to fully exploit their potential benefits. For instance, co-operative systems need a minimum number of equipped cars for the technology to function correctly (“critical mass“).
- Thirdly, market penetration determines the total costs of IVSS which are confronted with the benefits generated. Thus, in total, a thorough analysis and estimation of the penetration of IVSS within the vehicle fleet is performed.

3.3.5. Prediction of accident numbers and accident severity

Using the collision probability of an IVSS and its market penetration rate, and taking into account the vehicles´ driven mileage mostly available based on age of the vehicles, the corresponding number of accidents can be predicted. In this step the underlying collision probability is linked to general accident data. The target figure is the number of accidents that can be avoided by using IVSS.

Besides the avoidance of accidents, IVSS largely influence accident severity for those accidents which are remaining. This assessment differentiates between fatalities and accidents with severely, slightly and uninjured persons. The more time for driver or vehicle reaction the IVSS provides, the lower the impact severity will be. Since the severity of an accident depends on the impact energy which directly corresponds to the impact speed, the time-related pattern can be used to calculate the accident severity. In addition, the vehicle’s energy absorption potential which is determined by the passive safety system installed addresses the accident severity. The absorption potential can be translated into additional time for the specific accident type resulting in accident mitigation. Analogous to the prediction of the collision probability, collision severity figures can be determined by accident causation analysis. These can be used to predict each IVSS´ potential for accident mitigation (eIMPACT).

Literature is so extensive concerning the methodologies applied for evaluating the effectiveness of IVSS. While the common target of these methodologies is to evaluate the impact of IVSS on the road safety, different approaches are however available. The general framework below summarizes the differences found in the scientific papers dealing with the subject.

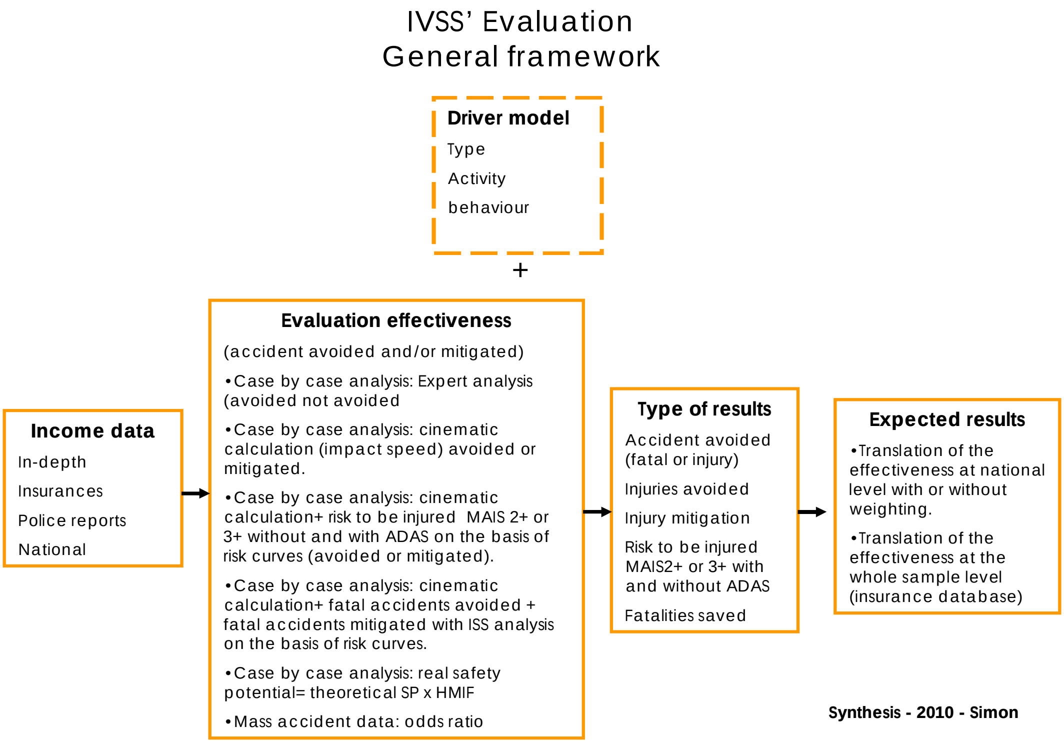

Figure 6 Methods for Predicting/ Evaluating the Impacts of IVSS on Accidents

This framework doesn't mean that all steps are performed by all evaluators. Some authors include the driver model (the driver type, activity, behaviour) before evaluating with case by case analysis the effectiveness of the IVSS[15] concerned and other authors assume that there is no adaptation, interaction of the driver on the evaluation. The following methodologies can be found in the literature:

- Vision safety for everybody – Situation interpretation - Multi object scenarios
- Driver model (Matlab environment)
- Safety potential (real/theoretic)
- Harm method
- Risk probability
- Odds ratio
- Simulation method
- Artificial Neural Networks

Since it is not possible to describe all these methods in detail the risk probability-approach (LAB) - for the example AEBS (Advanced Emergency Braking System) - shall be sketched very briefly[16]: This method is based on the knowledge of the correlation between crash severity and the probability to be injured (MAIS 2+ and 3+) in the accident. The relationship is illustrated by a risk curve built according to the type of impact and other parameters depending on the system evaluated (seat belt, location in the vehicle, EuroNCAP score…) and calculated with a logistic regression. Data come from the LAB in-depth database.

---

[15] In Figure 6 referred to as ADAS (Advanced Driver Assistance Systems).
[16] For the approach of Volkswagen AG see DaCoTA Deliverable 5.7.

As regards the type of IVSS and the type of relevant accidents, a type of impact is identified and becomes the basis of the analysis. Sometimes criteria related to the type of user, the place in the vehicle, and the seat belt are also included. Impact speeds are calculated on the basis of real accidents (available sample of relevant accidents according to the specifications of the IVSS). The calculation is re-run with the IVSS fitted the vehicle. New severity of crash (EES, delta-V) is derived by logistic regression and leads to an estimation of a new probability to be injured in the accident. Weighting methods correct the distribution of the severity criteria in the sample to meet the distribution of the representative database.

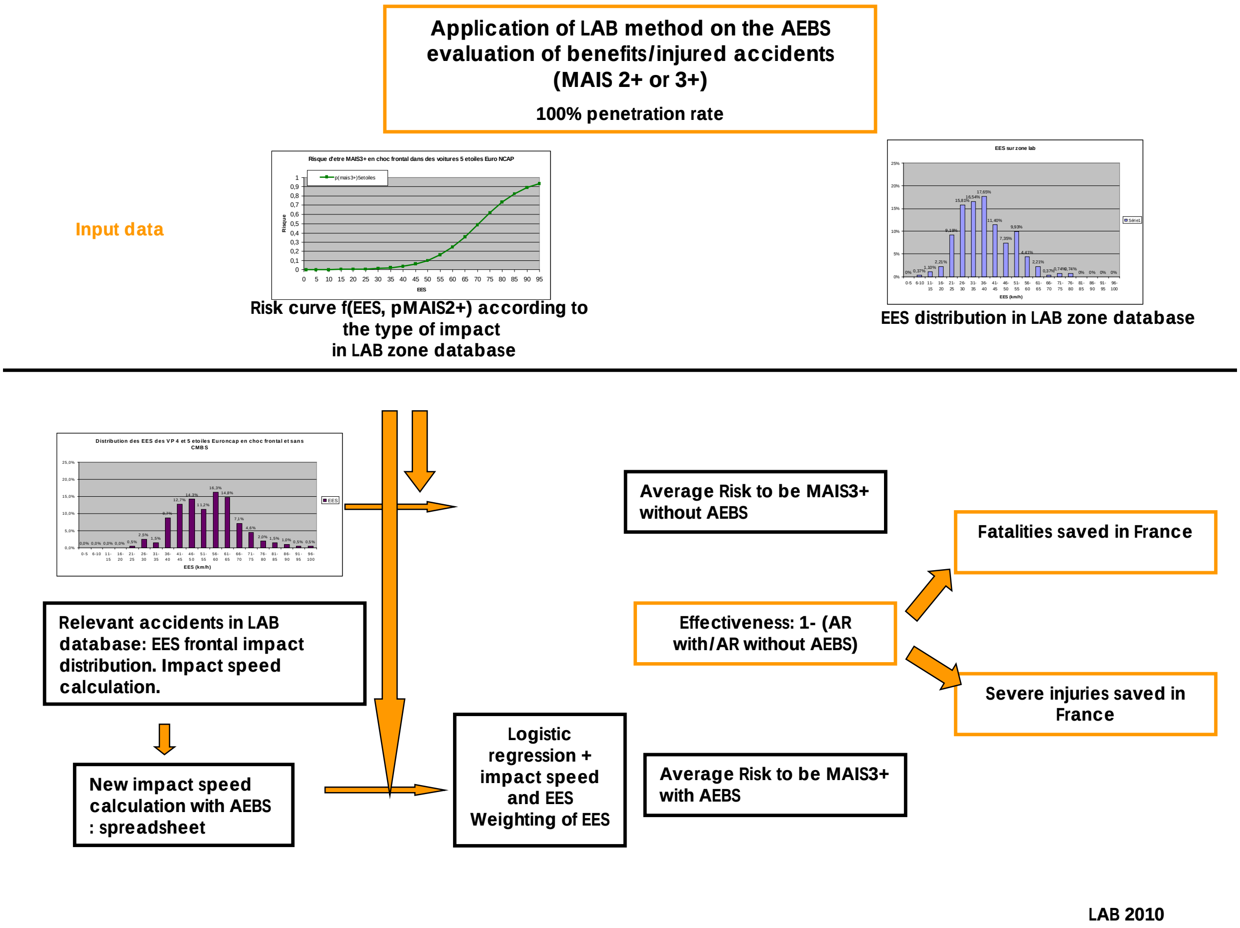

Figure 7 Risk Probability-approach (LAB, France)

Thus, in this case effectiveness is measured by the probability to be injured with or without the IVSS (here: AEBS) fitted.

