POSSIBILITIES AND LIMITATIONS OF ACCIDENT ANALYSIS.

S. Oppe

SWOV institute for road safety research

Leidschendam.

1. Introduction.

This paper is primarily based on personal experience concerning traffic safety, safety research and the role of accidents analysis in this research. These experiences resulted in rather philosophical opinions as well as more practical viewpoints on research methodology and statistical analysis. A number of these findings are published already elsewhere. References or (adjusted) citations will be given. The purpose of this paper is to describe my personal view on the possibilities and limitations of accident analysis.

The more philosophical remarks are primarily concerned with the role that accident analysis plays for policy makers and practitioners in the process of traffic safety improvement, as opposed to its role in traffic safety research.

The traffic safety problem was originally a national health problem, initiated by the societal concern with the consequences of modern motorized traffic as shown by the accident statistics. This data, together with the background of safety as a health problem, led to the epidemiological approach: the safety problem was some kind of social sickness, spreading like a disease. The main concern was to find the agent and to cure society from the disease. This background makes it clear that traffic safety was not a scientific problem in the first place, but a problem for practitioners.

The historical background of road safety is also the reason why so much of the safety research is based on national accident statistics. These statistics are not collected for research purposes, but to monitor the process of traffic development and its safety consequences. On this basis the safety problem in general was raised and many specific problems identified.

The scientific expertise was needed to justify actions and to decrease the uncertainty of decision making. This puts the scientist in an uneasy position. He becomes an advisor instead of a researcher. He hardly gets knowledge of results to check the scientific value of his advise, but at the most only its practical value.

It is not a contradiction to state that, although there has been a lot of progress in the field of traffic safety, there is little progress in traffic safety research. Most of the well-established facts in safety literature have more practical than scientific value. E.g. the statement: "the group of bicyclist between 11 and 14 years is most liable to accidents", although of great practical importance, has hardly any scientific value. It is a call for action: "What are we going to do about it?" It is not a fact that originates from a problem analysis of accidents, or
to falsify a hypothesis about their causes, but at most a trigger to stimulate thinking. This is contrary to the ordinary scientific approach.

Science raises from curiosity about phenomena in the real world. Each particular science focuses on a particular aspect of reality, called its object of study. Scientific research starts with systematic and repeated observation of this object. From reflection on these observations it develops a theory, to explain why the observations are as they are and to predict what will be observed in new situations.

Traffic safety continuously asks for social attention, although at the individual level, practitioners and researchers are rarely confronted with accidents. Contrary to regular scientific practice, traffic safety scientists will only rarely and surely not systematically observe their object of research, the accident, but almost exclusively investigate the documentation on accidents. This documentation is also not based on direct observation of the object itself but on data that has often been collected considerably long after the accident took place. In this respect, traffic safety research is more similar to the science of history than to the natural sciences.

From this lack of direct observation of accidents, a number of methodological problems arise, leading to continuous discussions about the interpretation of findings that cannot be tested directly. For a fruitful discussion of these methodological problems it is very informative to look at a real accident on video. It then turns out that most of the relevant information used to explain the accident will be missing in the accident record. In-depth studies also cannot recollect all the data that is necessary in order to test hypotheses about the occurrence of the accident.

For a particular car-car accident, that was recorded on video at an urban intersection in the Netherlands, between a car coming from a minor road, colliding with a car on the major road, the following questions could be asked:

Why did the driver of the car coming from the minor road, suddenly accelerate after coming almost to a stop and hit the side of the car from the left at the main road?

Why was the approaching car not noticed? Was it because the driver was preoccupied with the two cars coming from the right and the gap before them that offered him the possibility to cross? Did he look left before, but was his view possibly blocked by the green van parked at the corner? Certainly the traffic situation was not complicated. At the moment of the accident there were no bicyclists or pedestrians present to distract his attention at the regularly overcrowded intersection.

All the elements mentioned in this explanatory reasoning will not be registered in the documentary file and cannot be recollected afterwards by in-depth investigation or otherwise. The parked green van disappeared within five minutes, the two other cars that may have been important left without a trace.

It is hardly possible to observe traffic behaviour under the most relevant condition of an accident occurring, because accidents are very rare events, given the large number of trips.

Given the new video equipment and the recent developments in automatic incident and accident detection, it becomes more and more realistic to collect such data at not too high costs.

Additional to this type of data that is most essential for a good understanding of the risk increasing factors in traffic, it also important to look at normal traffic behaviour as a reference base.

The question about the possibilities and limitations of accident analysis is not lightly answered. We cannot speak unambiguously about accident analysis. Accident analysis covers
a whole range of activities, each originating from a different background and based on different sources of information: national data banks, additional information from other sources, specially collected accident data, behavioural background data etc.

