International Transport Forum

Sharing Road Safety

Developing an International Framework for Crash Modification Functions

Research Report

DECADE OF ACTION FOR ROAD SAFETY 2011-2020

OECD

Sharing Road Safety

Developing an International Framework for Crash Modification Functions

Research Report

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Please cite this publication as:
OECD (2012), *Sharing Road Safety: Developing an International Framework for Crash Modification Functions,* OECD Publishing.
http://dx.doi.org/10.1787/9789282103760-en

ISBN 978-92-821-0375-3 (print)
ISBN 978-92-821-0376-0 (PDF)

Photo credits: Cover © Adam Larsen

Corrigenda to OECD publications may be found on line at: *www.oecd.org/publishing/corrigenda*.

Revised version, November 2012.

© OECD/ITF 2012

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to *rights@oecd.org*. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at *info@copyright.com* or the Centre français d'exploitation du droit de copie (CFC) at *contact@cfcopies.com*.

INTERNATIONAL TRANSPORT FORUM

The International Transport Forum at the OECD is an intergovernmental organisation with 54 member countries. It acts as a strategic think tank with the objective of helping shape the transport policy agenda on a global level and ensuring that it contributes to economic growth, environmental protection, social inclusion and the preservation of human life and well-being. The International Transport Forum organizes an annual summit of Ministers along with leading representatives from industry, civil society and academia.

The International Transport Forum was created under a Declaration issued by the Council of Ministers of the ECMT (European Conference of Ministers of Transport) at its Ministerial Session in May 2006 under the legal authority of the Protocol of the ECMT, signed in Brussels on 17 October 1953, and legal instruments of the OECD.

The Members of the Forum are: Albania, Armenia, Australia, Austria, Azerbaijan, Belarus, Belgium, Bosnia-Herzegovina, Bulgaria, Canada, Chile, China, Croatia, the Czech Republic, Denmark, Estonia, Finland, France, FYROM, Georgia, Germany, Greece, Hungary, Iceland, India, Ireland, Italy, Japan, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Moldova, Montenegro, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Kingdom and the United States.

The International Transport Forum's Research Centre gathers statistics and conducts co-operative research programmes addressing all modes of transport. Its findings are widely disseminated and support policymaking in Member countries as well as contributing to the annual summit.

Further information about the International Transport Forum is available at
www.internationaltransportforum.org

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

TABLE OF CONTENTS

RECOMMENDATIONS

- Road safety policies should undergo performance and efficiency evaluations. Such evaluations cannot be undertaken without Crash Modification Functions (CMFs). Evaluation processes should be documented to ensure they are transparent.

- Research conducted to develop CMFs should follow the guidance provided in this report and, in particular, provide specific information that describes the countermeasure under consideration, the safety issue being addressed and the roadway environment in which it was tested.

- It is recommended that an international group be composed under an existing organization (e.g. Transportation Research Board, World Road Association, etc.) to foster dialog among researchers and practitioners on CMF research and reporting standards with the aim of increasing transferability of results. Coordination of research across countries on top priority countermeasures should be considered.

- International cooperation should aim to capture documentation and reporting of CMF research in a widely available transnational database.

- A concerted effort should be made to publicize the benefits of decision-making based on CMFs. This should take the form of: presentations and workshops at transport, injury prevention and health conferences; press releases; letters to political leaders and senior bureaucrats.

KEY MESSAGES

- The decision making process for safety interventions is complex, involving a number of actors (experts, public, politicians, etc.) and issues (environment, economy, congestion) competing for the scarce resources available. The risk of making poor decisions and the cost of making better decisions can be reduced by the use of reliable studies on how effective different safety measures are (i.e. Crash Modification Functions – CMFs).

- Road safety policy is increasingly dependent on sound indicators of the effectiveness of interventions. Policy makers need not only to justify expenditure on safety in terms of effectiveness but also to argue convincingly for measures in the face of sceptical and sometimes hostile lobbies. Crash modification factors and crash modification functions (CMFs) – the indicators that quantify the crash reductions that result from interventions – are persuasive in this context.

- CMFs are fundamental to identifying the most effective road safety countermeasures and for calculating safety benefits in economic analyses of safety policies when trying to make optimal use of resources.

- Demand for CMFs is growing in many jurisdictions as policy makers are increasingly required to demonstrate results and undertake cost-benefit and efficiency assessments.

- Lack of reliable knowledge of the effects of countermeasures is a key barrier to the advancement of many critical, life-saving initiatives. CMFs can be an effective tool in communicating that knowledge. Improved CMFs – in terms of presentation and dissemination, methodology and transferability between jurisdictions – will have tangible benefits for decision making.

- There is a need for more training and regular practical usage of CMFs to support the development of transferable CMFs. We are currently at a turning point, with the prospect of rapid advances and major cost savings through the transfer of results internationally.

- Transferability of CMFs relies first and foremost on analysing the extent to which a CMF is dependent on the circumstances in which it was developed.

- Variability in CMF research results is a major deterrent to transferability. Reducing variability through proper study design and reporting enhances transferability. Studies should control for the most important confounding factors related to the countermeasure analysed. Variability due to different circumstances can be reduced by making the CMF a function of the relevant circumstances. A key aim of the current report is to provide guidance for uniform screening and control procedures.

EXECUTIVE SUMMARY

The decision making process for safety interventions is complex, involving a number of actors (experts, public, politicians etc) and issues (environment, economy, congestion) competing for the scarce resources available. The risk of making poor decisions and the cost of making better decisions can be reduced by the use of reliable studies on how effective different safety measures are (ie. Crash Modification Functions – CMFs).

Road safety policy is increasingly dependent on sound indicators of the effectiveness of interventions. Policy makers need not only to justify expenditure on safety in terms of effectiveness but to argue convincingly for measures in the face of sceptical and sometimes hostile lobbies. Monitoring and analysis of effectiveness is not without cost, and indicators that relate safety improvements to interventions, "Crash Modification Functions", that are transferable from one situation to another are a valuable tool in spreading effective safety policies.

Crash Modification Functions (CMFs) are fundamental to identifying the most effective road safety countermeasures and for calculating safety benefits in economic analyses of safety policies when trying to make optimal use of resources.

Each year about 1.3 million people are killed and another 50 million people are injured on roads worldwide (WHO, 2010). These road crashes cost countries between 1 and 3 percent of their Gross Domestic Product (WHO, 2004). In addition, they cause great emotional and financial stress to the millions of families that are affected by these crashes. Many of these crashes can be prevented by implementing effective road safety measures. To be able to select the best measure, a decision maker needs information about the effectiveness of different measures. Moreover, information about the effectiveness of measures is needed to ensure governments invest appropriate amounts in road safety compared to the other demands on their budgets. In this light, many countries share the need for reliable estimates of the effectiveness of road safety treatments and strategies.

Many countries are moving toward the development of uniform criteria for establishing the effectiveness of road safety investments and infrastructural projects in general. For example, in the European Union, the Directive 2008/96/CE on "Road Infrastructure Safety Management" was published in November 2008 and will have to be implemented, at least on the Trans European Road Networks, in all the Member States. From both a scientific and policy point of view it is important to adopt a similar approach for determining the effectiveness of measures as this will lead to more reliable, credible and accessible tools and methods for the evaluation of safety effectiveness.

A crash modification function (CMF) allows a synthesis of diverse evaluation results that in turn allows for more universal understanding and application of safety effectiveness measures. The fundamental argument for a CMF is that it could allow more rapid adoption and dissemination of new life-saving safety measures. In the current political and administrative climate, the decision making process often demands a system of experimental local evidence and feedback before countermeasures are accepted as effective. This is the so-called principle of "learning by doing."

A properly developed CMF could facilitate this process and give local authorities more confidence in a particular measure, and allows earlier inclusion in strategies and guidelines sooner in the process.

Many decisions, when acted upon, affect road safety. CMFs facilitate the prediction of safety effect. So-called efficiency assessment tools (EATs) can help governments choose those measures that will likely maximize the social benefits of public investment. EATs have been defined as "a systematic assessment of the improvement in road safety that can be realised by means of various road safety measures" and comprise cost-effectiveness analysis and cost-benefit analysis. Cost-effectiveness analysis (CEA) seeks to compare the number of crashes/casualties prevented per unit of cost, for each of the available road safety measures. Cost-benefit analysis (CBA) addresses the question of integral efficiency, and aims at comparing the costs and benefits of different policy alternatives, measured in monetary units.

The main elements of EATs are:

1. A list of road safety measures available for solving a given safety problem.

2. An estimate of the effectiveness, i.e. the CMF, of each measure.

3. An estimate of the costs of each measure.

4. In CBA, a monetary valuation of impacts on safety, environment, travel time.

As noted, CMFs are used in point 2, and constitute an essential element of any efficiency assessment.

Demand for CMFs is growing in many jurisdictions as policy makers are increasingly required to demonstrate results and undertake cost-benefit and efficiency assessments.

Many countries set specific quantitative road safety targets and adopt road safety strategies to achieve these targets, within the constraints of the established priorities and the resources available. Within this framework, the efficiency assessment of road safety measures is considered to be an extremely useful tool in decision making. In particular, cost-benefit and cost-effectiveness analyses are carried out in several countries, in a more or less systematic way. These studies are based on some estimate of the safety effects of the examined measures following the implementation of the measure. However, a more widespread or fruitful use of efficiency assessment of road safety measure is in most cases limited by a lack of knowledge and data on the safety effects of road safety measures.

Nevertheless, the importance of efficiency assessment in road safety is widely recognised, and the need for more knowledge and best practice examples is becoming more and more pronounced. Existing best practice recommendations may cover the whole range of the efficiency assessment process, from the selection and application of appropriate and standardised methodologies to the interpretation of results and the identification of most efficient measures, especially in case different alternative measures need to be compared and ranked. However, the most important uncertainties involved in developing such best practice recommendations concern the adoption of appropriate values for the safety effects of road safety measures.

In the recent years, important research efforts have been made towards the standardization of the methods for estimating the safety effects of road safety measures by addressing some critical issues. The first issue examined concerns the accuracy of the estimation, so that potential bias or other confounders are eliminated. The second critical issue concerns the conditions and necessary adjustments required to allow the transferability of the safety effect estimates to different settings or countries.

This question has become very important at the international level, and particularly within the development of handbooks and manuals aiming to assist decision makers, researchers or other stakeholders involved in the efficiency assessment of road safety measures.

These sources are often used by countries within their national road safety efficiency assessment analyses, by adopting the values proposed (e.g. in terms of percentage reduction of crashes / fatalities, or CMFs), or by adjusting them to the local conditions. However, due to the important gaps in the knowledge concerning the transferability of such values across countries, several counties have developed their own methods and values for assessing the effectiveness of road safety measures.

The knowledge obtained from the international literature may prove very useful in the identification of good practice and cost-effective measure. However, thorough analysis on a case-specific basis is always necessary in order to produce a precise estimate of the effects of a measure in different countries or areas, taking into account the extent of the implementation, the implementation period, and specific national or local requirements. It is also necessary to ensure that such analyses are carried out in accordance with recognised standard methodologies.

Ultimately, efficiency assessment is an important part of the preparation of national, regional or local road safety plans. At the initial stage of evaluation, safety effects are usually unknown and in order to influence any decision making process, the efficiency assessment studies have to be prepared ex-ante, using impact data from previous programs using similar measures.

This stresses the need for strengthening the efforts for the estimation of appropriate values for the safety effects of the treatment examined. Moreover, it highlights the need for increasing the accessibility of this information, through the dissemination of efficiency assessment results on an international basis. Utilising information provided in this report can facilitate greater exchange of this information on an international level.

Lack of reliable knowledge of the effects of countermeasures is a key barrier to the advancement of many critical, life-saving initiatives. Improvement of our knowledge of CMFs will have tangible benefits for decision making.

In some instances, no efficiency assessment is carried out at all during the decision-making process. This is often due to a lack of knowledge about the expected impacts of available safety measures. This view is substantiated by a variety of experiences, most notably by Work Programme 2 of the European thematic network ROSEBUD. The main question of the questionnaire had to do with the reasons why efficiency assessment tools were not always performed. About 30 percent of the responses pointed to technical barriers, most of them connected with the lack of knowledge about impacts. A key conclusion from this example and other discussions is that any improvement in our knowledge of the effectiveness of safety measures, i.e. CMFs, will likely have tangible benefit on the way safety decisions are made. This report aims at providing guidance and support for overcoming this kind of technical barrier.

There is a need for more training and regular practical usage of CMFs to support the development of transferable CMFs. We are currently at a turning point, with the prospect of rapid advances and major cost savings through the transfer of results internationally.

While the understanding of CMFs among countries likely ranges from little knowledge of CMFs to a level of spreading knowledge and growing use of CMFs, it can generally be said that there is currently a lack of a full understanding of the value, importance and usage of CMFs in road safety decision making. At this time, CMFs may be integrated into guidelines and some state, provincial or other local governments may be using CMFs systematically to some extent in their decision making. However, there are currently no countries where CMFs are routinely used nationally in a direct manner by practitioners as part of the planning, design and management of roadways. As a result, there is not yet broad demand for a full library of CMFs from the international road profession. Lack of education, knowledge and practical usage of CMFs is currently the biggest obstacle to CMF development and transferability. However, because the underlying drive for effective safety and effectiveness analysis is taking place in a relatively universal fashion, the demand for reliable estimates of safety effects will continue to grow and the demand for knowledge and information on CMFs should grow in a corresponding fashion.

Transferability of CMFs relies first and foremost on analysing the extent to which a CMF is dependent on the circumstances in which it was developed.

International transferability of the results of road safety evaluation studies will take place most effectively in the ideal situation of studies being available from many countries over a long period, and when all these studies are of at least adequate and similar methodological quality. Many designs are used in road safety evaluation studies. The design is often dictated by the circumstances under which the evaluation was carried out and the skills and resources available. It is therefore unrealistic to expect that all studies have applied designs that are identical down to the finest detail. It is, however, reasonable to require that studies uniformly control for at least the most important potentially confounding factors.

While the ideal situation is difficult or impossible to attain, the report makes it clear that transferability of CMFs depends on knowing the circumstances under which different safety measures have been implemented. Two identical measures, implemented under two identical sets of circumstances, should have the same impact on the frequencies of accidents and casualties. Conversely, differences in circumstances are expected to induce differences in effectiveness.

It is essential that researchers disseminating the results of an effectiveness assessment provide as accurate and complete a description of circumstances as possible. This will allow researchers and practitioners from others regions and countries to evaluate the possibilities of successfully transferring the measure. As far as possible, the information about circumstances should be quantitative; only then can accident modification functions be developed. However, there is not any unique set of circumstances that might be said to be relevant to every research project.

Generally, documentation on a variety of supporting information related to the countermeasure, the development process and conditions under which the countermeasure was tested are valuable. The report presents a specific list of items that are considered essential for inclusion in any study presenting safety evaluation results. The report also provides a full list of all the information that would be desirable to have documented in all CMF reports.

Along with information on circumstances, any study should provide safety estimates by severity of the accidents, the standard error of the estimate of effectiveness, as well as some basic information about methods: study design, sample, data sources, biases, and others.

Variability in CMF research results is a major deterrent to transferability. Reducing variability through proper study design and reporting enhances transferability. Studies should control for the most important confounding factors related to a countermeasure analysed. Variability due to different circumstances can be reduced by making the CMF a function of the relevant circumstances. A key aim of the current report is to provide guidance for uniform screening and control procedures.

When past research indicates that whenever a particular safety countermeasure was implemented instead of some other action, approximately the same safety effect was found, especially under similar circumstances, the issue of transferability in most cases does not arise. Transferability concerns are justified when the same safety effect is not found when the same countermeasure has been applied, i.e. when the variability of the safety effects is large. This concern is valid irrespective of whether the future application is in a different country, city, project or time period.

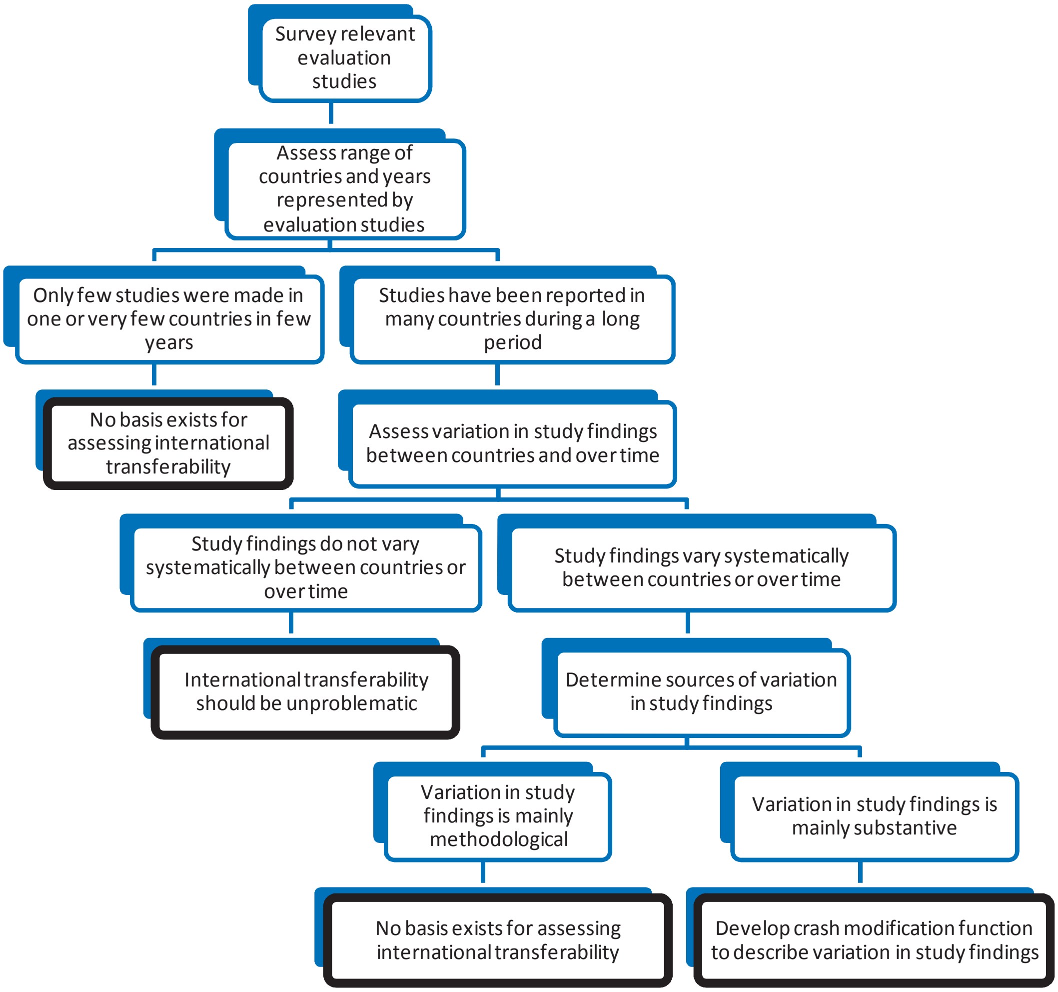
There are two groups of factors that affect the variability of CMFs. One group of factors pertains to the method by which the CMF estimates are obtained. If data is poor, if the sample size is small, if bias and confounding factors are not eliminated, the result will be unreliable. Most statistical attention is paid to this group of factors.

The other group of factors is less commonly examined but is equally important. It has to do with the fact that the same action or measure will have different safety effects in differing circumstances or accidents of different severity. Inasmuch as the CMF estimates we have come from studies conducted in differing times and circumstances, they are bound to differ. They would differ even if data were perfect, the sample size huge, and the experimental method without blemish. The only way by which this source of variability can be reduced is to make the CMF a function of the relevant circumstances. For example, a delineation treatment on curves may be expected to reduce crashes by different amounts depending on the approach speeds and curvature.

To make progress towards reducing the uncertainty about CMFs a two-pronged strategy has to be followed. First, the CMF estimates used to produce the probability distributions have to be reliable. Second, the dependence of the CMFs on the relevant circumstances has to be established. The report indicates one way of trying to answer this question. The answer proposed is that: (1) if there have been many studies of measure X, not just in country A, but in many other countries, and not just six years ago, but spanning three or four decades, and: (2) if these studies obtained highly consistent estimates of the effect of measure X, then: (3) it is more reasonable to conclude that the results of these studies can be applied in country B than to conclude the opposite. In other words, as long as history keeps repeating itself, it is more reasonable to expect it to continue repeating itself than to expect the opposite.

The report describes the range of replications technique and how it can give an indication of the stability of research results across countries and years. The report provides also preconditions that should be fulfilled before applying the range of replications technique. While the applicability of the technique is likely to be limited because of factors such as publication bias, it can be fruitfully applied to assess external validity when a large number of studies have been reported during a long period of time. This is the case with respect to many road safety measures, like road lighting, guard rails, traffic signals, speed limits and seat belts.

Figure 1. Flow-chart for assessing international transferability of road safety evaluation studies



The information contained in the report and summarized above indicate a number of actions that the authors believe could make a difference toward increasing the transferability of CMFs and, ultimately, speed up the process of improving safety on the world's highways.

Road safety policies should undergo performance and efficiency evaluation. Such evaluations cannot be undertaken without Crash Modification Functions (CMFs). Evaluation processes should be documented to ensure they are transparent.

A central element of any cost effectiveness study or cost-benefit study is the requirement to have a reliable or sound estimate of the safety effectiveness of a measure. CMFs are the most effective and supportable measure of safety effectiveness. As such, they offer the greatest opportunity to support a variety of decision making processes. Also, when objective information is used to make safety related decisions, the opportunity to make the processes transparent becomes much greater. The opportunity is greater because substantive, information driven decisions are highly defensible and understandable in the public realm. Non-substantive decisions are generally made in complex political environments behind the scenes and cannot be generally transparent for that reason. CMFs therefore provide a tremendous opportunity to open better and more constructive discussions on the ways and means to address the road safety problem.

Research conducted to develop Crash Modification Functions should follow the guidance provided in this report and, in particular, provide specific information that describes the countermeasure under consideration, the safety issue being addressed and the roadway environment in which it was tested.

Throughout this report, one of the primary issues that hinder the transferability of CMFs and the best practices to improve road safety is the lack of information on the countermeasure being considered and the circumstances under which it was analysed. Without this information, it is impossible to directly understand the safety effects of countermeasure that will be applied in a location different than that where it was previously implemented. Worse yet, it is possible that a lack of understanding can contribute to poor decision making that will lead to an ineffective use of funds or, potentially, the implementation of measures that will result in effective treatments or possibly even to treatments which increase crash frequency or severity. The report has spelled out specific information that should be provided in all reports that identify CMFs. This basic information will be valuable for researchers who want to build upon previous work and for practitioners and policy makers who want to identify countermeasures to address specific safety situations, sometimes in an urgent manner.

It is recommended that an international group be composed under an existing organization (e.g. Transportation Research Board, World Road Association, etc.) to foster dialog among researchers and practitioners on CMF research and reporting standards with the aim of increasing transferability of results. Coordination of research across countries on top priority countermeasures should be considered.

From the beginning of this effort, the group agreed that the final International Transport Forum report would not or could not be a stand-alone end product. Rather, it should establish a starting point for an ongoing process of cooperation and collaboration.

The next step should be in the international review and documentation of CMFs and related supporting information, and in assessment of their quality and potential for transferability. Such efforts would build upon the contents of the present report to continue to enhance and improve research methodologies and approaches for CMF development and reporting. Equally importantly, the report suggests that ideally there would also be efforts made to coordinate or collaborate at an international level on the development of CMFs among countries for high priority countermeasures that several countries have an interest in. Coordination of this type could potentially take many forms, from simply establishing a target countermeasure for research, agreeing on which countries would do independent studies and in what fashion ultimately to be brought back together in a single report. There would also be the possibility of such a group to foster shared research projects that could use “pooled funds” from several countries to develop a single product of value to all participating countries. There are some organizations in the world that would be best suited to convene a group to pursue this work on a sustainable basis. Efforts will be made to inform and enlist these groups as part of the outreach and marketing efforts that follow on the heels of this work.

International cooperation should aim to capture documentation and reporting of CMF research in a widely available transnational database.

International cooperation should advance on the assessment of CMF research results and documentation and reporting of these results should be captured in a widely available transnational database.

Getting the research right so results would be more readily transferable at an international level was the first aim of this work. A concomitant goal was to consider ways to increase availability of CMF information internationally. Ultimately the group believes that an easily available database might be best for people to gain access to this information. Such a database could build upon or be modeled after the work of the CARE database, IRTAD, or the U.S. CMF Clearinghouse. The approach, mechanisms and partnerships for building this database could be established by the group proposed in the previous recommendation.

A concerted effort should be made to publicize the benefits of decision-making based on CMFs. This should take the form of presentations and workshops at transport, injury prevention and health conferences; press releases; letters to political leaders and senior bureaucrats.

The Group believes that efforts should be undertaken to increase awareness, understanding and knowledge about CMFs to foster both greater usage and more international exchange. Members of the group will undertake efforts to promote heightened international information sharing at significant transportation events and elsewhere in the coming years. The Group also recommends that transportation leaders in International Transport Forum member countries support, encourage and promote the development, application and international exchange on CMFs to the fullest extent possible. Initiatives of this sort take the form of supporting appropriate research, creating policies that encourage application, and championing decision making based on reliable and quantitative safety information. Ultimately, all of these efforts can propel the International Transport Forum countries to a broader understanding of safety impacts and improve the effectiveness of investments in safety improvements specifically and road expenditures generally.

ISBN 978-92-821-0375-3
Sharing Road Safety
© OECD/ITF 2012

CHAPTER 1. Introduction

1.1. Background

Each year about 1.3 million people are killed and another 50 million people are injured on roads worldwide. These road crashes cost countries between 1 and 3 percent of their Gross Domestic Product (WHO, 2010). In addition, they cause great emotional and financial stress to the millions of families that are affected by these crashes. A lot of these crashes can be prevented by implementing effective road safety measures. To be able to select the best measure, a decision maker needs information about the effectiveness of different measures. Moreover, information about the effectiveness of measures is needed in case one (a politician) has to make a choice between expenses on road safety measures and other expenses, like measures to limit environmental effects of traffic or measures to improve traffic throughput.

Generally, countries try to establish a clear understanding of the costs and benefits of preventing road crashes. Decision makers need to choose between a variety of safety measures; understanding the effectiveness of different safety treatments and strategies (often referred to as Crash Modification Factors or Functions, CMFs) will help in making better decisions. Hence, CMFs are one of the tools to help professionals make safety decisions.

In this light, many countries share the need for reliable estimates of the effectiveness of road safety treatments and strategies. The development of reliable estimates is costly and time consuming, thus placing a burden on countries to develop full sets of measures independently of other countries. Therefore, it is important to share information between countries about the effectiveness of measures.

Many countries are moving toward the development of uniform criteria for establishing the effectiveness of road safety investments and infrastructural projects in general. From a scientific point of view it is important to adopt a similar approach for determining the effectiveness of measures as this will lead to more reliable, credible and widespread tools and methods for the evaluation of safety effectiveness.

The effectiveness of a measure depends on the local context in which a measure is implemented. The road system is managed under a regime that is adapted to its environment for regulating road safety risks. The capacity for change is constrained by strong and coherent forms that are intrinsic to the regime. These regimes vary from one country to another and will exhibit specific attributes that make sense for the local organisation of road safety in each individual country (Delorme, Lassarre, 2009). For example, the speed camera-based automated policing systems in France and Great Britain take different forms. Each organisation appropriates the technological innovation of associating a camera with a laser and a registration number recognition system according to its own rationale and structures.

The end-result is a centralised system overseen by a Prefect in France. In Great Britain, the camera network is distributed and managed at the local level by the police (Carnis, 2010).

A single crash modification factor applied to each speed camera site will depend on the characteristics of traffic flows and locations. In Great Britain, the overall impact can be assessed by aggregating local impacts as both the police and users consider speed cameras to be an ad hoc measure covering a limited portion of the highway or road which will catch delinquent drivers on local roads. In France, speed cameras are part of a whole which constitutes a system in the eyes of the State and users. Because of the diverse nature of the environments, applications and other influences that act upon an effectiveness measure, this report will focus on the development and use of crash modification functions (CMF).[1] A CMF allows a synthesis of diverse evaluation results that in turn allows for more universal understanding and application of safety effectiveness measures. The fundamental argument for a CMF is that it could allow more rapid adoption and dissemination of new safety measures. In the current context of safety countermeasure adoption, the local context often demands that the adoption requires a system of experimental local evidence and feedback that can serve as a localised impact evaluation. This is the so-called principle of "learning by doing." A properly developed CMF could facilitate this process and give local authorities more confidence sooner in the process of adopting a new strategy.

This report is intended to facilitate international collaboration to address this common need and to stimulate greater efficiency in the development and dissemination of reliable safety effectiveness estimates. The primary objective is therefore to evaluate opportunities for and obstacles to international collaboration on the development of CMFs. An examination of the international transferability of and access to the results of effectiveness assessments is included. Further, the report develops a theoretical basis for assessing countermeasure effectiveness and a framework for assessing the confidence that can be placed on crash reduction estimates based on the quality of individual studies and the consistency of their conclusions. This report therefore deals mainly with that part of the knowledge creation process which begins with the conduct of ex-post evaluations and leads to meta-analyses and to theorizing. The report also examines the availability of cost-effectiveness assessments of road safety interventions, and reviews the quality and transferability of the estimates available. Ultimately, the report provides recommendations that can improve and harmonise research methods and reporting standards, and thereby increase the potential for transferability and mutually beneficial ongoing international collaboration.

This report targets the road safety research community. Specifically, the research framework and other elements of the report are intended to provide guidance to researchers to ensure that their research results are more readily acceptable and transferable at the international level. Another target audience is policy makers in countries, states and provinces who can influence the development and application of CMFs. In particular, the report seeks to convince policy makers of the value and importance of CMFs as well as the need to make more and better CMFs available to practitioners.

Governments and organizations at different levels (transnational, national, regional, local) play an essential role in stimulating the demand for CMFs and creating a system of decision-making where scientific evaluation of the effectiveness of measures is an important variable in the final outcome.

1.2. Crash modification functions: Their role and nature

Many decisions affect road safety. The decisions may be whether or not to mandate the use of daytime running lights (DRL) in a region, whether to design a road curve with a 100m or a 150m radius, or whether to lower the legal BAC limit in a country from 0.08 to 0.05 mL/L.

The decision may even be one where the impact on safety is only a by-product of a contemplated action such as, for example, a decision about where to put a new sports stadium or whether to convert a traffic lane to a bus lane.

Decisions are influenced by various considerations, one of which is their impact on road safety. Road safety is usually measured in terms of the expected number of crashes classified by severity. For example, the decision about DRL implementation will be influenced by a comparison of the number of crashes by severity that one should expect with and without DRL. The difference between road safety with and without DRL is said to be the 'safety effect of DRL'.

Crash Modification Factors or Functions facilitate the prediction of safety effects. For example, if in a region of average latitude 43° one expects the introduction of year-round DRL to reduce daytime fatalities by 7 percent, then the corresponding CMF is 0.93. In this case, the expected 'safety effect of year-round DRL' for fatalities is the product: (Expected fatalities during daytime and twilight without DRL)×(1-the CMF of year-round DRL for fatalities). Similarly, if on a two-lane rural road where the speed limit is 90 km and where the angle between adjacent tangents is 123°, and if increasing the radius of the curve from 100m to 150m is expected to reduce injury crashes by 17 percent, then the CMF is 0.83 and the expected safety effect of this increase in radius is the product: (Expected injuries with 100m radius)×(1-0.83).

It follows that:

1. The CMF always pertains to some action or measure and two possible future states (with DRL legislation or without one; with a 150m radius or a 100m one; with the legal blood alcohol content (BAC) at 0.05 or at 0.08mL/L.)

2. The CMF always pertains to target crashes of a specific type and severity category (e.g. fatal crashes in a region for DRL, all injury crashes on the curve, crashes where alcohol was a causal factor for blood alcohol content (BAC) limits)

3. The CMF depends on various details and circumstances. For example, the CMF for DRL will depend on whether their use is required year-round or only in winter, on the extent of enforcement, on road user compliance, on the latitude of the region (duration of twilight), etc. Similarly, the CMF for the radius of horizontal curvature will depend on the two radii, on the approach speed, on the angle between the tangents, on whether the road is urban or rural, etc.

4. The expected safety effect depends not only on the CMF but also on the expected future number of target crashes, a quantity that is also imperfectly known.

1.3. CMFs come from the accumulation of knowledge

On what basis can we expect that implementing year-round DRL in a certain region will reduce fatalities by 7 percent? Estimates of this kind come from the gradual accumulation of research results. Initially, the promise of DRL was based on a general understanding of how detection depends on conspicuity, and how gap acceptance depends on headlight use, etc. Information from 'lights-on' road sections, from mandating the use of a headlight by motorcycles and from several 'fleet studies' followed. These already allowed the estimation of CMFs. The first wide-scale implementations of DRL came in the late 1960s leading to country-wide CMF estimates. As more countries introduced the use of DRL and conducted the corresponding evaluation research, more CMF estimates were obtained.

Eventually a meta-analysis could be done to obtain a characterization of what may be expected ‘on the average’ and how the CMFs vary from one implementation to another. A theory even emerged that endeavoured to explain how the CMF depends on the latitude of the region.

In this example, as in many others, the CMF for a certain action and crash type arises initially from a hunch that is based on general knowledge and small scale experiments. Further knowledge is obtained after the implementation of the action, and after ex-post evaluations. The results of ex-post evaluations are several different CMF estimates, which can then be examined by meta-analyses. A meta-analysis uses the diverse results of published research to provide an estimate of the ‘average CMF’ and a characterization of how widely the CMFs vary from one implementation to another. If the details of the action and the circumstances of the implementation are known for each ex-post evaluation, then the same results can be used for drawing conclusions about how the CMF depends on various variables.

1.4. Using CMFs in decision making

CMF estimates from past evaluation studies are the basis on which decisions about future actions are founded. For example, assume that one has to decide whether to implement a certain action or measure. Suppose that in four previous evaluation studies of that measure the estimates of the CMF were 0.83, 0.63, 0.63, and 0.54 with the corresponding standard errors of ±0.07, ±0.11, ±0.05, and ±0.09. On this basis one can say that the CMFs for this action or measure – based on meta-analysis – have a probability distribution with a (weighted) mean of about 0.66 and a standard deviation of about ±0.08. How these numbers were obtained will be explained in a later section. In short, past research results are the raw material from which the estimated probability distribution of the CMFs for an action or measure is built.

This leads to an important insight about the key feature of the situation: while the decision whether to implement an action or measure is best based on the estimate of the expected CMF (0.66 in the example), the actual CMF will most likely be different. In this example it will be somewhere in the range $0.66 \pm 2 \times 0.08$, i.e., between 0.48 and 0.84. In short, the actual CMF, which will determine the outcome of a future action, is not the average CMF of past estimates. Several observations follow.

First, because the actual CMF and the average of past results are not the same, there is always the possibility that the ‘implement’ or ‘do not implement’ decision will be wrong. Wrong decisions have real consequences. If the decision to ‘implement’ is wrong, money is wasted; if the decision ‘do not implement’ is wrong, an attractive opportunity to save life and limb is squandered.

Second, whether the decisions we make are right or wrong depends on the width (the standard deviation) of the probability distribution of the CMF. The narrower the CMF distribution, the larger the probability that decisions we make are correct. This observation goes to the core of this report; it declares that the role of research about CMFs is to reduce the width of the CMF distributions. The main question before us is how to conduct research about CMFs so that good progress is made in this direction.

Third, research that reduces the standard deviation of the probability distribution of a CMF has value; its value resides in reducing the chance of making an erroneous ‘implement or do not implement’ decision. Because erroneous decisions have real costs, such research has real value. This value can be estimated and can be used for determining research priorities.

1.5. Understanding variability in CMFs

There are two groups of factors that affect the width of the CMF distribution. One pertains to the method by which the CMF estimates are obtained. If data are poor, if the sample size is small, and if bias and confounding factors are not eliminated, the result will be unreliable. Most statistical attention is paid to this group of factors.