3.4. Evaluating Costs and Benefits of Safety Systems

Efficiency Assessment Tools or EAT (Cost Benefit Analysis and Cost Effectiveness) are more and more applied by policy decision makers because on the one hand the safety budget and the country resources are limited and on the other hand the decision makers require rational approaches to reach the target “saving life with the best measure" and to allocate the resources to the most beneficial use. In general, a socio-economic analysis aims at

- providing data to stakeholders to prioritize the actions,
- identifying the systems, programmes available and evaluate their effectiveness according to the social impact, the road safety impact, the environment impact, and the cost impact.

Socio-economic analysis rests on the assumption that all economically relevant impacts of a project are valued in monetary terms according to the principles of welfare economics (Hanley and Spash, 1993).

3.4.1. Types of socio-economic evaluation approaches

Three approaches to perform the evaluation of a socio economic impact of road safety measures can be distinguished:

- CBA (cost benefit analysis)
- CEA (cost effectiveness analysis)
- MCA (multi-criteria analysis)

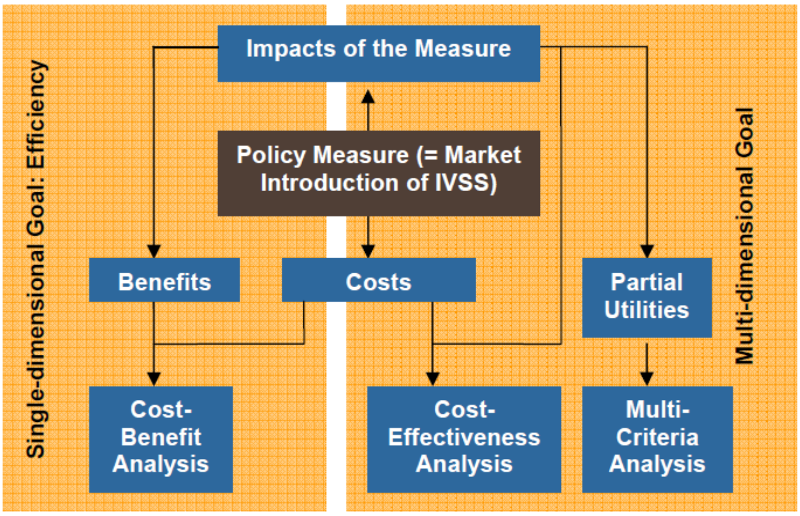

Figure 8 Types of socio-economic evaluation approaches (eImpact, 2007)

3.4.1.1. Cost-Benefit Analysis (CBA)

Cost-Benefit Analysis represents the traditional and most prominent methodology for determining the worth, value and feasibility of a policy measure. The main reason for doing Cost-Benefit Analysis of road safety measures is to help develop policies that make the most efficient use of resources, i.e., that produce the largest possible benefits at given costs.

CBA is based on welfare economics. Its benchmark is represented by the Kaldor-Hicks criterion: A policy measure is efficient when it makes some people better off without making other people worse off (this implies that winners can potentially compensate losers from their gain). In other words, the underlying question of CBA is whether it is profitable to the society to use productive resources (e.g. labour, capital) to achieve savings of resource consumption (e.g. savings of travel time, energy, casualties and environmental pollution).

Cost-benefit analysis, CBA, is a formal analysis of the impacts of a measure or programme, designed to assess whether the advantages (benefits) of the measure or programme are greater than its disadvantages (costs).

Both sides – the resource use (= costs) and the resource savings (= benefits) – are expressed in monetary terms and can be confronted to each other. The CBA results are expressed in terms of Benefit Cost Ratio, BCR, which allows a comparison between several systems or programmes.

Two measures of efficiency are used in CBA:

- **The Net Present Value (NPV) =**
  present value of all benefits - present value of all costs implementation costs of a measure

- **Benefit Cost Ratio (BCR) =**
  present value of all benefits / present value of implementation costs

The numerator of the Benefit Cost ratio, related to accidents prevented, is estimated as following:
- Number of accidents prevented (or expected to be prevented) by a measure = number of accidents expected to occur per year x safety effect of the measure

According to the eImpact project, resulting values of BCR can be interpreted as follows:
  - 0 < BCR < 1: “poor” (socio-economic inefficiency)
  - 1 < BCR < 3: “acceptable”
  - BCR> 3: “excellent”.

When the Net Present Value is positive, the BCR exceeds the value of 1.0.

An example of CBA application is the BCR of reducing 24 right-through opposing direction crashes (DCA 202) at an intersection by installing traffic signals which would be calculated as following:

BCR = Economic benefit of reducing 24 DCA 202 crashes / (costs of installing signal hardware and software adjustments + costs of additional ongoing maintenance to the signal - over the life of the facility)

The benefits can be determined by estimating the likely number of crashes prevented multiplied by the crash cost.

All in all, Cost-Benefit Analysis considers all relevant policy impacts and enables a direct comparison of costs and benefits. However, monetary valuation e.g. of human life is controversial and difficult (but inevitable). Moreover, not all effects can be assessed (e.g. human grief and suffering, distributional effects/ fairness) and information concerning accident costs and the effects of the safety measure on the mobility is needed.

One of the biggest problems in CBA is to obtain valid and reliable monetary valuations of all relevant impacts. It is therefore relevant to carry out a Cost Effectiveness Analysis in addition to a CBA.

CBA's are particularly useful if:

- multiple policy objectives exist (safety, environment, mobility)
- policy objectives are conflicting (e.g. safety vs. environment)
- CBA is necessary if different levels of injury severity are to be considered
- objectives refer to goods without market prices (safety, environment).

3.4.1.2. Cost Effectiveness Analysis (CEA)

"Cost Effectiveness Analysis" is a variant of the Cost Benefit Analysis which is based on two complementary goals:

- To identify the more efficient safety action for the same money.
- To identify the less expensive safety action for equivalent efficiency.

In contrast to CBA, the different effectiveness indicators are not transformed into monetary terms. This represents a substantial shortcoming of the cost-effectiveness analysis. Typically the CEA is expressed in terms of a ratio where the denominator is a gain in health from a measure and the numerator is the cost associated with the health gain. In the case of safety system evaluation the number of accidents prevented forms the denominator of the Cost-Effectiveness Ratio of a safety measure, consistent with the idea that one wants to minimise the cost-effectiveness ratio. Thus, the Cost-Effectiveness Ratio can be defined as:

Cost-Effectiveness Ratio = costs of the measure / number of accidents prevented

The accidents that are affected by a safety measure will be referred to as target accidents. In order to estimate the number of accidents prevented per unit implemented of a safety measure, it is necessary to:

- Identify target accidents (which may, in the case of general measures like speed limits, include all accidents),
- estimate the number of target accidents expected to occur per year for a typical unit of implementation,
- estimate the percentage effect of the safety measure on target accidents.

However, CEA examines only one effect of a measure on road safety and is expressed in a unit other than money such as the number of casualties.

3.4.1.3. Multi-Criteria Analysis (MCA)

"Multi-criteria Analysis" aims to identify the preferences of a safety action as regards to the whole objectives (difficulty to provide the best indicators to reach the objectives). The Multi-criteria Analysis uses several criteria of evaluation. A weighting is attached at each criterion according to the decision-maker preferences.

Both appraisal methods, Cost-Benefit Analysis and Multi-Criteria Analysis, represent evaluation techniques which are appropriate for assessing the socio-economic impact of IVSS. Cost-Effectiveness Analysis is inferior to both methods since it stops at the level of effectiveness without subsequent appraisal and without aggregating the different effectiveness contributions of a measure.

3.4.2. Socio-economic valuation methods

When Cost-Benefit Analysis of transport projects started in the 1960s, the only impacts that were included in the first analyses were travel time, vehicle operating costs and accidents. The benefits of preventing accidents were normally valued according to the so called “human capital” approach, which assigned a value to preventing a fatality or an injury proportional to the value of production lost. This approach is still used in some countries and is included as part of the valuation in countries that have adopted the Willingness-To-Pay (WTP) approach for the valuation of road safety (*SAFETYNET).*

There have been several reviews regarding societal costs of road traffic injuries. A major review was presented in 1994 by the European Commission: “Socio-economic cost of road accidents, final report of Action COST-313”[17]. A more recent survey was made as part of the ROSEBUD-project[18], where cost estimates for selected countries were presented.