To answer the question about the possibilities and limitations, we first have to look at the cycle of activities in the area of traffic safety. Some of these activities are mainly concerned with the safety management of the traffic system, some others are primarily research activities.

The following steps should be distinguished:
- detection of new or remaining safety problems;
- description of the problem and its main characteristics;
- the analysis of the problem, its causes and suggestions for improvement;
- selection and implementation of safety measures;
- evaluation of measures taken.

Although this cycle can be carried out by the same person or group of persons, the problem has a different (political/managerial or scientific) background at each stage. We will describe the phases in which accident analysis is used. It is important to make this distinction. Many fruitless discussions about the method of analysis result from ignoring this distinction.

Politicians, or road managers are not primarily interested in individual accidents. From their perspective accidents are often treated equally, because the total outcome is much more important than the whole chain of events leading to each individual accident. Therefore, each accident counts as one and they add up all together to a final safety result. Researchers are much more interested in the chain of events leading to an individual accident. They want to get detailed information about each accident, to detect its causes and the relevant conditions. The politician wants only those details that direct his actions. At the highest level this is the decrease in the total number of accidents. Accidents are no object of study for him, but units to manage in total. The main source of information is the national database and its statistical treatment. For him, accident analysis is looking at (sub-groups of) accident numbers and their statistical fluctuations.

This is the main stream of accident analysis as applied in the area of traffic safety. Therefore, we will first describe these aspects of accidents.

2. The nature of accidents and their statistical characteristics.

The basic notion is that accidents, whatever their cause, appear according to a chance process. Two simple assumptions are usually made to describe this process for (traffic) accidents:
- the probability of an accident to occur is independent from the occurrence of previous accidents;
- the occurrence of accidents is homogeneous in time.

If these two assumptions hold, then accidents are Poisson distributed.

The first assumption does not meet much criticism. Accidents are rare events and therefore not easily influenced by previous accidents. In some cases where there is a direct causal chain (e.g., when a number of cars run into each other) the series of accidents may be regarded as one complicated accident with many cars involved. The assumption does not apply to casualties. Casualties are often related to the same accident and therefore the independency assumption does not hold.
The second assumption seems less obvious at first sight. The occurrence of accidents through time or on different locations are not equally likely. However, the assumption need not hold over long time periods. It is a rather theoretical assumption in its nature. If it holds for short periods of time, then it also holds for long periods, because the sum of Poisson distributed variables, even if their Poisson rates are different, is also Poisson distributed. The Poisson rate for the sum of these periods is then equal to the sum of the Poisson rates for these parts.

The assumption that really counts for a comparison of (composite) situations, is whether two outcomes from an aggregation of situations in time and/or space, have a comparable mix of basic situations. E.g., the comparison of the number of accidents on one particular day of the year, as compared to another day (the next day, or the same day of the next week etc.). If the conditions are assumed to be the same (same duration, same mix of traffic and situations, same weather conditions etc.) then the resulting numbers of accidents are the outcomes of the same Poisson process. This assumption can be tested by estimating the rate parameter on the basis of the two observed values (the estimate being the average of the two values). Probability theory can be used to compute the likelihood of the equality assumption, given the two observations and their mean.

This statistical procedure is rather powerful. The Poisson assumption is investigated many times and turns out to be supported by a vast body of empirical evidence.

It has been applied in numerous situations to find out whether differences in observed numbers of accidents suggest real differences in safety. The main purpose of this procedure is to detect differences in safety. This may be a difference over time, or between different places or between different conditions. Such differences may guide the process of improvement. Because the main concern is to reduce the number of accidents, such an analysis may lead to the most promising areas for treatment.

A necessary condition for the application of such a test is, that the numbers of accidents to be compared are large enough to show existing differences. In many local cases an application is not possible. Accident black-spot analysis is often hindered by this limitation, e.g., if such a test is applied to find out whether the number of accidents at a particular location is higher than average.

The procedure described can also be used if the accidents are classified according to a number of characteristics to find promising safety targets. Not only with aggregation, but also with disaggregation the Poisson assumption holds, and the accident numbers can be tested against each other on the basis of the Poisson assumptions. Such a test is rather cumbersome, because for each particular case, i.e. for each different Poisson parameter, the probabilities for all possible outcomes must be computed to apply the test.

In practice, this is not necessary when the numbers are large. Then the Poisson distribution can be approximated by a Normal distribution, with mean and variance equal to the Poisson parameter. Once the mean value and the variance of a Normal distribution are given, all tests can be rephrased in terms of the standard Normal distribution with zero mean and variance one. No computations are necessary any more, but test statistics can be drawn from tables.

In practice the Poisson test is hardly ever carried out. Almost always the Normal approximation is used for testing. For small numbers of observations, the exact Poisson test is sometimes applicable if the Normal approximation is not. Since computers have taken over the computations, the labour involved in applying such tests is not a serious limitation any more.