The other group of factors is less commonly examined but is equally important. It has to do with the fact that the same action or measure will have different safety effects in differing circumstances. There is therefore no reason to think that DRL will have the same effect in Norway as in Italy; nor can one reasonably expect that the safety effect of a curve radius does not depend on the number of lanes, size and quality of shoulders, prevailing weather and speed, visibility distance, availability of ESC on most vehicles, and on whether the preceding tangent is 100m or 10,000m long. Inasmuch as the CMF estimates we have come from studies conducted in differing times and circumstances, they are bound to differ. They would differ even if data were perfect, the sample size huge, and the experimental method without blemish.

The only way to reduce this source of variability is to make the CMF a function of the relevant circumstances. Thus, if the safety effect of DRL depends on geographic latitude, then this functional dependence has to be established. Only then can a decision-maker for Italy make good use of information about what the DRL effect was in Sweden. The same is true for the function linking the safety effect of curve radius to the number of lanes and tangent length. The engineer designs a curve for a road with a certain number of lanes and tangent lengths and it is for this circumstance that the CMF is required. It is for this reason that the letter F in the acronym CMF can stand for both Function and Factor. When the CMF values in a handbook are tabulated for various circumstances or given by a mathematical expression in which the circumstances are variables, F stands for Function. When specific circumstances or variable values are used in the function, F stands for Factor.

To make progress towards reducing the uncertainty about CMFs a two-pronged strategy has to be followed. First, the CMF estimates used to produce the probability distributions have to be reliable. Second, the dependence of the CMFs on the relevant circumstances has to be established.

1.6. The benefits of transferability

The development of reliable CMFs is costly and time consuming. A typical project to develop a reliable CMF related to roadway features in the United States, for example, will cost about $US 200,000 and can take more than 10 years to complete. At this cost, the CMF will be based on a single study consisting of a set of unique circumstances. While the cost per countermeasure may be reduced by economies of scale – i.e. developing more than one at a time – it would currently not be less than $100,000. Additionally, some evaluations can cost considerably more. The ultimate cost of any single countermeasure will depend upon the magnitude of the expected effect, the statistical significance level desired, the expected number of crashes at the sites where the countermeasure is typically installed, and the sample size required (i.e. number of sites and duration of before and after periods).

A specific example from the United States includes one project to develop CMFs (factors) for four pedestrian countermeasures; this is estimated to cost about $500,000 or more. Another project designed to develop a CMF (factor or function) for intersection sight distances is estimated to cost about $450,000.

This level of effort places a burden on countries to develop full sets of measures independently of other countries. While it is natural that countries will often want to either verify CMFs from other countries or to develop their own CMFs before accepting the full scale adoption of countermeasures, it is not and should not be a requirement. Particularly at a time of economic uncertainty, governments want to find ways and means to maximize results with the limited funds that are available. The development of CMFs that can be readily understood, accepted and applied in multiple countries is something that all countries could benefit from.

Will the safety effect of DRL in Italy be similar to what it was in Sweden? Will reducing the mean speed by a certain amount have the same effect in Europe and in North America? Will paving shoulders on two lane roads in Ontario be the same as in Alabama? These are questions of 'transferability'.

One may reasonably expect that an action taken in identical circumstances will have identical results. However, circumstances are never identical. Italy differs from Sweden, Europe from North America, and Ontario from Alabama. To what extent these differences matter is an empirical question. Research shows that the safety effect of DRL depends, among other things, on geographic latitude. It is this kind of empirical evidence that enables us to make use of results from Sweden, Finland or Canada and tailor them for application in Italy. Research also shows that reducing the mean speed by a certain amount had a similar effect on safety in Europe as in North America. This kind of empirical evidence allows one to pool and apply research results across jurisdictions. In this sense, the question of transferability amounts to asking to what extent does the CMF of some action or measure depend on the circumstances that characterize a certain jurisdiction. The greater our understanding of how a CMF depends on circumstances and variables, the less acute the question of transferability will be.

1.7. CMFs and efficiency assessment tools

Policy makers are often faced with the question of how to get the most out of scarce resources. So-called efficiency assessment tools (EATs) can help governments choose those measures that will likely maximize the social benefits of public investment.

EATs have been defined as "a systematic assessment of the improvement in road safety that can be realised by means of various road safety measures" (ROSEBUD, 2006), and they comprise cost-effectiveness analysis and cost-benefit analysis.

Cost-effectiveness analysis (CEA) seeks to compare the number of crashes/casualties prevented per unit of cost, for each of the available road safety measures. Assuming a given budget, CEA identifies the set of measures that will maximize the total number of crashes/casualties prevented. Alternatively, if quantitative road safety targets are set as an initial constraint, CEA will help identify the measures that will minimize the cost of achieving such a target. A major drawback of CEA is that it cannot be applied when measures affect crashes of different severity in different ways, nor can it properly account for travel time and environment impacts.

Cost-benefit analysis (CBA) addresses the question of integral efficiency, and aims at comparing the costs and benefits of different policy alternatives, measured in monetary units. Measures for which benefits are greater than costs are called cost-effective, and ranked according to their benefit-cost ratio. Unlike CEA, CBA allows an explicit consideration of costs and benefits in domains such as travel time, energy consumption, gas emissions and noise. Also, crashes and casualties of different severities (e.g. deaths, serious injuries and slight injuries) can be handled in a single result.

In a CBA, the numbers of casualties in each category, prevented as a result of a safety measure, are transformed into monetary units, added and introduced in the numerator of the benefit-cost ratio. A CEA requires computing three different ratios.

The main elements of EATs are:

1. A list of road safety measures available for solving a given safety problem.

2. An estimate of the effectiveness, i.e. the CMF, of each measure.

3. An estimate of the costs of each measure.

4. In CBA, a monetary valuation of impacts on safety, environment, and travel time.

CMFs are used in point 2, and constitute an essential element of any efficiency assessment.

EATs can be applied to any kind of road safety measure, regardless of their cost and the size of the affected population. Thus, they cover a wide range of measures, from national road safety programmes to the safety treatment of a given junction. Ideally, EATs are one step —a major one, indeed— in a process of decision-making that would start with the identification of major road safety problems and the screening of available measures, then follow with the efficiency assessment and the choice of preferred alternatives, and end up with the implementation of measures and the assessment of their actual effectiveness. Guidelines for the application of CEA and CBA have been developed elsewhere (ROSEBUD, 2006; Highway Safety Manual, 2010).

Box 1.1. "Treatment life"

One of the key elements required for producing cost-benefit analysis (CBA) for treatments is knowledge of how long such treatments will continue to deliver safety outcomes. Sometimes termed 'treatment life', little accurate information is available on this topic. In a review of current Australian experience of this topic, Turner and Comport (2010) identified great variation in the treatment life values used in different states. In some cases the treatment life was almost double for a specific treatment when comparing one state with another. This has a substantial impact on the CBA for that treatment. Although far from robust, Turner and Comport (2010) provide some advice on this issue. However, further research is required on this subject to provide better guidance.

1.8. Applying efficiency assessment tools in the real world

So far, CBA and CEA have been applied to many varieties of road safety measures (ROSEBUD, 2006). These analyses prove that a great number of road safety measures have benefits greater than costs - far greater in some instances. Research suggests that policy-making based on a strict application of efficiency assessment can lead to a significant reduction in the number of crashes and casualties, even in those countries with the best safety records. For example, it has been estimated that the implementation of every cost-effective measure in Norway and Sweden might result in a 50-60 percent reduction in the number of fatalities over a 10 year period (Elvik, 2003). In the Netherlands, the potential benefit of cost-effective measures might be a 65 percent reduction in fatalities (SWOV, 2009).

Real political decisions, in many instances, are not knowledge based and do not conform wholly to the principles and results of efficiency assessments.

In some cases, CBA or CEA have indeed been carried out at the early stages of the process, but the final decision is not based strictly on these results. There may be a variety of reasons for this fact, and, even though in some instances political opportunism, vested interests or lobbying may play a role, it is important to realize that governments may have other legitimate interests different from overall efficiency. Regional disparities and inequities among groups of road users are some primary examples of safety problems that cannot be appropriately handled with efficiency assessments. Obviously, this does not mean that CMFs and information about the effectiveness of available measures are useless. Although the exact values of CBA or CEA ratios lose weight in the final decision, measures that have proved effective for treating the target problem will always be desirable and preferred to ineffective ones.

In some cases, no efficiency assessment is carried out at all during the decision-making process, due to a lack of knowledge about the expected impacts of available safety measures or due to preconceived assumptions. This report aims mainly at mitigating this kind of technical barrier.

While good and sound research can lead to reliable CMFs that can be used to predict the safety effects of specific countermeasures or groups of countermeasures, it is also sensible to understand the important role of solid practical experience in decision making. Specifically, studies may be carried out in a context or environment that generates a CMF that, for those conditions, may make sense. However, the practitioner with more experience may consider the CMF to be too low or too high for that context or situation.

While most CMFs should be consistent with practical experience more universally and reinforce the practitioner's experience, it is possible that there could be occasions when the CMF is counter to this experience in a positive or negative manner. In these situations it is advisable for the practitioners or policy makers to seriously revisit their perceptions of effectiveness and determine if they can accept a CMF or not. A decision not to accept a CMF based on a sound and reasonable study should be seriously considered and well supported.

In a similar vein, the topic of combining countermeasures in a single estimate using CMFs is discussed elsewhere in this report. There is no question that any estimates based on a combination of CMFs should be closely examined for their reasonableness. An experienced professional may or may not agree with these predicted values; s/he would not be expected to promote exaggerated predictions and would be expected to acknowledge where greater benefits could be achieved. Ultimately, the quantitative science can be balanced against good practical knowledge and experience in a reasoned manner.

In Work Programme 2 of the European thematic network ROSEBUD, a survey was used to explore the barriers that prevent efficiency assessment tools from being more often applied to real-life policy-making (Elvik and Veisten, 2004). Eighty-three persons from seven different countries responded to the questionnaire, representing the national and regional/local levels, and belonging to transport ministries, road administrations, regional and local institutions, research institutes and transport consultancies. They were involved in making decisions on road safety priorities, or in developing methodologies for efficiency assessments. The main question of the questionnaire had to do with the reasons why efficiency assessment tools were not always performed. About 30 percent of the responses pointed to technical barriers, most of them connected with the lack of knowledge about impacts.

A key conclusion of the previous discussion is that any improvement in our knowledge of the effectiveness of safety measures, i.e. CMFs, will likely have tangible effects on the way safety decisions are made.

1.9. Government has a critical responsibility

At present, the power to implement road safety measures resides mainly in public institutions. Governments at different levels (transnational, national, regional, local) must thus play an essential role in stimulating the demand for CMFs and creating a system of decision-making where scientific evaluation of the effectiveness of measures is an important variable in the final outcome. This by no means implies giving up other legitimate political interests that are hardly manageable by conventional efficiency assessment tools.

In many countries, it is mandatory to conduct a cost-benefit analysis for large scale road projects. For example, in the European Union, the Directive 2008/96/CE on “Road Infrastructure Safety Management” was published in November 2008 and will have to be implemented, at least on the Trans European Road Networks, in all the Member States. The Directive requires that the selection of safety treatments must account for the cost/benefit ratio and encourages the definition of guidelines to be shared among the Member States to assess the safety effectiveness and cost effectiveness of different treatments through before-after analyses.

Along with the elaboration of road infrastructure projects, the design of road safety programs increasingly requires scientific approaches based on effectiveness evaluations. In general, it is considered still too early to extend the binding nature of CBA to every field of road safety policy-making (Hakkert and Wesemann, 2005).

However, it is the task of governments to promote the use of safety assessments in as many areas as possible. Furthermore, governments can, by commissioning systematic and independent assessments of the effectiveness of every measure implemented, contribute to significantly improving our knowledge about the impacts of safety decisions.

Box 1.2. Decision making in road projects

Decision making related to road projects and programs is very complex with a variety of players and issues. As a result, safety may not always be a top priority and safety predictions based on valid research may not always make a difference. For example, roadway lighting is historically known to provide excellent safety benefits. One CMF for lighting that can be found on the FHWA CMF Clearinghouse is 0.51 – illustrating a very strong potential for crash reductions. However, there are many reasons why the lighting countermeasure might not be selected for a system or project, including the following.

1. Economics: While it is not too expensive to install roadway lighting, many communities choose not to install roadway lighting because of the ongoing electrical and other maintenance costs that they may not be able to afford over the long term.
2. Lack of Public Support: There could be a sizeable portion of the local community that is concerned about “light pollution” and its effects on the night-time sky.
3. Environment: Roadway lighting may have a negative impact on nocturnal habits of local wildlife and might not therefore be a good option in some scenarios.

While all of these decisions can be made in a vacuum, they can also be made with the informed knowledge that a number of crashes could be prevented. If that number is not sizeable, then some other priority might take precedence. If the number is sizeable, safety might take precedence. In either case, the final decision is much more highly informed with the advantage of reliable estimates.

KEY MESSAGES

- 1.3 million people die and 50 million people are seriously injured on roads every year in the world.
- Many of these crashes can be prevented by implementing effective road safety measures.
- At the same time, we need to make sure the limited funds we have available are used effectively.
- A significant barrier to the implementation of safety measures is a lack of knowledge about the expected impacts of safety measures.
- A properly developed Crash Modification Factor or Function (CMF) can facilitate effective decision making through understanding the safety effects and cost-effectiveness of our decisions.
- CMFs are the basis for evidence based safety policies.
- Two factors affect the width of CMFs: lack of data or lack of information on circumstances.
- An improvement in our knowledge of the effectiveness of safety measures will have tangible effects on the way safety decisions are made.
- It is mandatory in many countries to conduct cost-benefit analyses (that rely on good safety effectiveness measures) for large scale projects.
- It is the task of governments to promote the use of safety assessments (governments have a key role).
- It is important to have access to, and to use the best available information.
- The development of reliable CMFs is costly and time consuming. All countries could benefit from the development of CMFs that can be applied in multiple countries.
- CMFs help in the dissemination of good practices.
- International transferability of this information is crucial.

NOTES

1. The more commonly used term is Crash Modification Factor. The Working Group felt that this term was somewhat limiting inasmuch as the safety effect of some action is often a function of various circumstances and therefore the broader term is used here. The distinctions between 'factor' and 'function' are covered later in the report. Terms such as 'accident modification factor' and 'accident reduction factor' are deemed to be equivalent to crash modification factor or function.

REFERENCES

Carnis L. (2010), "A Neo-institutional Economic Approach to Automated Speed Enforcement Systems", *European Transport Research Review*, (2)1: 1-12.

Delorme R., Lassarre S., (2009), Les régimes français et britannique de régulation du risque routier : la vitesse d'abord. Synthèse INRETS No. 57.

Directive 2008/96/EC of the European Parliament and the Council of 19 November 2008 on road infrastructure safety management.

Elvik, R., Høye, A., Vaa, T., Sørensen, M. (2009) *The Handbook of Road Safety Measures, second edition.* Emerald Group Publishing Limited, Bingley.

Elvik, R. (2003) *How would setting policy priorities according to cost–benefit analyses affect the provision of road safety?* Accident Analysis and Prevention, 35, 557-570.

Elvik, R.; Veisten, K. (2004) *Barriers to the use of efficiency assessment tools in road safety policy.* ROSEBUD (*Road Safety and Environmental Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision-Making)*. WP2, report NºD4.

Hakkert, S. & Wesemann, P. (eds.) (2005). *The use of efficiency assessment tools: solutions to barriers. Workpackage 3 of the European research project ROSEBUD*. R-2005-2. SWOV, Leidschendam.

Highway Safety Manual (AASHTO, 2010; http://www.highwaysafetymanual.org).

SWOV (2009) Use of road safety knowledge by policy makers, Fact Sheet. Based on Wegman, F. (2001) *Veilig, wat heet veilig?* SWOV-visie op een nóg veiliger wegverkeer. R-2001-28. SWOV, Leidschendam.

Thematic network ROSEBUD (*Road Safety and Environmental Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision-Making*): Several authors (2006) Deliverable WP5: Recommendations (p.52).

Turner, B & Comport, L 2010, *Road safety engineering risk assessment part 4: treatment life for road safety measures*. Austroads report AP-T149/10. Austroads, Sydney, Australia.

WHO (2009), *Global status report on road safety: time for action*, World Health Organization.

ISBN 978-92-821-0375-3
Sharing Road Safety
© OECD/ITF 2012

CHAPTER 2. CHALLENGES AND OPPORTUNITIES FOR TRANSFERABILITY

This chapter reviews some of the technical challenges for international collaboration on crash modification functions. The chapter identifies opportunities to overcome obstacles for the transferability of road safety knowledge and encourage positive action to support transferability. It discusses the importance of the proper planning and documentation of research in improving international transferability of CMFs.

2.1. Introduction

Crash Modification Functions are growing in relative importance as their value in safety analysis has become more apparent. However, the high quality analysis that is required to develop a reliable CMF is costly to perform. Governments are therefore challenged to set policies or establish guidelines that will improve the overall quality of CMF research and results. The growing knowledge of, interest in, and application of the quantitative assessment of safety countermeasures is an opportunity to increase international cooperation in the development and sharing of CMFs. The benefits of this cooperation include the ability to maximize research investments among countries, and more rapid global dissemination and use of life saving countermeasures.

Given the potential benefits that would ensue from international collaboration, one might wonder why more has not been done to date. One reason is that the concepts related to the quantitative assessment of safety effects are still relatively nascent for practitioners and not widely understood. As these concepts have evolved, progressed and become much more refined, so too has our understanding of what makes a good CMF. In addition, a full understanding of the value of performing safety analyses utilising CMFs is only now truly beginning to permeate the highway community. In some regards, it is only now that quite a few OECD countries have passed a knowledge tipping point of sorts that invites the creation of a more cohesive approach to the development of CMFs internationally.

Another reason for the limited official initiatives at both national and international levels could be that road authorities may fear that ex-post evaluation of measures may prove that important road safety investments had little or limited impact with potential consequences for both the political and administrative authorities responsible for the programs. In addition, the comparison of a measure’s cost effectiveness between different regions and between different countries may reveal high discrepancies not only in the unit cost of the measure but also in the implementation effort, thus generating questions about the practices used not only by the authorities but also by the industry. While these concerns may exist, they are natural outcomes of our increased knowledge of and approach to estimating effectiveness. As a result, such concerns must be overcome.

This chapter reviews some of the technical challenges for international collaboration on CMFs and discusses the importance of the proper planning and documentation (including the circumstances under which the CMF was developed) of research in improving international transferability of CMFs. It also sets the stage for Chapters Three and Four, which explain how the transferability of road safety evaluation studies can be assessed and which propose a framework for enhancing the transferability of CMFs.

2.2. Technical challenges

The overarching challenge presented by the interest in international sharing of CMFs is to determine what government agencies in OECD and International Transport Forum member countries can do to make assessments of road safety measures systematically transferable and their results internationally acceptable.

Given that CMFs will often be generated through academic research or private sector project analysis, a challenging role for Government would likely be to develop and disseminate policies and guidelines that would influence the conduct of research and the appropriate documentation of results.

It is quite understandable that many factors influence the quality of research that is performed in any field. For CMFs, this is certainly true. For example, the U.S. CMF Clearinghouse shows that a study design that is statistically rigorous and includes a reference group will have an effect on the outcome of CMF-related research. Other factors such as the sample size, standard error, potential bias and the data source all have a great influence on the quality of a final result. It is in these areas and others of similar nature where the application of appropriate influence will help to both improve the final research products as well as their international acceptability (transferability?). The challenges associated with identifying the full set of essential factors and then educating appropriate audiences about them are large.

Communicating the value of certain countermeasures across international boundaries and seeking their rapid adoption is a challenging prospect if certain specific information is not presented. For example, an important element in the development and application of CMFs is the identification of the target crashes and crash severities for which the CMF is most applicable. Including such information in CMF research reports could increase the value of a report and ease the applicability of a CMF in a country other than where it originated. Other specific information that can enhance the perceived value or significance of a countermeasure is information on the circumstances in which the CMF was developed. For example, information on the type of area (e.g. urban, rural, both) for which the CMF is applicable is essential information that should be documented. Other circumstantial information that would help practitioners in one country more quickly adopt countermeasures from other countries includes roadway geometry, traffic volumes, roadway functional classification, or other safety treatments applied at the site. Without such information, the transferability of CMFs across borders is significantly hampered.

Furthermore, international recommendations and guidelines for the necessity of and the procedures of CMF analyses may prove to be very beneficial to countries with high inertia to change current practices that involve no evaluation and no accountability of road safety investment efficacy. Based on international recommendations, the first step may be to make CMF evaluation a required or recommended procedure for all road safety investments and, afterwards, to link any subsequent investments with the CMF results of the previous investments. The second step may well be to use a standard and uniform CMF evaluation procedure as established through continuous international cooperation in the field. Obviously, establishing a regular CMF development and usage procedure would be extremely challenging.

Assuming an ideal situation in which precise CMFs were available for all of the relevant measures, there would still be some technical difficulties - with ongoing research on how to handle them - that the practitioner should be aware of. The first kind of difficulty concerns the most appropriate way of combining CMFs. Most evaluation studies deal with the effectiveness of individual measures. Thus, when we say that the CMF for the channelization of junctions is 0.85, and the CMF for the substitution of yield signs into traffic lights is 0.70, we are implying that, on average, we expect a 15 percent reduction in the frequency of crashes, if *only* the channelization is implemented, and a 30 percent reduction, if *only* the conversion of yield signs is implemented. However, in many real-life cases, several measures are implemented simultaneously. Assume that, in a given set of junctions, channelization is carried out and, *at the same time*, traffic lights replace yield signs. The relevant question is how the two CMFs should be combined so as to best forecast the expected safety performance of the treated junctions.

Box 2.1. Combining CMFs

More than one example exists for combining multiple CMFs or for accurately predicting the effects of multiple countermeasures being used in one site. In the United States, discussion of combining CMFs for multiple treatments continues. Currently it is acceptable to use a calculated reduction that multiplies together the effects of each individual treatment. Because such an approach can reasonably be expected to overestimate the total effects (Mounce, 2005), a rule of thumb is to limit such multiplicative combinations to no more than 3 separate independent countermeasures.

Australian research identified that 4 out of 5 treated sites were treated with multiple treatments. It has therefore been recommended that efforts be taken to consider “packages” of countermeasures – i.e. multiple individual countermeasures that are typically used together to address a specific problem area. For example, chevrons, shoulders, markings and guardrail may typically be used together to address horizontal curve safety. For these packages of countermeasures, it is suggested that standalone CMFs for the entire group be developed and applied when estimating benefits.

The question of the impact of combined measures pertains to many safety decisions, and is particularly critical when assessing the outcome of road safety programmes and strategies, where packages of - usually many - measures have to be evaluated. Yet, our current knowledge about this technical issue seems to still be limited (Elvik, 2009). A common way of proceeding is to assume that the effectiveness of a certain measure does not depend on whether it is implemented as a stand-alone measure or as a component of a package of measures. Under this hypothesis, the CMF for a combination of measures is simply the product of all the CMFs. Considering the junction example above, if the channelization was implemented first, the number of crashes would be 85 percent of the initial value. If yield signs were then replaced by traffic lights, the final number of crashes would be 85 percent*0.70=59.5 percent of the initial value. However, other approaches are possible (Elvik, 2009) and it seems sensible to always conduct a sensitivity analysis. Of course this may overestimate the impact.

Another caveat should be made regarding the use of CMFs for assessing the impact of road safety programmes, which typically require a forecast of the number of crashes and casualties several years ahead from the base year. For example, we might be interested in setting a quantitative target for the year 2020, and in assessing the way several measures could help achieve such a target. For each measure considered, the key question is what the expected number of crashes would be in 2020 in two scenarios: if the measure is implemented, and if it is not.

A common method of dealing with this question consists of two steps: first, the determination of a base scenario of the evolution of the number of crashes, usually by extrapolating past trends (business-as-usual scenario); second, the multiplication of the number of crashes in the base scenario by the CMF for the measure considered. This method can be suitable for many situations. However, it should be borne in mind that many measures included in road safety programmes do not constitute truly new measures, but rather a continuation or intensification of already implemented policies. In those cases, there is a risk that the estimated reduction of crashes is biased by a double-counting of safety benefits, since the effects of the measure affect both the trend used in the determination of the base scenario and the CMF that is used to multiply the number of crashes. Correcting this bias may require the development and application of complex analytical methods.

A comprehensive review of the current state of knowledge and practices concerning the prediction of road crashes and the use of effectiveness assessments in road safety programs can be found in the special issue of the journal Safety Science, titled "Scientific Research on Road Safety Management" (Safety Science, Volume 48, issue 9).

Finally, there is a third issue related to the use of monetary valuation of safety benefits in cost-benefit analysis. Some international comparisons show that official estimates of the value of a statistical life vary by a factor of almost 60 between the countries with the highest and lowest estimates (European Road Safety Observatory, 2006). This fact reflects fundamental differences in the methods of evaluation, particularly regarding so-called human costs. Generally, countries with the highest estimates of the value of a statistical life based their evaluation of human costs (grief, pain...) on the 'willingness to pay' method, while countries with the lowest estimates usually apply the average compensations to victims and families dictated in courts. An unwanted consequence of these discrepancies is that measures regarded internationally as cost-effective and best practices may in some countries appear to have costs greater than benefits. Two possible ways of handling this are to revise the monetary valuation of crashes or to favour cost-effectiveness analysis, instead of cost-benefit analysis.

2.3. Challenges for transferability

One of the primary obstacles to international cooperation in the sharing of CMFs is the lack of a uniform understanding of the value, importance and usage of CMFs in road safety decision making. The understanding of CMFs among countries likely ranges from little knowledge of CMFs to increasing knowledge and growing use of CMFs. At this time, CMFs may be integrated into guidelines and some state, provincial or other local governments may be using CMFs systematically to some extent in their decision making. While some countries are applying CMFs, there are currently not many countries where CMFs are routinely used nationally in a direct manner by practitioners as part of the planning, design and management of roadways. There is a growing demand for a full library of CMFs from the international road profession. Lack of education, knowledge and practical usage of CMFs is currently the biggest obstacle to CMF development and transferability.

Among other obstacles that exist is the lack of uniformity in the performance of related research and the reporting of research results. While there are reasonably good channels of communication for the research and other communities that are responsible for developing CMFs, there is no international venue that promotes consistent, global approaches that will optimize the sharing of research practice and results. The Transportation Research Board in the United States may come closest to an international effort at this time through the TRB's Highway Safety Performance Committee. Although there was somewhat broader international participation in the Committee, by its nature and mission, the Committee is fundamentally a U.S. or North American focused group. Thus, the international dialog and leadership necessary to advance a broader global effort on research programs is missing and serves as a hindrance to greater success in this area.

One of the obstacles to CMF transferability is the nature of the road safety system, which is determined by a variety of interrelations between driver behaviour, road infrastructure and vehicle characteristics that make every road traffic system unique and necessitates a specific mix of road safety measures. For example, not all successful measures are suitable for all different road traffic environments. But it is also very possible that the same interventions may lead to significantly different results in two different road traffic environments.

However, scientists have identified several common safety patterns in various traffic systems, which when taken into account collectively by established CMF methodologies may result in appropriate tailor-made solutions applicable to several different traffic systems. Establishing common CMF analysis procedures is the first basic step towards the transferability of CMF experiences. The next steps will likely turn up through international cooperation in the field of progressive standardisation of CMFs.

Box 2.2. Data systems

Fundamental to the development of accurate CMFs is access to reliable crash data. Most countries have established crash database systems to record the location and circumstances of crashes. This data is essential to determine the number of crashes before and after the implementation of a treatment. The more accurate and comprehensive this data is, the more accurate CMFs will be. In some countries, treatment monitoring systems have been included as a module of the crash database system, allowing quick and comprehensive assessments of treatment effectiveness across parts or all of the network (Turner & Hore-Lacy, 2010).

Further information on the establishment of crash data systems can be found in WHO (2010).

Despite the considerable progress made in the evaluation of road safety measures at national and international levels, resulting in important questions being successfully dealt with (e.g. confounding effects, regression to the mean etc.), a major limitation of the existing efforts concerns the need to assess the particularities of setting, context, and implementation features of a specific measure.

As a consequence, the safety effects of even the most promising road safety measures cannot be guaranteed. For these reasons, a range of values is typically proposed in each study for the safety effects of each measure examined. Alternatively, the results are often labelled as "conservative", or "best" estimates.

The scientists' competition and quest for the "perfect" methodology, together with the inherent difficulties of CMF analyses, puts in question any CMF analysis with a consequence less than those found in other published results. Such results expose the researchers to the risk of criticism from their peers. It is true that scientific accuracy is difficult to obtain in the field of CMFs, not only because several assumptions are necessary in the process but also because it is very difficult to separate the safety effect of a measure from the effect of several other microscopic or macroscopic measures and phenomena (including statistical randomness) occurring at the same place. In addition, the requirements to publish original work and the resulting desire to be somewhat secretive about or restrictive with ongoing research efforts or products will deter cooperation. Similarly, language differences and lack of knowledge about who is performing such research and how one might find the results will prevent dialog that could encourage wider collaboration and information exchange.

Furthermore, it is a real possibility that CMF analyses invited by the authorities or their political leaders will tend to use faster and less rigorous CMF evaluation methodologies. Such approaches will generally favour prevailing opinions and decisions already taken, thus creating a wide variety of non-converging CMF results.

This non-convergence is a likely outcome due to the lack of appropriate data - especially the evolution of risk exposure data - and the variety of crash cost calculations. This scenario can facilitate the production of diverging CMF evaluation results at the international and national levels.

2.4. Opportunities Persist for International Cooperation and Transferability

While many of the obstacles described above are clearly real and have a negative effect on international cooperation in this field, there are a range of opportunities to overcome these obstacles. This report in its entirety is designed to identify these opportunities and encourage positive action to support transferability.

Appendix A provides a review of current global knowledge, experience and practices with CMFs. From this Appendix, we find that there is a foundation of cooperation among some select and respected researchers that can be drawn upon to involve others and expand cooperation. This report, prepared through international collaboration, is an indicator of that potential. The publication of the Highway Safety Manual by the American Association of State Highway and Transportation Officials builds on efforts in many parts of the world and is reflective of deep interest in the quantitative assessment of safety decisions utilizing CMFs.

Another significant opportunity that exists is the advancement of thinking about what research produces a good CMF (FHWA CMF Clearinghouse, Elvik, et. al., 2009). In this light, there are good examples of how to consider a study and assess its quality. This knowledge and experience serves as a foundation in this report for identifying in advance what qualities, characteristics and specific information CMF research reports should include if they are to have the potential to be shared internationally. Many of the items mentioned above are good examples of the knowledge we have in this regard and suggest how we can use this information to critically review studies.

Most European countries set specific quantitative road safety targets and adopt related road safety strategies towards these targets, within the established priorities and the resources available. Within this framework, the efficiency assessment of road safety measures is considered to be an extremely useful tool in decision making. In particular, cost-benefit and cost-effectiveness analyses are carried out in several countries, in a more or less systematic way, at the national, regional or local level. These studies are based on some estimate of the safety effects of the examined measures, in terms of crashes or a reduction in casualties following the implementation of the measure. However, a more widespread or fruitful use of the efficiency assessment of road safety measures is in most cases limited, apart from the various technical and institutional barriers, by a lack of knowledge and data on the safety effects of road safety measures.

Nevertheless, the importance of efficiency assessment in road safety is widely recognised, and the need for more knowledge and best practice examples is becoming more and more pronounced. Existing best practice recommendations may cover the whole range of the efficiency assessment process, from the selection and application of appropriate and standardised methodologies to the interpretation of results and the identification of most efficient measures, especially in case different alternative measures need to be compared and ranked. However, the most important uncertainties involved in developing such best practice recommendations concern the adoption of appropriate values for the safety effects of road safety measures.

In recent years, much research has been done to standardize the methods for estimating the safety effects of road safety measures. The first issue examined concerns the accuracy of the estimation, so that potential bias or other confounders are eliminated; these questions mainly concern analyses at the national level. The second critical issue concerns the conditions and necessary adjustments required to allow the transferability of safety effect estimates to different settings or countries; this question has become very important at the international level, and particularly within the development of handbooks and manuals aiming to assist decision makers, researchers or other stakeholders involved in the efficiency assessment of road safety measures.

A number of manuals, handbooks and other tools have been developed recently, which gather, harmonize and improve existing knowledge on the effectiveness of road safety measures. These comprehensive and helpful international information sources assist researchers and practitioners in assessing the effectiveness of road safety measures.

These sources are often used by countries within their national road safety efficiency assessment analyses, by adopting the values proposed (e.g. in terms of percentage reduction of crashes / fatalities, or CMFs), or by adjusting them to the local conditions. However, due to large gaps in the knowledge concerning the transferability of such values across countries, several countries have developed their own methods and values for assessing the effectiveness of road safety measures.

Although the international literature may prove very useful in the identification of several good practices of cost-effective measures, thorough analysis on a case-specific basis is always necessary in order to optimise the effects of a measure in different countries or areas, by taking into account the extent of the implementation, the implementation period, and specific national or local requirements. Furthermore, it is necessary to ensure that such analyses are carried out in accordance with recognised standard methodologies.

Several other methodological or technical problems are common in international and national evaluations of the effectiveness of road safety measures. These are mainly related to the correct application of the evaluation techniques, the identification of ways for validating the statistical significance of the evaluation results, the proper selection of side-effects to be considered along with safety effects and also the correct distinction between the implementation costs and negative side-effects of the measure.

Nevertheless, efficiency assessment is an important part of the preparation of national, regional or local road safety plans. At the initial stage of evaluation, safety effects are usually unknown and in order to influence any decision making process, the efficiency assessment studies have to be prepared ex-ante, using impact data from the application of similar measures. This stresses the need to estimate appropriate values for the safety effects of the treatment examined. Moreover, it highlights the need to increase the accessibility of this information, through the dissemination of efficiency assessment results on an international basis.

KEY MESSAGES

- Demand for safety effectiveness measures –i.e. CMFs - is increasing both regionally and country by country because they are used in multiple ways for cost-effectiveness and cost-benefit assessment.

- One of the primary obstacles to international cooperation in the sharing of CMFs is the lack of a uniform understanding of the value, importance and usage of CMFs in road safety decision making.

- While most countries use CMFs from other countries, the transferral process is imperfect and is hampered by research findings that are not well documented.

- Lack of education, knowledge, and practical usage of CMFs are currently the biggest obstacles to CMF development and transferability.

- Lack of uniformity in the performance of related research and the reporting of research results is another obstacle.

- Properly planned, conducted and documented research will improve the international transferability of CMFs.

- At the moment relatively few studies meet these standards.

- Communicating the value of certain countermeasures across international boundaries and seeking their rapid adoption will maximize research investments among countries and promote more rapid global dissemination and use of life saving countermeasures.

- Establishing common CMF analysis procedures is the first basic step towards the transferability of CMFs.

- International dialogue and leadership is necessary to advance global research programs, and international cooperation is needed for the progressive standardization of CMFs.

- Cooperation among selected researchers offers an opportunity to expand international dialog and collaboration on the development of CMFs.

REFERENCES

Elvik, R. (2009) *An exploratory analysis of models for estimating the combined effects of road safety measures.* Accident Analysis and Prevention, Volume 41, Issue 4, pp. 876-880.

Elvik, R., Høye, A., Vaa, T., Sørensen, M. (2009) *The Handbook of Road Safety Measures,* second edition. Emerald Group Publishing Limited, Bingley.

European Road Safety Observatory (2006) *Cost-benefit analysis*, retrieved January 18, 2008 from www.erso.eu.

FHWA CMF Clearinghouse, http://www.cmfclearinghouse.org/sqr.cfm.

Mounce, JM 2005, *Evaluation of hazard elimination safety program crash reduction factors*, Institute of Transportation Engineers (ITE) Annual Meeting, 2005, Melbourne, Victoria, Australia.

Scientific Research on Road Safety Management, Safety Science, Volume 48, issue 9, 2010

Turner, B & Hore-Lacy, W 2010, *Road Safety Engineering Risk Assessment* Part 3: Review of Best Practice in Road Crash Database and Analysis System Design. Austroads, Sydney, Australia.

WHO, 2010, *Data systems: A road safety manual for decision makers and practitioners*. World Health Organization, Geneva, Switzerland.