From 1989 to 1993, several countries have cooperated with the Action COST-313 in order to analyse and evaluate the differences in the various calculation methods to assess the range of elements of cost with a view to making recommendations about the categories of costs that should be taken into account and how they should be measured. As far as methods for estimating costs are concerned, the typology shown in Figure 9 was developed in COST-313:

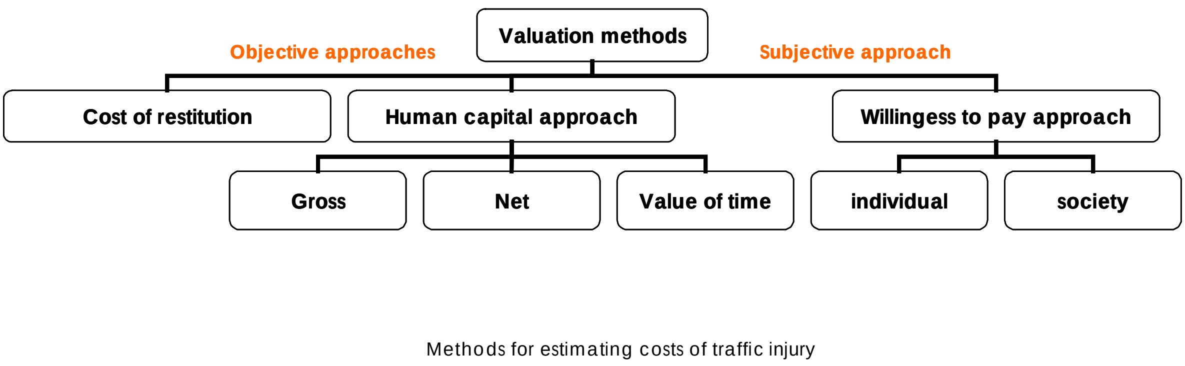

Figure 9 Methods for estimating costs of traffic injury (COST-313)

[17] COST, Socio-economic Cost of Road Accidents is an intergovernmental framework for European Cooperation in Science and Technology allowing the coordination of nationally-funded research on a European level.
[18] ROSEBUD (Road safety and environment Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision-Making). ROSEBUD is a thematic network funded by the European Commission to support users at all levels of government (European Union, national, regional, local) with road safety related efficiency assessment solutions for the widest possible range of measures (bringing together users, researchers, decision makers, policy makers and other stakeholders around efficiency assessment).

3.5. Application of Cost-Benefit Analysis

3.5.1. The steps of a CBA

This section is dealing with the practical application of a Cost-Benefit Evaluation, taking into account

- the nature of the criteria required to carry out a CBA
- the different steps to be performed.

In general, to be amenable to Cost-Benefit Analysis, a road safety measure should satisfy the following criteria:

- It should be known what category of accidents the measure affects (all accidents, accidents involving young drivers, accidents in the dark, etc.), preferably so that the number of "target" accidents can be estimated numerically. If possible, these estimates should state the severity of accidents or injuries they apply to.
- It should be possible to describe the use of the measure in numerical terms, e.g. number of junctions converted, number of cars equipped, number of drivers trained, man hours of police enforcement, etc. This information is needed in order to estimate marginal costs and benefits of the measure.
- Other impacts of the measure should be known, for example impacts on speed or the environment.
- Costs of the measure (implementation) should be known, and it should be known who pays the cost. This is because private expenditures and public expenditures are not treated identically in Cost-Benefit Analyses. An opportunity cost[19] of taxation is added to public expenditures, but not to private expenditures.
- Monetary valuations should be available for all impacts of the measure.

According to the SAFETYNET project the main steps of a CBA are as follows:

1. Develop relevant measures or programmes intended to help reduce a certain social problem (e.g. road accidents or environmental pollution).
2. Develop alternative policy options for the use of each measure or programme which can lead to compare alternatives (project alternative/null alternative). Comparison of project and null alternative over a long period, depending on the effectiveness duration of the measures. This involves weighing the effects during a period of years by which those effects that occur later weigh less than those that occur earlier (SWOV 2011).
3. Describe a reference scenario (sometimes referred to as business-as-usual or do-nothing alternative).
4. Identify relevant impacts of each measure or programme. There will usually be several relevant impacts.
5. Estimate the impacts of each measure or programme in "natural" units (physical terms) for each policy option.
    a. Impacts on safety
        i. Limiting material damage
        ii. Limiting medical costs
        iii. Limiting production lost
        iv. Limiting immaterial damage
        v. Limiting settlement costs

[19] Basic relationship between scarcity and choice. Cost related to the next-best choice available to someone who has picked among several mutually exclusive choices.

vi. Limiting traffic jam costs

b. Impacts on mobility

vii. Changes in travel time

viii. Changes in travel costs

c. Impacts on environment.

ix. Changes in emissions

x. Changes in noise nuisance (SWOV 2011)

6. Obtain estimates of the costs of each measure or programme for each policy option (investment costs, maintenance costs).
7. Convert estimated impacts to monetary terms, applying available valuations of these impacts (SWOV 2011). Impacts are expressed in terms of money[20].
   a. Using market prices (medical costs, business travel time)
   b. When there is no market price available, apply valuation methods such as Contingent Valuation Method (CVM) which uses previous surveys to determine a monetary value such as road safety's immaterial costs (UNITE recommendations)
   c. or Hedonic Pricing Method (HPM) which uses house prices, e.g., the price of noise nuisance is determined by the price difference between houses in areas with a lot and little noise nuisance.
8. Compare benefits and costs for each policy option for each measure or programme. Identify options in which benefits are greater than costs.
9. Conduct a sensitivity analysis[21] or a formal assessment of the uncertainty of estimated benefits and costs.
10. Recommend cost-effective policy options for implementation. Usual criteria are the balance and the ratio of the discounted benefits and costs.
11. Taking into account the uncertainties in the results of a CBA.

To identify relevant measures or programmes, a broad survey of potentially effective road safety measures should be conducted. A measure is regarded as potentially effective

- if it has been shown to improve road safety – and has not already been fully implemented or
- if there is reason to assume that it will improve road safety by favourably influencing risk factors that are known to contribute to accidents or injuries.

For each road safety measure, alternative options for its use should be considered. If, for example, the problem to be solved is bicyclist injuries, and the measure considered is bicycle helmets, alternative policy options could be:

- Do nothing; leave to each bicyclist to decide whether or not to wear a helmet.
- Conduct a campaign for bicycle helmets, while leaving their use voluntary.
- Make the use of bicycle helmets mandatory for children.
- Make the use of bicycle helmets mandatory for everybody.

[20] When there is no way to express the impacts in terms of money, impacts are included as a reminder item in the overview of costs and benefits (landscape impacts of a new infrastructure) (SWOV 2011).

[21] Investigates the robustness of a study: A Technique used to determine how different values of an independent variable will impact a particular dependent variable under a given set of assumptions.

Policy options in cost-benefit analysis are always compared to a reference scenario and represent changes from that scenario. Often the reference scenario will be “to do nothing”, i.e. not to introduce the road safety measure for which a cost-benefit analysis is performed. In some cases, however, one may foresee that a certain road safety measure will be introduced without any action from government. As an example, Electronic Stability Control is now rapidly becoming standard equipment on new cars and will spread in the car fleet during the next 10-15 years. In such cases, the foreseen rate of introduction should be regarded as the reference scenario.

The estimation of the impacts of the measures on safety will be based on a model that allows to account for

- the exposure (amount of travel) of each road user group,
- the risk of injury,
- the impact of each measure on exposure and risk of each road user group.

The basic model for estimating the number of injuries that can be prevented by each safety measure is:

Number of prevented injuries = Exposure x Risk x Effect of measure

This expression gives the expected number of injuries that can be affected by a measure.

The impacts of each measure will, to the extent available data make it possible, be partitioned into the following contributions:

Impacts on safety =
Change in exposure x Change in accident rate[22] x Change in injury severity

This equation allows the net change in the number of injuries to be broken down into contributions from changes in exposure, changes in the number of accidents per kilometre of travel (accident risk), and changes in the severity of injuries. A distinction can thus be made between measures aiming to reduce the number of accidents and measures aiming to reduce injury severity.

In general, the relevant measure of exposure for road related measures and police enforcement is vehicle kilometres of travel. For vehicle related and road user related measures, the relevant measure of exposure in most cases is person kilometres of travel.

A similar model can be applied to estimate the impact of a measure on the environment. Consider, for example, the impact of a measure on the emission of air pollution:

Impact on pollution = Change in exposure x Change in specific emission rate

With respect to environmental impacts, the term “exposure” denotes the population exposure to a certain concentration of ambient pollution. This, in turn, depends on the amount of traffic and the dispersion of emissions from traffic to people.