There is another reason why Poisson tests are seldomly applied. If a number of Poisson variables with different Poisson parameters, are compared, another characteristic of Poisson
variables can be used. As such, each of the Poisson processes under investigation results in a particular outcome. For many replications of the process, different values are found, that together follow the Poisson distribution. As said before, the sum of the Poisson outcomes for all the variables together in each replication is also Poisson distributed. The outcomes for this sum for the replications together follow again a Poisson distribution.

If a sum is known, then the outcome for each variable is restricted (for each variable it can never be larger than this sum). In such a case we speak of a conditional distribution: what is the probability of a particular outcome for a (Poisson distributed) variable, given that we know the (Poisson distributed) total?

This conditional distribution turns out to be the well-known multinomial distribution.

It is the distribution that describes the placement of a fixed number of balls (accidents) over a fixed number of boxes (accident conditions), each with its own probability of getting a ball placed. This probability is then equal to the ratio between the Poisson parameter for the condition and for the total.

This multinomial distribution plays an important role in testing, and is e.g. applicable to outcomes of questionnaires. The number of questionnaires is fixed but not the distribution of responses over the response categories. Because of the special relation to the Poisson distribution, tests that are based on this distribution can also be used for accident analysis. The well-known Chi-square test, used in testing accident outcomes, is applied on the assumption of multinomially distributed objects. Multinomial tests are also cumbersome to apply, but in practice, if the numbers are large enough, an approximation is possible again, using the Normal distribution. The Chi-square test is in fact a test based on the joint distribution of Normal variables. In case of small numbers exact tests can be applied, but if these numbers are too small, also exact tests will not do.

The Chi-square test on two-way tables of accident numbers is one of the most applied types of analysis. In this case the additional assumption to be made is that the Poisson probability for a cell is the product of that probability for the row and column to which this cell corresponds. If this is true for all cells, then the row and column variables are independent and cell-entries do not add extra information. The Chi-square test is therefore based on a multiplicative assumption for row and column counts.

Because of its restricted applicability, it was not popular with researchers, until the seventies. The popularity increased strongly after the extension of the Chi-square test to a log-linear test on multi-way tables.

3. The use of accident statistics for traffic safety policy.

The testing procedure described has its merits for those types of analysis that are based on the assumptions mentioned. The best example of such an application is the monitoring of safety for a country or region over a year, using the total number of accidents (eventually of a particular type, such as fatal accidents), in order to compare this number with the outcome of the year before. If sequences of accidents are given over several years, then trends in the developments can be detected and accident numbers predicted for following years. Once such a trend is established, then the value for the next year or years can be predicted, together with its error bounds. Deviations from a given trend can also be tested afterwards, and new actions planned. There is a long tradition of accident analysis that concentrates on the description of such trends. The most famous one is carried out by Smeed 1949. We will discuss this type of accident analysis in more detail later.

Another area of application is the disaggregation of the total number of accidents according to aspects of interest. The monitoring in general is not restricted to the total number of
accidents, but comparisons are made between groups of accidents. The dominant technique for this kind of accident analysis is contingency table analysis. The application of a Chi-square analysis to contingency tables is already mentioned.

Such comparisons, if based on absolute numbers are generally restricted in their use. It is not interesting to know that more accidents happen in daytime than at night, because most people travel during daytime. One wants to know whether a trip at night is more dangerous than the same trip during daytime. To make a sensible comparison, one should compare traffic risk at day and night; therefore, the accident numbers must be weighed for the amount of exposure to such risks.

A main break through in contingency table analysis took place at the end of the sixties and the beginning of the seventies. Then it was realised that the multiplicative model, used to analyse two-way tables, could be treated as a linear model of the log-counts.

Statistical properties of parameters, which are linear combinations of observed random variables, are easily defined in linear models. The nicest generalisation is found in Nelder and Wedderburn (1972), who made a distinction between (1) the (linear) model description of the data, (2) the link-function that relates the dependent variable to the model description and (3) the error function that takes care of the stochastic variation in the model. Different choices lead to different models, all united in the context of general linear models.

There are a number of standard textbooks available on log-linear analysis of contingency tables. Their application to road safety is treated in Fleischer 1981.

There are four main developments in recent contingency table analysis:

1. The application of the Chi-square test for interaction is generalised to higher order classifications. Foldvary and Lane (1974), in measuring the effect of compulsory wearing of seat belts, were among the first who applied the partitioning of the total Chi-square in values for the higher order interactions of four-way tables.

2. Tests are not restricted to overall effects, but Chi-square values can be decomposed regarding sub-hypotheses within the model. Also in the two-way table, the total Chi-square can be decomposed into interaction effects of part tables. The advantage of 1. and 2. over previous situations is, that large numbers of Chi-square tests on many interrelated (sub)tables and corresponding Chi-squares were replaced by one analysis with an exact partitioning of one Chi-square.