ISBN 978-92-821-0375-3
Sharing Road Safety
© OECD/ITF 2012

CHAPTER 3. ASSESSING TRANSFERABILITY OF ROAD SAFETY EVALUATION STUDIES

This chapter introduces a statistical technique for assessing the transferability (external validity) of road safety evaluation studies from one country to another. The technique can be used to calculate statistics showing the consistency in time and space of studies that have evaluated the effects of road safety measures. The method is illustrated through a number of examples.

The purpose of this chapter is to explain, and show examples of how the transferability of road safety evaluation studies from one country to another can be assessed systematically.

Is it possible to assess objectively whether the findings of road safety evaluation studies can be generalised in time and space, e.g. from one country to another or from one decade to another? This question refers to the external validity of road safety evaluation studies. External validity denotes the possibility of generalising the results of research to other contexts than those in which it was made. Context may be defined in terms of, for example, the year in which and the country where a study was done. Can the external validity of road safety evaluation studies be assessed objectively ?

A simple statistical technique is introduced for assessing the external validity of road safety evaluation studies. The technique can be used to calculate statistics showing the consistency in time and space of studies that have evaluated the effects of road safety measures. The technique is “objective” in the sense that the statistics estimated do not depend on decisions made by the analyst and are based on data that are easily reproduced. Informal assessments of external validity, made without the support of a statistical technique, are likely to be influenced by subjective judgements regarding, for example, the “similarity” of countries. Moreover, it is clearly not possible to demonstrate external validity with absolute certainty statistically. Generalising research to a new context is typically non-statistical. It therefore cannot rely on any statistical indicator exclusively, but will always contain an element of judgement. It is, however, possible to develop numerical indicators that may support the assessment of the external validity of research. Thus, while the statistical approach introduced in this chapter can never be conclusive by itself, it may serve as an input to, and a constraint on, a more informal assessment of external validity. The approach is presented here to provide a possible method for assessing transferability and to help stimulate discussion on this issue in the future.

3.1. A validity framework for assessing road safety evaluation studies

This chapter is based on the validity framework for evaluation studies developed by Shadish, Cook and Campbell (2002). They distinguish between four types of validity.

These are:

1. Statistical conclusion validity, which refers to sampling techniques and the appropriate use of statistical analysis in a study.

2. Construct validity, which may also be termed theoretical validity, and denotes the adequacy of operational definitions of theoretical concepts.

3. Internal validity, which refers to the basis for inferring causal relationships between the variables included in a study.

4. External validity, which is the possibility of generalising the findings of a study (or set of studies) to other contexts than those in which the study was made. In this study we will use transferability to equal external validity.

For each of these types of validity, Shadish, Cook and Campbell identify a number of threats to validity, i.e. types of error and bias that reduce validity. There is a difference between the first three types of validity and the fourth. Statistical conclusion validity, construct (theoretical) validity and internal validity all refer to the methodological quality of a study. External validity refers to whether results can be generalised to other contexts than those in which the studies were done. Studies can never have external validity unless they have high statistical conclusion validity, high theoretical validity and high internal validity. Assessing the external validity of studies of highly variable methodological quality is not very useful, since one cannot know whether differences in the findings of studies made in different contexts are attributable to methodological weaknesses of the studies or to real differences in effects. This chapter will not discuss how to assess the statistical conclusion validity, theoretical validity and internal validity of road safety evaluation studies. The analysis of external validity will be illustrated by means of studies that have an adequate, and homogeneous methodological quality.

3.2. Two approaches to assessing external validity

There are two main approaches to assessing external validity. These will be referred to as the deductive approach and the inductive approach. The deductive approach is based on a theoretical interpretation of the findings of road safety evaluation studies. A framework for interpreting road safety evaluation studies in theoretical terms has been proposed by Elvik (2004). There is no well-established theory which can explain the findings of road safety evaluation studies. The framework proposed by Elvik is not a theory; it should rather be seen as a conceptual scheme that can be used to develop arguments for or against imputing general validity to the findings of road safety evaluation studies. An example of how the framework can be applied is given later in this chapter.

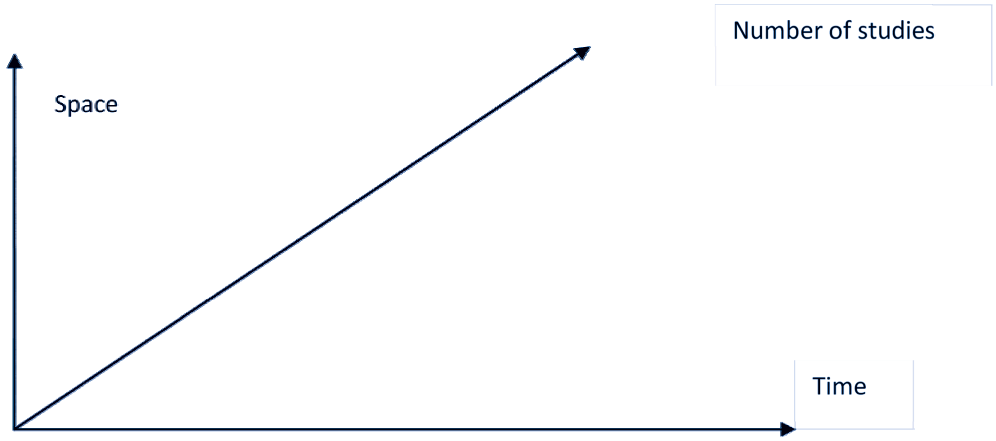
The inductive approach is based on assessing the stability of research results as these are accumulated in time and space. An example of how such an assessment can be made is given in a paper by Elvik (1996). The strength of the inductive approach depends on how many studies have been reported, the length of time covered by the studies, and the number of countries covered by the studies. Ideally speaking, the studies ought to rely on identical, or at least similar, study designs that control for at least the most important known confounding factors. When there are many methodologically strong studies, made over a long period of time in many different countries and different types of traffic environment, it is easier to apply the inductive approach than when there are fewer and poorer studies. From an epistemological point of view, assessing external validity inductively is like observing swans and concluding that: “All swans are white” when very many swans have been counted and all of them were white. As is well known, this inference is not logically valid and does not rule out that the next swan will be black.

Therefore, neither the deductive nor the inductive approach can provide absolute certainty that road safety evaluation studies are externally valid and can be applied to predict the effect of a road safety measure in a country and/or setting where it has not been evaluated before. There is, however, no alternative to the two approaches outlined above. One has to decide whether existing knowledge is applicable or not and introduce the measure if knowledge is trusted to be externally valid. This decision will always be made with uncertainty.

3.3. Key concepts describing external validity

A framework for developing concepts to describe the external validity of road safety evaluation studies is shown in Figure 3.1, which has three dimensions: time, space, and the number of studies. After the first study, all three dimensions take on the value of 1 (1 study in 1 country presented in year 1). The strength of an assessment of external validity is a function of the number and range of replications. A replication is any new study evaluating the effects of a given road safety measure and applying, at least in the main elements, the same study design as previous studies. Studies applying different study designs will not be regarded as replications, since different study designs tend to produce different estimates of effect (Gross, Persaud and Lyon 2010).

Figure 3.1. **Framework for assessing external validity of road safety evaluation studies**



The range of replications denotes the variation of the studies in time and space. Assume that time is represented by years and space by countries. A new study made three years after the first study and in a different country will then add a value of 4 to the range of replications – 3 for the number of years elapsed and 1 for the addition of a new country to the set of countries in which studies have been done. It follows that the more numerous the studies, the longer the period during which they are reported, and the larger the number of countries included, the greater the range of replications, and the number of tests that can be made of external validity will be.

Cumulative meta-analysis is well suited to the task of assessing external validity based on the range of replications. In a cumulative meta-analysis, one starts by synthesising the findings of the first two studies. Then the third study is added and a new summary estimate of the effect is obtained. Analysis proceeds by adding one study at a time and examining whether the summary estimate changes when a new study is added.

If the summary estimate of effect remains unaffected as the range of replications grows, this indicates high external validity.

If, on the other hand, a growing range of replications is associated with instability in the summary estimate of effect, this indicates low external validity, since such instability means that new studies do not reproduce the findings of earlier studies.

3.4. Preparatory analysis to assessing external validity

An analysis of external validity can only be informative if the studies included in it are of adequate and similar methodological quality in terms of statistical conclusion validity, theoretical validity and internal validity. Moreover, there should not be evidence of publication bias[1] in the set of studies available for analysis, as publication bias tends to reduce the heterogeneity of estimates of effect by suppressing estimates of effect that are inconsistent with the majority of estimates. Finally, the presence of heterogeneity, or systematic variation between studies in estimates of effect should be assessed. If there is no systematic variation in estimates of effect, one would expect external validity to be high. If there is systematic variation in estimates of effect, one potential source of such variation could be that study findings vary across years and countries. In that case, results obtained in one country in a certain year may not necessarily be transferable to another country and/or another year.

In this chapter, studies that have evaluated the effects of road lighting on injury accidents will be used as examples. The sample of studies is based on a previous meta-analysis (Elvik 1995) that included 37 studies. Two more studies (Griffith 1994, Wanvik 2009) have been added to the sample. There are several recent studies of road lighting in addition to these two, but completeness of coverage is not essential for the purpose of illustrating how to assess external validity. Analysis has been confined to studies estimating effects on injury accidents, as the effect of road lighting has been found to vary according to accident severity. In the previous meta-analysis, a comparison was made between studies that employed different study designs.

Summary estimates of effect were found to be robust with respect to study design. The sample of studies used in this chapter includes 30 studies; 25 of these employed a before-and-after design to evaluate the effects of road lighting. Applying the methodology for inclusion of estimates of effect in the Highway Safety Manual (Bahar 2010), most studies were rated as having medium low or low quality. Estimates of effect (crash modification factors) based on virtually the same set of studies as those used in this chapter were nevertheless included in the Highway Safety Manual (Highway Safety Manual 2010). It is therefore concluded that studies have sufficient methodological quality, and are sufficiently similar in terms of study design, for an assessment of external validity to make sense.

Road lighting varies considerably with respect to intensity, spacing of lighting poles, height above the road, etc. It is reasonable to expect the effects of road lighting to vary according to the standard of lighting. It seems reasonable to presume that the higher the standard, the larger the safety effect. Unfortunately, most evaluation studies provide little or no information at all about the standard of lighting. Studies surveyed so far most likely comprise lighting of varying standards, and this may be a source of heterogeneity in effects.

Thus, if effects of road lighting are found to vary between countries and over time, this does not necessarily mean that study findings are not transferable, but simply that effects are a function of lighting standard, which is not described in most studies.

The 30 studies contain 84 estimates of the effect of road lighting. An assessment of the possible presence of publication bias was made by applying the trim-and-fill technique developed by Duval and Tweedie (Duval and Tweedie 2000A, 2000B, Duval 2005). There was no evidence of publication bias.

A statistical test for heterogeneity in estimates of effect (Shadish and Haddock 1994) indicated that there was systematic between-study variation in estimates of effect. In the cumulative meta-analysis that was performed to assess external validity, a random-effects model of meta-analysis was therefore applied.

3.5. The range of replications and statistics describing external validity

3.5.1. The deductive approach

Elvik (2004) discusses whether theory can explain the results of studies that have evaluated the safety effects of road lighting. He argues that road lighting improves visibility and that by making observation easier, it may reduce the probability that road users will make errors of observation. On the other hand, lighting poles are a new hazard. Besides, it is likely that road lighting may induce behavioural adaptation among road users.

The net effect on accidents will depend on the relative strengths of the "engineering effect" and the "behavioural effect". This cannot be predicted very precisely, but the engineering effect of road lighting is quite large, adding a safety margin (increased detection distance) that is unlikely to be completely offset by behavioural adaptation. Based on theoretical considerations, the conclusion is therefore that road lighting on previously unlit roads is expected to reduce the number of accidents in darkness. Since darkness is a universal risk factor, there is also a presumption that the effects of road lighting will be rather similar everywhere.

3.5.2. The inductive approach

This report started out with the challenge of seeking to understand how countries could more effectively and systematically share the results of safety effectiveness measures and, through that process, more quickly adopt and implement good safety measures from other countries. While there is currently no evidenced-based practice or experience in achieving this level of transfer, the report seeks to propose ideas to challenge the reader and to suggest possible avenues to pursue for the future. In that light, the inductive approach is one possible approach.

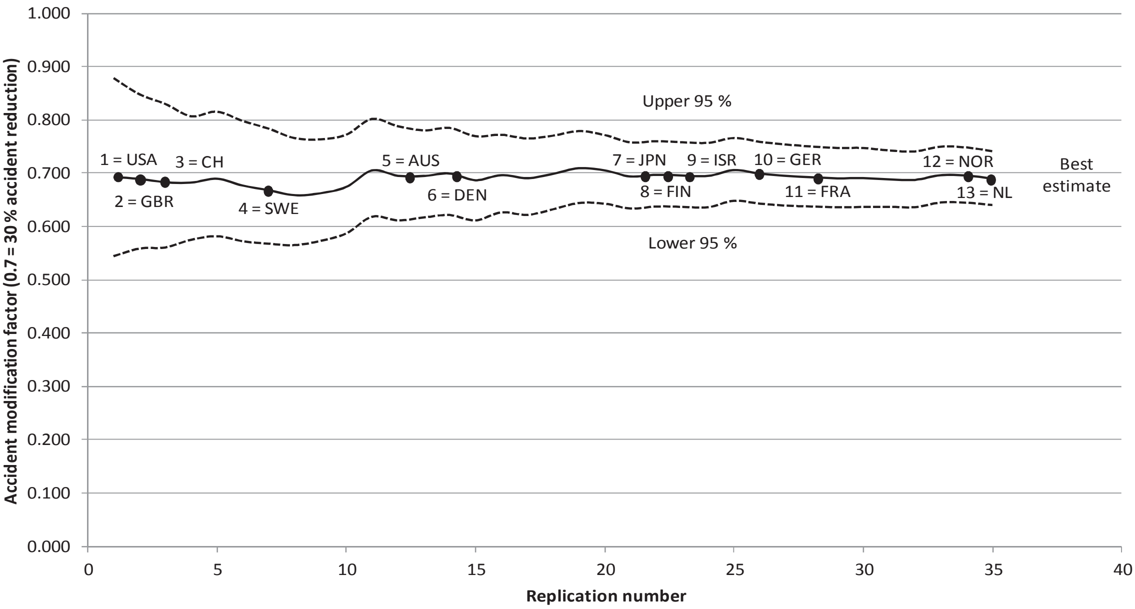
The 30 studies included in this study produced a total of 35 replications (counting the first study as "replication 1"). The first study was reported in the United States in 1948. As the first study, it was not a replication, but to correctly count the number of countries involved, it was coded as 1 for country. The next study was reported in Great Britain seven years after the first study. This study added the value of 1 for one additional country, and a value of 7 for the number of years elapsed between the first and second study. The total value of the range of replications for the first two studies is 9 (1 + (1 + 7) = 9).

The fourth study was from Great Britain in 1958, but it did not add to the range of replications, because it was from a country already included in the sample of studies and it was from the same year as another study. In general, if a study either: (1) was reported in a country already included, or (2) was reported in the same year as another study, it did not add to the range of replications. The fifth study was reported four years after the fourth study and therefore added a value of 4 to the range of replications, even if it was reported in a country already included. Studies that reported more than one replication have been listed for each replication contributed. The 30 studies produced a total of 35 replications.

The total value of the range of replications is equal to the sum of the number of countries in which studies have been reported and the numbers of years elapsed from the oldest to the newest study (in this case 61, since 61 years elapsed from 1948 to 2009). In this case, the range of replications is 74 (13 countries plus 61 years).

Figure 3.2. shows the summary estimate of effect based on a cumulative meta-analysis of the 35 replications. The summary estimate is remarkably stable and does not appear to be greatly influenced by the addition of new countries or new years to the set of replications. Figure 3.2. indicates when the first study was reported in each of the thirteen countries that have contributed studies regarding the effects of road lighting on accidents.

Figure 3.2. **Summary estimate of crash modification factor for road lighting**



The first studies, dating back to the 1950s, were made in the United States (USA), Great Britain (GBR) and Switzerland (CH). The most recent studies were reported in Norway (NOR) and the Netherlands (NL). The 95 % confidence interval for the summary estimate of effect narrows as more replications are added.

There are two kinds of replications produced by studies evaluating the effect on injury accidents of road lighting:

1. Replications within the same country or in the same year, not adding to the range of replications. These replications will be referred to as “replications within country or year”.
2. Replications in a different country or a different year, or both, adding to the range of replication. These replications will be referred to as “replications across countries or years”.

There are 10 of the first kind of replications and 25 of the second kind. It is only the second kind of replication that allows an assessment of external validity. External validity is described by means of four statistics that are derived from the findings of the cumulative meta-analysis.

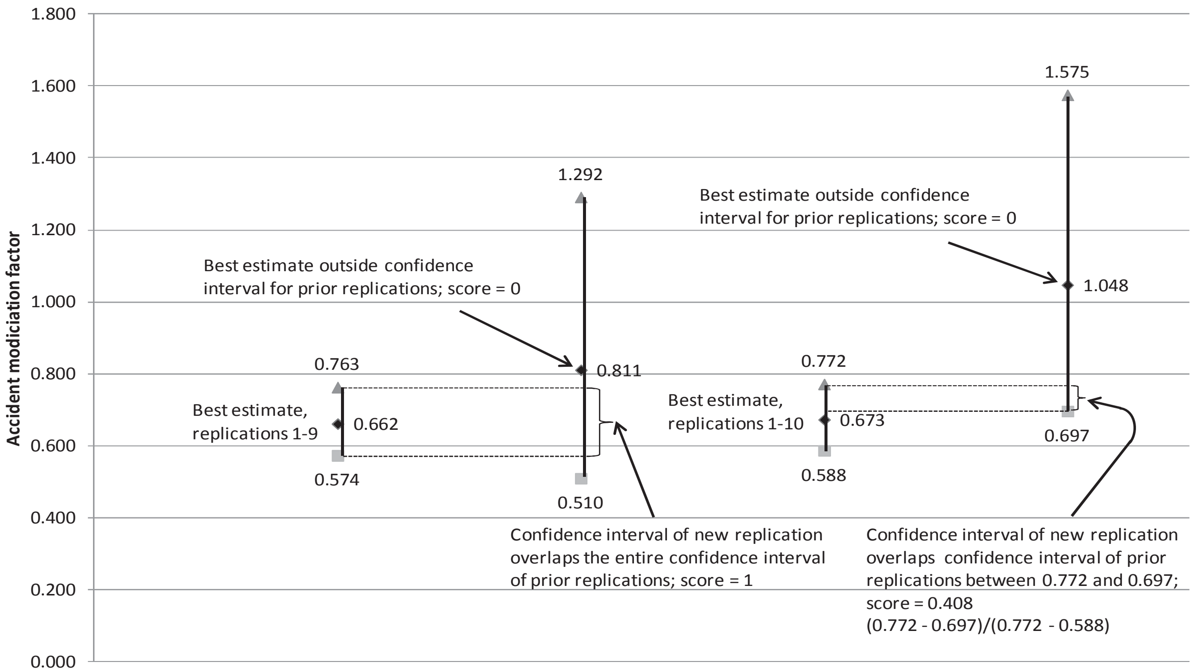
These four statistics are:

1. The proportion of best estimates accurately predicted
2. A simple consistency score

3. A consistency score weighted by study leverage
4. The linear trend in the relationship between study leverage and consistency score

For replications within country or year, only the first two of these statistics are applicable. For replications across countries or years, all four statistics are applicable. Figure 3.3. shows how the first two statistics are computed.

Figure 3.3. **Illustration of how consistency score is computed**



In Figure 3.3., the first nine replications have produced a summary estimate (crash modification factor) of 0.662 (i.e. 34 % accident reduction), with a 95 % confidence interval from 0.574 to 0.763. If the best estimate of effect in the tenth replication falls inside the confidence interval for the first nine replications, the best estimate will be regarded as accurately predicted and given a score of 1. As can be seen in Figure 3.3., the best estimate in the tenth replication was 0.811, which is outside the confidence interval of the first nine replications and therefore scored as 0, indicating an inaccurate prediction.

The simple consistency score is based on the degree of overlap between the confidence interval of a new replication and the confidence intervals of prior replications. In Figure 3.3., the confidence interval for the first nine replications spans from 0.574 to 0.763. The confidence interval of the tenth replication spans from 0.510 to 1.292. While wider than the confidence interval for the first nine replications, the smaller confidence interval is entirely contained within the greater; these two estimates are consistent (consistency score 1), although the estimate based on the tenth replication is considerably less precise than the estimate based on the first nine replications.

When the ten first replications are compared to the eleventh, it is seen that the confidence intervals only partly overlap, resulting in a consistency score of 0.408.

Two confidence intervals that do not overlap at all give a consistency score of 0. This way of estimating the consistency score assigns the same weight to all studies. However, it can be argued that some of the studies represent a stronger test of external validity than other studies and should therefore carry greater weight. A case can be made for assigning a larger weight to a study that contributes to a large increase in the value of the range of replications than to a study that only adds a little to this value.

The reasoning is as follows:

A study can contribute to a large increase in the value of the range of replications either by being reported a long time after the last previous study or by adding estimates for a large number of new countries. Since external validity refers to the stability of the findings of evaluation studies in time and space, a study representing a large addition of time or space will provide a stronger test of external validity than a study adding little to these dimensions.

The relative strength of a study as a test of external validity will be referred to as the leverage of the study. The leverage of a study is its relative importance and represents its contribution to a summary statistic. Since the first study of road lighting did not permit a test of external validity, its leverage is assigned a value of 0. All the other studies contribute, and their leverage is equal to their share of the maximum value of the range of replications, minus 1. Since the maximum value of the range of replications in this case was 74, leverage is estimated by computing the addition a study made to the value of the range of replications, divided by 73. Thus, the first of the two replications reported by study 30 (Wanvik 2009) added 15 to the value of the range of replications, giving it a leverage of 15/73 = 0.219. The sum of the leverage values is 1 and the weighted consistency score is computed by multiplying the simple consistency score of each study by its leverage and summing the values for all replications.

Table 3.1. reports the statistics computed to describe the external validity of studies that have evaluated the effects of road lighting. The proportion of accurately predicted best estimates ranges from 0 to 1. It has the value of 0.24 for replications across countries and/or years, suggesting that most new best estimates were outside the confidence intervals of prior replications and were thus not reproduced.

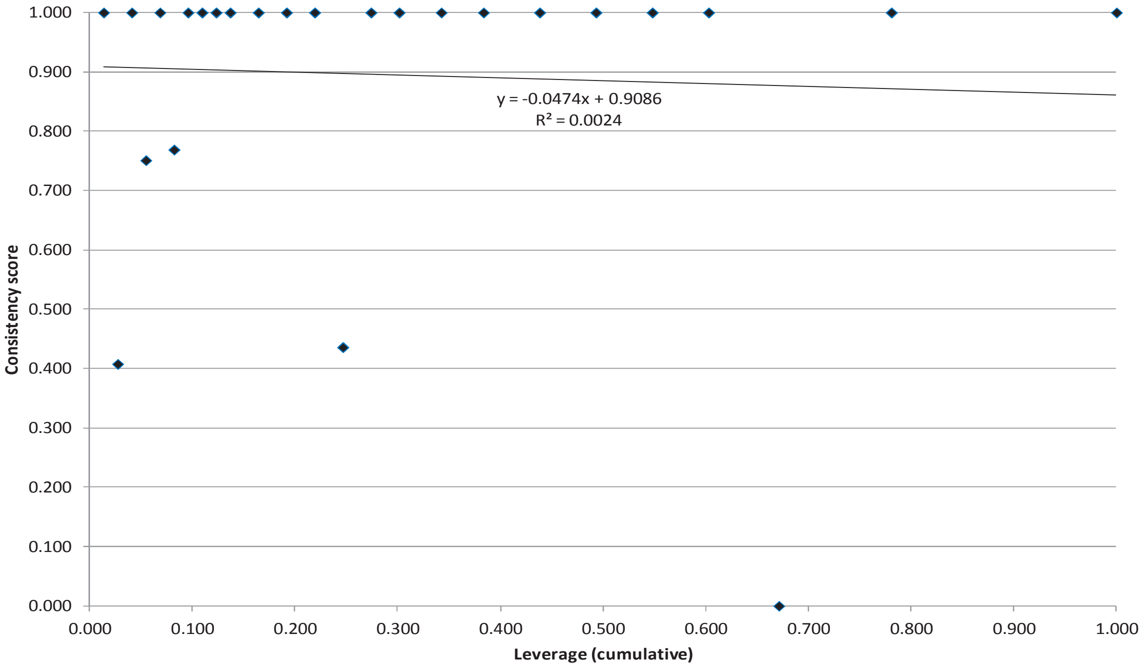
Table 3.1. **Statistics describing external validity of studies that have evaluated the effects of road lighting on injury accidents**

Type of replication	Proportion of best estimates accurately predicted	Simple consistency score	Weighted consistency score	Linear trend in relationship between leverage and consistency
Within country or year	0.500	1.000	Not applicable	Not applicable
Across countries and/or years	0.240	0.895	0.901	B = -0.047; P = 0.814

The simple consistency score (range 0 to 1) was 1.000 for replications within countries or years and 0.895 for replications across countries and/or years. The weighted consistency score was 0.901. The values indicate a very high consistency in estimates across countries and years.

Replications were sorted according to their leverage. For each value of leverage, replications were sorted chronologically. The simple consistency score was then plotted in a diagram with cumulative leverage on the horizontal axis and consistency score on the vertical axis. Figure 3.4. shows the resulting diagram.

Figure 3.4. **Relationship between leverage and consistency score**



There is almost no relationship between leverage and consistency score. A very weak tendency is seen for the consistency score to go down as leverage increases, but this tendency is very far from being statistically significant. In other words, replications that represent large steps in time or space are just as likely to be consistent with prior replications as replications that represent smaller steps in time or space.

3.6. A note on alternative and supplementary techniques

Testing the external validity of road safety evaluation studies can to some extent be done by means of meta-regression. One might then include year as a count variable and identify countries by assigning a dummy variable to each country. In a previous study (Elvik 2003), summarising the effects on non-US studies evaluating the effects on safety of converting junctions to roundabouts, the following conclusions were drawn:

"Meta-regression was run in two stages. The first stage included all variables, except year of study publication. No statistically significant effects were found of variables representing the country in which the study was reported. This is reassuring, because it means that it is defensible to generalize the results of evaluation studies across countries. In the second stage, country variables (each country was represented by a dummy variable) were therefore omitted." "Meta-regression was used, because it is in principle superior to a traditional subgroup analysis by being able to control for a large number of potentially confounding variables at the same time. Five variables were included in the main analysis (previous type of traffic control, number of legs, size of roundabout, study design, and accident severity).

Given the fact that there were 113 estimates of effect in total, using five variables in a meta-regression does not seem excessive. However, the number of combinations of values for the five variables is 400 (2 x 2 x 4 x 5 x 5). This number is well in excess of the number of observations (113) used to fit the multivariate model, indicating that the model is underdetermined by the data (since not all 400 logically possible combinations of values are found in the data set).

Most of the coefficients were indeed not statistically significant, and the coefficient of determination for the final model (not including country variables) was 0.375 (adjusted R-squared)."

In general, high correlation among the independent variables is a problem in meta-regression. In addition, countries must be represented by dummy variables, leading to a large number of variables when many countries are represented in the data set. A more serious problem, however, is that meta-regression does not produce any statistics indicating the level of external validity, such as the various consistency measures proposed in this chapter. It does not therefore really permit external validity to be tested the way it can be by means of the approach presented here. To test external validity in meta-regression, one approach would be to use half the data to fit a model and then use model coefficients to predict findings for the second half of the data. The predictive performance of the coefficients would then be tested. In most cases, however, such an approach is not feasible, because there are too few studies to use only half of them for model development.

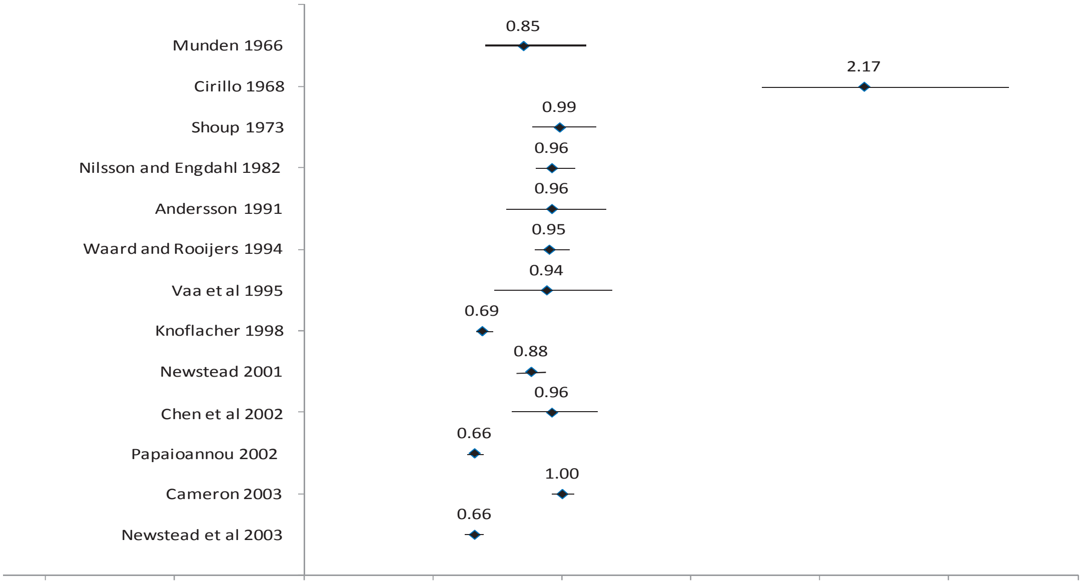
The results of a range of replications analysis can be tested by performing a multiple regression analysis, using replications as the unit of analysis, the estimate of effect as the dependent variable and year and country as independent variables. Such an analysis was performed for the 35 replications of road lighting, indicating country by a dummy variable for each country with at least three replications and putting countries with fewer than three replications in a group labelled "other countries". Each replication was weighted by means of the random-effects statistical weight applied in the cumulative meta-analysis. Somewhat surprisingly, all coefficients were statistically highly significant, indicating that a case could be made for applying country-specific and year-specific summary estimates of effect. However, unless a good reason is given for believing that estimates of effect can be expected to vary between countries and years, applying estimates of effect specific to each country and year would be a bit contrived in view of the striking stability of replications as indicated by Figure 3.2 and the consistency statistics estimated. The effect of road lighting on injury accidents would appear to be the same in all countries in which studies have been made so far and has remained remarkably stable over time.

3.7. Developing a Crash Modification Function based on heterogeneous studies[2]

The results of studies are not always as consistent as in the case of road lighting. Consider the results of thirteen studies that have evaluated the effects of speed enforcement on accidents, presented in chronological order in Figure 3.5. Each study produced between four and eight estimates of effect, referring to different levels of enforcement. In Figure 3.5, each study is represented by a summary estimate of effect. Each summary estimate is a crash modification factor (e.g. 0.85 = 15 percent accident reduction).

Although some of the studies produced findings that are highly consistent, such as the string of studies starting with Shoup in 1973 and ending with Vaa et al. in 1995, other studies produced findings that are clearly inconsistent with these studies. A statistical test indicated considerable between-study variation in estimates of effect. When the range of replications technique was applied to the studies, the simple consistency score was 0.648. Weighted by study leverage, the consistency score was 0.706. These values are lower than those found for road lighting.

Figure 3.5. **Forrest plot of summary estimates of effect in thirteen studies of speed enforcement**



It is reasonable to think that the more enforcement there is, the greater the effect will be on accidents (Elvik 2011A). Can such a dose-response pattern be found in the thirteen studies listed in Figure 3.5.? An initial attempt to answer this question was made by fitting functions to the data points of each study; as noted above there were between four and eight data points per study. This approach was not successful. The functions fitted to the data points of each study were highly inconsistent and many of them were implausible. The studies contained a total of 78 estimates of the effects of speed enforcement. Eleven of the thirteen studies were before-and-after studies. Two studies were cross-section studies. It was decided to try to develop a crash modification function for speed enforcement by pooling data points that referred to the same level of enforcement. The level of enforcement was stated as a relative level compared to a baseline level of 1. Thus a level of 0.5 represented a 50 percent reduction in enforcement and a level of 2 represented a doubling of enforcement compared to the baseline level. Both cross-section studies were methodologically weak and their findings were inconsistent with the before-and-after studies. These two studies were therefore omitted. The remaining eleven studies contained 63 estimates of effect.

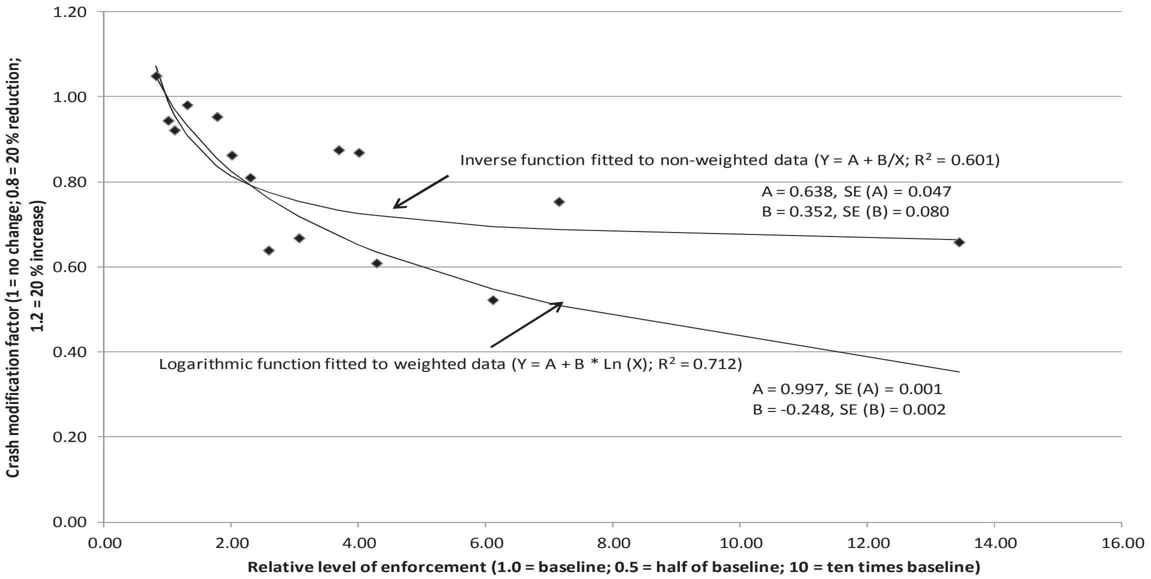
Many of these estimates referred to the same level of enforcement. Thus, there were eight estimates referring to the baseline level of enforcement. Data points referring to the same, or nearly the same level of enforcement were combined by means of meta-analysis. As a result, the number of data points was reduced from 63 to 15.

A crash modification function was fitted to these 15 data points by means of simple or weighted regression analysis. In the weighted analysis, each data point was weighted in proportion to its statistical precision, giving the largest weight to the most precise estimates of effect. The resulting crash modification functions are shown in Figure 3.6.

The inverse function was fitted to non-weighted data. The logarithmic function was fitted to weighted data. As can be seen, these functions are very close for relatively low levels of enforcement, but diverge at higher levels of enforcement.

This reflects the fact that there are fewer data points at high levels of enforcement than at low levels and that each data point referring to high levels of enforcement is less precise than each data point referring to lower levels of enforcement. The divergence between the curves can be taken as an indication of the uncertainty of the effects of high levels of enforcement.

Figure 3.6. Crash modification functions for speed enforcement



There are several lessons to be learnt from this study. One lesson is that a considerable number of estimates of effect are needed in order to reliably estimate a crash modification function. A single study will rarely contain enough data points to provide an adequate statistical basis for estimating a crash modification function. A second lesson is that when developing a crash modification function, several functional forms should be tested and compared. The functional forms considered should include both functions that are consistent with prior hypotheses and functions that are not consistent with prior hypotheses. A third lesson is that data may be consistent with more than one function; several functions may fit the data almost equally well, and selecting the preferred function must to some extent be based on its theoretical plausibility, not just the statistical goodness-of-fit.