Population exposure = Traffic volume x Dispersion of emissions in time and space

[22] Accident rate is defined as the number of accidents (all levels of severity) per million (vehicle or person) kilometre of travel: Accident rate =Number of accidents/ Million kilometres of travel

There is adequate knowledge about how noise generated by traffic spreads to the surroundings. Less is known about how various forms of pollution spread. Simplifying assumptions will therefore be made with respect to population exposure for pollution.

Concerning monetisation of the impacts, the value of life differs according to the country and the year of observation. The table below summarises the different VSL (Value of Statistical Life) estimates according to the author, decision maker and country:

Author	Year	Number or kind of studies	VSL
Dionne et al.	2004	28	>$3.5 million (2000 value)
De Blaeij et al.	2003	Europe + US	Average $4.4 million (1997 value)
De Blaeij Koetse Tseng Rietveld	2004	Official VSL in 7 countries	€ 1.4 to € 2.6 million (value 2000)
HEATCO and UNITE	2001-2006		€1.5 million (value 1998)
Miller	2000	68 studies in 13 countries scientifically	$2.7 million (Europe) $2.2 million (North America) $0.9 million for 49 countries (regression model) €3.0 million (value 2001)
ECMT	1998	Scientifically	€2.4+/-€1 million (value 1990)
ECMT	1998	Official (conservative approach) Average of 5 European countries	€1.5 million (value 1998) €1.6 million (value 2001)

Table 15 VSL estimates according to the author, the decision maker and the countries

These data show the differences between countries and the disparity according to the survey, project and thus with the different concerns and targets. However, the common remark is that the VSL estimates are higher in 2004 than 1998 and that the VSL is higher when a scientific approach is performed rather than a conservative approach (Human Capital Approach, often used by policy makers).

3.5.2. Standard values for accident costs

3.5.2.1. Comparability of different cost estimates

When comparing cost estimates between countries the following factors should be considered (*Victoria Transport Policy Institute 2010):*

- The purpose of the analysis, and therefore its perspective, such as whether it considers only short run marginal costs, long-run costs, and or total social costs.
- Categories of impacts considered, including vehicle costs, travel time costs, roadway costs, traffic services, parking costs, congestion impacts on other road users, delays to non-motorized travellers, accident costs, pollution emissions and other environmental impacts.

- Data sources and methodologies used to calculate costs, particularly non-market costs such as the costs of accident injuries and deaths, and environmental damages.
- How possible double-counting is addressed, such as whether taxes are counted as costs or economic transfers, and whether congestion costs are summed with travel time costs.
- Geographic scope, and the monetary exchange rates used if in different countries.
- The time period evaluated, what index is used for inflation.
- Driving conditions, such as whether the costs represent urban-peak, total urban, rural or overall average driving conditions.
- Differences in measurement units, such as between miles and kilometres, and between vehicle miles and passenger miles.
- The types of vehicles considered, such as whether cost estimates are for cars, automobiles, the fleet of personal vehicles, total roadway vehicles (including freight vehicles) or total motor vehicles (including train, air and marine vehicles)
- Whether cost estimates are point values or ranges.

3.5.2.2. Attempt of standardization

In 2004 the estimated annual costs, both direct and indirect, of traffic injury in the EU-15 countries exceeded 180 billion euros. At European Union level, the most frequently used “magic number” to put a value on the prevention of casualties was the “1 Million euro rule”[23].

The 1 Million euro value is frequently used as a test of the effectiveness of traffic safety measures and implies that a measure can be considered for implementation when for every million euros spent for a road safety measure at least one death is prevented. This amount takes into account the economic damage of a death, and also a certain proportion of the damage resulting from (serious) injuries and from accidents with only property damage, because, on average, for every prevented death there will also be a number of accidents with injuries and an even greater number of accidents with only property damage. This “magic number” estimation has not been updated since 1997 and is apparently still applied for EU-25.

- Harmonized data, cost-unit rate per casualty in EU-25:
  - accident with fatalities 1M€/fatality
  - accident with severe injuries 135.000€/severe injury
  - accident with slight injuries 15.000€/slight injury
  - property damage accident with fat. 12.000€/crash
  - property damage accident with injuries 3.500€/crash

But, there are legitimate sources of variation that could influence the results of CBAs with regard to the country:

- Varying economic valuation of road safety between countries (different incomes)
- Varying levels of cost of implementing a certain safety measure (according to the level of specification)
- Varying visions (of the road safety)

[23] This was introduced by the European Commission in its 17. Road Safety Programme 1997-2001 to help select traffic safety measures. Promoting road safety in the EU: The Programme for 1997-2001, Commission of the European Communities 1997

- Varying methods to reach the defined target
- Varying adjustment factors for incomplete accident and injury reporting
- Varying economic valuation of impacts other than safety of measures
- Varying discount rates used to convert future costs and benefits to present value (according to the currency and the reference years)
- Varying impacts of safety measures (according to the country, the target population)
- Varying exposure data (according to the country).

The costs of road accidents were first estimated in the 1950's in Great Britain and the United States. Today, all the highly motorised countries try to estimate these costs, but the cost items included and the methods used in estimating them differ between countries. A detailed survey of the methods used in 20 motorised countries to estimate road accident costs, has been made by Elvik (1995). The survey considered the valuation of fatalities only. Differences between countries with respect to the definition and level of reporting of non-fatal injuries make it difficult to compare the costs of non-fatal injuries. Three cost elements were identified (Figure 10):

- The valuation of lost quality of life (welfare)
- The costs of lost output
- Direct outlays

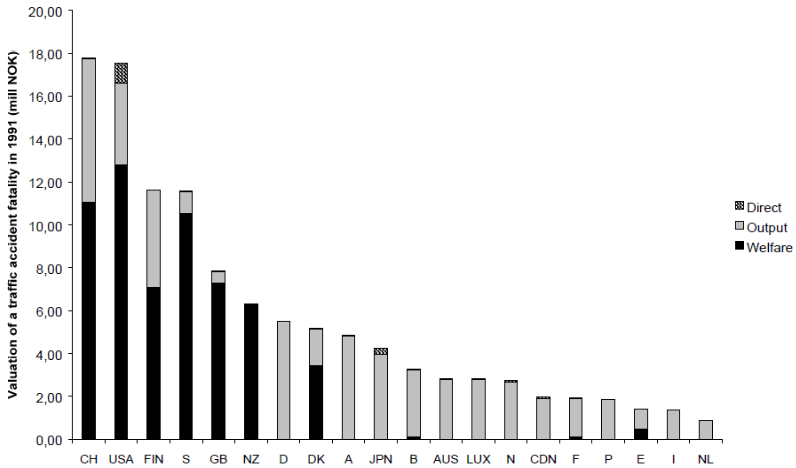

Figure 10 Official economic valuation of a traffic accident fatality in 20 motorised countries in 1991. *Million NOK per fatality (Elvik 1995)*

Not all countries estimated the costs of lost quality of life in 1991. In countries that did include this cost, it represented more than half of the total costs.

Elvik (2000) compared the costs of 8 European and 4 non-European countries that have an estimate of human costs. The data referred to one year in the 1988–1997 period. The (non-weighed) average share of human costs in the total costs was 44%, but ranged from 8% in Germany to 80% in New Zealand. The method used to estimate human costs probably explains a large part of the variation between countries.

The most recent recommendations for the monetary valuation of road safety are given in a report delivered by the HEATCO-project 2006 (Developing Harmonised European Approaches for Transport Costing and Project Assessment) within a set of monetary values for the prevention of traffic injuries:

Country	Fatality	Severe injury	Slight injury	Fatality	Severe injury	Slight injury
	(€2002, factor prices)			(€2002 PPP, factor prices)		
Austria	1,760,000	240,300	19,000	1,685,000	230,100	18,200
Belgium	1,639,000	249,000	16,000	1,603,000	243,200	15,700
Cyprus	704,000	92,900	6,800	798,000	105,500	7,700
Czech Republic	495,000	67,100	4,800	932,000	125,200	9,100
Denmark	2,200,000	272,300	21,300	1,672,000	206,900	16,200
Estonia	352,000	46,500	3,400	630,000	84,400	6,100
Finland	1,738,000	230,600	17,300	1,548,000	205,900	15,400
France	1,617,000	225,800	17,000	1,548,000	216,300	16,200
Germany	1,661,000	229,400	18,600	1,493,000	206,500	16,700
Greece	836,000	109,500	8,400	1,069,000	139,700	10,700
Hungary	440,000	59,000	4,300	808,000	108,400	7,900
Ireland	2,134,000	270,100	20,700	1,836,000	232,600	17,800
Italy	1,430,000	183,700	14,100	1,493,000	191,900	14,700
Latvia	275,000	36,700	2,700	534,000	72,300	5,200
Lithuania	275,000	38,000	2,700	575,000	78,500	5,700
Luxembourg	2,332,000	363,700	21,900	2,055,000	320,200	19,300
Malta	1,001,000	127,800	9,500	1,445,000	183,500	13,700
Netherlands	1,782,000	236,600	19,000	1,672,000	221,500	17,900
Norway	2,893,000	406,000	29,100	2,055,000	288,300	20,700
Poland	341,000	46,500	3,300	630,000	84,500	6,100
Portugal	803,000	107,400	7,400	1,055,000	141,000	9,700
Slovakia	308,000	42,100	3,000	699,000	96,400	6,900
Slovenia	759,000	99,000	7,300	1,028,000	133,500	9,800
Spain	1,122,000	138,900	10,500	1,302,000	161,800	12,200
Sweden	1,870,000	273,300	19,700	1,576,000	231,300	16,600
Switzerland	2,574,000	353,800	27,100	1,809,000	248,000	19,100
United Kingdom	1,815,000	235,100	18,600	1,617,000	208,900	16,600

PPP: purchasing Power. The PPP adjusted values account for differences in income and prices between countries.