3. More attention is put to parameter estimation. E.g., the partitioning of the Chi-square made it possible to test for linear or quadratic restraints on the row-parameters or for discontinuities in trends.

4. The unit of analysis is generalised from counts to weighted counts. This is especially advantageous for road safety analyses, where corrections for period of time, number of road users, number of locations or number of vehicle kilometres is often necessary.

The last option is not found in many statistical packages. Andersen 1977 gives an example for road safety analysis in a two-way table. A computer programme WPM, developed for this type of analysis of multi-way tables, is available at SWOV (see: De Leeuw and Oppe 1976).

The accident analysis at this level is not explanatory. It tries to detect safety problems that need special attention.

The basic information needed consists of accident numbers, to describe the total amount of unsafety, and exposure data to calculate risks and to find situations or (groups of) road users with a high level of risk.
Smeeds analysis was concerned with these two aspects of the problem. His approach was descriptive: he was looking for regularities in accident numbers and corrected these numbers for exposure, using population and car park figures. And although Smeed himself did not claim to have found more than an empirical relation, there was a strong tendency to interpret it as a model for predictions.

From the beginning onwards his model got a lot of criticism, because it was purely descriptive and did not give an explanation for the relations and trends that were found. It was argued that such an explanation is necessary to justify a model even at this highest level of analysis. His model failed to describe the decrease in accident numbers after the second half of the seventies. This fact justified the criticism and made his model also lose its practical value.

Oppe [1991] and related publications show that at this highest level of aggregation it is possible to state a model that describes and predicts developments in traffic safety. It claims not to be merely descriptive, but to have a simple theoretical basis, only using the two fundamental concepts of traffic safety: exposure and risk. The outcomes for the Netherlands are given in figure 1.

The model is able to cope with the increase in accidents before, as well as with the decrease after the mid-seventies. Its fundamental assumption is, that the development of risk at the highest level is exponential, with a negative (learning rate) parameter. The explanation given is that the improvement of safety is a societal process of learning, comparable to learning processes of individuals, that are also described by exponential functions. It is the end product of a large number of collective learning events. In the beginning there is a lot to be learned, but the number of learning events decreases with time. Given a constant learning capacity, the learning result is an exponential decay function for the number of learning events and the corresponding safety improvement. A large amount of evidence for exponential learning, as well as the theoretical basis for it, is found in the psychological literature.

The model has been applied at SWOV to describe disaggregated data as well. Exposure and risk developments were modeled for groups of road users, divided by age and traffic mode. Detailed accident and exposure data, collected over a ten year period, combined with the long term trends for the total and information about observed and predicted population growth per age category, were used to make predictions. These predictions were then compared to target values for each group as aimed at in the national safety programme for the years 2000 and 2010.

Figure 2. shows the results for four age groups of bicyclists. The observed accidents, together with the model predictions and their error bounds are given, together with the target values. It looks as if the target will be easily met for the younger road users, but not for the older ones (Oppe 1993).
The examples show that an accident analysis is very dependent on the aims one has and on the structure of the problem. The accident figures at the highest level hardly reflect any specific influences. From figure 1c, in which the number of accidents is described using the amount of motor vehicle kilometres as a measure of exposure, and the two parameters of the exponential curve for risk, it is seen how little of the variance in accident numbers remains to be explained by other factors. The total figures hardly reflect anything else but the total effort of society as described by the exponential risk reduction over time. The disaggregated figures show much more variation, that need not all be stochastic noise.

Apart from these more descriptive analyses. Several attempts have been made to explain the development of traffic safety at this highest level, using all kinds of variables other than the amount of traffic. Many of these variables (e.g., economic factors such as fuel consumption, fuel prices, amount of unemployment etc.) are again related to exposure or traffic volume. Others are supposed to cover particular risk related factors. An overview of some of these attempts is given in the special edition of Accident Analysis and Prevention, 1991, No 2. In one of these studies, Gaudry 1984 claims to have found a satisfactory model description for the accidents in Québec, on the basis of a large number of explanatory variables. Statistically speaking, his approach is sound. The model is nicely structured and statistically well based. Its weakness seems to be the large number of variables and thus parameters involved.

In spite of this and other claims, it is my opinion that this type of explanation is not suited for this aggregated level of analysis. The disadvantage of explanatory models as used by Partyka and Gaudry is, that they are based on a description of relations between variables, without an explicit theoretical foundation. If relatively many variables are used, a fairly good description may result that seems to explain the fluctuations, but still results in bad predictions.