3.8. Other Issues in Applying the Range of Replication Technique

The range of replications technique should not be applied when there is evidence of publication bias. Publication bias tends to reduce the heterogeneity of estimates of effect and may thus produce inflated estimates of the consistency of findings across years and countries. A recent re-analysis of a meta-analysis of bicycle helmet efficacy is a case in point (Elvik 2011B). A trim-and-fill analysis suggested the presence of publication bias. To test the potential effect of such bias on the results of a range of replications analysis, studies that have evaluated the effect of bicycle helmets on head injury were used. In the re-analysis, 17 studies containing 18 estimates of effect were included.

These studies permitted 17 tests of external validity, of which 5 were within countries or years and 12 were across countries and/or years. The replications across countries and years indicated highly consistent findings. The simple consistency score was 0.891 and the weighted (by leverage) consistency score was 0.918.

Six fictitious studies, generated by a trim-and-fill analysis were then added. Each of these studies was assumed to add one year and one country to the range of replications; in total these studies added a value of 12 to the range of replications. The consistency statistics were then re-estimated. The addition of the fictitious studies to adjust for publication bias reduced consistency considerably. The simple consistency score was now 0.594 and the weighted consistency score 0.649.

Thus, the presence of publication bias may inflate measures of consistency and give a false impression that study findings are transferable when they are not.

It is, unfortunately, very common that road safety studies are of highly varying quality, and that there is a relationship between study quality and study findings. Often, many studies have low statistical conclusion validity, low theoretical validity or – most commonly – low internal validity. Assessing external validity then hardly makes sense, as it is impossible to know if differences in the findings of studies made in different countries or in different years are real or just methodological artefacts.

Studies evaluating roundabouts are a case in point. Elvik (2003) included 28 studies in his meta-analysis, but only three of those studies had controlled for long-term trends and regression-to-the-mean. Most of the remaining before-and-after studies did not control for any confounding factors, and their results differed from those of the more well-controlled studies. There were also a number of cross-section studies. Their findings differed from the before-and-after studies and it is likely that these findings were influenced by confounding factors not controlled for.

Is it at all possible to assess external validity in a case like this? The best option is probably to select the best studies and base the assessment of external validity on those studies exclusively. A major drawback of this approach is that it greatly reduces the number of studies that can be used to assess external validity.

If only before-and-after studies that controlled for long-term trends and regression-to-the-mean are included in assessing the external validity of studies that have evaluated the safety effects of converting junctions to roundabouts, only six studies remain. Three of these studies were included in the meta-analysis reported in 2003 (Elvik 2003), and three are new studies. These six studies were reported during a period of 20 years in 5 countries. The range of replications was 25. The studies permitted 5 tests of external validity. The simple consistency score was 0.599, and the weighted consistency score was 0.543.

The statistics obtained in this case indicate a fairly low external validity. However, since only five tests were possible, and the studies were quite small, one should not necessarily conclude that transferring results across countries and years is impossible, at least not before accounting for the heterogeneity of study findings.

3.9. Discussion

How can we know if the results of a study reported, say, six years ago in country A regarding the effects on road safety of measure X can be applied today in country B, i.e. used to predict what the safety effect of introducing measure X will be in country B? This chapter has indicated one way of trying to answer this question.

The answer proposed is that: (1) if there have been many studies of measure X, not just in country A, but in many other countries, and not just six years ago, but spanning three or four decades, and: (2) if these studies obtained highly consistent estimates of the effect of measure X, then: (3) it is more reasonable to conclude that the results of these studies can be applied in country B than to conclude the opposite. In other words, as long as history keeps repeating itself, it is more reasonable to expect it to continue repeating itself than to expect the opposite.

Generalising by counting and classifying repeated observations of a phenomenon is, however, never perfectly reliable. The range of replications technique can give an indication of the stability of research results across countries and years, but that is all. It does not, by itself, provide a non-refutable basis for inferring external validity. Moreover, it can only be applied when the following preconditions are fulfilled:

1. The studies reported should be of consistently high methodological quality, controlling for at least the most important potentially confounding factors that can influence research results. If studies vary in methodological quality, it is impossible to know if results that differ between countries do so because the true effects of a road safety measure vary between countries or because different study methods were used in different countries.

2. There should not be a trend over time in research findings. The presence of a trend over time suggests that the findings of older studies cannot be applied and that one should rely on the most recent studies only.

3. There should not be evidence of publication bias in a set of evaluation studies. Publication bias tends to reduce the heterogeneity of study findings. If there is publication bias, study findings are likely to vary less than if there is no such bias, producing inflated estimates of consistency across time and countries.

4. A comparatively large number of evaluation studies reported in different countries during a long period of time should be available for analysis.

Before applying the range of replications technique, one should assess if these preconditions are fulfilled. In one of the cases used to illustrate the technique in this chapter, there was evidence both of a time trend and of publication bias. Applying the range of replications technique in such a case is potentially misleading. While it is in principle possible to adjust for publication bias by adding “missing” studies, these studies are fictitious and have no publication date or country of origin. Moreover, analysis that doesn’t adjust for publication bias indicated that the range of replications technique does not capture a trend over time.

One might expect the technique to capture the effects of a trend by assigning low consistency values to new studies that are inconsistent with prior studies. In the case of bicycle helmets, however, the confidence intervals of both prior and new studies were often too wide to statistically conclude that the estimates of effect were inconsistent. The statistical power of the technique is reduced when studies rely on small samples that are associated with wide confidence intervals.

The applicability of the technique is therefore likely to be limited. It can, however, be fruitfully applied to assess external validity when a large number of studies have been reported during a long period of time. This is the case with respect to many road safety measures, like road lighting, guard rails, traffic signals, speed limits and seat belts. What if the range of replications technique suggests that the findings of studies reported in different countries and different years are not consistent? Does that mean that the findings of a study made in one country cannot be applied in a different country? Yes, it does show that a straightforward application would be problematic. It is possible, however, to account for the heterogeneity of study findings.

If sources of heterogeneity are found, it may be possible to model systematic variation in effects in terms of a crash modification function, i.e. a mathematical function predicting the effect of a road safety measure as a function of characteristics of the measure or of the context in which it is applied. Developing such a function could demonstrate the transferability of research by showing that differences between countries with respect to the effects of a road safety measure are explicable in terms of differences with respect to the factors producing systematic variation in the effect of the measure.

KEY MESSAGES

- There are still large gaps in knowledge and there is a need for many more studies from a number of countries.
- In the absence of these studies, we may be able to increase the transferability of CMFs by creating functions based on the circumstances of the CMF development.
- A considerable number of estimates of effects is needed in order to reliably estimate a crash modification function.
- When developing crash modification functions, several functional forms should be tested and compared.

NOTES

1. Publication bias is difficult to estimate as typically studies that do not report significant effects are not published in the open literature.

2. Illustrative examples.

REFERENCES

Andersson, G. 1991. *Effekter på hastigheter av intensifierad övervakning med radar*. Transportforskningberedningen (TFB) og Statens väg- och trafikinstitut (VTI) forskning/research nr 6. Stockholm.

Bahar, G. 2010. *Methodology for the development and inclusion of crash modification factors in the first edition of the Highway Safety Manual.* Transportation Research Circular E-C142. Transportation Research Board, Washington D. C.

Cameron, M., Newstead, S., Diamantopoulou, K., Oxley, P. 2003. *The interaction between speed camera enforcement and speed-related* mass media publicity in Victoria. Report 201. Monash University Accident Research Centre, Melbourne.

Chen, G., Meckle, W., Wilson, J. 2002. *Speed and safety effect of Photo radar enforcement on a highway corridor in British Columbia*. Accident Analysis and Prevention, 34, 129-138.

Cirillo, J. A. 1968. *Interstate System Accident Research Study II*, Interim Report II. Public Roads, 35, 71-75.

Duval, S., Tweedie, R. 2000A. *Trim and fill: a simple funnel plot based method of testing and adjusting for publication bias in meta-analysis*. Journal of the American Statistical Association, 95, 89-98.

Duval, S., Tweedie, R. 2000B. *A non-parametric trim and fill method of assessing publication bias in meta-analysis.* Biometrics, 56, 455-463.

Duval, S. 2005. *The trim and fill method.* In Rothstein, H., Sutton, A. J., Borenstein, M. (Eds): Publication bias in meta-analysis: *Prevention, assessment and adjustments*, 127-144. John Wiley and Sons, Chichester.

Elvik, R. 1995.*Meta-analysis of evaluations of public lighting as accident countermeasure*. Transportation Research Record, 1485, 112-123.

Elvik, R. 1996. *Does prior knowledge of safety effect help to predict how effective a measure will be?* Accident Analysis and Prevention, 28, 339-347.

Elvik, R. 2003. *Effects on Road Safety of Converting Intersections to Roundabouts. Review of evidence from Non-U.S.* Studies. Transportation Research Record, 1847, 1-9.

Elvik, R. 2004. *To what extent can theory account for the findings of road safety evaluation studies?* Accident Analysis and Prevention, 36, 841-849.

Elvik, R. 2011A. *Developing an accident modification function for speed enforcement*. Safety Science, 49, 920-925.

Elvik, R. 2011B. *Publication bias and time-trend bias in meta-analysis of bicycle helmet efficacy*. A re-analysis of Attewell, Glase and McFadden (2001). Accident Analysis and Prevention, 43, 1245-1251.

Griffith, M. S. 1994. *Comparison of the safety of lighting options on urban freeways*. Public Roads, online (Autumn 1994). Federal Highway Administration, Washington D. C.

Gross, F., Persaud, B., Lyon, C. 2010. *A guide to developing quality crash modification factors*. Report FHWA-SA-10-032. U. S. Department of Transportation, Federal Highway Administration, Washington D. C.

Highway Safety Manual. 2010. American Association of State Highway and Transportation Officials (AASHTO), Washington D. C.

Knoflacher, H. 1998. *Eine Kosten-Nutzen-Analyse der Verkehrsüberwachung*. Zeitschrift für Verkehrssicherheit, 44, 12-18.

Munden, J. M. 1966. *An experiment in enforcing the 30 mile/h speed limit. RRL Report No 24.* Harmondsworth, Road Research Laboratory.

Newstead, S. V., Cameron, M. H., Leggett, L. M. W. 2001. The *crash reduction effectiveness of a network-wide traffic police deployment system.* Accident Analysis and Prevention, 33, 393-406.

Newstead, S., Cameron, M. 2003. *Evaluation of the crash effects of the Queensland speed camera program. Report 204*. Monash University Accident Research Centre, Melbourne.

Nilsson, E., Engdahl, S. 1982. *Effekter av trafikövervakning på trafikolyckor*. TOS ab, Stockholm.

Papaioannou, P., Mintsis, G., Taxiltaris. C., Basbas, S. 2002. *Enforcement and traffic accidents: recent experience from Greece*. Proceedings of 15th ICTCT Workshop, Brno.

Shadish, W. R., Cook, T. D., Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston, Houghton Mifflin Company.

Shadish, W. R.; Haddock, C. K. 1994. *Combining estimates of effect size*. In Cooper, H.; Hedges, L. V. (Eds): *The Handbook of Research Synthesis*, Chapter 18, 261-281. Russell Sage Foundation, New York.

Shoup, D. C. 1973. *Cost effectiveness of urban traffic law enforcement*. Journal of Transport Economics and Policy, 12, 32-57.

Vaa, T., Christensen, P., Ragnøy, A. 1995. *Politiets fartskontroller: Virkning på fart og subjektiv oppdagelsesrisiko ved ulike overvåkingsnivåer*. TØI rapport 301. Oslo, Transportøkonomisk institutt.

Waard, D. de, Rooijers, T. 1994. *An experimental study to evaluate the effectiveness of different methods and intensities of law enforcement on driving speed on motorways*. Accident Analysis and Prevention, 26, 751-765.

Wanvik, P. O. 2009. Road lighting and traffic safety. *Do we need road lighting?* Doctoral theses at NTNU, 2009:66. Norwegian University of Science and Technology, Trondheim.

ISBN 978-92-821-0375-3
Sharing Road Safety
© OECD/ITF 2012

CHAPTER 4. QUANTITATIVE FRAMEWORK FOR ENHANCING TRANSFERABILITY OF CRASH MODIFICATION FACTORS OR FUNCTIONS

This chapter provides a quantitative framework for assessing and enhancing the international transferability of crash modification functions. Variability in CMF research results is identified as a major deterrent to international transferability. The chapter provides detailed guidance on how to assess variance. It further describes how researchers can build studies to reduce variance from this perspective and how practitioners can better understand CMFs that they would like to apply.

The purpose of this chapter is to present a framework for assessing and enhancing the international transferability of crash modification factors or crash modification functions. The framework is indicated in Figure 4.1. Figure 4.1. shows a flow-chart describing how to systematically assess the international transferability of the findings of road safety evaluation studies. It is useful to start by conducting a systematic literature review, designed to identify all studies that have evaluated the effects on safety of a certain road safety measure. Once all relevant studies have been obtained, one should assess the range of countries and time covered by the studies, as shown in Chapter 3. There are two possible outcomes of such an assessment. If there are fewer than five studies and all or nearly all of these studies have been made in the same country, or during a short period, for example less than ten years, no basis for assessing the international transferability of study findings exists. This outcome is shown in the box to the left with thick lines. When there are only a few studies, there will often be no statistical basis for reliably determining whether study findings vary systematically. Should the findings of a few studies vary substantially, identifying the sources of this variation in a way that will support transferability will not always be possible.

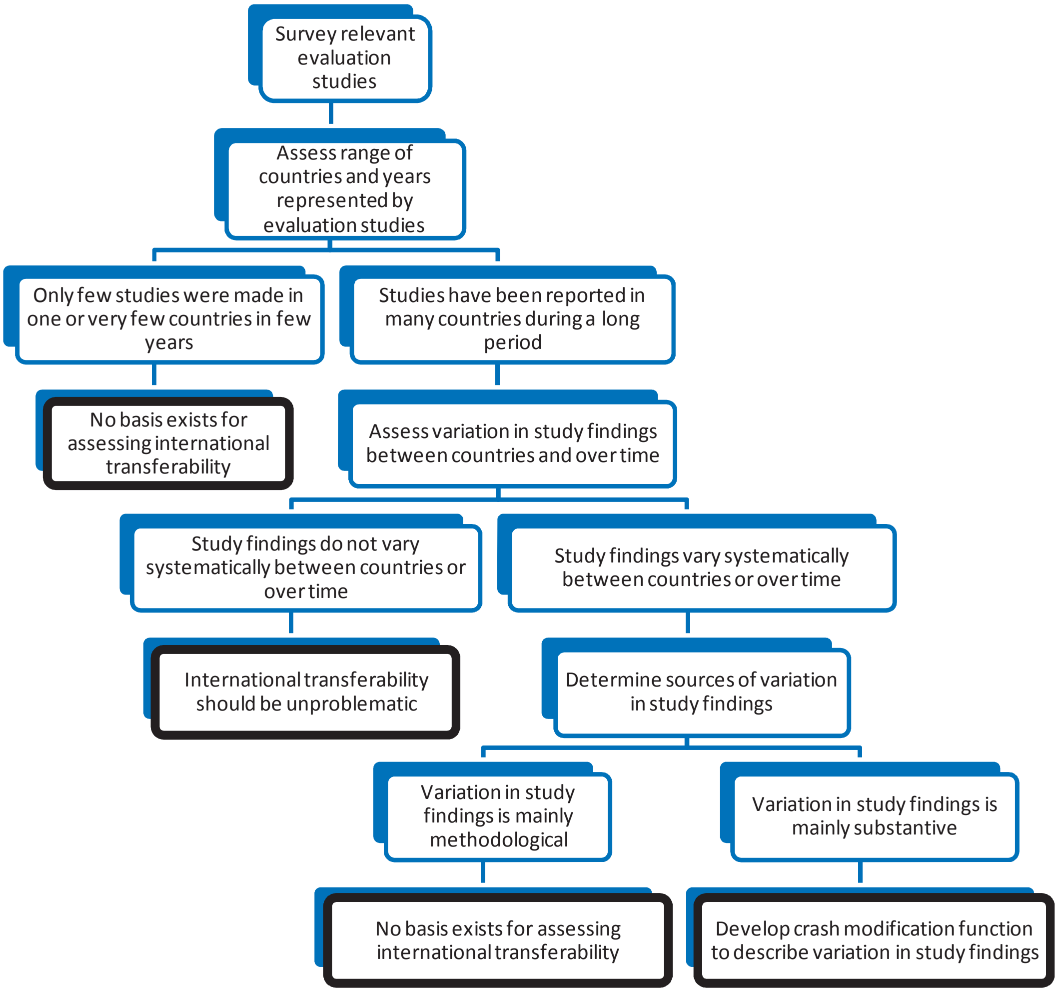
It is not possible to state a precise rule regarding the minimum number of studies, the minimum number of countries or the minimum number of years needed to statistically assess international transferability. As a rough guideline, however, if there are fewer than five studies, made in fewer than three countries and spanning less than ten years, trying to perform a more formal assessment of international transferability or developing a crash modification function will often not be possible.

If this first hurdle is cleared, i.e. there are a sufficient number of studies from different countries and different years (possibly also differing in terms of other characteristics, such as the traffic environment in which they were made), the next stage of analysis is to assess variation in study findings between countries, over time, and between different types of traffic environment, according to the design or standard of the safety measure – or according to any other measurable characteristic for which information is available for all studies. A statistical technique for identifying systematic variation in study findings is presented later in this chapter. If there is no, or only minor, systematic variation in study findings, international transferability should not be a problem.

This outcome is indicated in the box with thick lines on the left side of the flow-chart. The studies of road lighting discussed in Chapter 3 are an example of this outcome.

The stability of the findings of these studies between countries and over time is striking and supports the use of a uniform Crash Modification Factor for the effect of road lighting in all countries.

Figure 4.1. **Flow-chart for assessing international transferability of road safety evaluation studies**



In most cases, however, one must expect study findings to vary more than they did in the case of road lighting. If there is systematic variation in study findings, the next stage of analysis is to identify the sources of that variation. There are two main sources of variation in the estimates of effect produced by road safety evaluation studies: methodological and substantive. Ideally speaking, methodological variation in study findings ought to be eliminated. It is a source of error and bias, not a source of genuine knowledge. Unfortunately, many road safety studies have applied flawed study designs that produce potentially misleading findings. Examples of this will be given later in this chapter. If different studies have used different methods and come to different results, it is really not possible to generalise the results of these studies, as it is impossible to know whether the variation in their findings is real or attributable to differences in method. Hence, if it is found that the variation in study findings is mainly methodological, no basis exists for assessing international transferability. This outcome is indicated in the box to the bottom left side of the flow-chart.

If, on the other hand, all studies have applied similar and defensible methods, but still produce different estimates of effect, analysis should proceed by trying to develop a Crash Modification Function to describe variation in the effects of a road safety measure. Section 4.1 will give examples of cases that show how to use the flow-chart in Figure 4.1 to assess the transferability of study findings. Section 4.2 presents a statistical technique for developing Crash Modification Functions and illustrates the use of the technique. Finally, section 4.3 discusses how to improve the quality of road safety evaluation studies and promote the use of more uniform methods.

4.1. Research and implementation

Looking to the future, there may be possibilities for countries to accelerate or advance the transferability of safety effectiveness measures. The following sections provide examples of how multiple studies from various countries could potentially be combined to form crash modification functions that more universally describe the safety benefits of particular countermeasures. It should be emphasized that these examples are merely illustrative for the purposes of showing the long term potential of international collaboration and cooperation in transferring road safety effectiveness experiences.

4.1.1. A taxonomy of factors that can produce variation in research findings

In the previous sections, variance in the estimates of the effects of safety measures is a primary issue affecting transferability. This section describes the main factors that produce variation in the findings of studies evaluating these safety effects.

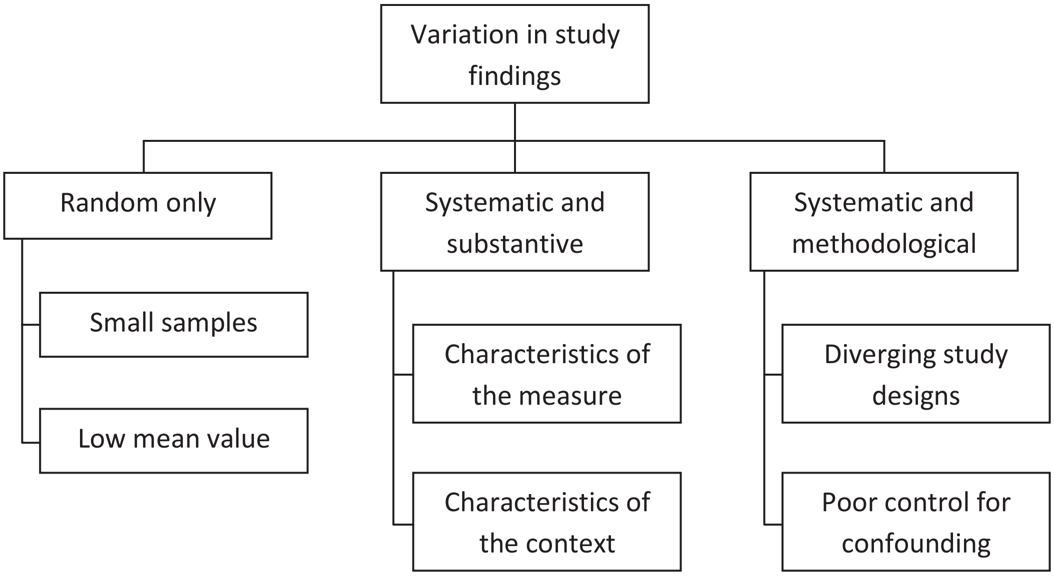
There are three main groups of factors. The first is random variation in study findings. This is most likely to occur if studies are based on small samples with a low mean number of crashes. If there is only random variation in the findings of evaluation studies, there is, in a sense, no point in performing a further analysis to determine transferability or developing a crash modification function. Results can be regarded as transferable by virtue of the fact that they do not differ systematically from one another. Moreover, any function intended to describe systematic variation in study findings would be entirely spurious and would merely capitalise on chance.

There is, however, a possibility of discovering a systematic pattern even in data that are dominated by random variation. To do so, it may be necessary to perform a continuity correction of the data. The reasons for making such a correction, and its impact will be briefly discussed later in this chapter.

When there is systematic variation in estimated effects of a road safety measure, it usually has two sources: substantive and methodological. The two main substantive sources of variation in effects are characteristics of the measure and characteristics of the context in which the measure is introduced. These are both broad concepts that are intended to refer to a range of phenomena, examples of which will be given later in the chapter.

Systematic variation attributable to methodological sources can be specified in considerably greater detail than shown in Figure 4.2., where an initial distinction has been made between diverging study designs and poor control for confounding factors. These two categories are often closely related, but it may still be fruitful to treat them as conceptually distinct. Examples of studies where a diversity of study designs or poor control for confounding factors prevent transferability are given later in the chapter.

Figure 4.2. **Sources of variation in the findings of road safety evaluation studies**



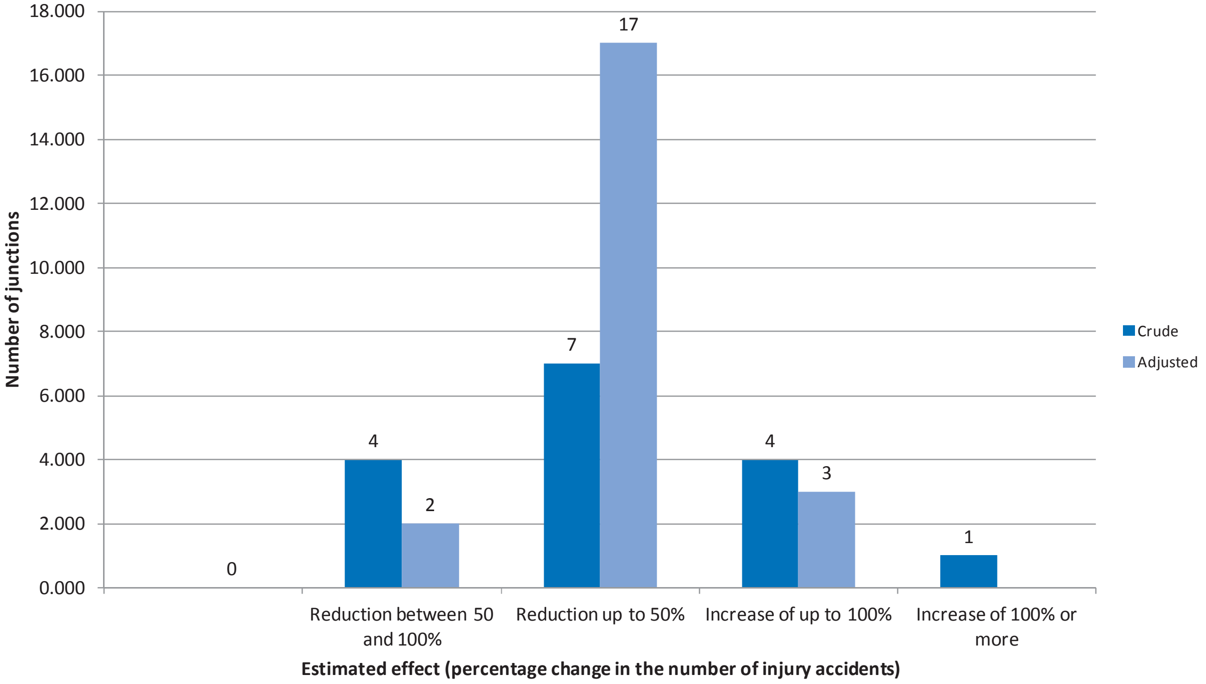
4.1.2. Finding patterns in results based on small samples and low mean values

Developing crash modification functions requires a clear picture of the pattern of changes caused by the treatment in relation to one or more key attributes of the road environment (or other features of the treatment program). Identifying patterns when samples are small or mean values are low is challenging.

Elvik (2009) performed an exploratory study for the purpose of developing crash modification functions. Two road safety measures were used as examples in the study. For one of them, bypass roads, a crash modification function was developed based on before-and-after studies reported in Denmark, Norway and Sweden. For the other, converting junctions to roundabouts, a crash modification function based on two Norwegian studies was proposed, but it was concluded that it was highly implausible and uncertain. Despite this, it is referred to in this chapter as an illustration of a possible method for combining the results of studies.

One reason for the lack of success in developing a crash modification function for roundabouts is probably the very low count of crashes in the junctions that were converted to roundabouts. This resulted in large, and largely random, fluctuations between junctions in the estimate of the effect of conversion to a roundabout. A smoothing procedure – or continuity correction – derived from the empirical Bayes method for estimating safety was applied to the data (Elvik 2011A). This resulted in considerably less dispersion in estimates of effect, as indicated in Figure 4.3.

Figure 4.3. **Effect of continuity correction on dispersion of estimates of effect in 22 junctions converted to roundabouts.**



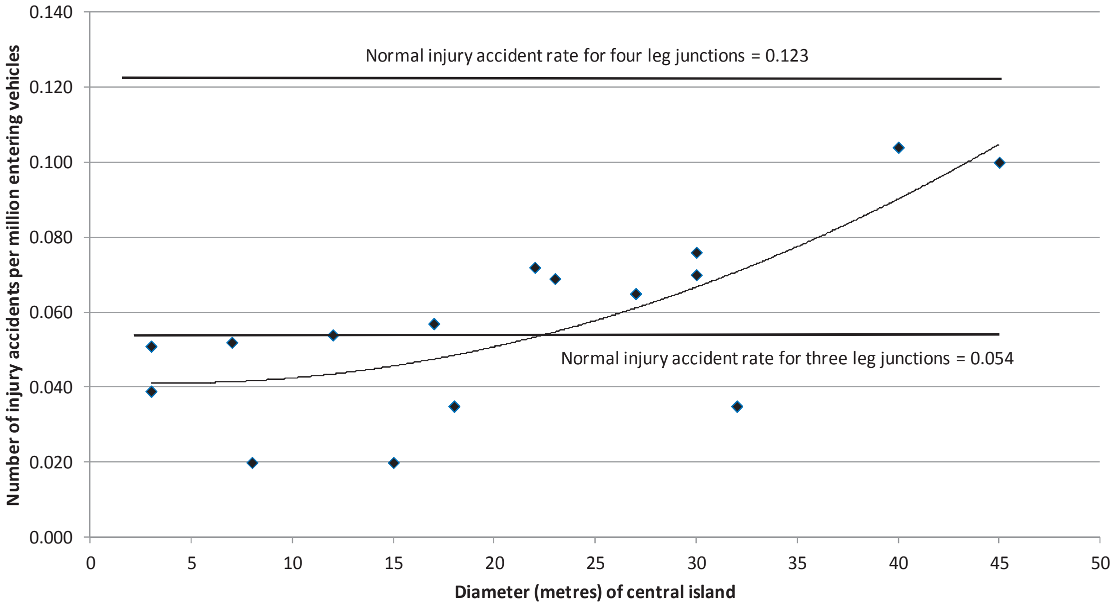
Based on Elvik 2011A.

Before continuity correction, there was an apparent crash reduction of 100 percent in six junctions. This estimate is most likely an artefact, attributable to the fact that these junctions had a zero count of crashes in the after period. Before continuity correction, an increase in the number of crashes of more than 100 percent was observed in one junction. After continuity correction, estimates of effect were considerably less dispersed.

The main reason for performing the continuity correction was that many intersections recorded zero accidents either before conversion to a roundabout, after conversion, or during both periods. A count of zero accidents cannot be regarded as an unbiased estimate of the long-term expected number of accidents, as it is theoretically highly implausible that an accident could never happen. Therefore, the long-term expected number of accidents must always be positive. The continuity correction did not adjust the total number of accidents recorded after conversion to roundabouts, but corrected for the fact that the recorded number of accidents after conversion was clearly below the long-term mean in some intersections (i.e. those that recorded zero accidents) and possibly above the long-term mean in other intersections.

Assuming that the long term safety performance of a roundabout is indicated by the rate of crashes per million entering vehicles, it is possible to develop a crash modification function based on the relationship between the diameter of the central island and the crash rate. Figure 4.4 shows this relationship based on three Scandinavian studies (Tran 1999, Brüde and Larsson 1999, Jørgensen and Jørgensen 2002).

Figure 4.4. **Relationship between diameter of central island in roundabouts and injury crash rate**



The crash rate increases as the central island becomes bigger. The normal injury crash rate in Norway for three leg and four leg junctions has also been inserted in the figure. It can be seen that the effect on safety of converting a junction to a roundabout, all else being equal, is expected to be greater in four leg junctions than in three leg junctions. The effect on safety increases as the diameter of the central island decreases.

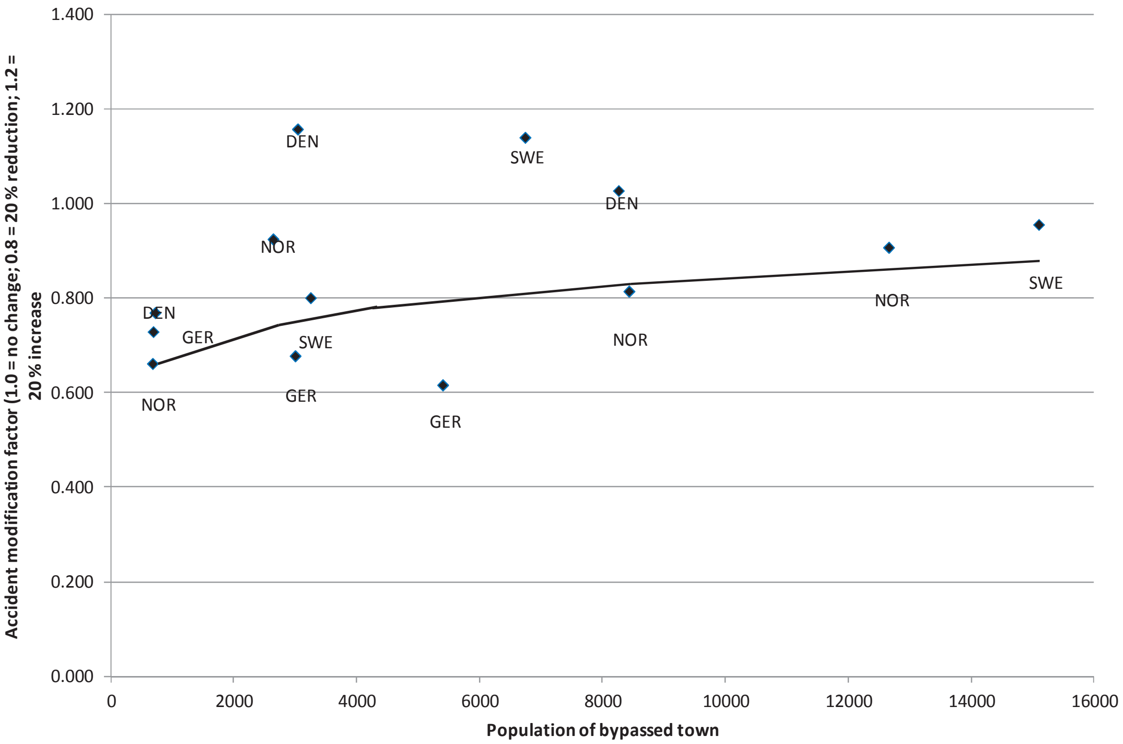
A wide range of roundabout diameters is included in the study, which is likely to include roundabouts in different road and traffic environments, varying in lane number, carriageway width and operating speeds. Future developments of the model should address these issues to avoid the CMF for roundabouts being influenced by confounding variables as discussed later in this chapter.

This model is obviously very simple and is still exploratory. However, it shows that by pooling studies from different countries, a pattern may emerge which is not always seen clearly within each country, and which may permit more accurate predictions to be made regarding variation in the effect of converting junctions to roundabouts.

4.1.3. Contextual variables influencing the effect of road safety measures

The crash modification function fitted to studies evaluating the effects of bypass roads by Elvik (2009) has been updated by adding a German study (Weissbrodt 1984). Figure 4.5. presents an updated crash modification function, fitted to data for four countries. It should be noted that data were grouped in each country in order to reduce random fluctuations. Thus, in Norway, data for 20 bypassed towns were merged to form four groups with maximum between-group differences in population, which is the independent variable of the crash modification function.

Figure 4.5. **Crash modification function for bypass roads**



In Figure 4.5., data points referring to each country have been identified (DEN = Denmark; GER = Germany; NOR = Norway; SWE = Sweden). Data points for the various countries are mingled throughout the range of the function. Norway has the first, fourth, eleventh and twelfth data points. The perceptive reader will notice that the data points for Germany deviate from the trend seen in the other three countries, in which the effects of bypass roads get smaller as the population of the bypassed town gets bigger. However, if data are aggregated even further, by merging data points that refer to almost the same size of population, the trend shown by the curve (a power function) is maintained. For example, the data point for Germany at a population of about 5,000 is counterbalanced by the data point for Sweden at a population of about 6,500.

In this case, it was possible to generalise the findings of evaluation studies by identifying a contextual variable that influenced the effect of the safety measure. However, the wide dispersion of data points in Figure 4.5. suggests that the effect of bypass roads is most likely influenced by other characteristics in addition to the population of the bypassed town. Note also that the CMF for each group of towns depends on the safety performance of the pre-existing road networks as much as the safety performance of the by-passes.

4.1.4. Lack of consistency in study designs precluding transferability

Periodic motor vehicle inspection is a road safety measure that has been evaluated in several countries over a long period. Table 1 lists evaluation studies that have been reviewed as part of the research for the Handbook of Road Safety Measures (Elvik, Høye, Vaa and Sørensen 2009). A total of thirteen studies are listed in Table 4.1. Many of these studies provided more than one estimate of effect, but in Table 4.1. only a single summary estimate of effect is presented for each study. The studies were reported during a period of 44 years in four different countries.

In principle, this should provide a sufficient range of replications to assess external validity (transferability) statistically, although few countries are included.