Table 16 Recommended values of safety *(Source: Bickel et al. 2006) HEATCO project* (Developing Harmonised European Approaches for Transport Costing and Project Assessment)

In the second set of values denoted PPP (Purchasing Power), factor prices are adjusted according to the differences in purchasing power and are intended to be more directly comparable across countries (*HEATCO-report – SAFETYNET).*

The HEATCO report indicated that these values should be used if the country doesn't adopt the Willingness-To-Pay approach (WTP). Some countries which applied this WTP approach have higher values than spread official data (WTP approach results are not applied in the official monetary valuation of road safety (conservative interpretation due to the numerous sources of error associated with the WTP approach)). WTP approaches are usually performed in Belgium, Denmark, France, Great Britain, Greece, the Netherlands and Sweden.

With regard to the methodology applied, Human Capital Approach or Willingness-To-Pay approach, the respective valuation could result in large differences. In 2009, the European status report on road safety (WHO report 2009) compiled some countries' reported estimated costs per death and the related method in an attempt to express the costs of a human life in monetary terms:

Country	Costs per death in thousands of euros[a]	Method used	Year
Austria – including (excluding) human suffering	2676 (1399)	Willingness to pay	2006
Netherlands	2427	Gross output method; Willingness to pay	2003
United Kingdom	2137	Willingness to pay	2005
Finland	1752	Gross output method; Willingness to pay	2007
France	1194	Gross output method	2006
Germany	1162	Gross output method	2004
Hungary	769	Willingness to pay	2002
Estonia	767	Gross output method	2007
Cyprus	480	Gross output method	2008
Latvia	398	Gross output method	2006
Slovakia	287	Gross output method	2007

[a] Currencies were converted into euros using the appropriate mid-year exchange rate.

(Gross output method = human capital approach)

Table 17 Estimated economic costs of one death in selected countries in the WHO European Region (*European status report on Road safety*)

According to the method applied, the valuation of death costs could differ leading to the recommendation of using national or local data when available to avoid valuation bias and assessment bias.

3.5.2.3. Adjustment factors: Incomplete and inaccurate official road accident statistics

The percentage of injury accidents reported in official road accident statistics varies substantially between countries. In most countries, the level of reporting has been determined by comparing the number of injured road users treated in hospitals (including outpatients not staying in hospital overnight) to the number of injured road users recorded by the police. For some of the new member states of the European Union (Poland, Slovakia, the Baltic states), the level of accident reporting in official statistics is not or only partially known.

An update survey performed by ETSC in 2007 reported the following accident reporting rate in different European countries:

Country	Year of most recent study of accident reporting	Percent of injured road users reported in official statistics
Czech Republic	2005	66
Denmark	2004	21
France	2006	42
Germany	1992	39
Great Britain	1996	56
Hungary	2005	76
Netherlands	1990	43
Norway	1995	37
Spain	2000	18
Sweden	1987	54
Switzerland	1990	25
Mean of listed countries		41

Table 18 Level of accident reporting in 8 European countries (ETSC 2007)

That means that it would be incorrect to base analyses on an average reporting level for all European countries. If such an average were to be used, it would underestimate safety problems and the benefits of safety measures in countries with a low level of reporting (below average) and overestimate them in countries with a high level of reporting. The level of reporting for various categories of road users and types of accident in some European countries shows that the reporting is lowest for bicycle accidents, in particular those that do not involve other road users. Very few of these accidents are found in official accident statistics. Single vehicle accidents involving motorcycles also have a very low level of reporting (ETSC 2007).

A recommendation for adjusting incomplete accident reporting in official statistics is summarised below:

	Fatality	Serious injury	Slight injury	Average injury	Damage only
Average	1.02	1.50	3.00	2.25	6.00
Car	1.02	1.25	2.00	1.63	3.50
Motorbike/moped	1.02	1.55	3.20	2.38	6.50
Bicycle	1.02	2.75	8.00	5.38	18.50
Pedestrian	1.02	1.35	2.40	1.88	4.50

Table 19 Recommendations for European average correction factors for unreported road accidents *(Source: Bickel et al. 2006)*

The factors listed are multipliers, by which the officially recorded number of injured road users should be inflated. A small correction factor is applied to fatalities, due to the 30-day definition of a fatality which is more accurate than the record of injured whatever is the country. Whenever national correction factors are available, these should be used rather than the European average values.

3.5.2.4. Resume

Differences are shown (CBA effects) between countries in Europe and could be noted inconsistent due to the differences of incomes, of evaluation of the road safety effects according to the network, the traffic, the different impacts likely to the country (*SAFETYNET)*.

For example, in the Netherlands the Benefit-Cost Ratio of alcohol interlock systems was estimated to 4.1. A similar value (4.5) was found in Norway whereas for the Czech Republic the Benefit-Cost Ratio amounts to 1.6 and for Spain the costs even exceed the benefits (0.7; Vlakfeld et al. 2005). Thus, the findings of cost-benefit analyses of safety measures may be somewhat inconsistent.

Finally, the details of the data required to perform a CBA according to a Human Capital Approach or a Willingness-to-Pay approach are summarised in the following table:

	Cost of	Detailed costs	Availability of the data	Generic Method
Direct impacts or accident costs	Medical care	At scene care	Economics institutes (UNECE, EUROSTAT in Europe) Studies (Europe or national level)	HCA[24]
		Transport		
		In hospital stay		
		Out-patient treatment		
		Drugs		
		Prosthetics		
		Emergency organisation		
		Insurance taxes	National federation of insurances European insurance and reinsurance federation: CEA	
		Public Welfare incomes and outcomes	National welfare system databases (CLEISS in Europe)	
		Private health subscriptions and costs	National private health federation Economic and social council	
		Funeral costs	Surveys National federation of insurances European insurance and reinsurance federation: CEA	
	Disability care	Medical assistance costs	Economics institutes (UNECE, EUROSTAT in Europe) National welfare system databases (CLEISS in Europe) National private health federation	HCA
		Physiotherapy	Economics institutes (UNECE, EUROSTAT in Europe)	
		Psychology		
		Wheelchair-prosthetics		
		Insurance taxes	National federation of insurances European insurance and reinsurance federation: CEA	
		Public Welfare system	National welfare system databases (CLEISS in Europe)	
		Private health subscriptions	National private health federation WHO: European Observatory on Health systems and policies	
		Family support	Economics institutes (UNECE, EUROSTAT in Europe)	
		Home adaptation		
		Work adaptation		

[24] HCA: Human Capital Approach

	Cost of	Detailed costs	Availability of the data	Generic Method
Direct impacts or accident costs	Road, environment maintenance	At scene support	National road management databases	
		Maintenance organisation	Road management databases (ERF in Europe)	
		Public land damage	Road management databases (ERF in Europe) National federation of insurances European insurance and reinsurance federation (CEA)	
		Road Pollution	Road management databases Environment federation (EEA European Environment Agency)	
		Property prices	Market price Statistics and Economics institutes (UNECE, EUROSTAT in Europe)	
	Vehicle support	At scene support	National federation of insurances European insurance and reinsurance federation: CEA	
		Emergency organisation		
		Property reparation	National federation of insurances European insurance and reinsurance federation: CEA	
		Insurance taxes		
		Vehicle destruction		
		Vehicle recycling		
	Loss of production	Unemployment support: state/employer	Statistics and Economics institutes (EUROSTAT in Europe)	HCA
		Loss of incomes (state/family)	Surveys Statistics and Economics institutes (EUROSTAT in Europe)	
		Loss of consumption (family/society)		
	Decrease of production	Decrease of incomes		
		Work time adaptation (State/employer/public and private health insurance)	Surveys Public welfare databases Private health insurance databases Statistics and Economics institutes (EUROSTAT in Europe)	
		Decrease of consumption (family/society)	Surveys Indicators Statistics and Economics institutes (EUROSTAT in Europe)	

	Cost of	Detailed costs	Availability of the data	Generic Method
Indirect impacts	Research		National and European authorities stakeholders	WTP[25]
	Air pollution		EEA European Environment Agency	
	Noise		EEA European Environment Agency	
	Time (travel time)		EUROSTAT in Europe	
	Pain, grief, harm		Insurances Jurisprudence	
	Family pain		Jurisprudence	

Table 20 General data required to perform a CBA - direct and indirect impacts/costs (without implementation costs)

In order to make the costs and benefits comparable, a conversion of the values to a certain time reference is required. Such an action needs a definition of the economic frame, i.e. the duration of impact (length of service life of the project) and the interest rate, which are those commonly used for the performance of economic evaluations in the country.