Generally speaking predictions of the total number of accidents, based on data up to, or from the late seventies onwards, will not show large deviations. As in the case of Smeeds model, a nice test for these models is to make a description of traffic safety up to the beginning of the seventies, and to use the model parameters together with the explanatory variables to predict the number of accidents after that period, up to the nineties.

My expectation is, that unless those variables are used which together predict the exponential decay of risk (e.g. the special safety effects of all measures taken), no model will explain the fall of the accident figures from the second half of the seventies onwards.

My general conclusion is, that accident analysis at the highest level of aggregation can be used to describe and even model traffic safety developments, but can hardly be used to explain these developments in causal terms.

Similar attempts have been made to show that particular safety measures were effective, e.g. by applying a kind of time-series analysis called intervention analysis. Again, figure 1c. shows how difficult it will be to detect such effects at a macroscopic level. The explanation based on exposure and the total risk reduction leaves not much left to be explained by additional factors. Only over rather short periods of time such effects may be isolated from the effects of all other newly taken safety measures. The possibilities for showing such effects are larger for more specific accident types. The effect of the safety belt law is a good example. Its effect is supposed to be found for car occupants and not for other road users, although some researchers claim side-effects for these groups too, because the risk taking behaviour of car drivers may increase if their own risk is reduced. Monthly data collected over a short period before and after the initiation of the law may then prove a direct effect.
Figure 3. shows such a result for the U.K., based on monthly data (see Harvey and Durban 1986). Application to the Dutch situation, using the same procedure failed, because the safety belt law effect was confounded with a number of other safety measures taken at the highest level, such as an alcohol law and a general speed law.

****************Insert figure 3 here****************

4. Accident analysis for research purposes.
Traffic safety research is concerned with the occurrence of accidents and their consequences. Therefore, one might say that the object of research is the accident. The researchers interest however is less focused at this final outcome itself, but much more at the process that results (or does not result) in accidents. Therefore, it is better to regard the critical event in traffic as his object of study. One of the major problems in the study of the traffic process that results in accidents is, that the actual occurrence is hardly ever observed by the researcher.

Investigating a traffic accident, he will try to reconstruct the event from indirect sources such as the information given by the road users involved, or by eye-witnesses, about the circumstances, the characteristics of the vehicles, the road and the drivers.

As such this is not unique in science, there are more examples of an indirect study of the object of research.

However, a second difficulty is, that the object of research cannot be evoked. Systematic research by means of controlled experiments is only possible for aspects of the problem, not for the problem itself.

The combination of indirect observation and lack of systematic control make it very difficult for the investigator to detect which factors, under what circumstances cause an accident. Although the researcher is primarily interested in the process leading to accidents, he has almost exclusively information about the consequences, the product of it, the accident.

Furthermore, the context of accidents is complicated. Generally speaking, the following aspects can be distinguished:

- Given the state of the traffic system, traffic volume and composition, the manoeuvres of the road users, their speeds, the weather conditions, the condition of the road, the vehicles, the road users and their interactions, accidents can or cannot be prevented.

- Given an accident, also depending on a large number of factors, such as the speed and mass of vehicles, the collision angle, the protection of road users and their vulnerability, the location of impact etc., injuries are more or less severe or the material damage is more or less substantial.

Although these aspects cannot be studied independently, from a theoretical point of view it has advantages to distinguish the number of situations in traffic that are potentially dangerous, from the probability of having an accident given such a potentially dangerous situation and also from the resulting outcome, given a particular accident.

This conceptual framework is the general basis for the formulation of risk regarding the decisions of individual road users as well as the decisions of controllers at higher levels. In the mathematical formulation of risk we need an explicit description of our probability space, consisting of the elementary events (the situations) that may result in accidents, the
probability for each type of event to end up in an accident, and finally the particular outcome, the loss, given that type of accident.

A different approach is to look at combinations of accident characteristics, to find critical factors. This type of analysis may be carried out at the total group of accidents or at sub-groups. The accident itself may be the unit of research, but also a road, a road location, a road design (e.g. a roundabout) etc.

This type of analysis is often called analysis of relational data, because the various factors that are involved are not independent from each other. Volume, speed, road width etc. are considerably correlated. Analysis techniques should take these correlations into account.

Sometimes the data are collected in multiway contingency tables and analyzed with the previously mentioned log-linear model for contingency tables. The model structure of these models however, assumes independent factors, as in the case of an analysis of variance. In laboratory experiments the researcher is able to design his study in such a way (e.g. by using an orthogonal design) that interrelations between explanatory factors do not exist. In field experiments and for accident data that are classified afterwards instead of forwards, such an analysis may be misleading (see Oppe 1992 and the examples in section 5).

Techniques such as multiple linear analysis correct for such intercorrelations, but the disadvantage is, that they are meant for measurements and do not apply to counts. A better choice for accident counts is a log-linear model, but then a regression type of model instead of the model for contingency table analysis. GLIM does have both possibilities.