Table 4.1. **Studies that have evaluated the effects of periodic motor vehicle inspections.**

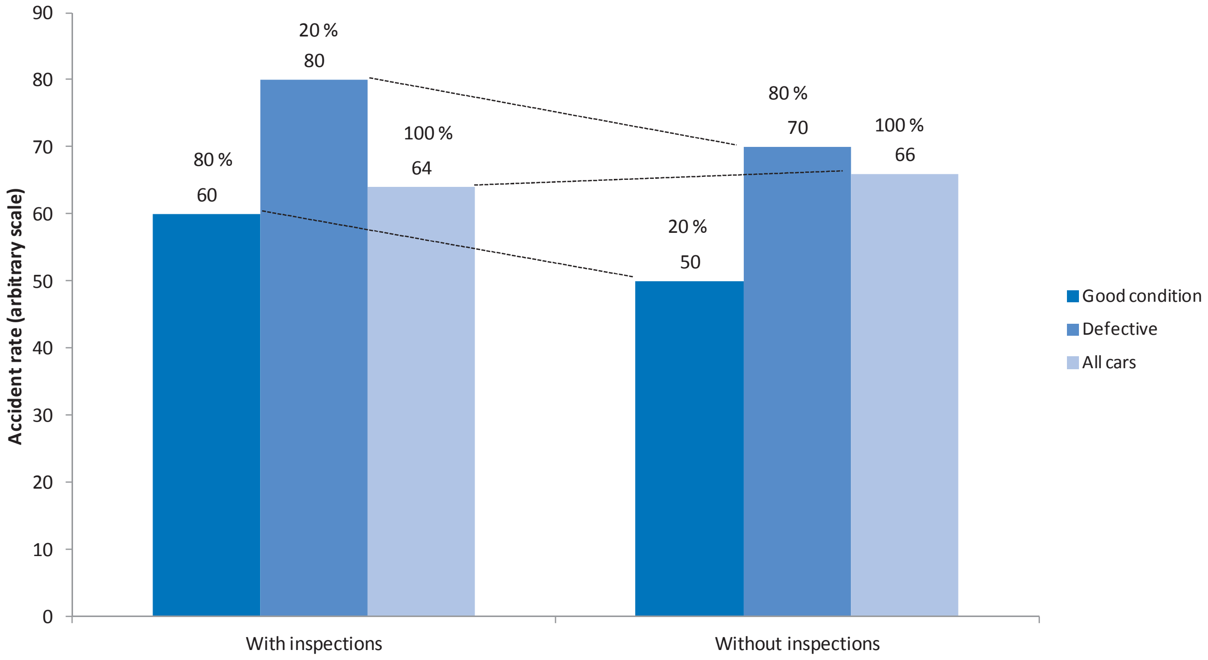
Authors	Year	Country	Level of data	Study design	Crash modification factor
Mayer and Hoult	1963	United States	Aggregate	Cross-section	0.83
Buxbaum and Colton	1966	United States	Aggregate	Cross-section	0.61
Fuchs and Leveson	1967	United States	Aggregate	Multivariate model	0.94
Foldvary	1971	United States	Aggregate	Regression discontinuity	0.73
Little	1971	United States	Aggregate	Before-and-after	0.96
Colton and Buxbaum	1977	United States	Aggregate	Multivariate model	0.83
Schroer and Peyton	1979	United States	Individual	Multivariate model	0.91
Crain	1980	United States	Aggregate	Multivariate model	0.97
Van Matre et al.	1981	United States	Aggregate	Multivariate model	0.89
Berg et al.	1984	Sweden	Aggregate	Multivariate model	0.79
White	1986	New Zealand	Individual	Trend analysis after	0.91
Fosser	1992	Norway	Individual	Experiment	1.01
Christensen and Elvik	2007	Norway	Individual	Multivariate model	1.05

Based on Elvik et al (2009)

When the studies are examined more critically, the basis for assessing transferability becomes less convincing. The first nine studies were all reported in the United States. The oldest studies were simple comparisons of crash rates (per inhabitant or per million vehicle kilometres of driving) between states that had periodic motor vehicle inspections and states that did not. It was soon realised that these comparisons did not control very well for potentially confounding factors. To better control for potentially confounding factors, multivariate models were developed in several studies, but most of these studies continued to use states as the unit of analysis. This is referred to as an "aggregate" level of data in Table 4.1. These studies did not have data on the crash rates of cars that were inspected, and the studies did not determine whether the crash rate of cars changed following an inspection. This is a fatal flaw of these studies.

While the overall crash rate is lower in states with periodic motor vehicle inspections than in states without them, it is entirely conceivable that cars in good technical condition as a result of inspections in states with periodic motor vehicle inspections could have higher crash rates than cars in the same condition in states without such inspections,. Figure 4.6. shows how this can occur.

Figure 4.6. **Potential aggregation bias in studies evaluating the effects of periodic motor vehicle inspections**



In Figure 4.6., it is assumed that 80 percent of cars in states with inspections are in good technical condition, and 20 percent are defective. Defective cars are assumed to have a higher crash rate than cars in good technical condition. In states without inspections, it is assumed that 20 percent of cars are in good technical condition, and 80 percent are defective. The mean crash rate will be a weighted average of the crash rates for cars in good technical condition and for defective cars. As can be seen in Figure 4, it has been assumed that both groups of cars have a higher crash rate in states with periodic motor vehicle inspections than in states without it. Despite this, the overall crash rate will be slightly lower in states with periodic motor vehicle inspections than in states without it (64 versus 66). If only the aggregate crash rates are known, the estimate of effect is misleading. If the odds ratio of crash rates is used as the estimator of effect, the estimate is: (60/80)/(50/70) = 1.05.

Although the example is hypothetical, it is nevertheless sufficient to undermine confidence in the studies relying on aggregate data only. If these studies are discarded, only four studies are left: one in the United States, one in New Zealand and two in Norway. These studies were all based on data referring to individual cars, but they differ with respect to study design. From a methodological point of view, the strongest study by far is the experiment from Norway (Fosser 1992). This was a very large randomised controlled field trial, in which (see later in the chapter) all important potentially confounding factors that may influence such trials were eliminated. It is, in fact, one of the methodologically best studies ever reported in the road safety evaluation literature.

The diversity of study designs employed in studies evaluating the effects of periodic motor vehicle inspections, and the poor quality of many of these studies precludes a quantitative assessment of the transferability of study findings.

4.1.5. Poor and variable control for confounding factors precluding transferability

The identification, analysis and treatment of hazardous road location, also referred to as black spots (Elvik 1997) or sites with promise (Hauer 1996) has a long tradition in traffic engineering. This approach to crash prevention has been widely applied and has been regarded as a rational and evidence-based approach to crash prevention. Thus, the Institution of Highways and Transportation noted (1990):

“It is well established that considerable safety benefits may accrue from application of appropriate road engineering or traffic management measures at hazardous road locations. Results from such applications at “black spots” demonstrating high returns from relatively low cost measures have been reported worldwide”.

In other words, there have been many studies of the effects of treating hazardous road locations. These studies have been reported in all parts of the world, and therefore their results can be applied to predict the effects of such treatments.

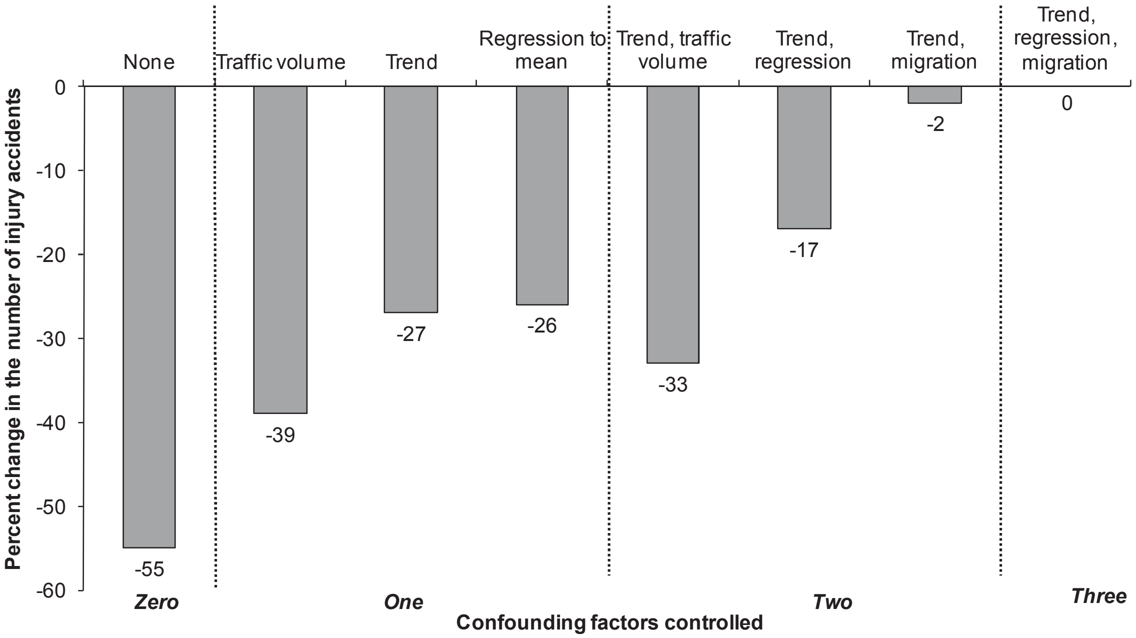
Unfortunately, the claim made by the Institution of Highways and Transportation does not withstand critical scrutiny. Elvik (1997) performed a meta-analysis of 36 evaluation studies reported during 28 years in six countries. In principle, these studies represent a sufficient range of replications for a formal assessment of their external validity to make sense. All the studies were before-and-after studies. Diversity of study designs was therefore not an issue.

When the studies were examined critically with respect to how well they controlled for potentially confounding factors, it was found that studies varied greatly in this respect. Four potentially confounding factors in before-and-after studies of treatment of hazardous road locations were considered:

1. Regression-to-the-mean
2. Long term trends in the number of crashes
3. Changes in traffic volume
4. Crash migration

It should be noted that these potentially confounding factors will not always actually confound study results. However, the only way to find out whether a potentially confounding factor actually does influence study results, is to control for the factor by means of study design or statistical estimation. Regression-to-the-mean is very likely to be present when a site has been selected for treatment based on a bad crash record. Figure 4.7. shows the effect on injury crashes for studies that controlled for various confounding factors.

Figure 4.7. **Effects attributed to black spot treatment in studies controlling for various confounding factors.**



Based on Elvik (1997)

The effect attributed to treatment varies greatly depending on which confounding factors were controlled for. Simple before-and-after studies, not controlling for any confounding factors, reported an impressive 55 percent crash reduction. Studies controlling for regression-to-the-mean, long term trends and crash migration found no effect of treatment. No study controlled for all four of the potentially confounding factors considered. There were only four studies that controlled for at least three of the four potentially confounding factors. Three of these studies had been reported in Great Britain, the fourth in the United States. These four studies had not evaluated the same types of treatment. Therefore, the studies do not provide any meaningful basis for assessing the international transferability of the results of studies that have evaluated the effects of treating hazardous road locations.

The conclusion is that when there are few studies of adequate methodological quality, and these studies do not evaluate the same road safety measure, there is no basis for assessing the international transferability of study findings.

4.2. Fundamental issues for a transferability framework [1]

When the implementation of an intervention, countermeasure, treatment, or decision is contemplated one of main questions is what change in safety it is likely to cause. An action (intervention, countermeasure, treatment, or decision) is said to have caused a change in safety if the change would not have occurred without it.

Consider some actions such as illuminating an unlit stretch of road, reducing the legal blood alcohol content (BAC) from 0.08 to 0.05 ml/l, or using an 800 m radius for a horizontal curve instead of a 600 m radius. In each case the comparison is of (at least) two actions, to be denoted ‘a’ and ‘b’. In the examples above, ‘a’ could stand for ‘illuminate’, ‘reduce’, and ‘use 800m radius’ while ‘b’ could stand for ‘leave unlit’, ‘keep at 0.08’, and ‘use 600 m radius’. The expected target crash frequency with the action implemented, denoted by μ_a, is compared with the expected target crash frequency prevailing under identical conditions but without the action, denoted by μ_b. Research results usually report estimates of the ratio μ_a/μ_b. This ratio is the Crash Modification Function (CMF) of implementing ‘a’ instead of ‘b’ and is denoted by[2] θ(a,b).

$$CMF\ for\ implementing\ 'a'\ instead\ of\ 'b' \equiv \theta(a,b) \equiv \frac{Expected\ (target)\ crashes\ with\ 'a'}{Expected\ (target) crashes\ in\ identical\ conditions\ but\ with\ 'b'} \equiv \frac{\mu_a}{\mu_b} \tag{1}$$

If the expected crash frequency with an intervention implemented is, for example, 61.1 crashes/year while the expected crash frequency under identical conditions but without the intervention is 73.2 crashes/year then the CMF for that intervention is θ(with intervention, without intervention) = 61.1/73.2=0.83. When the numerator and denominator are equal, then doing ‘a’ instead of ‘b’ has no effect on the expected crash frequency and θ(a,b)=1. When doing ‘a’ instead of ‘b’ reduces the expected target crash frequency, then the CMF for that intervention is less than 1 (θ(a,b)<1). Conversely, when doing ‘a’ instead of ‘b’ increases the expected crash frequency, then the CMF for the intervention is greater than 1 (θ(a,b)>1).

Estimates of the CMF (θ(a,b)) come from research and more will be said about this in the next section. The main use of an estimate of the CMF ($\hat{\theta}(a,b)$) of θ(a,b) is for predicting what we expect to be the safety effect of doing ‘a’ instead of ‘b’ in some specific circumstances. Thus, transposing terms in equation 1 and adding the caret to signify ‘estimate’,

$$\hat{\mu}_a = \hat{\mu}_b \times \hat{\theta}(a,b). \tag{2}$$

For example, if research indicates that the estimated CMF ($\widehat{\theta}$) for implementing an action is 0.89 and we estimate that without implementation there would be 271.3 target crashes, then with the action implemented we should expect 271.3×0.89=241.5 target crashes.

The safety effect of an action is usually measured in terms of the expected change in the number of target crashes (classified by severity) which is caused by that action. The estimate of this expected change is

$$\hat{\mu}_b - \hat{\mu}_a = \hat{\mu}_b \times [1 - \hat{\theta}(a,b)]. \tag{3}$$

Therefore, an estimate of the CMF, $\hat{\theta}(a,b)$, is needed for the prediction of the safety effect. What should be used for CMF estimate ($\hat{\theta}(a,b)$) and how reliable it is will be discussed next.

4.2.1. From past research to future action[3]

The need to use a CMF estimate ($\hat{\theta}(a,b)$) arises when the potential safety benefit of a future action is contemplated. Guidance about what estimated CMF ($\hat{\theta}(a,b)$) to use in decision-making comes from research about the safety consequences of similar actions taken in the past.

For example, suppose that the question is whether to illuminate a certain stretch of access-controlled road in Estonia. We need the estimated CMF, or $\hat{\theta}(\text{illuminate, do not illuminate})$, that will be applied to this stretch of Estonian road, with its climate, vehicles, users, and illumination design. Imagine that there are only two past research studies about the safety effect of illumination on access-controlled roads[4]. One is from Arizona with data from 1992-1995 in which the estimated CMF for night-time injury accidents was 0.75 with a standard error (s) of ±0.04; the other study is from British Columbia (B.C.) with data from 2001-2006 in which the estimated CMF was 0.62 with s=± 0.06. If 'A' stands for Arizona, 'B' for British Columbia, and '^' for "estimate", we have an estimated CMF for Arizona $(\hat{\theta}_{A)} = 0.75 \pm 0.04$ and an estimated CMF for British Columbia $\widehat{(\theta_{B)}} = 0.62 \pm 0.06$. What should the estimated CMF $(\hat{\theta}(\text{illuminate, do not illuminate}))$ be in equation 3 to determine the safety effect in Estonia?

Illumination[5] does not have the same effect everywhere and at all times. For some kinds of actions one may find a small variability of the safety effect from one application to another; for other actions the variability may be large. That is, a CMF (θ(a, b)) is not a constant but is instead a variable that has a probability distribution. Whether the variance of the CMF (θ(a, b)) is small or large depends on how sensitive it is to the details of the actions 'a' or 'b' and to the variety of circumstances in which it is implemented. The size of the variance is an empirical question and ways to answer it will be described later.

Thinking of a CMF (θ(a,b)) as a random variable allows us to correctly frame the question of 'transferability'. The issue is this: in a cost-effectiveness or cost-benefit framework, decisions are based on expected consequences. This is why, to predict the future safety effect by equation 3 we use an estimate $(\bar{\theta})$ of the expected value of the CMF (θ) from past research, as the estimated CMF $(\hat{\theta}(a, b))$**.** The problem is that the estimated CMF $(\bar{\theta})$ based on past implementations in Arizona and British Columbia is not the CMF (θ) of the future safety effect of illuminating the road in Estonia. It is the difference between the CMF (θ) and the estimated CMF $(\bar{\theta})$ that determines whether the decision to illuminate or not to illuminate the road in Estonia is right. Thus, concern about transferability amounts to concern about how well the estimated CMF $(\bar{\theta})$ based on past implementations predicts the CMF (θ) of a future implementation.

When past research indicates that whenever 'a' was done instead of 'b' approximately the same CMF (θ) was found, especially under similar circumstances, the issue of transferability in most cases does not arise[6]. Transferability concerns are real when the variance of a CMF for doing "a" instead of "b" (θ(a,b)) is large.

In this sense the question of 'transferability' is not only about how well the past experience of one country predicts the future safety effect of a similar action in another one. Concern about transferability arises whenever the variance of a CMF for doing "a" instead of "b" (θ(a,b)) is large, irrespective of whether the future application is in a different country, city, project or time period. If the variance of the CMF (θ(a,b)) is large, one should ask why this is so, and how it can be reduced.

4.2.2. What we can observe and what we need to know

Past research produces the estimates of a CMF - the $\hat{\theta}$s. The $\hat{\theta}$s usually show a degree of diversity as shown in Figure 4.8.

Figure 4.8. **What is observed and what we want to know**

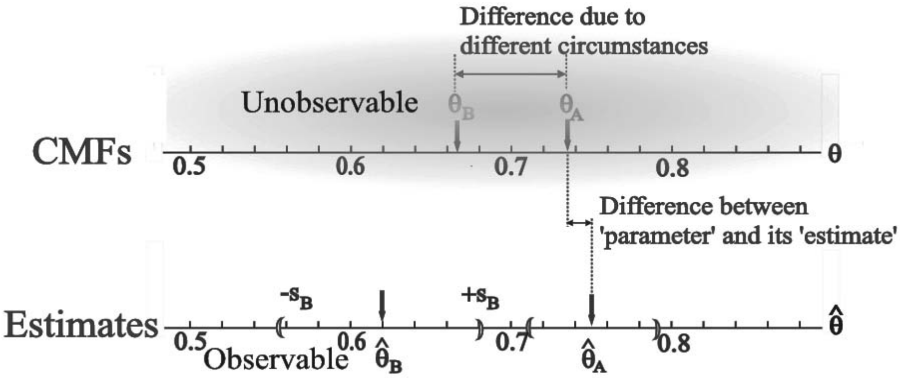


Figure 4.8. consists of two tiers. In the lower tier are the CMF estimates ($\hat{\theta}$s): the Arizona estimate ($\hat{\theta}_A$) at 0.75, and the British Columbia estimate $\hat{\theta}_B$ at 0.62. The estimates are surrounded by brackets of ±one standard error. The values on this axis are computed from data and are therefore 'observable'. The upper tier shows the parameters θ_A and θ_B; those are what the estimates in the lower tier would converge to if estimation could be repeated many times under identical conditions. Since this cannot be done, these values are never known and are therefore 'unobservable'; this is why they are shown as if they were behind a cloud.

The difference between the unobservable parameters θ_A and θ_B is due to the many differences in the circumstances of the implementations in Arizona (A) and British Columbia (B). These circumstances could include the procedures followed, the units used, the environment, the road users, and the years under consideration. The other difference in Figure 4.8 is shown between the two tiers. This is the statistical difference between the parameter θ_A and its estimate $\hat{\theta}_A$; a similar difference exists between θ_B and $\hat{\theta}_B$. These 'statistical' differences are due to the limitations of data, method, and sample size.

There is no reason to think that the specifics of illumination design (see footnote 5) or the traits of roads, road users, traffic, twilight duration, in Arizona and British Columbia were the same. Nor is there reason to think that the safety effect of illumination does not depend on implementation details and unit traits. This is why the two CMFs (θs) in the upper tier are shown as different[7]. For simplicity, Figure 4.8 shows only two CMFs (θs) from two past research studies. In principle, however, there could be many such research studies, each with its θ. These θs would have a probability distribution with a mean ($E\{\theta\}$) and a standard deviation ($\sigma\{\theta\}$).

In order to make a decision about a future action, we need to predict the θ for that future action. The assumption is that the future will be similar to the past. If so, the θ for the future action will be one of the values from the probability distribution[8] of past θs. For decision-making it is best to assume that the θ for the future action is the estimate $\bar{\theta}$ of $E\{\theta\}$, the mean of past experiences. The $\bar{\theta}$ has a standard error to be denoted by $s\{\bar{\theta}\}$.

If $s\{\bar{\theta}\}$ and $\sigma\{\theta\}$ are small, equation 3 can be used confidently to predict the safety effect of doing 'a' instead of 'b'. If $s\{\bar{\theta}\}$ and/or if $\sigma\{\theta\}$ are large, prediction by equation 3 will be inaccurate and decisions based on such a prediction will be in danger of being wrong[9]. To judge the accuracy of predicting the safety effect by using equation 3, we need estimates of three elements: $\bar{\theta}$, $s\{\bar{\theta}\}$ and $\sigma\{\theta\}$.

4.2.3. Estimating the CMF, standard error and standard deviation for a future action

In this section we describe a way of estimating $\bar{\theta}$ and its standard error. Suppose that from past research we have the unbiased estimates $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_i, \cdots, \hat{\theta}_n$ and their standard errors $\pm s_1$, $\pm s_2$,, …, $\pm s_i$, …,$\pm s_n$. The weighted average of these estimates is the linear combination $\sum_1^n (w_i/\sum_1^n w_i)\hat{\theta}_i$ in which w_i is the raw weight of the i-th estimate. The variance of this linear combination is $\sum_1^n (w_i/\sum_1^n w_i)^2 \mathrm{VAR}\{\hat{\theta}_i\}$. This variance is smallest[10] when w_i is proportional to $1/\mathrm{VAR}\{\hat{\theta}_i\}$. When this weight is used, the variance of the weighted average is $1/\sum_1^n (1/\mathrm{VAR}\{\hat{\theta}_i\})$. Replacing $\mathrm{VAR}\{\hat{\theta}_i\}$ by s_i^2

$$\text{Weighted Average} \equiv \bar{\theta} = \sum\nolimits_1^n \frac{1/s_i^2}{\sum_1^n 1/s_i^2}\hat{\theta}_i \tag{4}$$

$$\widehat{\mathrm{VAR}}\{\bar{\theta}\} \equiv s^2\{\bar{\theta}\} = \frac{1}{\sum_1^n 1/s_i^2} \tag{5}$$

Computations using the Arizona and British Columbia data are shown in Table 4.2.

Table 4.2. **Computations (invented data)**

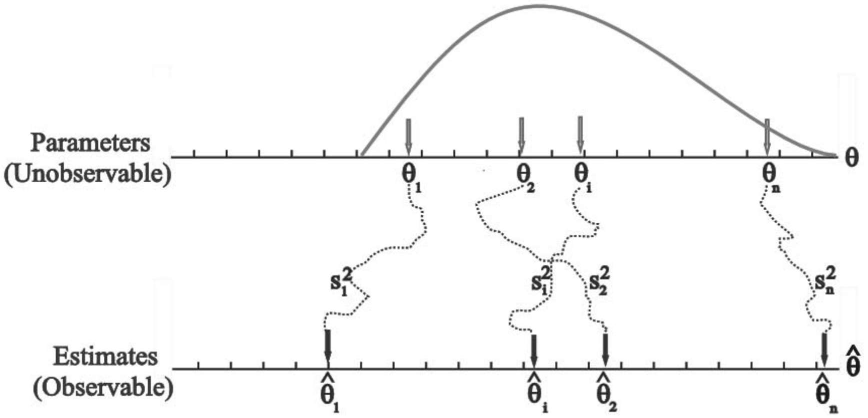
	$\hat{\theta}$	**S**	**s^2**	**$1/s^2$**	**Weights**	**Contributions to mean**
Arizona	**0.75**	**±0.04**	**0.0016**	**625**	**0.69**	**0.519**
B.C.	**0.62**	**±0.06**	**0.0036**	**278**	**0.31**	**0.191**
Sums				**903**	**1**	**$\bar{\theta}$ =0.710**

The shaded part of the table contains the data. The next three columns show the computation of the weights. The weighted contribution of each $\hat{\theta}$ is in the last column; their sum, the weighted average, is 0.71. The variance of $\bar{\theta}$ is estimated to be 1/903=0.0011 and $s\{\bar{\theta}\} = \sqrt{1/903} = \pm 0.03$.

If one could reasonably assume that the θ of illuminating limited access roads was the same in Arizona in the late nineties as in British Columbia in the early two thousands and, furthermore, that the θ will be the same in Estonia in the future, then one could take the ±0.03 as describing the uncertainty about the θ for Estonia. However, it is unreasonable to do so. The uncertainty about θ for Estonia is not only the uncertainty about the $\bar{\theta}$; it is also the question of how variable θ is from one application to another. This variability is measured by σ{θ}. Thus, the next task is to estimate σ{θ}.

Figure 4.8. was closely linked to the Arizona-B.C. example. Figure 4.9. is a more generic visualization of the situation.

Figure 4.9. **Parameters and their estimates.**



The parameters, θs, form the upper tier of the figure. For each unobservable parameter (θ) in the upper tier there is an observable estimate ($\hat{\theta}$) in its lower tier. The parameters and their estimates are linked with wiggly lines representing random variation. The nature of this link is that the estimates are unbiased and have finite variances. Estimates of these variances are shown in the figure as s_1^2, s_2^2, …, s_i^2, …, s_n^2. The 'n' parameters in the upper tier can be thought of as a random sample from a distribution with a mean and a variance. In a previous section we discussed how to estimate the E{θ}. In this section the question is how to estimate VAR{θ}from the $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_i, \cdots, \hat{\theta}_n$ in the lower tier.

Because the θs of the upper tier are linked to the $\hat{\theta}$s in the lower tier, VAR{θ}is related to the $\mathrm{VAR}\{\hat{\theta}\}$. To understand the relationship between the two, imagine first that estimation is without error. If so, the wiggly lines in Figure 4.9 would be straight and vertical and VAR{θ) would equal $\mathrm{VAR}\{\hat{\theta}\}$. However, estimation always involves error and the $\hat{\theta}$s tend to be more widely dispersed than the θs. This is why $\mathrm{VAR}\{\hat{\theta}\} > VAR\{\theta\}$. The larger the s_i, the larger the difference $\mathrm{VAR}\{\hat{\theta}\} - \mathrm{VAR}\{\theta\}$ will be. Using the law of total variance[11] it can be shown that

$$\mathrm{VAR}\{\theta\} = \mathrm{VAR}\{\hat{\theta}\} - \mathrm{E}\{\mathrm{VAR}\{\hat{\theta}|\theta\}\} \tag{6}$$

If the weighted average $\bar{\theta}$ is the same as E{θ} we can estimate VAR{θ} by $\hat{\mathrm{V}}$:

$$\hat{\mathrm{V}} = \begin{cases} \dfrac{\sum_1^n (\hat{\theta}_i - \bar{\theta})^2}{\mathrm{n\ or\ (n-1)}} - \dfrac{\sum_1^n s_i^2}{n} & \text{if positive} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

However, inasmuch as $\bar{\theta}$ has a positive variance (the estimate of which is in equation 5), it has to be included. Denote the expected value of the squared difference θ-$\bar{\theta}$ by VAR*(θ). We estimate VAR*{θ} by:

$$\widehat{\mathrm{VAR}}^*\{\theta\} = \hat{\mathrm{V}} + \widehat{\mathrm{VAR}}\{\bar{\theta}\} = \hat{\mathrm{V}} + \frac{1}{\sum_1^n 1/s_i^2} \tag{8}$$

The Arizona-British Columbia data is used again (in the shaded part of Table 4.3).

Table 4.3. **Computations**

	$\hat{\theta}$	S	$(\hat{\theta}-\bar{\theta})^2$	s^2
Arizona	**0.75**	**±0.04**	**0.0016**	**0.0016**
B.C.	**0.62**	**±0.06**	**0.0081**	**0.0036**
Averages			$\widehat{\mathrm{VAR}}\{\hat{\theta}\} = 0.00485$[12]	$\hat{E}\{\mathrm{VAR}\{\hat{\theta}\|\theta\}\}=0.0026$

From the data in Table 4.3 we have $\hat{V} = 0.0048\text{-}0.0026\text{=}0.0022$. Since $\widehat{\mathrm{VAR}}\{\bar{\theta}\} = 1/903 = 0.0011$, $\widehat{\mathrm{VAR}}^*\{\theta\}\text{=}0.0033$ and $\hat{\sigma}^*\{\theta\} = \sqrt{0.0033} = \pm 0.06$.

At this point, we can summarize how to use the Arizona and British Columbia experience to predict the safety effect of illumination in Estonia.

To predict the safety effect (by equation 3) we should use $\hat{\theta}(a,b) = \bar{\theta} = 0.71$. Considering that the θ is likely to be somewhere within ±2 standard errors, the corresponding range is about (0.59-0.83). If to break even[13] in this project θ has to be less than 0.59, then the decision should be not to illuminate. If to break even it is enough for θ to be above 0.83 the decision should be to illuminate. For break-even values in the (0.59-0.83) range there is a good chance of the decision being wrong. The narrower the range $\bar{\theta} \pm 2\sigma^*\{\theta\}$, the smaller that chance of making a wrong decision.

The Arizona-British Columbia-Estonia story was only an example. The general framework is now clear. Equation 3 is used to predict the safety effect of doing 'a' instead of 'b'. For the best predictions, for $\hat{\theta}(a,b)$ in equation 3 one should use a weighted average of past research findings, $\bar{\theta}$. How $\bar{\theta}$ and its standard error $s\{\bar{\theta}\}$ can be computed is shown in equations 4 and 5. The $\bar{\theta}$ which we use to predict the safety effect is not the θ that will materialize in the future. Whether the decision based on our predicted safety effect will be right or wrong will depend on the difference between θ and $\bar{\theta}$. How large the difference might be is measured by $\sigma^*\{\theta\}\text{=}\sqrt{\mathrm{VAR}^*\{\theta\}}$. Estimators are in equations 7 and 8.

When $\pm 2\sigma^*\{\theta\}$ is in a narrow range around $\bar{\theta}$, decisions can be made confidently and no new research is needed. When the range is so wide that it covers the break-even value of θ, there is a good chance for the decisions based on $\bar{\theta}$ to be wrong. In this case, new research to reduce $\sigma^*\{\theta\}$ may be needed.

4.2.4. How to improve the prediction of safety effect

When the variance of θ is large we are uncertain what the θ of some future action will be, and therefore cannot confidently predict the safety consequences of that action. When the safety effect is uncertain it is difficult to make good decisions. To reduce the chance of making bad decisions one must reduce the variance of θ.

As shown in equation 8, $\mathrm{VAR}^*\{\theta\}$ is the sum of two elements. One element, $\mathrm{VAR}\{\bar{\theta}\}$, measures the uncertainty about how the weighted mean of past results ($\bar{\theta}$) differs from E(θ). The more studies are done, and the more $\hat{\theta}$s go into the determination of $\bar{\theta}$, the smaller this source of uncertainty becomes. The other element of $\mathrm{VAR}^*\{\theta\}$, V, reflects that part of the variability of θs which is due to differences in the 'circumstances' of implementation.

That is, how different will the θs tend to be from one instance of implementation to another? The variance due to this source cannot be reduced by doing more and more studies; it can be reduced only by determining how the θs depend on this or that circumstance of implementation. These two options for reducing the VAR*{θ} will be examined separately.

Reducing VAR{θ} by doing more studies*

The estimate of E{θ} is the weighted mean $\bar{\theta}$ (equation 4). If the number of past study results is small and their standard errors are large, VAR{$\bar{\theta}$} will be large. Doing additional research to produce new estimates $\hat{\theta}_{n+1}, \hat{\theta}_{n+2}, \cdots, \hat{\theta}_{n+m}$ with standard errors $\pm s_{n+1}, \pm s_{n+2}, \ldots, \pm s_{n+m}$ will reduce this variance by[14]:

$$\text{Reduction in VAR}\{\bar{\theta}\} = \frac{1}{\sum_1^n 1/s_i^2} - \frac{1}{\sum_1^{n+m} 1/s_i^2} \tag{9}$$

The magnitude of standard errors (the $\pm s_{n+1}, \pm s_{n+2}, \ldots, \pm s_{n+m}$) is to some extent in our hands; it is based on what data and resources are available for the new research. Consider again the data in Table 4.2 where $\sum_1^2 1/s_i^2$ was 903. If a proposed new study (m=1) allows one to estimate the θ with $s_3 = \pm 0.03$ then the new sum ($\sum_1^3 1/s_i^2$) will be $903+1/(0.03)^2= 903+1111= 2014$. If so, the reduction in VAR{$\bar{\theta}$} would be 1/903-1/2014=0.0011-0.0005=0.0006.

The marginal benefit of conducting more research declines rapidly. If a few past unbiased study results already exist, and there is still wide uncertainty about what the θ in a future application will be, the only way to reduce the uncertainty is to make θ a function of circumstances.

Reducing VAR{θ} by making θ a function of circumstances*

Suppose that several past study results already exist but VAR{θ} is still too large. Now one must ask: "Why is it that the same kind of action has diverse safety effects? Can one divide the extant study results into groups that have a circumstance in common such that within each group VAR{θ} will be small? The Arizona-B.C. hypothetical example will be developed further to show how making θ a function of a variable can reduce VAR{θ}.

The latitude of Phoenix is 33.5º N and of Vancouver is 49.2º N. Could this difference partly account for the difference between $\hat{\theta}_{\text{Arizona}}$=0.75 and the $\hat{\theta}_{\text{B.C.}}$=0.62 (Table 4.3)? Two data points are insufficient for speculation about why the θ estimates differ. However, there is other research to suggest that the safety effect of illumination may depend on geographical latitude[15]. This question could be answered by estimating θ at a few more latitudes. If the resulting relationship seems regular, then one can estimate how θ depends on latitude and thereby reduce VAR{θ}.

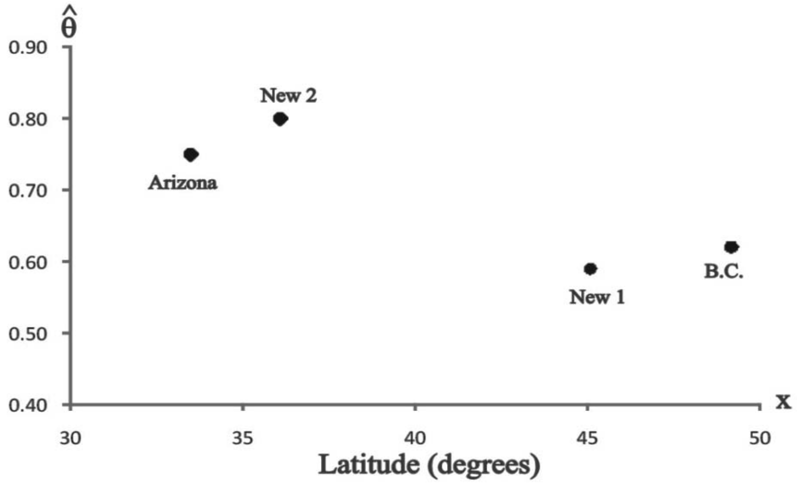
Suppose that new research was done at two more locations with both old and new results shown in Table 4.4.