Remind that the Benefit Cost ratio is defined as:

Benefit-cost ratio =
present value of all benefits/ present value of implementation costs

In a basic case, where the benefits come from the accidents saved only (and no influences on travel expenses and the environment are expected), the numerator of the benefit-cost ratio will be estimated as:

Present value of benefits =
number of accidents prevented by the measure * average accident cost * the accumulated discount factor,

where the accumulated discount factor depends on the interest rate and the length of life of the measure.

[25] WTP: Willingness-To-Pay

4. EXPANSION OF ACCIDENT DATA TO EU27

In this chapter a method for adjusting a multidimensional table of counts to some external marginal totals based on the so-called „iterative proportional fitting procedure" is presented. In this context the method is used to expand accident data from selected regions or countries to the EU27-level. After the description of the underlying statistical background an example for the expansion of the accident cause "failure to observe priority rules" is shown.

4.1. Introduction

Accident data are typically available at best on a national level and in most cases for a few countries only. Moreover, differences in variable definitions and the like preclude simple combination into one database.

Assuming that a team of experts has agreed on a grouping of countries into homogeneous classes (where of course the grouping may depend on the specific topic under investigation), a method has been developed for combining the available country-specific knowledge into an estimate of the overall European picture. The basic idea is visualized in Figure 11: Suppose the EU-27 countries have been grouped into four classes, with class representatives A, B, C, and D having a complete data table available. For other countries, only table totals (i.e. the bottom right corner of the table) or selected table margins (i.e. the row and/or column totals of the table) are available. Given this situation, a method was developed that allows to

- combine the information from all countries within a class of countries into an overall table for this class of countries
- combine the thus-created tables for classes of countries into an overall EU-27 table, allowing for the possibility that certain table margins are known on a EU-27 level.

The problem outlined above may be considered as a problem of adjusting (or raking) a multidimensional table of counts to some external marginal totals. The work therefore heavily draws on the so-called „iterative proportional fitting procedure" which was proposed by Deming and Stephan (1940). Like always, the algorithm can of course not generate information that is not available. However, the method allows to make use of the complete information available at European level; only the missing parts are taken from the national databases. As can be expected, the less information is actually available, the stronger are the assumptions that need to be made for conclusions to be viable.

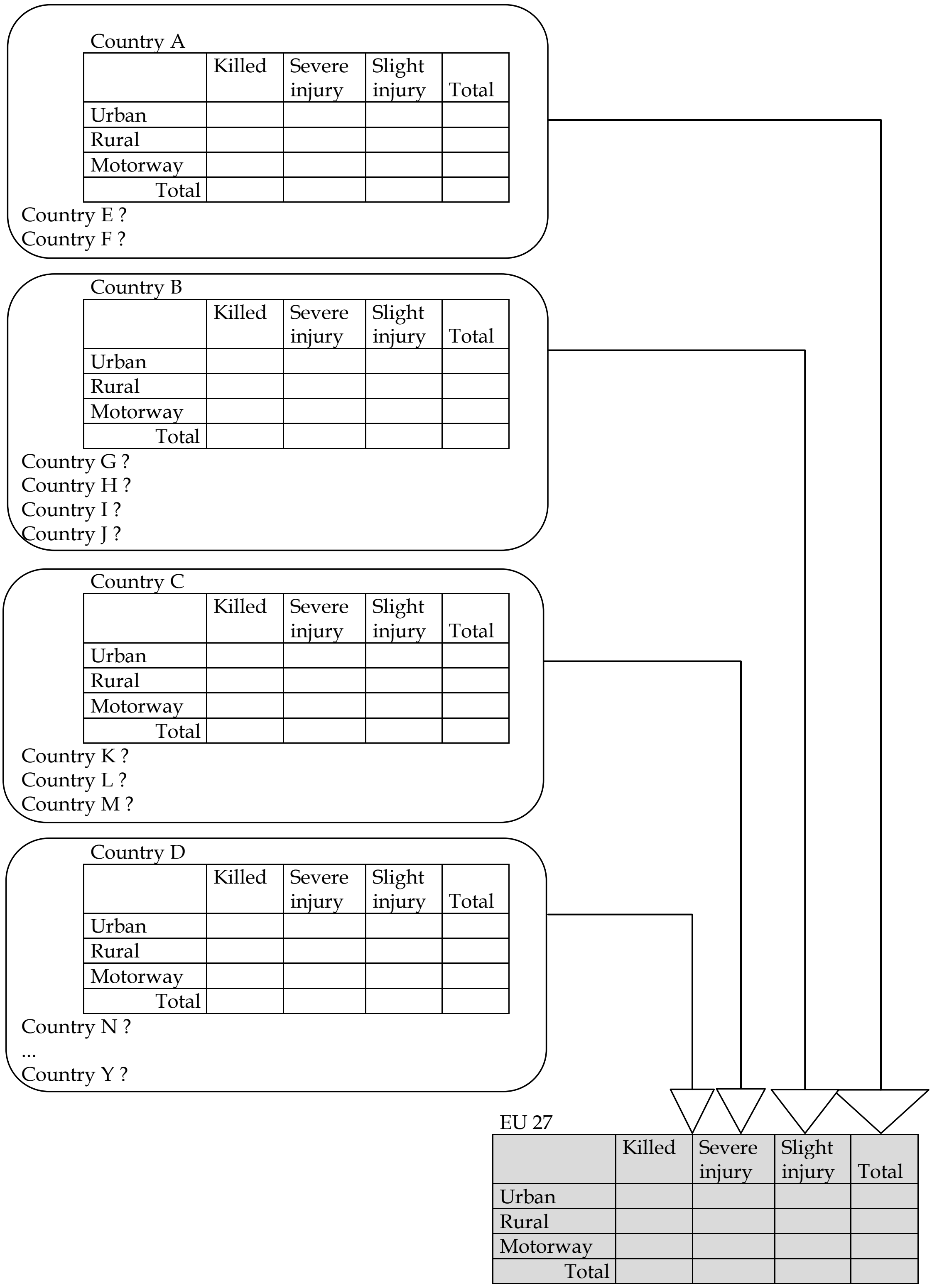

Figure 11 Framework for drawing inferences on the European situation

4.2. Expanding Data to EU-27 Level as a Statistical Adjustment Problem

Valid identification of e.g. accident causes requires in-depth traffic accident investigations or at least data from national road traffic accident statistics. As in-depth accident studies are rare and in no way cover complete countries or even EU-27, direct estimation of EU population totals of accident characteristics using standard procedures of sampling theory is not possible. Likewise, national traffic accident data cannot be expanded to EU level using standard techniques as country-specific datasets can simply not be considered as random samples from the population of accidents occurring in Europe as a whole.

If traffic accident data from selected regions or countries shall be expanded to the European level at least some basic auxiliary information on traffic accidents at EU-27 level is necessary for this purpose. Such information can be found in sources like ECE, IRTAD, ERSO and ETSC.

As already said before, the necessity exists to expand traffic accident data from a few countries (where the data is available) to EU-27 level under certain assumptions which, of course, should be sufficiently realistic and acceptable. In this situation an important research target is to create synthetic tables of accident and road user counts at EU-27 level by combining (1) data from regional in-depth accident studies or national traffic accident statistics with (2) some coarse structural accident and road user information available at the European level under an appropriate statistical model.

To illustrate the problem we consider the following example. A three-dimensional frequency table from an in-depth study or from national accident statistics ("initial table") displays the annual number of fatalities broken down by

- type of traffic participation,
- age group of road user and
- accident cause attributed to road user.

The task is now to estimate the entries of the corresponding three-dimensional table for EU-27 ("estimated table") as this information does not exist at the European level.

In doing so, the following two one-dimensional frequency tables for EU-27 available from IRTAD can be used as auxiliary ("external") information:

- EU-27 fatalities by type of traffic participation
- EU-27 fatalities by age group of road user

Naturally, the estimated three-dimensional EU-27 table should be consistent with the two external IRTAD tables. This means that

- aggregating the estimated three-dimensional EU-27 table over age group and accident cause should yield the IRTAD table on EU-27 fatalities by type of traffic participation and
- aggregating the estimated three-dimensional EU-27 table over type of traffic participation and accident cause should yield the IRTAD table on EU-27 fatalities by type of traffic participation.

From a statistical point of view we are faced with the problem of adjusting a multi-dimensional contingency table to satisfy some external information about the margins of the table. This problem can be solved using the so-called iterative proportional fitting procedure.

4.3. Adjusting a Table of Counts to Satisfy some Marginal Constraints

A multi-dimensional contingency table of observed count data serves as initial or starting table. Then, the iterative proportional fitting procedure (IPFP) is applied to adjust the starting table to certain one- or higher-dimensional marginal distributions which represent the external information. The adjusted table is an easily calculated solution to a table which satisfies the marginal constraints and preserves those main and interaction effects for which no external margins are available (see Fienberg and Meyer, 1982). Applied to our expansion problem, the adjusted table produced by the IPFP combines different data sources in a way that all information available at the European level is used and only the missing information is taken from the regional or national data bases.