The regression type of log-linear analysis is used by TRRL (Maycock and Hall 1984), to find combinations of risk factors for various types of roads.

There are other techniques developed at the Leiden State University, that are specially suited for this type of exploratory research (Gifi 1990). A comparison between these techniques and GLIM, applied to the TRRL-data is also described by Oppe 1992. The techniques can be found in the recent versions of the statistical package SAS-Stat, under the headings of PRINQUAL and TRANSREG. These types of analysis can be used for exploratory analysis, to find potential problems. For a further (causal) analysis of such problems, more information is necessary.

As said before, accident records do not cover much of this necessary information. Therefore, researchers have tried to get additional information about accidents, in order to study these more in-depth. In other areas of transport (air/rail/water) in-depth accident investigations are much more standard. Many attempts have been made to borrow in-depth techniques from these areas. However, there are a number of reasons why this transfer is not straightforward. Accidents in road traffic are much more frequent, the outcomes are often less serious and get less public attention. Therefore, the cost/benefit rate per accident is much higher. Furthermore, the accidents are of a different nature. In the air, on railways and at the water more time is available for critical decisions to be made (predominantly) by professionals. Procedures for all kinds of conditions are specified and often monitored so that they can be traced afterwards. This makes the decisions taken more rational and easier checked retrospectively. Especially in air traffic, but also on rail or at sea, technical failures are often the cause of an accident and much effort is put in changing the responsibility of the humans involved from the tactical level of actual manoeuvring to the strategic level of monitoring the process.

In road traffic the situation is completely different. The driver has to carry out most tasks himself at the tactical as well as strategic level. His decisions have to be made in a very short time, often in rather complicated situations. The decision rate is on average much higher, less supported by special equipment or controlled by procedures and almost completely not monitored.
Recent developments in road traffic, particularly those in the area of traffic and transport telematics as in PROMETHEUS and DRIVE, are concerned with changes of the drivers task and the driving environment. The implementation of such changes request a detailed insight into this task, to avoid unforeseen accident risks. At the moment, there is a strong tendency to experiment and to find acceptable solutions by trial and error. The evaluation of these systems is thought to be carried out by before and after studies, based on accident numbers. The possibilities for applying such evaluation studies are limited. The use of accident analysis for these purposes in general is very limited. The type of accident analysis that may become useful, in combination with additional behavioural studies, is the video-based type of analysis mentioned in the introduction, but only for systems that are implemented on a sufficiently large scale.

Although the in-depth accident study as carried out in other transport modes is difficult to use for road traffic, many attempts have been made to carry out in-depth accident studies in that area as well. These studies turned out to be rather costly. Most studies focus on the outcomes of accidents, either to improve medical care or vehicle design. Sometimes special teams are trained to carry out such studies on-the-spot, as soon as possible after the accident. Others are primarily based on extra information from participants and witnesses, collected through questionnaires. Research at INRETS, carried out by Fleury, Malaterre and others, focused primarily on the pre-crash aspects of the accidents and less on the outcomes.

In all these studies the aim is to reconstruct the accident retrospectively. As described in the example at the beginning, it will be very difficult to detect the relevant cues, especially the information about the last five or ten seconds. Reports from participants and witnesses turned out to be not very reliable. The most promising examples are studies, such as those carried out by INRETS, focusing on particular types of accidents, testing well defined hypotheses.

Because of the unreliability of the information collected afterwards, reconstruction of accidents and their causes is limited. Such analyses cannot be carried out without additional information about what drivers do in comparable situations, not resulting in accidents. Behavioural studies will give this additional information. The main disadvantage of this approach is, that the risk itself cannot be observed, but must be measured indirectly, on the basis of theoretical assumptions about risk. The missing link between both methods is the observation of the accident. Careful study of the actual accident process may tell us which factors in which situations are responsible for the loss of control. It may also tell us what information is necessary to reconstruct an accident.

This type of combined accident and behavioural research is in my opinion the most promising tool for safety analysis. Without this type of information it is hardly possible to test behavioural assumptions as well as conclusions drawn from (in-depth) accident studies.

A rather popular type of accident research is accident black-spot analysis. It is indeed one of the most promising areas of traffic safety improvement, because it is directed to a very important aspect of safety: the architecture of the traffic environment. Accident black-spot analysis is also the best example to show the limitations of accident analysis. This regards detection and selection, problem analysis and evaluation.

The idea is to select and improve those locations (road sections or intersections) in a certain area or class of roads that are the most dangerous.