Table 4.4. **Two new estimates (invented data)**

	$\hat{\theta}$	s	Latitude °N (x)
Arizona	**0.75**	**±0.04**	**33.5**
B.C.	**0.62**	**±0.06**	**49.2**
New location 1	**0.80**	**±0.02**	**36.1**
New location 2	**0.59**	**±0.02**	**45.1**

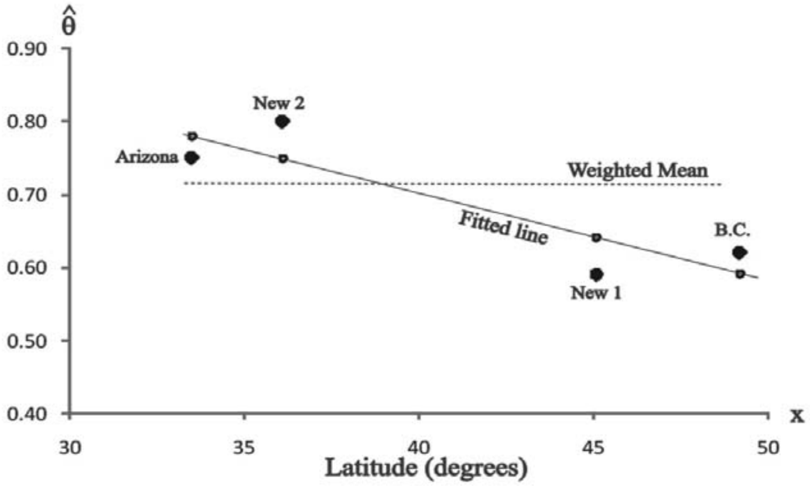
Figure 4.10 is based data from Table 4.4. and shows how the $\hat{\theta}$s vary with latitude. The two new estimates seem to confirm the speculation that the further north you go, the larger the safety effect of illumination is (seen as a smaller numerical value of θ).

Figure 4.10. **$\hat{\theta}$ versus Latitude (invented data)**



Assuming that the relationship is approximately linear, a regression line[16] was fitted to these data points as shown in Figure 4.11.

Figure 4.11. **Linear regression to the data in Table 4.4.**



The benefit in variance reduction is now plain to see. In Figure 4.11., the initial estimation of VAR{θ} was based on the average squared difference between the solid circles (the $\hat{\theta}$s) and the dotted horizontal line (weighted mean). However, when the influence of latitude is considered, the squared differences are now those between the solid circles and the empty circles (the $\bar{\theta}_i$s) on the fitted regression line. Because the fitted line comes closer to the data points than the horizontal line, the squared differences will now be much smaller. The change from computing the squared differences around the weighted mean to computing them around values on the fitted regression line requires modifications in equation 7 and 8. The $\bar{\theta}$ in equation 7 has to be replaced by $\bar{\theta}_i$ (the empty circles in Figure 4.11.) and equation 8 is replaced by

$$\widehat{VAR}^{*}\{\theta_i\} = \hat{V} + \widehat{VAR}\{\bar{\theta}_i\}$$
$$\text{Where } \widehat{VAR}\{\bar{\theta}_i\} = \frac{\sum_1^n \left(\hat{\theta}_i - \bar{\theta}_i\right)^2}{n-2}\left[\frac{1}{n} + \frac{(\bar{x}-x_i)^2}{\sum_1^n (\bar{x}-x_i)^2}\right] \qquad (10)$$

Table 4.5. shows the estimates of VAR*$\{\theta_i\}$ based on the modified equations and data from Table 4.4.

Table 1.5. **Estimates of VAR*$\{\theta_i\}$**

	$\widehat{VAR}^{*}\{\theta_i\}$	$\pm\hat{\sigma}^{*}\{\theta_i\}$
Arizona	**0.0023**	**0.05**
B.C.	**0.0025**	**0.05**
New location 1	**0.0016**	**0.04**
New location 2	**0.0014**	**0.03**
Average	**0.0020**	**0.04**

The estimate of σ*{θ} is ±0.08, based on the data in Table 4.5., and not taking into account any dependence on latitude. However, when θ is a function of latitude, the $\hat{\sigma}^{*}\{\theta_i\}$ are between ±0.03 and ±0.05. The moral is that when the available $\hat{\theta}$s exhibit a regular dependence on some variable, making that dependence explicit is an effective way to reduce the VAR*{θ} and to improve the quality of decision making. In the "illumination" example, decisions everywhere in the world would be based on the same $\bar{\theta}$ if latitude is not taken into account. Decision making is greatly improved, though, when the decision is tailored to the latitude of the project by using $\bar{\theta}_i = 1.18 - 0.0120 \times \text{Latitude}$.

4.3. Classifying and rating study designs to enhance transferability

The examples given in the previous sections show that when there is a sufficient number of studies with acceptable methodological quality, it may be possible to develop a crash modification function that describes systematic variation in the effect of a road safety measure. By applying the function, one may predict the effect of the measure in countries where no evaluation studies have been made.

If, on the other hand, available studies are of poor or highly variable methodological quality, assessing the transferability of their findings is not possible. What needs to be done in these cases is to embark on a programme of research in as many countries as possible, performing studies by means of the same study design in all countries in order to develop a sufficient number of estimates of effect to permit a formal assessment of international transferability.

The purpose of this section is to give some short guidelines regarding the development of study designs that adequately control for potentially confounding factors and thus permit the effects of road safety measures to be estimated. These guidelines will be given for three study designs:

1. Experiments (randomised controlled trials)
2. Before-and-after studies
3. Multivariate models

Experiments (randomised controlled trials)

Although considered the ‘gold standard’ in evaluation research, experiments are rare in the field of road safety. An experimental study design ought to be applied more often, since in theory an experimental design would control for all confounding factors. However, this may not be the case and attention must be paid to the following potentially confounding factors in randomised controlled trials (Elvik 2008):

4. Failure of randomisation to ensure pre-trial equivalence between groups
5. Incomplete treatment implementation and/or treatment diffusion
6. Differential rates of attrition
7. Unintended side effects (“Hawthorne” effect)

The random assignment of experimental subjects to one or more treatment groups and a control group should ensure that these groups are identical or very similar with respect to all characteristics except the treatment whose effects the experiment is designed to measure. However, if the sample size is small, randomisation may not always be successful. In addition, a matched pair design, in which pairs are matched with respect to variables that are believed to influence the number of crashes, might not be the best choice if one member of each pair assigned at random to the treatment group is vulnerable to bias.

If a treatment is not fully implemented, or if the treatment is introduced in the control group to almost the same extent as in the treatment group, the experiment is compromised, since the groups do not then differ with respect to the treatment.

Differential rates of attrition refer to the fact that withdrawal from an experiment differs between the treatment group and the comparison group. This problem is perhaps most likely to occur when the treatment is felt to be onerous, and those assigned to the treatment group want to escape from it. Systematic differences between the groups may then arise as a result of withdrawal from the experiment.

Unintended side effects are all effects that are not intended by the treatment, but that may arise as a result of behavioural adaptation among subjects. Such effects were observed for the first time in industrial experiments designed to enhance productivity and were named “Hawthorne” effects. In medical trials, a well known unintended effect is the placebo effect. To identify unintended side effects, one should describe as precisely as possible the intended effects of a treatment in order to recognise those effects that are not intended.

It should be noted that unintended side effects, such as those first observed in the Hawthorne factories, are considerably more likely to occur in experiments involving human subjects than in experiments not involving human subjects. One way of guarding against unintended behavioural adaptation to different experimental conditions is to perform double blind experiments; i.e. keep both experimenters and subjects ignorant of the treatment condition to which a subject has been assigned.

This design is very common in medical trials of new drugs. It may, however, be more difficult to implement when a safety treatment is being tested.

A high quality randomised controlled trial will test formally whether any of the sources of confounding factors are present or not. Thus, the experimental study of periodic motor vehicle inspections in Norway tested for all confounding factors and found that none of them confounded the study (Fosser 1992). The results of this test are reported in Table 4.6.

As can be seen from Table 4.6., there were minimal differences between the three groups (cars inspected annually, cars inspected once, cars not inspected) with respect to percentage with collision insurance, mean deductible amount on collision insurance, mean annual distance driven, mean birth year of car owner and the percentage of car owners who were male. Randomisation was, in other words, successful with respect to these variables.

Close to 70 percent of the cars in the two treatment groups were inspected; 14 percent of cars in the control group were also inspected. Thus, the groups did differ with respect to implementation of treatment, although the actual rate was not 100 percent in the treatment groups and 0 percent in the control group.

The rate at which cars were scrapped was very similar in the three groups and not likely to generate any bias (no differential attrition).

Finally, owners might adapt to the experiment by selling their car or by repairing it on their own if they knew it was going to be inspected frequently. No evidence of such unintended side effects was found.

Table 4.6. **Tests for confounding factors in an experimental study of periodic motor vehicle inspections. Based on Fosser 1992**

	Experimental condition		
	Inspected annually	**Inspected once**	**Not inspected**
	Check of pre-trial equivalence		
Percentage with collision insurance	**75.2**	**75.0**	**75.1**
Mean deductible amount on collision insurance (NOK)	**1836**	**1841**	**1828**
Mean annual insured driving distance (km)	**13370**	**13370**	**13280**
Mean year of birth of car owner	**1947.2**	**1947.1**	**1947.0**
Percentage of male car owners	**77.0**	**77.0**	**77.8**
	Check of treatment implementation and diffusion		
Percentage of cars treated	**67.8**	**68.6**	**14.2**
	Check of differential rates of attrition		
Percentage of cars scrapped at the end of the period	**6.1**	**6.1**	**6.6**
	Check of unintended side-effects		
Percentage of cars changing owners in last year of study	**4.8**	**5.0**	**5.0**
Defects repaired by owner before inspection	**0.60**	**0.71**	**0.69**

To make sure an experiment actually goes well, researchers should always check pre-trial equivalence between the groups, monitor the rate of treatment implementation and diffusion, monitor rates of attrition and investigate if unintended side effects are present. Should an experiment fail in any of these respects, it may sometimes be possible to adjust statistically for the differences. However, the study should then be treated as an observational study, not an experiment.

Before-and-after studies

Observational before-and-after studies are very commonly used in road safety evaluation. The techniques for doing such studies have developed considerably in recent years. The empirical Bayes (EB) approach is currently recommended as the state-of-the-art approach; a detailed textbook-like description of this approach, with worked numerical examples is given in the Highway Safety Manual (2010). The following potentially confounding factors are relevant in before-and-after studies:

1. Regression-to-the-mean
2. Long term trends in crashes
3. Exogenous changes in traffic volume
4. Co-incident events
5. Use of multiple measures at the same site
6. Crash migration

There are many versions of the EB approach. As presented in the Highway Safety Manual, it will control for regression-to-the-mean, long term trends and changes in traffic volume.

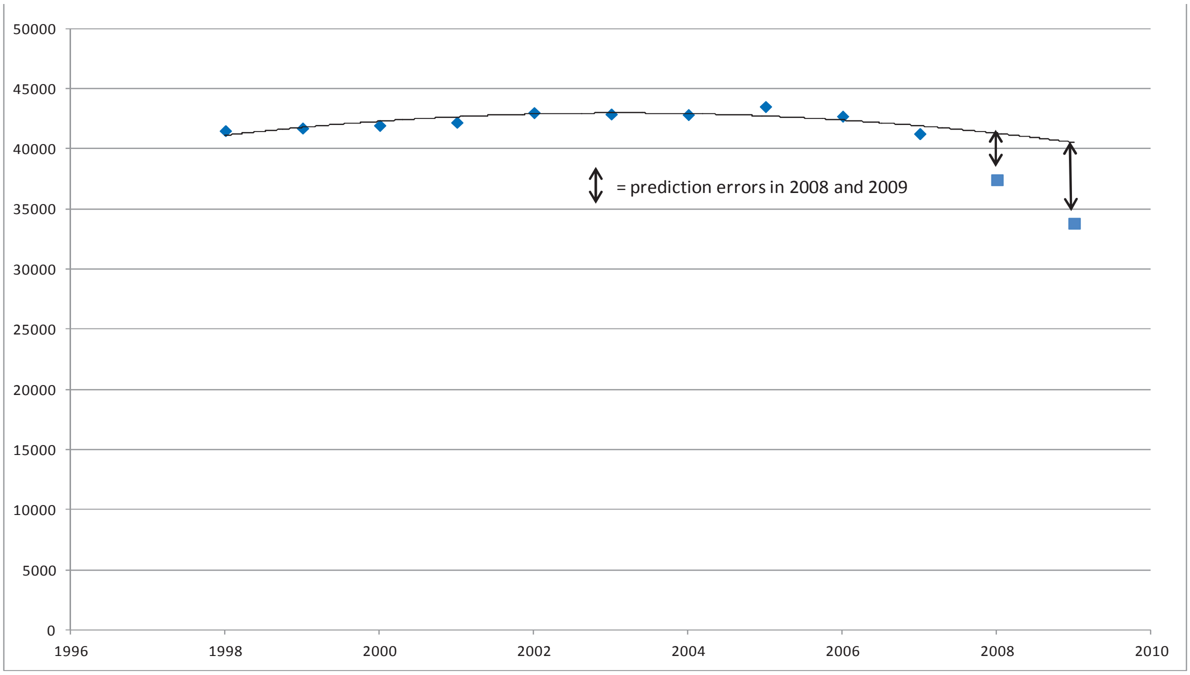
Changes in traffic volume should only be controlled for if they are not the result of the measure introduced. If, for example, traffic calming measures are introduced in order to discourage traffic in residential areas, a reduction in traffic volume is an intended effect of the measure and should not be controlled for. One should only control for those effects that would otherwise have happened, i.e. those effects that would be expected to occur even if the measure had not been introduced.

A distinction should also be made between long term trends and co-incident events. A co-incident event is an event that takes place at the same time as the introduction of a road safety measure, and that leads to a clear breaking point in the long term trend, meaning that prolonging the trend would not accurately predict the number of crashes expected to occur in the after-period.

The distinction between a long term trend and a co-incident event can be illustrated by the effect of the recent economic recession on the number of traffic fatalities in the United States. Figure 4.12. shows the number of traffic fatalities from 1998 to 2009.

The trend during the period 1998-2007 is well described by a second degree polynomial. This function captures the fact that a decline in the number of fatalities started in 2006. However, the declines in 2008 and 2009 were considerably greater than predicted by the trend line, and represent a break in the trend. The sharp decline in the number of fatalities in 2008 and 2009 is unlikely to be attributable mainly to road safety measures implemented in those years.

Figure 4.12. **Number of traffic fatalities in the United States 1998-2009**



In an evaluation study using 2007 as the before-period and 2009 as the after-period, one should therefore regard the large drop in fatalities as a confounding factor that should be controlled for. A road safety measure would only be regarded as effective if it was associated with a larger reduction of the number of fatalities than the reduction seen in Figure 4.12.

Sometimes, more than one road safety measure is implemented at a site. An evaluation study can then only estimate the combined effect of all measures, not the effect of each of them. In general, heterogeneity of the treatments implemented makes both the development of crash modification functions and the transferability of study results more difficult. While it may be impossible to completely avoid this problem, researchers should take care to describe in detail the measures that are evaluated.

Crash migration denotes a tendency for crashes to "migrate" to neighbouring sites close to treated sites, i.e. the number of crashes increases at non-treated sites as a result of treatment introduced at sites nearby. This topic attracted a lot of interest during the period from about 1985 to about 1995. Several papers have been published, including a recent report from Austroads (2010). The report provides an overview of the effect known as crash risk migration (CRM) and is intended to build an understanding of the potential for CRM to occur as a result of road safety treatments. It is also intended to inform current approaches to evaluating road safety risk. In particular, evaluation approaches generally examine the extent to which a safety issue at a treatment site can be addressed, without considering the possibility that some treatments may impact crash rates at other locations.

The report focuses on situations where CRM may occur as a result of traffic redistribution. Some examples of treatment types that Austroads thought would be most susceptible to migration include the following:

- turn controls or bans
- major changes to a route such as parking changes

- bridge closure
- localised speed limit changes
- intersection changes e.g. signalisation, turn phase timing change, turning lanes
- traffic calming
- lane additions
- addition of overtaking lanes
- pedestrian treatments at intersections and at mid-block locations
- railway crossing control
- mid block turning provision

Even with thorough reports such as this, the discussion about how common this phenomenon is has not been completely resolved. It is clear that regression-to-the-mean is a potentially very important confounding factor in many before-and-after studies. It is not equally clear that crash migration is always a threat to study validity. There are many unresolved issues with respect to crash migration. When is it likely to occur and what causes it? How far from the treated sites are crashes likely to migrate? Will the effect arise only very close to treated sites or will it extend to sites far away?

Is the effect likely to be a novelty effect only, i.e. occurring only in a short period after treatment, or will it be a lasting effect? Is crash migration likely to lead to an increase in the number of crashes that fully offsets the effect of treatment, or will it only lead to a small, barely detectable increase in the number of crashes?

It is unfortunate that interest in studying crash migration abated before these questions had been thoroughly investigated. Current knowledge is insufficient to conclude that a study which did not control for crash migration should be rejected on methodological grounds. Before-and-after studies should, at a minimum, control for regression-to-the-mean, long-term trends and exogenous changes in traffic volume. If there is a reason to believe that co-incident events have led to a break in long-term trends, studies should control for the effects of such events. If there is reason to believe that crash migration will occur, studies should try to control for it. The reasons for believing that crash migration is likely to occur must be stated explicitly and should identify an observable behavioural mechanism that could generate crash migration. This requirement is essential for making a hypothesis about crash migration testable at all. Unless a mechanism leading to it can be identified, it is easy to invoke any increase in the number of crashes at non-treated sites as evidence supporting crash migration, when in fact such increases could have entirely different reasons, such as increase in population, changes in land-use, or a reduction in other safety measures, such as reduced police enforcement.

Multivariate crash models

The development and statistical estimation of multivariate crash models has been a very active and productive field of research in recent years. Such models are increasingly applied to estimate the effects of road safety measures. Using multivariate crash models to estimate the effects of road safety treatments is, however, not without problems (Hauer 2010, Elvik 2011B).

In principle, multivariate statistical analysis is a very attractive technique for estimating the effect of road safety measures, since it controls for all variables that are included in the model estimation.

In practice, multivariate modelling is associated with a number of confounding factors of its own, the most important of which are:

1. Small sample and/or low mean value bias
2. Bias due to aggregation, averaging or incompleteness in data
3. Presence of outlying data points
4. Inappropriate choice of dependent variable
5. Endogeneity of safety treatment
6. Wrong functional form for effects of independent variables
7. Co-linearity among explanatory variables
8. Omitted variable bias
9. Misspecification of the structure of systematic variation in crashes and residual terms
10. Mixing levels of crash severity
11. Inappropriate model form

State-of-the-art crash modelling is characterised by the following approach to dealing with these confounding factors:

1. The development of a model is based on a data set that predominantly contains systematic variation in the number of crashes. Models should not be based on small samples with a low mean number of crashes (Lord 2006, Lord and Miranda-Moreno 2008).
2. Data are recorded at the lowest available level of aggregation and homogeneous road sections are formed on the basis of key explanatory variables to ensure maximum between-section variation and minimum within-section variation (Cafiso et al. 2010).
3. If variables representing safety treatments are included, analysis should be designed to control for a potential endogeneity bias attributable to such variables. Endogeneity refers to a statistical tendency according to which abnormal values in the dependent variable, i.e. crashes, influences the use of safety measures. The problem is analogous to regression-to-the-mean bias in before-and-after studies, but the bias can often go in the other direction, suggesting that a road safety measure is ineffective or has adverse impacts when it is in fact effective. For an instructive example, see Kim and Washington (2006).
4. The functional form used to describe the relationship between an explanatory variable and the dependent variable is explicitly chosen based on an exploratory analysis. Guidelines for choosing the functional form are given by Hauer and Bamfo (1997).
5. Potential bias due to co-linearity among explanatory variables is addressed.
6. Potential bias due to omitted variables is addressed.
7. Potential bias due to outlying data points is addressed.
8. The structure of systematic variation in the number of crashes and in residual terms is specified as accurately as possible. Residual terms are described statistically in a way that permits the use of model output in the empirical Bayes approach to road safety estimation.
9. Crashes at different levels of severity are modelled separately. If possible, different types of crashes should also be modelled separately.

10. The choice of model form is made explicitly. A dual-state model should only be chosen if prior knowledge suggests that it is superior to a single-state model, given the purpose of developing the crash prediction model. A dual-state model is a model that postulates that there are two modalities for the crash generating process, differing with respect to the expected number of crashes per unit of time.
11. The dependent variable should preferably be the number of crashes at a given level of severity.

Not all multivariate crash models that are found in the literature satisfy all these requirements. In particular, the following confounding factors are likely to be present quite often:

1. Bias due to aggregation, averaging or incompleteness of data. In particular AADT as a measure of traffic volume may be biased both because it is an average, it is an aggregate (of the various types of vehicles that make up traffic) and it is very often incomplete (pedestrians and cyclists are rarely included). Crash reporting is always incomplete; however this is not a problem that can be solved by statistical estimation only.
2. Wrong functional form for effects of independent variables. Most models tend to rely on the assumption that all relationships are monotonic. Functional forms ought to be tested in an exploratory analysis.
3. Omitted variable bias. Pedestrian and cyclist volumes are very often omitted. Variables describing road user behaviour are also very rarely included in crash models.
4. Mixing levels of crash severity. Mixing levels of crash severity can produce results that are almost impossible to interpret. If separate models cannot be fitted for crashes at each level of severity, then at least crash severity ought to be included as a variable in the model.
5. Inappropriate model form. A dual state model should not be used merely because it happens to fit the data better than a single state model. A reason should always be given for choosing a dual state model. Models implying a zero-state, i.e. a state in which the expected number of crashes is zero or very close to it, have no substantive meaning and should never be used.

4.4. Conclusions

This chapter identified how variability in CMF research results is a major deterrent to international transferability. In particular, when multiple research reports indicate small variance in a CMF, in similar circumstances, the issue of transferability should not arise and the research results are expected to be broadly applicable in many if not all countries. On the other hand, when the variance reported or found in the research supporting a CMF is large, transferability concerns arise that need close attention. The chapter discusses how to either consider the degree to which variance is an issue in particular studies, or how variance can be addressed to increase transferability.

Generally speaking, reducing variability through proper study design and reporting enhances transferability. Variance can be reduced by carrying out more studies or by making the CMF a function of circumstances. The chapter provides detailed guidance on how to assess variance.

It further describes how researchers can design studies to reduce variance from this perspective and how practitioners can better understand CMFs that they would like to apply. Of specific interest to readers of this report, the chapter describes how studies made in different countries need to employ identical or at least similar methods so that they are more comparable.

The chapter provides an approach that can serve as a checklist when performing a systematic review of road safety studies. It can also serve as a guide or framework for the development and conduct of research so that international transferability can more readily be accomplished. This chapter can therefore serve as a useful and practical guide for the road safety research community and the safety practitioners as well.

KEY MESSAGES

- Variation in CMF study findings is a central concept for considering the international transferability of CMFs.
- When multiple research reports indicate small variance in a CMF, in similar circumstances, the issue of transferability should not arise.
- Concern about transferability arises whenever the variance is large.
- Variance can be reduced by carrying out more studies or by making the CMF a function of circumstances.
- Studies made in different countries need to employ identical or at least similar methods so that they are more comparable.
- This report provides detailed guidance on how to assess variance.
- It is recommended that this framework be used as a checklist when performing a systematic review of road safety studies.

NOTES

1. Much of the text is based on: E. Hauer, Crash modification functions in road safety. Proceedings of the 28th Annual Conference of the Canadian Society for Civil Engineering, London, Ontario, 2000.

2. Or by θ when the context is clear.

3. The following sections draw heavily on the ‘Second Report’ prepared for the Highways Safety Research Center at the University of North Carolina within “NCHRP Project 17-48 – Development of a Strategic National Highway Infrastructure Safety Research Agenda.”

4. In fact many such research studies exist. I have invented two such studies to keep the example simple. In practice, more studies are needed.

5. Illumination designs differ by pole placement and spacing, how they are protected, and illumination intensity, etc.

6. If it still arises, the motivation behind it is of the 'not invented here' kind and has no place in cost versus effect discourse.

7. Some might be tempted to ask whether the null hypothesis of equality can be rejected. While it is true that the statistical hypothesis that the two CMFs are the same cannot be rejected, the alternative hypothesis that the two differ by 0.75-0.62=0.13 is much more likely and even more difficult to reject. Although one cannot rule out the possibility that the θs in Arizona and British Columbia were (nearly) the same, if they were, this would be an unlikely special case.

8. When the θ of a certain future implementation can be assumed to be a realization from the probability distribution of θ from past research is not easy to say. In terms of the illustration used here the question might be whether the differences in circumstance between the illumination projects in Arizona and British Columbia are of the same kind as those between Estonia and the other two.

9. In this case further research may be needed. The following section discusses the kind of research that reduces the standard error of $\bar{\theta}$, and that reduces $\sigma\{\theta\}$.

10. See Hauer (1997), Observational before-after studies in road safety. Pergamon. P. 193.

11. The law of total variance (or, equivalently, the variance decomposition formula) follows logically from the axioms of probability. For more details see: http://en.wikipedia.org/wiki/Law_of_total_variance.

12. Using the sample variance, not the bias corrected sample variance (dividing by n, not by n-1).

13. To break even, the value of the crash reduction predicted by equation 3 must be equal to the opportunity cost of doing 'a' instead of 'b'.

14. It will also give us a better idea about V, but may or may not reduce it.

15. See M. Koornstra, F. Bijleveld, M. Hagenzieker, (1997) The Safety Effects of Daytime Running Lights. SWOV Institute for Road Safety Research, R-97-36. Leidschendam, The Netherlands.

16. Since the standard errors in Table 3 are not the same, a weighted regression would be in order. However, in the present context such pedantry is unnecessary. Thus we are assuming that $E\{\theta_i\} = \alpha + \beta x_i$t where α and β are unknown constants and x_i is the value of a variable (here Latitude) for study i. The estimate of $E\{\theta_i\}$ is $\bar{\theta}_i = \hat{\alpha} + \hat{\beta} x_i$ where $\hat{\alpha}$ and $\hat{\beta}$ are estimates of α and β. With the data in Table 3, $\bar{\theta}_i = 1.18 - 0.0120 x_i$.

REFERENCES

Austroads, *Road Safety Engineering Risk Assessment, Part 2*: *Crash Risk Migration*, Austroads Technical Report AP-T147/10, Sydney, Australia, 2010.

Berg, G.; Danielsson, S.; Junghard, O. 1984. *Trafiksäkerhet och periodisk fordonskontroll*. VTI-rapport 281. Linköping, Väg- och TrafikInstitutet (VTI).

Brüde, U., Larsson, J. 1999. *Trafiksäkerhet i cirkulationsplatser avseende motorfordon*. VTI meddelande 865. Linköping, Väg- och Transportforskningsinstitutet (VTI).

Buxbaum, R. C.; Colton, T. 1966. *Relationship of Motor Vehicle Inspection to Accident Mortality.* Journal of the American Medical Association, 197, 31-36.

Cafiso, S., Di Graziano. A., Di Silvestro, G., La Cava, G., Persaud, B. 2010. *Development of comprehensive accident models for two-lane rural highways using exposure. geometry. consistency and context variables*. Accident Analysis and Prevention. 42, 1072-1079.

Christensen, P., Elvik, R. 2007. *Effects on accidents of periodic motor vehicle inspection in Norway.* Accident Analysis and Prevention, 39, 47-52.

Colton, T.; Buxbaum, R. C. 1977. *Motor vehicle inspection and accident mortality*. In Fairley, W. B.; Mosteller, F. (Eds) Statistics and Public Policy, 131-142. Reading, Mass, Addison-Wesley.

Crain, W. M. 1980. *Vehicle Safety Inspection Systems. How Effective?* AEI studies 258. Washington DC, American Entreprise Institute for Public Policy Research.

Elvik, R. 1997. *Evaluations of road accident blackspot treatment: a case of the Iron Law of evaluation studies?* Accident Analysis and Prevention, 29, 191-199.

Elvik, R. 2008. *Making sense of road safety evaluation studies. Developing a quality scoring system.* Report 984. Oslo, Institute of Transport Economics.

Elvik, R. 2009. *Developing accident modification functions*. Exploratory study. Transportation Research Record, 2103, 18-24. Washington D. C. Transportation Research Board.

Elvik, R. 2011A. *Treatment of zero counts in before-and-after road safety evaluation studies: an exploratory study of continuity corrections.* TRB paper 11-0133. Available on TRB 2011 Annual Meeting paper website.

Elvik, R. 2011B. *Assessing causality in multivariate accident models. Accident Analysis and Prevention*, 43, 253-264.

Elvik, R., Høye, A., Vaa, T., Sørensen, M. 2009. *The Handbook of Road Safety Measures.* Second edition. Emerald Publishing, Bingley.

Foldvary, L. A. 1971. *A Review of Vehicle Inspection in relation to road safety*. Report NR/9. Canberra, Australian Department of Transport.

Fosser, S. 1992. *An experimental evaluation of the effects of periodic motor vehicle inspection on accident rates*. Accident Analysis and Prevention, 24, 599-612.

Fuchs, V. R.; Leveson, I. 1967. *Motor Accident Mortality and Compulsory Inspection of Vehicles.* Journal of the American Medical Association, 201, 657-661.

Hauer, E. 1996. *Identification of sites with promise*. Transportation Research Record, 1542, 54-60. Washington D. C. Transportation Research Board.

Hauer, E. 2010. *Cause, effect and regression in road safety: A case study*. Accident Analysis and Prevention, 42, 1128-1135.

Hauer, E., Bamfo. J. 1997. *Two tools for finding what function links the dependent variable to the explanatory variables.* Paper presented at ICTCT workshop in Lund. Sweden, November 5-7, 1997.

Highway Safety Manual. First edition. American Association of State Highway and Transportation Officials (AASSHTO). Washington D. C., 2010.

Institution of Highways and Transportation. 1990*. Guidelines for accident reduction and prevention*. International edition. London.

Jørgensen, E., Jørgensen, N. O. 2002. *Trafiksikkerhed i rundkørsler i Danmark*. Rapport 235 2002. København, Vejdirektoratet.

Kim, D-G., Washington. S. 2006. *The significance of endogeneity problems in crash models: an examination of left-turn lanes in intersection crash models*. Accident Analysis and Prevention, 38, 1094-1100.

Little, J. W. 1971. *Uncertainties in evaluating periodic motor vehicle inspection by death rates.* Accident Analysis and Prevention, 3, 301-313.

Lord, D. 2006. *Modeling motor vehicle crashes using Poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter.* Accident Analysis and Prevention, 38, 751-766.

Lord, D., Miranda-Moreno, L. F. 2008. *Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective*. Safety Science, 46, 751-770.

Mayer, A. J.; Hoult, T. F. 1963. *Motor Vehicle Inspection*. A Report on Current Information, Measurement, and Research. Wayne State University, Institute for Regional and Urban Studies, January 1963.

Schroer, B. J.; Peyton, W. F. 1979. *The effects of automobile inspections on accident rates*. Accident Analysis and Prevention, 11, 61-68.

Tran, T. 1999. *Vegtrafikkulykker i rundkjøringer – 1999*. TTS rapport 2, 1999. Oslo, Vegdirektoratet, Transport- og trafikksikkerhetsavdelingen.

VanMatre, J. G.; Overstreet, G. A. 1981. *Motor Vehicle Inspection and Accident Mortality: A Reexamination.* Journal of Risk and Insurance, 48, 423-435.

Weissbrodt, G. 1984. *Auswirkungen von Ortsumgehungen auf die Verkehrssicherheit.* Heft 48, Unfall- und Sicherheitsforschung Strassenverkehr. Bergisch-Gladbach, Bundesanstalt für Strassenwesen.

White, W. T. 1986. *Does periodic vehicle inspection prevent accidents?* Accident Analysis and Prevention, 18, 51-62.

ISBN 978-92-821-0375-3
Sharing Road Safety
© OECD/ITF 2012

CHAPTER 5. OVERCOMING BARRIERS TO IMPLEMENTATION

This chapter discusses some of the challenges of safety decision making and provides an illustration on how Crash Modification Functions can support decision making to overcome some of these barriers. One of the hindrances to the widespread use and transfer of CMS is the lack of supporting documentation related to the countermeasure, the development process, and conditions under which the countermeasure was tested. The chapter provides a list of essential reporting elements for inclusion in any study presenting safety evaluation results. Finally, this chapter also discusses the underlying conditions required for sharing knowledge of effective safety policies with developing countries.

Road program and project decision making is a complex and challenging process. It usually involves such things as political considerations, public needs and acceptance, cost and the need to balance competing demands for safety, operations, economic development and other factors that all come to bear in making decisions. Safety predictions generated by CMFs are therefore obviously not the only information available nor will they by themselves be the single impetus for a final decision. This complex process presents barriers to the implementation and use of CMFs in the decision making process. Generally speaking, there are two types of barriers to consider when implementing CMFs: barriers associated with the implementation of the safety measures, and barriers associated with the application of the CMF itself.

Efficiency assessment tools can help governments choose those measures that will likely maximize the social benefits of public investment. An economic evaluation will help in identifying whether a proposed change will increase economic welfare. One of the aims of the decision making process is to ensure that resources are distributed in a way which provides the maximum level of utility (economic efficiency). Cost-benefit analysis is widely recognised as a useful tool which offers the best value for money.

Crash Modification Functions (CMFs) are fundamental to identifying the most effective road safety countermeasures and for calculating safety benefits in economic analyses of safety policies when trying to make optimal use of resources for improving road safety. It follows then that as a part of all road safety policy processes CMFs should be used as a core evaluation measure.

Despite the clear value of CMFs, this does not necessarily mean that they will immediately be implemented and endorsed by decision makers. Political decisions do not wholly conform to the principles and results of efficiency assessments. In some cases, CBA or CEA has been carried out at the early stages of the process, but the final decision is not based strictly on those results. It is also important to realize that governments may have other legitimate interests different from overall efficiency.

Another factor that can affect the implementation of CMFs is the extent of behavioral adaptation. For example, the implementation of a new speed limit will depend on the level of available enforcement and the willingness of policemen to enforce the new legislation (Ross 1981). Imposing an obligation to wear safety belts can yield either a more favorable, or a less favorable outcome depending on the willingness of drivers to follow this new regulation.

In the following sections, we describe some of these challenges that need to be met and provide ideas on how better CMFs can be useful. The key point is that any improvement in our knowledge of the effectiveness of safety measures, i.e. CMFs, will have tangible effects on the way safety decisions are made. It is crucial to use this information to improve the quality and widespread use of evidence based safety decisions.

Reducing the variance of individual CMFs and enhancing their transferability are key ways to assure the widest possible use of CMFs.

5.1. Using CMFs to overcome barriers to countermeasure implementation

While CMFs by themselves cannot overcome all barriers, they can contribute to a dialogue about the implementation of countermeasures and they can become influential in the decision making process. The most common obstacle to implementation of safety countermeasures centers on political or economic constraints, or has to do with public acceptance. The following sections illustrate how CMFs can support decision making in these constraining environments.

5.1.1. Political constraints

A CMF provides evidence about the likely impact of a certain countermeasure. For any policy to be adopted, however, it needs to be both technically and politically desirable (Rose 2001, p. 15). If the program is both technically and politically desirable, the probability of its implementation will be greater. On the other hand, a technically feasible program which is politically undesirable will be easily rejected for political reasons. Hence, the adoption of a safety initiative depends on political judgment.

The decision maker also has to have enough confidence that the policy intervention will actually reduce accidents. Safety decisions may require political will and sometimes political courage, especially when the proposed safety measure is considered socially unacceptable.

Finally, the decision maker has to know about the safety impacts of different countermeasures. This means that the relevant CMFs have to be both available and accessible (Rose 2001, p. 7).

Better knowledge of the expected impacts of different types of measures is crucial to reducing the risk of making bad decisions. Therefore, after safety measures are implemented, evaluations should be undertaken as a matter of course to improve the accuracy of CMF values, and to reduce related uncertainties. This will add to the knowledge base and reduce the risk of making wrong decisions. This could be especially important when the proposed measure is socially unpopular, thus helping the decision maker to have confidence in the expected results.

The decision making process on any safety initiative may be made without full knowledge of the impacts. This may have negative outcomes in terms of the effectiveness of the safety measure. Therefore, it is crucial to impose a process of decision making that requires the replication of CMFs.

The safety decision making process needs to be transparent. A requirement for the political process to apply CMFs in the evaluation of different safety measures and to document these correctly will provide crucial information and will greatly increase the transparency of the decision making process.

5.1.2. Economic constraints

Economic constraints can be a significant barrier to implementing safety policies. Lack of resources may limit the number of CMFs that can be developed or the available policies that can be implemented.

Implementing a public policy always requires resources and an important barrier to implementation is the absolute cost of the intervention.

This barrier is often reinforced by the uncertainty of outcomes, which can indeed be a crucial factor for adopting a policy measure, especially when the cost of that measure is high. The cost of a countermeasure also needs to be compared with available resources. The national government may have more resources and possibilities than local government and cities, while a wealthy nation may face fewer constraints than a developing country.