To illustrate the algorithm we consider a three-way starting table $\mathbf{x} = \{x_{ijk}\}$ which is to be adjusted to three two-dimensional margins denoted by X_{ij+}, X_{i+k} and X_{+jk}.

The IPFP takes the initial table

(0) $m_{ijk}(0) = x_{ijk}$ for all i, j, k.

As the initial table is to be adjusted to three margins, the r-th iteration (r=1, 2, …) consists of three steps which form:

(1) $m_{ijk}(r|1) = m_{ijk}(r-1|3) \bullet X_{ij+}/m_{ij+}(r-1|3)$,

(2) $m_{ijk}(r|2) = m_{ijk}(r|1) \bullet X_{i+j}/m_{i+j}(r|1)$,

(3) $m_{ijk}(r|3) = m_{ijk}(r|2) \bullet X_{+ij}/m_{+ij}(r|2)$.

Steps 1, 2 and 3 are repeated until the change in the adjusted counts at the end of a cycle is sufficiently small. Clearly, after the final iteration also the adjusted and external margins are sufficiently close. For a detailed discussion of convergence and some other properties of the algorithm see, for instance, Bishop, Fienberg and Holland (1975).

The IPFP is not only a computational technique for adjusting tables of counts. Rather, the IPFP is also a commonly used algorithm for maximum likelihood estimation in log-linear models. The IPFP forms the core of several log-linear computer packages (see McCullagh and Nelder 1992, p. 183).

As an adjustment procedure the IPFP goes back to Deming and Stephan (1940, pp. 427-444). Therefore, the IFPF is sometimes also called Deming-Stephan algorithm. The application of the IPFP as a tool for expanding regional accident data to the European level is demonstrated in the next section.

4.4. Example: Accident Cause “failure to observe priority rules” in Europe

The specific traffic accident cause “failure to observe priority rules” is of considerable importance as many fatalities and substantial economic losses are resulting from this type of misconduct in road traffic. Therefore, in an empirical investigation on accident causation in Europe, a research team might be especially interested in accidents which had been mainly caused by the failure to observe priority rules (e.g. not observing the rule “right has priority over left” or traffic signs regulating the priority).

For brevity, this specific accident cause is termed “priority/precedence” in the sequel. Among other things, the following topics could be addressed:

- association between accident cause and injury severity
- association between accident location and accident cause

Obviously, this type of information is not available at the European level. In order to generate such information the research team might proceeded as follows:

Initial table

From the GIDAS files for the period 1999-2005 a 4-dimensional empirical starting table can be built displaying the number of casualties (n=13 064) broken down by

1. injury severity of road user (injured, killed),
2. accident location (within built-up area, outside built-up area),
3. lighting conditions (daylight, not daylight)
4. main accident cause (priority/precedence, other).

As each of the four variables has two outcomes, the GIDAS starting table contains 2 x 2 x 2 x 2 = 16 cells (“groups”):

Group	Severity	Lighting	Location	Cause	n of casualties
1	killed	daylight	built-up	priority	8
2	killed	daylight	built-up	other	46
3	killed	daylight	other	priority	6
4	killed	daylight	other	other	119
5	killed	other	built-up	priority	6
6	killed	other	built-up	other	59
7	killed	other	other	priority	2
8	killed	other	other	other	95
9	injured	daylight	built-up	priority	1’395
10	injured	daylight	built-up	other	4’935
11	injured	daylight	other	priority	234
12	injured	daylight	other	other	2’320
13	injured	other	built-up	priority	515
14	injured	other	built-up	other	2’014
15	injured	other	other	priority	70
16	injured	other	other	other	1’240

In total, 17.1% of the casualties recorded in the GIDAS files have been involved in accidents with main cause “priority/precedence”.

External margins

After some data preparation work the following two empirical 2-dimensional margins for EU-27 can be built from ECE statistics[26]:

- EU-27 casualties 2004 by injury severity and lighting conditions
- EU-27 casualties 2004 by injury severity and accident location

It appears that at EU-27 level the number of casualties 2004 amounts to N = 1 797 224. As can be seen, information is available at the European level for three of the four variables considered in the accident causation analysis. The following two 2x2-tables display this information:

- External margin 1: “Severity x Lighting”

Injury severity	Lighting conditions		Total
	daylight	other	
killed	25 433	20 721	46 154
injured	1 226 222	524 848	1 751 070
Total	1 251 655	545 569	1 797 224

- External margin 2: “Severity x Location”

Injury severity	Accident location		Total
	Built-up area	other	
killed	15 716	30 438	46 154
injured	1 068 329	682 741	1 751 070
Total	1 084 045	713 179	1 797 224

Estimated 4-dimensional EU-27 table

Using the IPFP, the estimated 4-dimensional table at EU-27 level can be calculated. This table contains 16 cells displaying the N = 1 797 224 casualties in Europe 2004 broken down by (1) injury severity of road user, (2) accident location, (3) lighting conditions, and (4) main accident cause (priority/precedence, other). As in our context the 4-dimensional table as such is of no specific interest, the table is not presented here.

From the estimated EU-27 table one can derive the result that in Europe as a whole the specific accident cause “priority/precedence” accounts for 15.8% of all casualties. The deviation of this value from the corresponding GIDAS figure (17.1%) can be

[26] Source: United Nations, Economic Commission for Europe: Statistics of Road Traffic Accidents in Europe and North America, Vol. LI, 2007

explained by structural differences between the GIDAS survey region and EU-27 as regards road traffic accidents.

Estimated 2-dimensional EU-27 tables

In order to answer the two research questions formulated above, the following two 2-dimensional EU-27 tables can be created by proper aggregation of cells:

- <u>Estimated EU-27 table A</u>: “Severity x Cause”

Injury severity	Main accident cause		Total
	priority / precedence	other	
killed	2 989	43 165	46 154
injured	280 922	1 470 148	1 751 070
Total	283 911	1 513 313	1 797 224

According to the expansion results, the specific accident cause “priority/precedence” in Europe as a whole (year 2004) accounts for 6.5% of all casualties among persons killed and for 16.0% of casualties among persons injured in road traffic accidents. Accidents mainly caused by failure to observe priority rules are less severe compared to accidents mainly caused by other forms of misconduct of road users.

- <u>Estimated EU-27 table B</u>: “Location x Cause”

Accident location	Main accident cause		Total
	priority / precedence	other	
built-up	228 876	855 169	1 084 045
other	55 035	658 144	713 179
Total	283 911	1 513 313	1 797 224

This table shows that at EU level the specific accident cause “priority/precedence” accounts for 21.1% of the casualties in built-up areas but only for 7.7% of the casualties outside built-up areas. Failure to observe priority rules as a specific accident cause is typical for built-up areas: 80.6% of the casualties due to this form of misconduct have been injured or killed within built-up areas (among the casualties due to other accident causes only 56.5% were in an accident within built-up areas).

The estimated 2-dimensional EU-27 tables presented above have been obtained from the 4-dimensional EU-27 table by appropriate aggregation over two of the four dimensions of the table. It is important to note that if this 4-dimensional table is analysed by a log-linear model for contingency tables, the following effects stem from the two external margins tables i.e. reflect the accident situation at the EU-27 level:

- grand mean (reflecting the average number of accidents per cell in the 4-dimensional table)
- main effects “severity”, “lighting” and “location”
- bivariate interaction effects “severity x lighting” and “severity x location”.

In contrast to this, the following effects are "borrowed" from the GIDAS starting table , i.e. reflect conditions and dependencies as to be found in the GIDAS study area:

- main effect "cause"
- bivariate interaction effects "severity x cause", "lighting x location", "lighting x cause", "location x cause"
- all higher interaction effects (which, however, are less important)

From this property of the estimated cell frequencies (accident counts at EU level) obtained using the IPFP it follows that one should be cautious when interpreting the estimation results. Differences between countries in usage of priority signs, traffic lights or roundabouts may, for instance, be a limiting factor for validity of the results that are just an expansion of German conditions with regard to all effects related to the characteristic "cause". This becomes clear in the following thought experiment: If outside Germany no priority rules in traffic were existent at all, road users in these countries had no opportunity for causing an accident by "failure to observe priority rules". Despite this, the IPFP would yield an estimated table indicating that in EU-27 the accident cause "failure to observe priority rules" is roughly of the same importance as in Germany.

4.5. IPFP Software

Simple applications - e.g. adjusting a given 2-dimensional initial table to one or two external (1-dimensional) margins - can be performed by using MS-Excel.

More powerful statistical software is needed for adjusting higher-dimensional tables to arbitrary external margins. The adjustment described in the preceding Section has been performed by means of SAS software (IML). Recently, a SAS macro RAKING has been introduced which offers additional diagnostic features (information on speed of convergence).