The total number of accidents in one or more years is often chosen as the safety criterion. However, these numbers are mostly rather low and therefore they give no reliable estimates of danger. E.g., for a location with nine accidents on average per year, it is not very unlikely
to have yearly accident figures that range from three to fifteen. An ordering of the locations on the basis of such small numbers is rather unstable, and locations will be selected that are not particularly dangerous, while much more dangerous locations are left out. This need not be a serious problem if the procedure is carried out at a yearly basis, because then the dangerous locations will show up in the end anyway.

Increasing the number of accidents by extending the observation period to three or five years may be a solution, but then the assumption is that nothing is changed over that period and that the past number of accidents still is a valid estimator for the number of expected accidents at the moment of analysis. Also increasing the number of accidents by using all accidents, including material damage only accidents, may have drawbacks, because the degree of registration for this type of accidents is very low and also selective. If the set of locations is homogeneous and/or the accident type is restricted, this need not be much of a problem.

Sometimes the problem is solved by using conflict observation techniques, to find out whether a location is dangerous or not. This procedure, if carried out by trained observers, has a number of advantages: the observation period can be kept rather short and still representative; additional countings for exposure can be carried out for cars as well as pedestrians and bicyclists and preparatory accident analysis with regard to infrastructure and behaviour can be carried out by this systematic observation of ordinary traffic flow. Critics of this approach point to the fact that conflicts are no accidents and that it should be proved first that they can replace them. This argument as such is correct, and much effort is put in showing this, but these critics often forget to be equally critical with regard to the above mentioned problems concerning recorded accidents.

A second problem is, that these absolute numbers are not the most appropriate criterion, because they do not measure the risk for passing such a location. To measure risk, a correction for exposure (vehicle kilometres driven or encounters between road users) is necessary. This type of information is often missing. This is one of the most important reasons for choosing the absolute number of accidents as a criterion. It decreases the effectiveness of treatment further, because many locations with high numbers of accidents will be highly occupied (inter)sections probably already with a high quality design. Further improvements will be costly and therefore less cost/effective. A reasonable criterion seems to be the percentage of risk reduction per unit of money, times the absolute number of expected accidents.

The most difficult part always is the accident analysis, in order to select remedial measures. Accident data are far too restricted for these purposes. There is little or no hope to find regular patterns in the accidents. Each selection will be mixed with other selections and soon result in one, two, three examples per selected combination, e.g. for the most basic selection according to conflict type, age group, weather condition, time of day.

As said before, the available information from the accident records is also limited. A mixed procedure, where locations for potential treatment are selected on the basis of accident records and investigated by additional conflict and behavioural analysis seems the most efficient.

The last problem to be met, is the evaluation of safety measures. Apart from the fact that accident numbers are too small to find reasonable effects of measures per location, there is the statistical problem of the regression-to-mean. If all locations were equally dangerous, then the actual accident numbers for a particular period would still fluctuate by mere chance. An ordering according to the outcomes, followed by a classification in classes for low, medium and high accident numbers, will show an average increase, no increase and decrease in the next period respectively. Because the classification was based on the differences that were caused by chance only, the expected outcomes are the same for all
locations within the three groups in this second period. The group of selected locations will therefore on average show a decrease in accident numbers, even when no actions are taken at all. For examples of this effect in actual situations, see Hauer 1986. There are procedures to correct for this effect (see e.g. Maher 1990). It is also possible to test for homogeneity within a group of locations. If the variance of the accident numbers is larger than the mean, this is an indication that the Poisson parameters differ from each other. The negative binomial distribution describes the accident distribution better in this case.

An alternative approach, based on the analysis of the characteristics of all the locations together, that circumvent the small numbers problem as well as the regression-to mean problem is given in Oppe 1982.

A before and after study on the basis of behavioural observations or conflict techniques both carried out after selection, need no correction for the regression-to-mean effect either. Such an evaluation in terms of behaviour has also the advantage that the intermediate variables which the safety measures are supposed to change can be checked for their effectiveness. E.g., if speed reduction measures after implementation do not show any effect on the observed speed, then an evaluation of the safety effect by means of accident reductions is superfluous.

5. Evaluation studies.

The last area of accident analysis that will be treated is evaluation research. Problem detection and problem analysis, as covered in 3. and 4. may be carried out rather independently. As we saw in the last example, this is not the case with evaluation. Here the two approaches meet. Evaluation is carried out to prove effectiveness of safety management, but also to test the underlying assumptions on which the applied measures are based. The safety manager is primarily interested in the safety outcome, as measured by accidents. The researcher in the intervening variables the changes of which are assumed to produce this outcome. We gave already some examples in which especially this process evaluation is important. A detailed treatment of this issue is without the scope of this paper. An excellent overview is found in the proceedings of "Evaluation 1985".

A complete review of the possibilities and limitations of accident analysis for product evaluation will also lead us too far. Haight 1986 and Hauer 1986 give detailed treatments of methods to be used. A recent overview of the most important issues for small scale evaluation projects is given in Oppe c.s. 1993.