Some policies though, can be implemented at a very low cost. The introduction of the yellow jacket and the red triangle for emergency situations in France was a relatively cheap measure for road safety. The cost of buying these items was less than 30 euros per car while the gains associated with the use of them were directly felt by the driver.

The question of sharing costs and gains is also essential to consider. Measures with widespread benefits and concentrated costs may be more likely to face opposition whereas widespread costs and concentrated gains could potentially benefit from the support of a minority group (Olson, 1987).

Similarly, cost recovery issues need to be taken into account. Often it is a governmental agency that bears the costs of adopting certain measures. The assumption is that the government's budget is sustainable and is able to pay for adopting these measures. However, a measure can be economically beneficial but financially unsustainable. A good example is the netting-off scheme adopted by the British authorities to expand the safety camera program (Carnis, 2007).

In the opposite scenario, renouncing financial benefits can be helpful for implementing a safety policy. For instance, Queensland authorities have created a dedicated fund for revenue generated by the speed camera program. Revenue is used for communication campaigns, rehabilitation programs for road accident victims and safety related road infrastructure modifications. This requirement has been essential for the popularity of the speed camera program. Speed cameras are deployed for reducing road fatalities, and not for raising revenue, but the program is possible because it is financially sustainable.

Because economic constraints can be difficult to overcome, it is essential to ensure optimal outcomes. A key policy challenge is to create decision making structures that encourage the selection of the most cost-effective safety policies, those that deliver optimal safety outcomes for the lowest cost. Governments need to get the greatest return possible, especially in times of tight budgets.

Therefore, road safety policies should undergo performance and efficiency evaluations, which cannot be undertaken without CMFs. The better our knowledge is on the effectiveness of different road safety measures, the smaller the risk of making wrong (or bad) decisions. More reliable and transferrable CMFs reduce the risk of wasting scarce resources for non-effective policies. Evaluations of the effectiveness of safety measures should be undertaken as a matter of course to improve the accuracy of CMF values and reduce related uncertainties.

There are also gains to be made in terms of saving money in producing CMFs. If results can be made more easily transferrable, this reduces the need for individual countries to carry out expensive research on the impacts of different countermeasures. The development of reliable CMFs is time consuming and costly; the development of one CMF alone can cost up to $200,000.

5.1.3. Public acceptance

Public acceptance can be a major issue in the successful implementation of safety measures. Any information that can help the community to understand that the proposed measure will lead to improved road safety can help in fostering acceptance. Communicating the benefits of safety measures through different media is therefore important.

A measure can yield an expected and direct effect, but can also generate an indirect and undesirable effect. Implementing a speed camera program can be helpful in reducing speed and the number of road accidents. However, if it leads to a reduction of police enforcement, the picture is less clear concerning the net effects (Carnis 2010). A safety measure can also be diluted by behavioral adaptation from drivers or by poor implementation. Adopting a speed camera program does not imply they will be installed at the correct place or that drivers will obey blindly the new regulation.

Evidence based on past experience needs to be communicated. CMFs are a way to measure the effectiveness of past experience. There is also a clear need to carry out follow-up studies to validate results from previous CMFs. These studies will help to calibrate results when applying safety measures in different countries or locations. More importantly, however, they will help to communicate results to the public later on, while simultaneously adding to the knowledge base.

Researchers and policy makers need to have an information or dissemination plan that provides the public with objective information on the impacts of a countermeasure. This type of communication plan is important also for the public to understand what the actual safety problem is, and what the benefits (and related costs) of countermeasures are. There is a need to "celebrate successes and remember failures".

Finally, cost-benefit assessment provides a framework for assessing societal pros and cons of policy interventions. CMFs, as a part of economic assessment, can help in presenting a fair and balanced view of outcomes that can be used to balance safety issues with other agendas, such as CO2 emissions or labour issues. Informed decision making helps in increasing public acceptance.

5.2. Overcoming barriers to the implementation of CMFs

As the adoption of a new safety countermeasure is complicated for a number of reasons, the decision to use CMFs potentially challenges professionals and organizations that are not currently using them. From the fundamentals of simply knowing what a CMF is to the more difficult challenge of understanding its relevance in a particular context, the challenges described in the following sections illustrate the need for more professional development in the understanding and use of CMFs for more rapid advancement of this science.

5.2.1. Contextual challenges

The adoption of a 55 mph speed limit in the United States was an adequate measure for many states and large urban areas. However, it may be less adequate for some rural states with low population density and little law enforcement. Similarly, is it meaningful to generalize day-light running regulations for all EU countries? It would be adequate for Nordic countries with a low level of luminosity and long winters but could be less useful for southern European countries.

Different levels of government can also lead to conflicting objectives. It may not be possible to implement a safety policy because some aspects are not legally recognized at a higher political and governmental level. For example, a city may not be in a position to implement a specific radar device because federal law forbids the use of such a device.

Another example is the impossibility of prosecuting foreign drivers for speeding when enforcement is carried out by automatic devices. For example, the French automated speed enforcement system cannot catch a Spanish driver.

Finally, some ideological constraints can be obstacles to transferring and implementing an appropriate policy measure (and use of CMF). This could be interpreted as a failure in translating general knowledge to the particular situation. Some measures might be either adopted or rejected because they were implemented in countries with which there are strong political agreements or disagreements. For instance, Anglo-Saxon countries are more inclined to exchange and to share similar solutions because of cultural familiarity (Rose 2001, p. 17).

Safety decisions are made within the constraints of institutional barriers. The decision to transfer an intervention from other countries should be based on sound knowledge, and has to take into account contextual challenges. Importing a CMF assumes that the new social and legal context will be similar enough to that of the original country to ensure successful implementation of the intervention.

In order to overcome these types of challenges, it is crucial that research on the effectiveness of the countermeasure (and related CMFs) document in detail the context in which it was developed. When using CMFs developed elsewhere, users can then take into account the context in which they were originally developed. As part of the adaptation process, it is also essential that the rationale for the modified CMF be documented and applied consistently. Chapter 5.3. recommends items essential for inclusion in any study presenting safety evaluation results.

Qualitative knowledge and investigations emphasising cultural and other institutional differences can be useful in transferring safety measures (and CMFs) from one context to another. For instance, roundabouts could be appropriate for reducing crashes at intersections in urban areas. However, they might not be the best solution for intersection crashes in rural areas. A potentially better solution might be a modified intersection that uses some aspects of roundabout technology. This kind of solution involves the adaptation of general knowledge to a local problem. A wide knowledge base that includes both technical and non-technical methods to address traffic safety issues would be useful for designing appropriate public policies. However, it is important to highlight that assumptions still need to be documented and applied consistently without changing them constantly according to subjective needs.

5.2.2. Knowledge

Transfer and implementation processes assume that the decision maker clearly understands the safety problem under consideration. Are the traffic fatalities explained by excessive speed? Or is drunk driving more of a problem? The decision maker needs to have a certain amount of information available to conceptualize the problem and potential solutions. Therefore, it is important to ensure that an adequate crash database is available in the first place to accurately identify the true nature of the safety problem.

Capacity building is an important element in the transfer of knowledge from one context to another. If users don’t understand CMFs or how they are used, this will prevent the implementation of effective safety programs and measures. Traffic safety knowledge includes information about treatments and their effects, countermeasure development and related CMFs, but also includes statistical details such as sample size, etc. Through capacity building and education, it is possible to move from subjective towards objective knowledge.

Countries should make efforts to ensure that practitioners become familiar with CMFs and their use. This will require knowledge transfer and capacity building through international efforts, possibly in the form of workshops, the preparation of guides and best practice manuals, and twinning projects.

There should also be minimum methodological standards for the development of CMFs. Previous chapters have provided guidance on how CMFs should be constructed and evaluated. As noted earlier, after safety measures are implemented, evaluations should be undertaken as a matter of course. The results of these evaluations, along with an indicator of how reliable the results are, should be stored in a transnational database.

In order to ensure the quality and homogeneity of safety assessments carried out at various levels, including national and regional, some possible actions are:

- The development of good practices guidelines and recommendations. The new US Highway Safety Manual, the European thematic network ROSEBUD and the European Project SUPREME 1 are prominent examples of such initiatives.
- The creation and maintenance of transnational databases, with sound estimates of the effectiveness of safety measures. These comprehensive databases would ensure that practitioners do not overlook any significant measure when screening for available options.
- The establishment of a system for quality control, based on independent and impartial reviews of efficiency assessments.

Finally, the reinforcement of international co-operation and communication between policy-makers and the scientific community about effective safety measures is an important step to overcoming knowledge related barriers for implementing safety policies. The transfer of CMFs and knowledge about current practices will enable authorities to make decisions even if resources are unavailable for them to make their own estimates.

5.3. Essential reporting elements for safety studies

The question of transferability concerns the circumstances under which different safety measures are implemented. Ideally, two identical measures, implemented in identical circumstances, should have the same impact on the frequencies of accidents and casualties. Conversely, discrepancies in circumstances are expected to induce discrepancies in effectiveness.

It is essential that researchers disseminating the results of an effectiveness assessment provide a description of the circumstances as accurately and completely as possible. This will allow researchers and practitioners from other regions and countries to evaluate the possibility of successfully transferring the measure. Whenever possible, information about circumstances should also be quantitative; only then can crash modification functions be developed.

There is no unique set of circumstances that is relevant to every research question. For example, while the initial speed of traffic flow can (and most probably will) influence the effectiveness of speed cameras, it is hardly relevant to the assessment of the impact of technical inspections.

Along with information on circumstances, every study should provide the standard error of the estimate of effectiveness, as well as some basic information about methods: study design, samples, data sources, and biases.

International transferability of the results of road safety evaluation studies is only possible when studies are available from many countries, over a long period, and when all these studies are of at least adequate and similar methodological quality. Many different designs are used in road safety evaluation studies – it may therefore be too restrictive to require that all studies be identical down to the finest detail. It is, however, reasonable to require that studies have uniformly controlled for at least the most important potentially confounding factors.

As documented previously in this report, one of the hindrances to the transfer of CMFs is the lack of supporting information related to the countermeasure, the development process, and conditions under which the countermeasure was tested. The group therefore recommends that all studies related to CMFs should provide as much documentation as possible in order to facilitate the exchange, transfer and application of the CMF to the fullest extent possible. Box 5.1. presents those items considered essential for inclusion in any study presenting safety evaluation results.

While the items presented in Box 5.1 are essential, they are not all inclusive. Box 5.2. provides a full list of information that would be desirable to have documented in CMF reports.

Box 5.1. Essential reporting elements

Countermeasure description
1. Detailed description
2. Baseline & future
3. Range of application (e.g. specific and general determinants (use of speed cameras))

Safety
1. Crash target group; type and severity (if infrastructure-related countermeasure)
2. Target risk factors (e.g. speeding, red light running, drunk driving, dangerous curves etc.)

Environment
1. Speed environment (speed limit)
2. Urban /rural (outer urban, semi-rural, suburban)
3. Geometric elements (e.g. divided/undivided, alignment, shoulders - hard/sealed)
4. Volume by key road user type

Box 5.2. Desirable reporting elements

Measure

Provide a description of the road safety measure with as much detail as possible, including information about the population of treated entities:

1. Nature and type of treated entities:
    - Road user: drivers, occupants, pedestrians, etc.
    - Vehicles: motorcycles, passenger cars, trucks, etc.
    - Road entities: segments, intersections, etc.
2. Traits of treated entities:
    - Road user: age, gender, etc.
    - Vehicles: age, mass, load capacity, etc.
    - Road entities: type of road, layout, accesses, configuration of intersection, speed limit, etc.
    - Criteria for selection of treated entities (e.g. frequency of accidents/casualties?)
    - Environmental conditions

Outcomes

Define outcome variables: accident, fatality, seriously injured. Report frequencies before and after.

Exposure

Provide information about exposure (average daily traffic or kilometres driven) in the population of treated entities, both before and after the implementation of the measure.

Other information about traffic conditions, if relevant

Enforcement level, average speeds, etc.

Control/comparison group

Clearly describe control/comparison group, if applicable.

Methods

Clearly describe methods, including:

- Samples
- Periods
- Study design (experimental, before-after, cross-section, case-control, multivariate models, time series)
- Treatment of potential biases (diffusion of treatment to control/comparison group, regression to the mean, long-term trends, changes in traffic volume, co-incident events, accident migration, etc.)
- Definition of independent variables
- Sources of information, and their accuracy (e.g. underreporting, errors of measurement)

Results

Estimates and standard errors

Other results: analysis of subgroups, sensitivity analysis

5.4. Sharing CMFs with Lesser Developed Countries

A CMF, although predictable in one or more developed countries, may or may not be fully realized in lesser developed countries. There are many contextual elements that can affect the eventual actual crash reduction that could occur if conditions are not the same as in the country or countries where the

CMF was developed. This is perhaps a more critical consideration when CMFs are applied in a developing country. For example, a paved or hard shoulder in most developed countries can predictably reduce crashes by a certain amount. However, a paved or hard shoulder in a lesser developed country might encourage improper use of the road environment – e.g. the erection of stalls for selling items to travelers – that could decrease the overall safety of the roadway environment.

Conversely, it is also possible that a CMF that has proven reliable in developed countries could have higher benefits when applied in a developing country. Returning to the paved or hard shoulder example, if there is highly mixed motorized and non-motorized traffic on a roadway, the paved or hard shoulder might default to a lane that accommodates the non-motorized traffic. Such a separation could potentially have greater safety benefits than if the shoulder is only used for traditional or primary purposes.

The simple examples provided above highlight the need to identify as much as possible any unintended consequences that could occur when a countermeasure is applied. Therefore, as one considers the application of a CMF developed in a more industrialized country to a less developed country, certain underlying conditions should be considered when evaluating the success of that application in the new environment. If the underlying conditions will prevent the full benefits of a countermeasure to be realized as predicted by a CMF, then consideration should be given to how the underlying condition can be managed or mitigated. If it cannot be managed or mitigated, the proposed countermeasure may not be the best choice for the situation.

The underlying conditions described below are typical conditions found in the developed world that can have an effect on the results of the implementation of a CMF if these conditions are not found in a lesser developed country. Efforts should be made to understand these factors in any country where a CMF is to be applied with the intent of considering if the environment or context is substantially different from that in which the CMF was developed. If it is substantially different, then one must make an effort to estimate the impact on the safety outcome of the application of the CMF in the new context.

i. An effective government

A government that produces effective legislation and provides appropriate regulation can have an influence on the successful implementation of a CMF. Without an effective government in place, not only will the decisions about which countermeasures should be applied be more complicated, but underlying legislation that establishes appropriate behaviors on the roadway may not exist. Likewise, regulatory enforcement may not be possible, thus creating a situation in which safety benefits from certain countermeasures may not be realized as other safety problems caused by lack of regulatory control are overwhelming. The lack of an effective government can also lead to the lack of sufficient financial resources that may limit the achievement of results predicted by a CMF. For example, the quality of the implementation of a countermeasure, if not constructed or installed properly, could prevent the full realization of safety results.

Finally, having accurate and reliable crash and traffic information - including communication and information systems - at the right place and time facilitates the application of CMFs. These will likely be non-existent without an effective government in place.

ii. Governance

The existence of both laws and the enforcement of those laws is perhaps the most significant issue for obtaining predicted safety results. In the shoulder example above, there should and likely will be a law that does not allow stalls or vendors on shoulders.

If, however, that law is not enforced, then the lack of good governance will prevent the full achievement of safety results. Similarly, the lack of the establishment of truck size and weight in many countries does create a scenario in which the outcomes predicted by CMFs may not be realized.

iii. A homogeneous, motorized population

For the most part, the vehicle fleets in developed countries are fully motorized and reasonably homogeneous in their design, mix and operational characteristics. While this may be true in some developing countries, it is also likely that there is a far wider spread of vehicle types, both motorized and unmotorized, in many developing countries. This creates an entirely different environment from that in which the CMF was developed. As a result, the impact of a particular countermeasure as predicted by a CMF may not be realized.

KEY MESSAGES

- As a part of all road safety policy processes, CMFs should be used as a core evaluation measure and they should be documented, thus ensuring transparency.
- To ensure optimal outcomes, road safety policies should undergo performance and efficiency evaluations, which cannot be undertaken without CMFs.
- After safety measures are implemented, evaluations should be undertaken as a matter of course to improve the accuracy of CMF values and reduce related uncertainties.
- The results of these evaluations, along with an indicator of how reliable the results are, should be stored in a transnational database.
- As part of the evaluation process it is necessary to document key evaluation elements such as context (see Chapter 5.3.).
- There should be minimum methodological standards for the development of CMFs.
- Researchers and policy makers need to have an information or dissemination plan that provides the public with objective information on the impacts of a countermeasure. There is a need to “celebrate successes and remember failures”.
- Countries should make efforts to ensure that practitioners become familiar with CMFs and their use. This will require knowledge transfer / capacity building through international efforts (e.g. workshops, guides, best practice manuals, and twinning projects).
- When using CMFs developed elsewhere, users need to take the context into account.
- When transferring CMFs, as part of the adaptation process it is essential that the rationale for the modified CMF be documented and applied consistently.
- It is important to reinforce international co-operation and communication between policy-makers and the scientific community about effective safety measures.

NOTE

1. Highway Safety Manual (AASHTO, 2010; http://www.highwaysafetymanual.org).

   ROSEBUD (Road Safety and Environmental Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision-Making), financiada por la Comisión Europea (http://partnet.vtt.fi/rosebud/).

   SUPREME project: Summary and publication of best practices in road safety in the EU member states. http://ec.europa.eu/transport/road_safety/projects/doc/supreme.pdf.

REFERENCES

Carnis L. (2010), *A Neo-institutional Economic Approach to Automated Speed Enforcement System*, European Transport Research Review, (2)1: 1-12. http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s12544-010-0023-4.

Carnis L. (2009), *Automated Speed Enforcement: What The French Experience Can Teach Us*, Proceedings of the 2009 Australasian College Road Safety Conference: Road Safety 2020: Smart Solutions, Sustainability, Vision, 5-6 November 2009 Peer-Reviewed Paper, Perth, Western Australia, pp. 1-11.

Carnis L. (2007), *L'automatisation des contrôles en Grande-Bretagne : entre révolution technique et continuité administrative* Revue Internationale de Sciences Administratives, Décembre, (73)4 : 597-610 (English version avalable).

Carnis L. (2006), “Montana's Basic Rule: The End of a Myth”, Paper presented for the Annual Conference of the European Public Choice, Turku.

Elvik R. (2006), *Laws of Accident Causation*, 18th ICTCT Workshop, Road Safety Issues.

Elvik R. (2004), *To What Extent Can Theory Account for the Finding of Road Safety Evaluation Studie*, 15th ICTCT Workshop, Theoretical Issues and Strategies.

Guba E. G. and Yvonna S. L. (1994), *Competing Paradigms in Qualitative Researc*, in N. Denzin and Y. Lincoln (eds.), Handbook of Qualitative Research, 1st edition, pp. 105-117, Thousands Oaks, California, Sage.

Jones B. D. (2001), *Politics and the Architecture of Choice*, Bounded Rationality and Governance, University of Chicago Press.

Olson M. (1987), *Logique de l'action collective*, Presses Universitaires de France : Paris, 2nde édition.

Pressman J. L. and Aaron W. (1984 [1973]), Implementation, *The Oakland Project*, University of California Press: Berkeley.

Rose R. (2001), *Ten Steps in Learning Lessons from Abroad*, Future Discussion Governance Papers 1, Future Governance: *Lessons from Comparative Public Policy*, Economic and Social Research Council.

Ross L. R. (1981*), Deterring the Drinking Driver*, An Insurance Institute for Highway Safety Book, Lexington Books, Lexington, Massachusetts, Health.

Sabatier P. A. (1991), *Towards Better Theories of the Policy Process*, Political Science and Politics, Vol. 24, n°2, June, pp. 147-156.

Stone D. (2001), *Learning Lessons, Policy Transfer and the International Diffusion of Policy Idea*, CSGR Working Paper n°69/01, Center for the Study of Globalisation and Regionalisation, University of Warwick.

Stone D. (1999), *Learning Lessons and Transferring Policy across Time, Space and Disciplines*, Politics, (19)1: 51-59.

Yin R. K. (2009), Case Study Research, *Design and Methods,* 4th Edition, Applied Social Research Methods Series, Vol. 5, Sage.

ISBN 978-92-821-0375-3
Sharing Road Safety
© OECD/ITF 2012

CHAPTER 6. CONCLUSIONS

This report has highlighted the complex nature of decision making for sound investments in road safety. Among other items, crash modification factors or functions (CMFs) that relate safety effectiveness to interventions and are transferable from one situation to another are a valuable tool in spreading effective safety policies. CMFs are fundamental to identifying the most effective road safety countermeasures and for conducting economic analysis of safety policies. Demand for CMFs is growing in many jurisdictions as policy makers increase their requirements to demonstrate results and undertake cost-benefit and efficiency assessments and as managers seek to ensure they are making the best possible decisions for safety in their projects.

The report has identified the lack of substantiated knowledge of the effects of countermeasures as being a key barrier to the advancement of many critical, live-saving, initiatives. By contributing to our knowledge of safety effects, CMFs, when used in balance with practical local knowledge, can have a tangible effect on the way safety decisions are made. The report showed that we are at a turning point in this regard and that there is a very real prospect of rapid advances and major cost savings through the transfer of results internationally. In this reference, transferability of CMFs relies first and foremost on analysing the extent to which a CMF is dependent on the circumstances in which it was developed. The report has documented ways to address this issue in research reports.

While the potential is high for safety gains through transferability of CMFs, the report also cautions that variability in CMF research results is a major deterrent to transferability. Reducing variability through proper study design and reporting enhances transferability. The report provides a framework that illustrates how studies can control for the most important confounding factors related to the countermeasure analysed and thus provides guidance for uniform screening and control procedures. In this regard, the report will serve as a useful guide for transferring road safety measures and in supporting countries in their efforts to collaborate on essential road safety research.

Among other identified key messages, the report described the complexity of the decision making process for safety interventions. The risk of making poor decisions or the cost of making good decisions can be reduced by the correct use of reliable studies on how effectively different safety measures perform (i.e. CMFs). In this regard, the report points out that CMFs are fundamental for calculating safety benefits in economic analyses of safety policies when trying to make optimal use of resources for improving road safety.

The report provided examples of how the demand for CMFs is growing in many jurisdictions as policy makers increase their requirements to demonstrate results and undertake cost-benefit and efficiency assessments. This increasing demand has brought to light the fact that lack of substantiated knowledge of the effects of countermeasures is a key barrier to the advancement of many critical, life-saving initiatives.

The improvement, therefore, of CMFs and the provision of more training and regular practical application of CMFs will have tangible effects on the way safety decisions are made. We are currently at a turning point, with the prospect of rapid advances and major cost savings through the transfer of results internationally.

The report provides support for the suggestion that road safety policies should undergo performance and efficiency evaluations and provides guidance on the application of CMFs and related transparency benefits.

The report also recommends that an international group be created to foster dialog among researchers and practitioners on CMF research and reporting standards with the aim of increasing transferability of results. It also promotes better coordination of research across countries on top priority countermeasures. Ultimately, the report suggests that international cooperation should aim to capture documentation and reporting of CMF research in a widely available transnational database.

In summary, the working group that prepared this report found that CMFs offer a great opportunity for the OECD and International Transport Forum countries to continue to individually advance road safety initiatives on their road systems. It is also an area that is ripe for international collaboration that has the potential to increase the efficiency and quality of safety effectiveness measurements and, ultimately, to promote international exchange. All of this translates to fewer crashes, fatalities and serious injuries on our roads.

APPENDIX A. REVIEW OF EXISTING APPROACHES AND INITIATIVES

This appendix provides a review of national and international approaches for the efficiency assessment of road safety measures. In particular, the most important international initiatives for providing standardized and accurate methods or tools for the estimation of safety effects of road safety measures are presented. Furthermore, examples from the use and application of these (or other) methods and tools at national level are analyzed.

A.1. International approaches and initiatives

A.1.1. Handbooks and manuals

The Handbook of Road Safety Measures - Elvik et al (2009)

The handbook aims to provide a systematic overview of current knowledge regarding the effects of road safety measures, by presenting state-of-the-art summaries of current knowledge regarding the effects of 128 road safety measures. The types of measures that are included are road design and road equipment, road maintenance, traffic control, vehicle design and protective devices, vehicle and garage inspection, driver training and regulation of professional drivers, public education and information, police enforcement and sanctions, post-crash care and general purpose policy instruments.

More specifically, the handbook aims to provide answers to the following questions related to the type of measures which can be used to reduce the number of traffic crashes or the severity of injury in such crashes, the crash problems and types of injury that may be affected by the different measures, the effects on crashes and injuries of the various road safety measures, as well as their related effects on mobility and the environment. Moreover, the costs of road safety measures are examined, and the potential for cost-benefit evaluations of the measures is demonstrated.

It is also demonstrated that the safety effect of a measure may vary from place to place, depending on the design of the measure, the number of crashes at the spot, any other measures that have been implemented, etc. An attempt has been made to identify sources of variation in the findings of different studies and to try to form as homogeneous groups as possible when presenting estimates of the effects of measures on road safety.

In particular, the handbook seeks to develop objective knowledge about the effects of road safety measures by relying on an extensive and systematic search of literature and by summarising this literature by means of formal techniques of meta-analysis that minimise the contribution of subjective factors that are endemic in traditional, narrative literature surveys. A systematic framework has been used to assess the validity of the studies that are quoted. Moreover, the need to develop crash modification functions in order to describe systematic variation in the effects of road safety measures is stressed.

The criteria of study quality that have been applied to assess the road safety evaluation studies referred to in this handbook are to a great extent based on the validity framework of Cook and Campbell (1979). According to this framework, the quality of a study can be assessed in terms of four types of validity:

- Statistical conclusion validity; sampling technique, sample size, reporting of statistical uncertainty in results, measurement errors, specification of crash or injury severity
- Theoretical validity; identification of relevant concepts and variables, hypotheses describing the relationships between variables, knowledge of causal mechanisms
- External validity; generalisability of the results of a study
- Internal validity; basis for inferring a causal relationship between treatment and effect, statistical association between treatment and effect, clear direction of causality, dose-response pattern, specificity of effect, control of confounding factors.

The Highway Safety Manual

The Highway Safety Manual (HSM) aims to introduce a science-based technical approach and to provide tools for conducting quantitative safety analyses. This allows safety to be quantitatively evaluated alongside other transportation performance measures such as traffic operations, environmental impacts, and construction costs. In particular, the HSM provides a method to quantify changes in crash frequency as a function of cross-sectional features. With this method, the expected change in crash frequency of different design alternatives can be compared with the operational benefits or environmental impacts of these same alternatives.

The HSM provides the following tools:

- Methods for developing and evaluating a roadway safety management program, including the identification of hazardous sites, the diagnosing of conditions at the site, the evaluation of conditions and the identification of potential treatments, the prioritization and programming of treatments, and subsequently the evaluation of the crash reduction effectiveness of programmed treatments.
- A predictive method to estimate crash frequency and severity.
- A catalogue of crash modification factors (CMFs) for a variety of geometric and operational treatment types, developed using before/after studies that account for regression to the mean.

The HSM is organized into four parts:

Part A explains the relationship of the HSM to planning, design, operations and maintenance activities. Part A also includes fundamentals of the processes and tools described in the HSM.

Part B presents suggested steps to monitor and reduce crash frequency and severity on existing roadway networks. It includes methods useful for identifying improvement sites, diagnosis, countermeasure selection, economic appraisal, project prioritization, and effectiveness evaluation. Several new network screening performance measures are introduced to shift the safety analysis focus away from traditional crash rates, in order to deal with the major limitation associated with crash rate analysis i.e. the incorrect assumption that a linear relationship exists between traffic volume and the frequency of crashes.

Part C provides a predictive method for estimating expected average crash frequency of a network, facility or individual site, and it introduces the concept of safety performance functions (SPFs). The methods are provided for road segments and intersections for different facility types (rural two-lane roads, rural multilane highways, urban and suburban arterials).

Part D provides Crash Modification Factors (CMFs), allowing one to quantify the change in expected average crash frequency as a result of geometric or operational modifications to a site that differs from set base conditions. These concern roadway segments, intersections, interchanges, special facilities and road networks. These CMFs are claimed to be readily applicable to any design or evaluation process where optional treatments are being considered.

The HSM incorporates many, but not all geometric features. Moreover, the proposed models (SPFs) cannot explain crash causes. For instance, weather and driver behavior are not explicitly addressed in these models. Finally, in several cases the proposed models are very sensitive to the availability and quality of the necessary data.

The Rosebud Handbook

Within the activities of the Rosebud thematic network, a handbook titled "Examples of assessed road safety measures - a short handbook", was issued in July 2006 as the main outcome of the Rosebud project. The handbook includes information about various assessed road safety measures. The assessment methods used are cost effectiveness analysis (CEA) or cost-benefit analysis (CBA). In CEA, the costs of a measure are confronted with its effects; the effects of the measures are not expressed in monetary terms. On the contrary, in CBA the result of the evaluation is obtained by comparing costs with benefits. Economic evaluation of road safety measures using cost-benefit analysis is based on the costs incurred as a result of road crashes. Avoiding such costs represents the economic benefit of road safety measures. The benefit-cost ratio represents the economic advantage of the safety measures.

According to the benefit-cost ratio, measures are ranked as poor, acceptable and excellent. Measures from Rosebud consist of user related, vehicle related and infrastructure related measures.

The Handbook is available at
http://partnet.vtt.fi/rosebud/products/deliverable/Handbook_July2006.pdf

The CEDR Report

The Conference of European Directors of Roads (CEDR) has been promoting collaboration and exchange of information and expertise amongst its members since 1998. In particular, it aims to provide support to the activities of the Road Directors and their national road administrations, to promote a high level of common information, and to give assistance to the European Commission in the preparation of reports concerning the development of the Trans European Road Network (TERN).

The source of CEDR measures is the Final report of "Best Practice on Cost Effective Road Safety Infrastructure Investments", April 2008
(http://www.cedr.fr/home/fileadmin/user_upload/Publications/2008/e_Road_Safety_Investments_Report.pdf).

According to the report, the five most promising investments were identified (as a result of preliminary assessment and related ranking of investments) and selected for further in-depth analysis using the existing literature in conjunction with the results of Questionnaire 2 of the CEDR task group O7 (Road safety). These investments concern the following measures:

- Roadside treatment
- Speeding
- Junctions layout

- Junction traffic control
- Traffic calming

The FHWA Clearinghouse CMFs

The CMF Clearinghouse is home to a web-based searchable database of CMFs along with supporting documentation to help transportation engineers identify the most appropriate countermeasure for their safety needs. In addition to search functions, one can also submit CMFs to be included in the clearinghouse. CMFs are rated using a five star system for quality; five stars are needed for a CMF to be included in the HSM. Stars are applied based on a review of five criteria: study design; sample size; standard error; potential bias; and data source. The clearinghouse can be found at: http://www.cmfclearinghouse.org/sqr.cfm.

The Cochrane reviews

Cochrane Reviews are systematic reviews of primary research in human health care and health policy. They investigate the effects of interventions for prevention, treatment and rehabilitation. They also assess the accuracy of a diagnostic test for a given condition in a specific patient group and setting.

The Cochrane Injuries Group has been preparing Cochrane Reviews on the effectiveness of interventions for road safety, including slowing traffic speed, wearing helmets, and driver education. The findings of these Cochrane Reviews provide guidance on the effectiveness of interventions for road safety in the hope that governments, urban planners, and individuals will be encouraged to improve road safety as a matter of urgency.

http://www.thecochranelibrary.com/details/collection/691655/Safety-on-the-road.html

Countermeasures that work: A Highway Safety Countermeasure Guide For State Highway Safety Offices - 6th Edition (2011)

This Guide is intended to be a key reference to assist State Highway Safety Offices (SHSOs) in the USA selecting effective, evidence-based traffic safety countermeasures for major road safety problem areas. The Guide describes strategies and countermeasures that are relevant to SHSOs, summarizes their use, effectiveness, costs, and implementation conditions and includes references to the most important publications (research summaries and individual studies) in the field.

The Guide includes countermeasures related to the following road safety problems and research areas:

- Alcohol-impaired driving
- Seat-belt use and child restraints
- Aggressive driving and speeding
- Distracted and fatigued driving
- Motorcycle Safety
- Young drivers
- Older drivers
- Pedestrians

- Bicycles

Each section starts with a brief literature review on the road safety problem (e.g. the reader is often referred to the Cochrane reviews), followed by a presentation of the related strategies and countermeasures. More than 115 individual countermeasures are examined and typically one page is devoted to each countermeasure. In each case, the countermeasures are ranked in terms of their effectiveness on the basis of a rating in stars; the use, costs and time needed for implementation are also assessed. Effectiveness is measured by reductions in crashes or injuries:

- 5 stars - The measures are demonstrated to be effective by several high-quality evaluations with consistent results
- 4 stars - Demonstrated to be effective in certain situations
- 3 stars - Likely to be effective based on balance of evidence from high-quality evaluations or other sources
- 2 stars - Effectiveness still undetermined; different methods of implementing this countermeasure produce different results
- 1 star - Limited or no high-quality evaluation evidence

The use of the measures is ranked high (i.e. more than two-thirds of the states, or a substantial majority of communities), medium, or low (i.e. fewer than one-third of the states or communities). The implementation costs are ranked high (i.e. requires extensive new facilities, staff, equipment, or publicity, or makes heavy demands on current resources), medium, or low (i.e. can be implemented with current staff, perhaps with training; limited costs for equipment, facilities, and publicity). Finally, the time to implementation is ranked long (i.e. more than one year), medium, or short (i.e. three months or less). A 'varying' option for the above rankings is also used in several cases.

Source: http://www.nhtsa.gov/staticfiles/nti/pdf/811444.pdf

Austroads Road Safety Engineering Toolkit (www.engtoolkit.com.au)

The Road Safety Engineering Toolkit is a reference tool for road engineering practitioners in state and local governments in Australia and New Zealand. It outlines best-practice, low cost, high return road environment measures to achieve a reduction in road trauma. The Toolkit seeks to reduce the severity and frequency of crashes involving road environment factors. The Toolkit draws together existing road safety engineering knowledge as far as possible into one Toolkit for easy access by practitioners. The presented knowledge has been updated with recent experience from local and state government agencies, and with the results of comprehensive road safety research reviews. The Toolkit is considered a ‘living’ document including updates and revisions, so that more recent safety ‘wins’ are captured and disseminated.

International Road Assessment Programme (iRAP) Road Safety Toolkit (http://toolkit.irap.org/)

The Road Safety Toolkit provides information on the causes and prevention of road crashes that cause death and injury. Building on decades of road safety research, the Toolkit helps engineers, planners and policy makers develop safety plans for car occupants, motorcyclists, pedestrians, bicyclists, heavy vehicle occupants and public transport users. It is aimed primarily at users in developing countries. It has been translated into French, Spanish and Mandarin.

A.1.2. Research projects

The PROMISING project

The PROMISING project is aimed at developing measures that reduce the risk of injury to vulnerable and young road users as much as possible in a non-restrictive way. It was commissioned by the European Union and was coordinated by the SWOV Institute for Road Safety Research.

PROMISING project measures come from the WP5 “Cost-benefit analysis of measures for vulnerable road users”, July 2001. Cost-benefit analysis was carried out for a number of measures.

The ROSEBUD thematic network

ROSEBUD (Road Safety and Environmental Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision-Making) 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.

ROSEBUD examined the factors affecting the quality of the efficiency assessment of a safety measure, which in turn depends on the quality of the available values of safety effect. The latter were found to depend on a number of factors, including the availability of values, the validity of the data, the variability of the effect, the local versus general effects, and the changeability of the effect.

The recommendations include ways to systematize the values of safety effects, mainly by documenting the effects on the basis of either a meta-analysis or traditional literature surveys, and by providing for theoretical effects based on known relationships between risk factors and crashes. They also include criteria for examining the local findings on safety effects of road infrastructure improvements.

http://ec.europa.eu/transport/road_safety/pdf/projects/rosebud.pdf

The SUPREME project

The SUPREME research project was funded by the European Commission and its goal was to collect, analyse, summarise and publish best practices in road safety in the Member States of the European Union as well as in Switzerland and Norway. The target audiences of the project are decision and policy makers at all levels, from European to local, as well as the scientific community and practitioners in the field. The aim was to provide users with specific information on outstanding safety measures with a view to implementation in other countries or at the European level.