Summarising, the approach presented and illustrated above can be characterised as follows:

- The proposed method enables researchers to estimate figures relevant for traffic safety analyses at EU level under clearly defined assumptions. The method can even be used for expanding results on the (a priori) evaluation of vehicle safety systems coming from a simulation tool, for instance. Here, for each case it can be determined whether or not the presence of the system would have avoided or mitigated the accident (suppose, the variable "cause" in the example above is replaced by "accident prevented – yes/no"). Thus, the distributions (regarding e.g. injury severity, lighting conditions, etc.) of both the affected and unaffected accidents are known and can be expanded to a wider accident population.
- The proposed method uses the complete information available at EU level. Only the missing parts are taken from national or regional data bases.

5. CONCLUSION

The rapid growth of intelligent systems fitted to vehicles and the road infrastructure has raised the need to systematically evaluate the impact on safety and to give guidance on the most valuable functionalities of these systems.

The safety benefits of systems can either be assessed on the basis of real-world accident data using epidemiological approaches or by a priori evaluation methods based on simulation tools or case-by-case analyses. The application of epidemiological methods necessitates that the system under investigation is on the market long enough to exert an influence visible in real-world accidents. Only then it is possible to gain information on its efficiency based on accident statistics. Many of these systems, however, take more than a decade to achieve a sufficient penetration rate. As a rule it is not possible to wait e.g. 10 years until the assessment of a system is feasible. Thus, the application of simulation tools can be a helpful instrument. Quite naturally these tools require detailed accident analyses and are based on certain assumptions, e.g. on the extent the system reduces impact speed. In order to verify these assumptions and the resulting predicted efficiency it could be beneficial to assess the outcomes of the tools by a posteriori methods as soon as the system shows a sufficient market penetration.

When using a posteriori or epidemiological methods it has to be determined if the evaluation is based on routine data or if a special survey should be conducted. Although the usage of routine data generally cause less costs it is often not possible to perform the evaluation on this basis since information on the equipment of vehicles with the safety system under investigation are not available in these data. Thus, in many cases the best way to perform an (a posteriori) evaluation of vehicle safety systems is to conduct a cohort study, possibly under application of a matched-pairs concept (pairing an equipped vehicle with a - similar but unequipped - reference vehicle). In any case the accumulation of safety systems has to be thoroughly looked at when the efficiency of a certain system is to be assessed.

If the evaluation results shall be expanded from one or a few countries to the EU-27 the iterative proportional fitting procedure can be applied as far as some basic auxiliary information at EU-27 level are available. This is especially relevant for results coming from an a priori evaluation because here for each case it can be determined whether or not the presence of the system would have avoided or mitigated the accident. Thus, the distributions (regarding e.g. injury severity, light conditions, etc.) of both the affected and unaffected accidents are known and can be expanded to a wider accident population. However, one should be cautious when interpreting the estimation outcomes since differences between countries e.g. regarding vehicle fleet may be a limiting factor for the validity of the results.

Concerning socio-economic evaluation of systems, the application of a cost-benefit-analysis should be aimed at. In order to estimate the benefits (cost reduction due to the mitigation or prevention of accidents) standard accident cost schemes can be used.

REFERENCES

Abele, J., Kerlen, C., Krueger, S., Baum, H., Geißler, T., Grawenhoff, S., Schneider, J., and Schulz, W.H. (2005). Exploratory Study on the potential socio-economic impact of the introduction of Intelligent Safety Systems in Road Vehicles. SEiSS Final Report

Agresti, A. (2002). Categorical data analysis, 2nd ed. New York: Wiley Interscience

Bickel, P., Friedrich, R., Burgess, A., Fagiani, P., Hunt, A., Jong, G. de, Laird, J., Lieb, C., Lindberg., G., Mackie, P., Navrud, S., Odgaard, T., Ricci, A., Shires, J., and Tavasszy, L. (2006). Proposal for harmonised guidelines. Deliverable 5 of the EU project HEATCO (Developing Harmonised European Approaches for Transport Costing and Project Assessment). European Commission, Brussels

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975). Discrete Multivariate Analysis. Cambridge, Mass.: MIT Press

Böhning, D. (1998). Allgemeine Epidemiologie und ihre methodischen Grundlagen. München/ Wien: Oldenbourg Verlag

Cook, T.D. and Campbell, D.T. (1979). Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally

Cummings, P., McKnight, B., and Greenland, S. (2003). Matched cohort methods in injury research. Epidemiologic Reviews 25: 43–50

Cummings, P., McKnight, B., and Weiss, N.S. (2003). Matched-pair cohort methods in traffic crash research. Accident Analysis and Prevention 35 (1): 131-141

Cummings, P. and McKnight, B. (2004). Analysis of Matched Cohort Data. The Stata Journal 4 (3): 274–281

De Blaeij, A.T., Florax, R.J.G.M., Rietveld, P., and Verhoef, E. (2003). The value of statistical life in road safety; a meta-analysis. Accident Analysis and Prevention 35 (6): 973-986

De Blaeij, A.T., Koetse, M., Yin-Yen, T., Rietveld, P., and Verhoef, E. (2004). Valuation of safety, time, air pollution, climate change and noise: Methods and estimates for various countries. Draft March 2004, Amsterdam, Department of Spatial Economics, Vrije Universiteit Amsterdam

Deming, W.E. and Stephan, F.F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. Ann. Math. Statist. 11: 427-444

Dionne, G. and Lanoie, P. (2004). Public choice and the value of a statistical life for cost-benefit analysis: the case of road safety. Journal of Transport Economics and Policy 38 (2): 247-274

eIMPACT (Socio-economic Impact Assessment of Stand-alone and Co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe) (2008). Deliverable D2, Stand alone and co-operative Intelligent Vehicle Safety Systems - Inventory and Recommendations for in-depth socio-economic impact assessment

Elvik, R. (1995). An analysis of official economic valuations of Traffic accident fatalities in 20 motorized countries. Accident Analysis and Prevention 27 (2): 237-247

Elvik, R. (2000). How much do road accidents cost the national economy? Accident Analysis and Prevention 32 (6): 849-851

Elvik, R. (2001). Cost-benefit analysis of road safety measures: Applicability and controversies. Accident Analysis and Prevention 33 (1): 9-17

Elvik, R. and Vaa, T. (2004). The Handbook of Road Safety Measures. Amsterdam: Elsevier Science

ETSC (2007). Social and Economic consequences of road traffic injury in Europe

Fienberg, S.E. and Meyer, M.M. (1982) Iterative proportional fitting. In: Johnson, N.L. and Kotz, S. (Ed.) Encyclopedia of Statistical Sciences, New York: Wiley, 275-279

Grömping, U., Pfeiffer, M., and Stock, W. (2007). Statistical Methods for Improving the Usability of Existing Accident Databases. Deliverable 7.1 of the EU-project TRACE: Traffic Accident Causation in Europe

Hanley, N. and Spash, C.L. (1993). Cost-benefit analysis and the environment. Aldershot: E. Elgar

Hauer, E. (1980). Bias by Selection: Overestimation of the Effectiveness of Safety Countermeasures Caused by the Process of Selection for Treatment. Accident Analysis and Prevention 12 (2): 113-117

Hautzinger, H. (2006). Design and Analysis of Matched Studies in Empirical Car Safety Research. SARAC 2: Quality Criteria for the Safety Assessment of Cars Based on Real-World Crashes. Final Report

Hautzinger, H., Pastor, C., Pfeiffer, M., and Schmidt, J. (2007). Analysis Methods for Accident and Injury Risk Studies. Deliverable 7.3 of the EU-project TRACE: Traffic Accident Causation in Europe

McCullagh, P. and Nelder, J.A. (1992). Generalized Linear Models. Second Edition, London: Chapman & Hall

Miller, T.R. (2000). Variations between countries in values of statistical life. Journal of Transport Economics and Policy 34 (2): 169-188

OECD (1990). Behavioural adaptations to changes in the road transport system. Paris: OECD expert group

Robinson, T. and Knight, I. (2009). A common approach to understanding strengths and limitations of different cost benefit analysis techniques, ESV 2009 Paper Number 09-0395

Schlag, B. (2008). Behavioural Adaptation. Internet publication (http://www.scitopics.com/Behavioural%20Adaptation.html)

SWOV (2011). Cost-benefit analysis of road safety measures, SWOV fact sheet

SWOV (2012). The valuation of human losses of road deaths, SWOV fact sheet

Vlakveld, W., Wesemann, P., Devillers, E., Elvik, R., and Veisten, K. (2005). Detailed cost-benefit analysis of potential impairment countermeasures. Deliverable D-P2 of the IMMORTAL project

WHO (2009). European status report on road safety, Towards safer roads and healthier transport choices

Woodward, M. (2005). Epidemiology – Study Design and Data Analysis, 2nd ed. Boca Raton/London: Chapman & Hall/CRC

Zangmeister, T., Kreiß, J.-P., Schüler, L., de Vries, Y., and Ruijs, P. (2007). Methods for Safety Functions Effectiveness Evaluation and Prediction. Deliverable 7.4 of the EU-project TRACE: Traffic Accident Causation in Europe