Two examples of possible pitfalls will be treated here, because they are rather specific for accident analysis, relate to techniques that are widely used and to limitations mentioned before.

The most simple example of an evaluation is the test of the hypothesis of no-effect (no change, no difference). Simple a test of this kind may be, both conceptually and statistically, its simplicity is its main pitfall. We will illustrate the necessity of a sound theory and experimental design as fundamental conditions, by means of a theoretical example.

Table 1 represents the imaginary outcome of an investigation regarding the effect of safety belts. In many studies we find the results reported in a number of two-way tables. Inspecting this table, it is obvious that there is a highly significant interaction between the use of safety belts (A1) and the probability of having an injury (B2). We can even estimate the effectiveness of safety belt use and compute error bounds for this effect.
Table 1. Relation between the use of safety belts (A) and seriousness of accidents (B); theoretical example.

<table>
<thead>
<tr>
<th></th>
<th>A1: belts</th>
<th>A2: no belts</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1: no injury</td>
<td>91</td>
<td>20</td>
</tr>
<tr>
<td>B2: injury</td>
<td>14</td>
<td>51</td>
</tr>
</tbody>
</table>

However, if table 1 turns out to be the aggregated result of the two sub-tables in table 2 according to rural or urban location (C1 and C2 respectively), then we are confronted with a serious problem. The conclusion drawn from table 1 does not count for either of the two sub-tables. The safety belt effect on accidents turns out to be an artifact of the relations between the location and the use of safety belts and of location and injury. Of course this is a theoretical example (we are not seriously stating that safety belts are ineffective), but it illustrates how wrong obvious conclusions may be.

Table 2. Disaggregation of table 1 according to variable C; the interaction between A and B disappears.

<table>
<thead>
<tr>
<th></th>
<th>C1: rural</th>
<th>C2: urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1: no injury</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>B2: injury</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

The conclusion should not be that we have to investigate three-way tables instead of two-way tables, but to analyse the situation and to understand its (possible) structure, before we state or test our hypotheses. The confirmation should be based on a sound theoretical framework. If we expect interactions with other variables to be of influence, then we either have to adapt our study design or else our analysis, in order to take these variables into account.

The next example concerns design problems. Imagine that table 3 results from an investigation of accidents in relation to characteristics of the location, showing the effect of the relation between the type of road (T) and the type of street lighting (L) on safety. Again, there is an obvious effect, suggesting a strong safety interaction between the type of road and the type of lighting. The road types seem to have safety problems under special lighting conditions.

Table 3. Number of accidents for road types (T) and types of lighting (L).

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>20</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>L2</td>
<td>2</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>L3</td>
<td>3</td>
<td>7</td>
<td>25</td>
</tr>
</tbody>
</table>

However, if the road administrator has a rather strict policy regarding the installation of types of lighting depending on the type of road (preferring the combinations T1-L1, T2-L2, T3-L3), then the safety effect is an artifact of this policy and does not say anything about the safety effect of combining these conditions. If all types of lighting were equally spread over all types of roads, the effect could disappear completely. This problem, in its extreme form known as the problem of structural zero's, is a serious problem in contingency table analysis. Technically speaking, it can be solved by means of a correction for exposure, but this makes
a modification necessary for the analytical technique as described in the beginning of section 3. This problem also clarifies the advantages of GLIM and TRANSREG analyses over contingency table analysis for this type of data. We will no go into those details, but we only want to state that a theoretical framework is necessary for applying significance tests to evaluate effects.

It is the main concern of exploratory data analysis to investigate the structure of the problem. It often helps to investigate the structure of relations between a large number of variables, possibly connected with the safety problem under investigation.

In laboratory experiments a lot of effort is put in the design of the study, in order to control the experimental and control situation. In traffic and traffic safety research this is hardly ever possible. However, although we cannot control the situation, this should not mean that we have to ignore the possible effects. In many, highly dedicated log-linear analyses this design problem is seriously overlooked.

6. Summary conclusions:

Accident statistics, especially collected at a national level are particularly useful for the description, monitoring and prognosis of accident developments, the detection of positive and negative safety developments, the definition of safety targets and the (product) evaluation of long term and large scale safety measures.

The application of accident analysis is strongly limited for problem analysis, prospective and retrospective safety analysis on newly developed traffic systems or safety measures, as well as for (process) evaluation of special short term and small scale safety measures.

There is an urgent need for the analysis of accidents in real time, in combination with background behavioural research. Automatic incident detection, combined with video recording of accidents may soon result in financially acceptable research. This type of research may eventually lead to a better understanding of the concept of risk in traffic and to well-established theories.
Literature


19 Oppe, S. Models for the development of traffic and traffic safety at a disaggregated level; SWOV, Leidschendam, 1993.