SUPREME measures come from the final report and are mainly from Part C of "Handbook for measures at the Country level", and Part D of "Handbook for measures at the European level", both from June 2007. The evaluated safety measures described are ranked as best, good or promising practices in the following areas:

- Licensing
- Policy
- Enforcement
- Campaigns

- Infrastructure interventions
- Safety equipment
- Data analysis
- Post impact care

The measures within the SUPREME research project were collected through a questionnaire sent to experts working for international or European organisations, NGOs, interest groups and industries. The information collected through the questionnaires was supplemented by additional research from the authors.

According to the SUPREME project, "best practice" measures were scientifically proven to lead to a reduction of road crashes and/or deaths and serious injuries, had a positive cost-benefit ratio and were expected to lead to the effects' sustainability and/or to public acceptance. Measures were rated as "good" when the available information on the above criteria was not sufficient to assess if they were the best practice in their category because of a clear lack of systematic evaluations of implemented measures. Measures that have not yet been implemented at the European or international level but have proven to be successful in one or more Member States were rated as "promising".

The Road Safety Engineering Risk Assessment Project

Funded by Austroads, this program of research began in 2002 on a relatively limited scale, but formed a substantial part of the Austroads road safety research program from 2004 to 2007. The results were intended to provide road authorities with more effective methods and tools to reduce road crashes and injuries. This has included assessment of the effectiveness of commonly used road safety measures, as well as reviews of other associated issues required as part of an economic evaluation for road safety. A total of 11 reports were published based on this research (available from www.austroads.com.au), as well as 15 newsletters (the 'risk reporter' newsletter series is available from www.arrb.com.au).

A.1.3. Other

The IRTAD Annual Reports

The source of IRTAD measures is the "IRTAD -Road Safety Annual Report 2009", which includes road safety data from 27 member countries. This report summarizes the recent road safety measures (2007-2009) as well as the National Road safety targets and strategies without always indicating their effectiveness.

Technical Assistance in support of the Preparation of the European Road Safety Action Programme 2011-2020

COWI is a northern European consulting group which undertakes studies within the fields of engineering, environmental science and economics.

The source of COWI study measures is the final report of "Technical Assistance in support of the Preparation of the European Road Safety Action Programme 2011-2020", January 2010, carried out for the European Commission DG-TREN. The main sources of this report are ERSO, ETSC, EuroRAP, The Global Status Report on Road Safety -Time for Action of WHO, IRTAD and national sources.

According to the report, the enforcement of rules on speed, drunk driving, and seat belt wearing in EU25 is provided as a quantitative ranking (scale 0-10) and a qualitative ranking (good/ improving/ need to do more). The effectiveness of helmet wearing enforcement and child restraints is also provided. In addition, infrastructure interventions concerning engineering actions such as formal audits on new roads, regular inspections on existing roads, and EuroRAP assessment (risk mapping or star rating) are also included.

Roads are assessed according to: the separation of directions (how well the medians are treated), the design standard and frequency of intersections, how well the roadsides are protected, how the edge of the carriageway is treated, and the occurrence of fatalities for pedestrians and cyclists.

Concerning the education and campaign measures of the safety measures, the COWI study presents the most common campaigns on speed, seatbelts, alcohol, helmets, young drivers and school children education.

As for trauma management, the COWI study presents the performance for several countries according to the Safetynet study ranking (high level, medium level, low level, relatively low level).

I-cars network, thematic group on impact assessment measures

http://www.icarsnetwork.eu/en/thematic_groups/tg3_-_impact_assessment_methods/ This project provides an overview of impact assessment studies (all types of studies, from expert judgement to crash studies) for intelligent vehicle systems. The objective of the project is to exchange experience regarding the use of different methods of impact assessment and socio-economic evaluation with the goal of leading to more reliable methods with higher predictive validity.

A.2. Existing national approaches and initiatives

Australia

Crash Reduction Factors (CRFs) form the basis of decision making for many policy and infrastructure improvements within Australia. A solid body of evidence has been established over many years on this topic. This is particularly strong for road infrastructure improvements, with less information available for behavioural and vehicle-based improvements.

National guidance has been produced on the treatment of crash locations (Austroads 2009) and this provides information on the use of CRFs. This guidance is used across Australia and New Zealand.

Various robust evaluations have been conducted in Australia on the effect of infrastructure improvements, particularly those treatments at high crash locations (i.e. blackspot programs). These evaluations have occurred at national and state levels, and cover a range of CMFs. Numerous studies on individual treatments have also been conducted on an ad-hoc basis. The predominant methodology for Australian research has been to use before and after studies with a control group. A 3-5 year before and after period is generally selected to minimise the effect of regression to the mean.

Numerous studies have also been undertaken on safety initiatives relating to vehicle improvements and behavioural programs. Information on these aspects of safety is less coordinated than for infrastructure safety improvements.

Various summaries on this research have been published within international journals. However, much of the research is not, and resides in published or unpublished reports and conference proceedings.

Several relevant research products on this topic from Australia have been developed as part of the Road Safety Engineering Risk Assessment research program. The first is titled *Road Safety Engineering Risk Assessment, Part 6: Crash Reduction Factors* (Turner et al., 2010). The objectives of the project were to improve crash reduction estimates associated with various road design features or safety measures (termed 'issues') in different environments, and to provide more objective analysis methods for use by road authorities in assessing relative risk. The report examines 47 different treatments. A second report titled *Road Safety Engineering Risk Assessment, Part 2: Crash Risk Migration* (Styles et al., 2010) provides an overview of the effect known as crash risk migration (CRM). The report is intended to build an understanding of the potential for CRM to occur as a result of road safety treatments. It is also intended to inform current approaches to evaluating road safety risk. In particular, evaluation approaches generally examine the extent to which a safety issue at a treatment site can be addressed, without considering the possibility that some treatments may impact crash rates at other locations. The report focuses on situations where CRM may occur as a result of traffic redistribution. A further study examines the issue of treatment life, or how long engineering treatments can be expected to deliver a road safety benefit (*Road Safety Engineering Risk Assessment Part 4: Treatment Life for Road Safety Measures*, Turner & Comport, 2010).

The research program also led to the development of a rating scale that assigns a level of robustness to individual studies according to the research methodology that was used. This scale ranks research from 1 (poor quality research) to 5 (high quality). A level of confidence in the robustness for the CRFs associated with various treatments has been developed (a rating of high, medium and low). This rating is based on the number of studies, the robustness of their methods indicated by the rating scale, the consistency of the results, the recentness of the study, and the transferability of the research.

Research needs where there is either a low level of confidence or where there is no adequate research have been identified. A method to prioritise these gaps in knowledge has been developed. This is based in part on the economic cost of not knowing accurate information about the treatment.

In addition, research has been conducted on the cumulative effect of more than one engineering treatment at a site. A formula has been developed to estimate the aggregated benefit. This research found that only 1 in 5 treated sites had just a single treatment. A combination of treatments had been used at most of the sites examined.

A national guidance document is currently being prepared on evaluation methodology for the development of CRFs. This will cover different issues to consider when evaluating road safety engineering measures, as well as methodology options. This document has been produced with reference to this current OECD/ITF report to ensure that future research is compatible with international efforts on this topic.

Austria

In general, the efficiency results from the Handbook of Road Safety Measures are used.

A large examination was carried out from the KfV, concerning the efficiency analysis of remedial measures at high risk sites. The measures were classified in four different groups. A statistical assessment showed the appropriate significance of three different testmodes. Eventually, a cost-benefit analysis showed the potential of the different measures for Lower Austria.

It was found that possible savings by reconstruction measurements are 3 percent − thus, in Lower Austria €23 million of social costs could be saved.

Moreover, the KfV published a report in 2009 titled "Good and Promising Interventions for the Prevention of Injuries to Pedestrians and Two-wheelers - Inventory and Guidebook for the Health Sector". The aim of this "inventory and policy guide" was to build common knowledge on what is known to work in injury prevention when targeting vulnerable road users and falls in pedestrians (inventory), as well as to guide politicians, administrators and other stakeholders in the health care sector in the implementation of policies and practices (policy guide).

http://www.euroipn.org/apollo/reports/ANNEX%205.1%20ApolloGPGuideandInventory.pdf

Canada

Throughout Canada, several provinces use Collision Reduction Factors to evaluate the effectiveness of countermeasures. There are several resources that are used by practitioners to obtain these factors. These include:

- In Service Safety Reviews
- FHWA Reports
- Sources of general information such as ITE, TAC, TRB and AASHTO
- Specific provincial lists
- The Canadian Guide to In-Service Road Safety Review (TAC- January, 2004); Section 6.3 Countermeasures and Effectiveness (Tables 6.1 to 6.12)
- CRF's produced by the ICBC and the British Columbia Ministry of Transportation and Infrastructure
- Highway Element Investment Review (HEIR) Guidelines

Noteworthy is Ontario's Highway Element Investment Review (HEIR) Guidelines. HEIR are guidelines accompanied by an excel spreadsheet. The purpose of HEIR Guidelines is to help practitioners determine how to make the best investment in highway infrastructure. In this document there are equations whose purpose is to determine the safety benefits of highway improvements. The guidelines are accompanied by an excel spreadsheet that helps perform cost/benefit analyses of potential improvements to the highway.

CMF's are used within the HEIR Guideline calculations to help determine the overall benefit of safety measures. The change in frequency for different collision severities is calculated; this collision reduction value is then multiplied by a corresponding collision cost. The results of these calculations are used to determine total collision reduction benefits. This process is performed automatically in the HEIR accompanying excel spreadsheet.

Finland

Since 1995, practically all traffic safety effects of road improvements on public roads have been evaluated using an evaluation tool (TARVA) whose major feature is the use of CRFs. TARVA was created for the Finnish road administration by VTT.

In addition to using relevant CRFs, the emphasis is on appropriate evaluation of current safety situations and taking into account overlapping safety measures. The CRFs used are based on a Norwegian traffic safety handbook modified for local conditions. Cost/benefit ratios have been calculated for different safety measures - it was also found that the AADT on the site of improvement heavily influences it. The CRFs are also utilized while evaluating safety effects of traffic safety programs.

Literature: Peltola, H. (2007) Evaluating measures in order to achieve safety targets. Road Safety on Four Continents, Bangkok, Thailand 14-16 Nov 2007; Peltola, H. (2009) Evaluating road safety and safety effects using Empirical Bayesian method. 4th IRTAD Conference, Seoul, Korea, 16 - 17 Sept. 2009 http://internationaltransportforum.org/irtad/pdf/seoul/8-Peltola.pdf

TARVA (2010) - home pages at web: http://www.tarva.net/tarvaintro.asp

Germany

Road traffic safety work is an interdisciplinary task. In Germany the current practice in safety assessment incorporates infrastructure and vehicle technology as well as behaviour related measures. Safety assessment is done in both ex-post evaluations as well as in ex-ante evaluations.

Vehicle:

Recently, a lot of scientific work was done to evaluate the possible safety effects of vehicle systems. Research development started from simple vehicle safety systems, moved on to intelligent safety systems which work autonomously, and then included communication based cooperative systems to link vehicles with each other or together with infrastructure devices. Much of this work was done in the scope of EU-projects encompassing work on the methodological foundations of CBA, EU-level standardization of the assessment methods used, and socio-economic assessments of road safety measures and programs. The most important projects concerning vehicle safety systems are the following:

- The EU-project ROSEBUD (2003 – 2006) supported users at all levels of government (European Union, national, regional, local) with information about road safety related efficiency assessment (see section 3.2.2)
- The IMPROVER project funded by the EC (2004 – 2006) examined the impacts of light goods vehicles (LGV) on road safety, emissions and fuel consumption. The safety measures considered were speed limiters, electronic stability control, social rules (using digital tachographs), seat belt reminders or seat belt locks, professional fleet safety management, and an increase in the minimum age of drivers to 21 years.
- In the EU-project eIMPACT (2006 - 2008) the socio-economic effects of stand-alone and cooperative Intelligent Vehicle Safety Systems (IVSS), and their impact on traffic safety and efficiency were assessed. In addition to including a CBA, the project also addressed policy options to improve market deployment of the IVSS and included the views of stakeholders such as users, OEMs and insurance companies concerning the profitability of the systems.
- The iCars Network funded by the EC (2008 – 2010) (see section 3.2.3)

- The project SAFESPOT (2006-2010) aimed to develop co-operative safety systems (IVSS) with communication and information platforms. Within the subproject BLADE a comprehensive assessment and impact evaluation of IVSS was done. In this context, an integrated framework was used, which consisted of a CBA, sensitivity analysis and a complementary stakeholder analysis. The impacts considered in the assessment were safety, traffic flow and environmental effects of cooperative safety systems.
- The EU-project ASSESS (2009 – 2012) aims to develop a relevant set of test and assessment methods applicable to a wide range of integrated vehicle safety systems considering driver behavioral aspects, pre crash sensing performance and crash performance with a special focus on emergency brake assist systems.

Infrastructure:

Within the national project "Possibilities of faster realisation and prioritization of structural measures to improve road safety at black spots" a catalogue was developed which contains black spots with examples of measures for their elimination. In this catalogue, the efficiency (benefit-cost analysis) as well as the effectiveness (crash costs and impact of measures) were exemplarily described for different measures. In addition, every example was judged and examined, whether or not the effect of the measure is significant. The examples are to support crash commissions and decision makers assessing advantages and disadvantages of planned measures. An evaluation of the effectiveness and efficiency for 110 measures taken at black spots was possible.

Human behaviour:

In the past, different measures were scientifically accompanied and evaluated by ex-post analyses. Two behavioural measures were introduced recently: "Accompanied Driving from the Age of 17" and the "Ban on Alcohol for Beginners".

- Accompanied driving from the Age of 17: Germany has had no previous experience in preparing beginners for driving from an early age. Before introducing the model "Accompanied Driving from the Age of 17", there was an analysis of its effects in foreign countries and of the possibilities to adopt this model in Germany. In particular, experiences from Sweden, Canada and United States were investigated. The introduction of this model and its safety effects were evaluated in three studies that differed regarding their period and location of investigation. All three studies showed a double-digit decrease in driving offence and crash risk. Based on these results the model project will be converted into common law.

- Ban on Alcohol for Beginners: Even at low alcohol concentrations, young drivers and beginners already have a significantly higher crash risk. The introduction of a "Ban on Alcohol for Beginners" in Germany in 2007 was accompanied by a before and after comparison as well as by an analysis of international experience. The results of this study showed a significant decrease in alcohol related crashes in beginner drivers. Furthermore, there was also a decrease in the risk of committing driving offences. To strengthen these results the time period of the evaluation is to be expanded as soon as actual crash and offence data are available.

Greece

The evaluation of road safety measures in Greece is not routine within road safety decision making. A number of evaluations are available, mainly on the basis of research projects, but these are only occasionally used. The assessments mainly include infrastructure improvements, such as an evaluation of the upgrade of the interurban road network to motorways, and an evaluation of traffic calming and low-cost traffic engineering measures in urban areas (Yannis et al. 2005). An evaluation of speed and alcohol enforcement was also carried out, confirming that the intensification of enforcement was one of the main reasons for the significant improvement of road safety in Greece from 2003 onwards (Yannis et al. 2008). Some results from the evaluation of safety campaigns are also published by non-profit organizations (www.ioas.gr).

Furthermore, there are no official figures for crash costs but work is on-going to make the evaluation process more efficient. A complete assessment of the social costs of road crashes and related casualties, with particular emphasis on the estimation of human cost (i.e. the value of statistical life) is presented in Yannis et al. (2005), suggesting that the social cost of fatal crashes and their victims in Greece is considerably higher compared to the estimates of international publications (e.g. ROSEBUD, UNITE).

Ireland

In the past, safety assessment in Ireland was carried out mainly from the engineering perspective while behavioural assessments were left out. Current research is trying to correct this by also incorporating behavioural measures.

Japan

A policy referred to as “Management to produce results” is now being implemented in Japan. As part of the policy, road administrators are establishing a database with integrated data of traffic accidents, traffic and road conditions, so that a kind of CMF could be introduced, according to the process outlined below.

Collecting data

In Japan, separate administrative bodies are in charge of traffic safety. Initiatives taken to reduce traffic accidents are implemented by both road administrators and police agencies. Police agencies, which manage traffic, collect crash data. Road administrators consider countermeasures using crash data provided by police agencies in addition to traffic and road condition data. Two databases are being established. The first one is the “Integrated Accidents Database” and concerns trunk roads. The contents of this database are aggregated crash data, traffic and road conditions for 700 road sections, totalling 180 thousand kilometers. The second one is the “Database of Locations Implemented Measures”, and includes descriptions of countermeasures implemented, crash data, and basic data concerning the road traffic environment of the locations where the countermeasures were implemented. This database concerns national roads and trunk local roads.

Analyzing data and establishing CMFs

Based on the “Database of Locations Implemented Measures”, the effect on reducing traffic accidents for each type of countermeasure and the general effectiveness of the countermeasures implemented at each location were analyzed.

The establishment of CMFs requires adequate data collection. If additional data is needed during the CMF estimation process, further data collection will be initiated.

Continuous reviewing of CMFs

As the policy includes the PDCA (Plan, Do, Check and Action) management cycle, data will continue to be accumulated after the establishment of CMFs. These CMFs will be reviewed to improve precision after data is collected for several years in many locations.

Netherlands

The SWOV institute of road safety research uses crash reduction factors (CRFs) mainly to estimate the effects of future traffic safety measures. The effect of a certain measure is estimated on the basis of the crash reduction factor, the penetration level and the target group. A recent study discussed the CRFs of 30 traffic safety measures. The following criteria were applied to assess the quality of the evaluation studies that were analysed:

- Is the crash reduction factor based on a before and after study?
- Is there a control group?
- Are the appropriate statistical analyses performed and are they performed correctly?
- Is the number of crashes high enough?
- Is the crash reduction factor based on a Dutch evaluation study?

When available, CRFs from Dutch before and after studies are used (preferably with a control group). However, the number of 'good' evaluation studies is limited, because of the costs of data collection. When no CRFs based on Dutch studies are available, factors from other countries may be used, and in some cases these factors are adjusted (based on expert judgement) to the Dutch situation.

In some cases, CRFs from other countries cannot be applied. For example, this is the case in the evaluation of roundabouts, since the number of cyclists is much higher in the Netherlands compared to other countries.

Norway

Norway applies an abbreviated version of the Highway Safety Manual with new numerical estimates of measures. Norway also applies a five star system in the manual to describe the quality of the assessment / study.

Spain

In Spain, different steps are being taken towards the goal of establishing a procedure for systematically assessing the effectiveness and efficiency (cost-effectiveness) of road safety measures. This process is mainly led by the Ministry of Interior, responsible for the coordination of road safety policies, traffic management and enforcement, driving licences, vehicle registrations, education, training and awareness, and the Ministry of Public Works, responsible for the construction and maintenance of roads and the public transport of goods and people. For any action to succeed, the involvement of departments at different levels —regional, provincial and local— shall prove essential.

Within the Ministry of Interior, the National Road Safety Observatory has, as one of its main tasks, the evaluation of the impact of road safety policies. This evaluation is generally carried out by external research centres and universities. During the past few years, the Observatory has promoted the evaluation of measures such as the penalty point system or the signposting of black spots. This will contribute to developing a set of national estimates of effectiveness (CMFs).

The Observatory has also promoted research about the value of preventing both fatal and non-fatal accidents in Spain. This research has been conducted by a team of University experts, and has been based on the so-called 'willingness to pay' approach, which is now regarded as the best option for computing public preferences for public investments. As a result of this research, the value of preventing a fatality in Spain has been estimated at 1.4 million euros, a value that is consistent with those from other countries also using the 'willingness to pay' approach. The publication and dissemination of the value of preventing a non-fatal accident will follow shortly. This research will allow Spain to comply with the prescriptions of the Directive 2008/96 on road infrastructure safety management, and will make it possible to use cost-benefit analysis for those measures where a crash modification factor is available.

Along with the Ministry of Interior, the Ministry of Public Works is also promoting the use of effectiveness estimates in the construction and operation of roads. At present, road safety must be explicitly considered during the planning, design, construction, operation and maintenance of roads. At the planning stage, the designer must analyse all factors that influence safety, such as the continuity of layout, and the density and design of junctions or facilities for pedestrians and cyclists. An assessment of the expected safety performance of the different options must be made, and safety must be used as a criterion for the final choice. At the stages of layout design and construction, the designer must consider the expected safety impact in the so-called 'safety appendix'.

Different tools are now being used for the safety assessment of roads. The Ministry of Interior published a manual called 'Recommendations for the economic evaluation and cost-benefit analysis of road projects'. These recommendations offer methods for estimating the expected frequency of accidents and victims for a given option. In addition, the Ministry has developed a system for monitoring the evolution of safety in those sites where safety measures have been implemented. This process comprises the registration of data such as the: identification and description of the site, description of the measure, cost of the measure, frequency of accidents during the 'before' period, and frequency of accidents during the 'after' period (typically five years).

The Ministry of Public Works has recently adopted Directive 2008/96, for the safety management of the Spanish trans-European network. According to the prescriptions of the Directive, several procedures for safety management shall be introduced, including road safety impact assessments and road safety audits for infrastructure projects, as well as safety ranking, management and inspections for the road network in operation.

In conclusion, a number of significant initiatives are now being undertaken in Spain regarding the evaluation of the impact of road safety measures and the application of efficiency assessment tools, such as cost-benefit analyses. This process should lead, in the medium and long term, to the consolidation of procedures for systematically assessing the impact of safety measures, both ex-ante (i.e. during the stage of planning) and ex-post (i.e. once the measure has been implemented). There are no national guidelines, nor is there a handbook for systematically collecting CMFs in Spain; the improvement of efficiency assessment tools is thus necessary.

Sweden

The Swedish Transport Administration requires all projects to apply the EVA-model to analyse safety effects in the preliminary and final design stages. The EVA-model operates as follows:

- Networks with and without proposed countermeasure are defined
- Traffic volumes including route choice are applied (car and truck volumes, pedestrian and bicycle flows only for intersections)
- Networks are divided into intersections and links
- Links are divided into homogenous sections by cross-section and speed limit. More specifically, in rural conditions they are classified per alignment class, roadside area type and type of wildlife area, and in urban areas per road function, traffic type and degree of separation of vulnerable road users
- "Normal" traffic safety results, SPFs, are given by the EVA-model for homogenous links (fatalities, injuries etc.)
- Crash cost is weighted with Swedish socioeconomic cost values
- Intersections are classified per number of arms and intersection type (lane layouts, degree of separation, traffic control, grade-separation)
- "Normal" traffic safety results, SPFs, are given in the same way as for links
- Fatalities are assumed to decrease by 2 percent per year and severe accidents by 1.5 percent due to improved cars, increase in safety belt use etc.
- The safety effect is given as the difference between the networks

The EVA link model gives injury type rate SPFs (e.g. fatalities per million vehicle-km) independently of traffic flow, mainly because traffic flow intervals for different cross-section types are rather narrow in Sweden.

EVA intersection models give injury type numbers (e.g. fatalities per year) due to major road traffic flow, secondary road traffic flow and flow of pedestrians and bicyclists. The weighting factors used include traffic volume for the crash data and a regression to the mean factor.

Normal safety performance could be changed using crash reduction / modification factors applied for car, wildlife or pedestrian/bicycle accidents. These crash reduction factors should be picked based on an engineering judgement from a countermeasures catalogue. The catalogue gives a mixture of quantitative and qualitative recommendations for a number of measures, based on research and crash statistics in Sweden and elsewhere (often the Norwegian Safety Handbook). A procedure similar to the HSM with statistical meta tests is not used.

Swedish safety research normally applies "with/without" comparisons (typically 5 to 10 years) at the same speed and traffic flow conditions or before-after studies with regression to the mean and control group corrections.

Literature: Manual cost-benefit analysis for investment projects in Transport Plan preparations 2008 (in Swedish), Swedish Road Administration 2008-06-28; Investment and rehabilitation measures - effects Chapter 6 Safety (in Swedish) Swedish Road Administration.2008:11 updated.

United Kingdom

The Department for Transport provides general guidance on evaluation http://www.dft.gov.uk/pgr/evaluation/evaluationlinks and specific evidence based guidance on road safety to local authorities and local level practitioners. For example, see the road safety good practice guide http://www.dft.gov.uk/pgr/roadsafety/laguidance/roadsafetygoodpracticeguide?page=2 and manual for streets http://www.dft.gov.uk/pgr/sustainable/manforstreets/

In addition to the above, TALs (traffic advisory leaflets) provide summaries of research on specific road related safety interventions http://www.dft.gov.uk/pgr/roads/tpm/tal/

Research on specific interventions is also available. For example, the effectiveness and impact of 20mph zones and limits is well evidenced over a long period of time using standard evaluations of case control before and after measures of impact on casualty numbers up to 3 years before and after. Increasingly, these evaluations are looking beyond impacts on casualty numbers alone and are also considering impacts on the environment, including health and social impacts, e.g. changes in modal choice with increased walking and cycling.

Individual measures are usually part of an integrated approach – the impact on users, vehicles and the road environment. It is also essential to consider the impacts of programmes of activity, e.g. the contribution of Local Safety Schemes to Casualty Reduction
http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme5/rsrr108findings.pdf

Local authorities deliver the bulk of engineering interventions via funding provided mainly through Local Transport Plan allocations. This study was commissioned to help understand overall performance amongst peer authorities and across England as a whole. This study examines the performance of a large sample (408 in number) of the schemes from 2004/05 and quantifies what contribution these schemes made to casualty reduction.

Also, when looking at programmes and the transferability of measures and strategies, it is necessary to contextualize the geographical, socio-economic, cultural, political and legislation framework. This was the approach taken by the SUNFLOWER initiative in which the UK participated.

In addition to engineering focussed measures there are also education, training and publicity (ETP) interventions focussing on behaviour change. These can be evaluated at an individual intervention level or as part of a holistic programme.

For example, individual education interventions may be evaluated to assess impacts on skills, knowledge, attitudes and behaviours. In the UK, the child pedestrian training intervention Kerbcraft was piloted nationally in deprived areas. One hundred and fifteen pilot schemes were funded in 75 authorities across England and Scotland, in areas with high child pedestrian crash rates and high levels of deprivation. Effective ways of establishing and sustaining practical child pedestrian schemes were identified.
http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme1/childpedestrianprojects/networkchildpedestrianhtml

One of the main challenges in undertaking impact evaluations of transport interventions is the ability to demonstrate that the observed outcomes and impacts have been caused by the intervention, thus ruling out the influence of external factors. DfT has produced guidance for evaluators to help choose an evaluation approach to achieve better attribution.
http://www.dft.gov.uk/pgr/evaluation/evaluationguidance/transportimpact/

In order to improve the access to and use of evidence and evaluation in local safety measures, a number of initiatives are underway:

- TAP - the Transport Advice Portal is a joint venture involving the Department for Transport (DfT) and the Chartered Institution of Highways and Transportation (CIHT). The technical library is a repository of core documents covering: legislation; TALs, good practice guidelines, general information, research and useful websites.

  http://tap.iht.org/

- DfT are in the process of developing a web based toolkit for the evaluation of education training and publicity type interventions.

United States

There are several initiatives in the U.S., most notably the production of the Highway Safety Manual (HSM). Because high quality CMFs are essential for effective use of the HSM, the FHWA manages a CMF Pooled Fund Study in which funds from multiple states are combined to invest in the research on CMFs for particular countermeasures. Because good quality research on CMFs is generally expensive, the selection of countermeasures to include in the pooled fund effort is very competitive. Additionally, many studies (recently completed or underway) sponsored by the National Cooperative Highway Research Program are developing CMFs.

In this framework, the Transportation Research Board published Circular E-C142, Methodology for the Development and Inclusion of Crash Modification Factors in the First Edition of the Highway Safety Manual. This circular can be found at: http://onlinepubs.trb.org/onlinepubs/circulars/ec142.pdf.

To ensure practitioners have access to as much information as possible, the Federal Highway Administration manages the CMF Clearinghouse online. The Federal Highway Administration also developed the *Guide to Developing Quality Crash Modification Factors* that is intended to provide clear, concise guidance to agencies interested in developing CMFs. The guide discusses the process for selecting an appropriate evaluation methodology and the many issues and data considerations related to the various methodologies. It is available online at the CMF Clearinghouse.

REFERENCES

AASHTO - American Association of State Highway and Transportation Officials (2010). *Highway Safety Manual*. Washington, D.C.

Abele, J.; Baum, H.; et al (2005) *Exploratory Study on the potential socio-economic impact of the introduction of Intelligent Safety Systems in Road Vehicles,* Germany, January 2005.

Austroads, 2009, *Guide to Road Safety Part 8: Treatment of Crash Locations*, by ARRB Project Team, Austroads report AGRS08/09, Sydney, Australia.

Turner, B Imberger, K, Roper, P, Pyta, V & McLean, J, *Road Safety Engineering Risk Assessment,* Part 6: Crash Reduction Factors, Austroads Technical Report AP-T151/10, Sydney, Australia, 2010

Styles, T Houghton, N, Styles, E, Roper, P & Tay, J, *Road Safety Engineering Risk Assessment*, Part 2: Crash Risk Migration, Austroads Technical Report AP-T147/10, Sydney, Australia, 2010.

Baum, H., Höhnscheid, K.-J. (2001) *Economic Evaluation of Road Traffic Safety Measures: Germany*, in: ECMT, Round Table 117, Paris 2001; P. 5 – 40.

Bliss T (2004). *Implementing the Recommendations of the World Report on Road Traffic Injury Prevention*, Transport Note No. TN-1, World Bank, Washington DC.

Bristow, A.L. et J. Nellthorp (2000). *Transport project appraisal in the European Union*. Transport Policy 7. pp. 51-50.

CEDR - Conference of European Directors of Roads (2006). *Most Effective Short-, Medium- and Lon-Term Measures to Improve Safety on European Roads*. Final Report prepared by the Conference of European Directors of Roads.

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

COWI (2010) *Technical Assistance in support of the Preparation of the European Road Safety Action Programme 2011-2020*. European Commission DG-TREN, January 2010.

ECMT (2001*) Economic evaluation of road traffic safety measures*. Report of the 117 round table on transport economics; Paris, October 2000. European Conference of Ministers of Transport.

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

Elvik R., Vaa T. (2009) *The Handbook of Road Safety Measures*. Elsevier.

Elvik, R. (1996) *Does prior knowledge of safety effect help to predict how effective a measure will be?* Accident Analysis and Prevention 28, no. 3, 339–347.

Elvik, R. (2003) *Assessing the validity of road safety evaluation studies by analysing causal chains*. Accident Analysis and Prevention 35, 741–748.

ETSC (2003*) Cost effective EU Transport Safety Measures*, Brussels 2003

Gramlich, E.M. (1994). *Infrastructure Investment: A Review Essay*. Journal of Economic Literature, 32, 1176-1196.

Hauer, E., (1997). *Observational Before-After Studies in Road Safety*, Pergamon.

Hedman K. O.; Stenborg L. (1991) *Traffic safety measures and Cost- Effectiveness Ratios*. Swedish Transport Research Board & Swedish Road and Transport Research Institute. Linköping.

Institute of Transportation Engineers (1999). *Traffic Engineering Handbook Fifth Edition*. Washington, D.C., Institute of Transportation Engineers.

IRTAD (2009) *IRTAD Annual Report 2009*. OECD/ITF 2010.

Lamm, R., Psarianos, B., Mailaender, T.(1999). *Highway Design and Traffic Safety Engineering Handbook.* McGraw Hill.

NHTSA (2010). *Countermeasures that work: A Highway Safety Countermeasure Guide For State Highway Safety Offices* - 5th Edition.

OECD (1999): *Safety strategies for rural roads*, OECD, Paris.

Ogden, K.W. (1996) *Safer Roads: A Guide to Road Safety Engineering*. Avebury Technical, England.

Peden M, Scurfield R, Sleet D, Mohan D, Hyder A, Jarawan E, Mathers C (Eds.) (2004). *World Report on Road Traffic Injury Prevention,* World Health Organisation, Geneva.

Peltola, H. (2007) *Evaluating measures in order to achieve safety targets*. Road Safety on Four Continents, Bangkok, Thailand 14-16 Nov 2007.

Peltola, H. (2009) *Evaluating road safety and safety effects using Empirical Bayesian method*. 4th IRTAD Conference, Seoul, Korea, 16 - 17 Sept. 2009 http://internationaltransportforum.org/irtad/pdf/seoul/8-Peltola.pdf

PIARC Technical Committee on Road Safety (2003). *Road Safety manual*. World Road Association.

PIARC Technical Committee on Road Safety (2007). *Road Accident Investigation Guidelines for Road Engineers.* World Road Association.

Planitzer S., Körmer C. (2009) *Good and Promising Interventions for the Prevention of Injuries to Pedestrians and Two-wheelers* - Inventory and Guidebook for the Health Sector. Kuratorium für Verkehrssicherheit (Austrian Road Safety Board), Vienna, Austria, February 2009.

PROMISING (2001) *Promotion of Measures for Vulnerable Road Users*, Deliverable D5 *Cost-benefit analysis of measures for vulnerable road users*, July 2001.

ROSEBUD Consortium (2004). *The Use of Efficiency Assessment Tools: Solutions to Barriers*. Available on-line at http://partnet.vtt.fi/rosebud/products/deliverable/Report_WP3_final_june_2004.pdf

ROSEBUD Consortium (2005). *Testing the efficiency assessment tools on selected road safety measures.* ROSEBUD WP4 Final Report. Available on-line at http://partnet.vtt.fi/rosebud/products/deliverable/ROSEBUD_WP4_REPORT_FINAL.pdf

ROSEBUD Consortium (2006). *Examples of assessed road safety measures - a short handbook*. Available on-line at http://partnet.vtt.fi/rosebud/products/deliverable/Handbook_July2006.pdf.

SUPREME (2007) Part C, *Handbook for measures at the Country level*, June 2007.

SUPREME (2007) Part D, *Handbook for measures at the European level*, June 2007.

Swedish Road Administration (2008) *Investment and rehabilitation measures - effects chapter 6 Safety* (in Swedish) Swedish Road Administration.2008:11 updated.

Swedish Road Administration (2008) *Manual cost-benefit analysis for investment projects in Transport Plan preparations 2008* (in Swedish), Swedish Road Administration 2008-06-28.

TARVA (2010) - home pages at web: http://www.tarva.net/tarvaintro.asp

The Cochrane Injuries Group Reviews *(2011) Prevention of road traffic injuries*. Available on-line at: http://injuries.cochrane.org/injuries-group-reviews

Transportation Research Circular E-C142, *Methodology for the Development and Inclusion of Crash Modification Factors in the First Edition of the Highway Safety Manual*. Transportation Research Board, Washington DC. Available on-line at: http://onlinepubs.trb.org/onlinepubs/circulars/ec142.pdf.

Yannis G., Evgenikos P., Papadimitriou E. (2005). *Cost - Benefit assessment of selected road safety measures in Greece.* Proceedings of 13th International Conference on Road Safety Road Safety on four continents, Warsaw, 5-7 October.

Yannis G., Papadimitriou E., Evgenikos P., (2008) *Cost-benefit assessment of the intensification of road safety enforcement in Greece*. Proceedings of the 4th International Congress on Transport Research in Greece, Athens, May 2008.

Yannis G., Evgenikos P., Papadimitriou E. (2008). *Best Practices for Cost-Effective Road Safety Infrastructure Investments.* CEDR - Conference of European Road Directors, Paris, April 2008. Available on-line at: http://www.cedr.fr/home/fileadmin/user_upload/Publications/2008/e_Road_Safety_Investments_Report.pdf

GLOSSARY

Symbols

$\sigma(\theta)$	Standard deviation of the random variable θ
$\theta(a,b)$	The CMF for implementing ‘a’ instead of ‘b’
^	Caret above denotes ‘estimate’
$\hat{\theta}$	An estimate of a CMF - what we expect to be the safety effect of doing ‘a’ instead of ‘b’ in some specific circumstances.
$E\{\theta\}$	The mean value of a distribution of θ
$\bar{\theta}$	Estimated CMF for a future action.
μ	Expected number of (target) crashes.
a and b	Action (intervention, countermeasure, treatment, decision etc.)
$s\{.\}$	Standard error of $\{.\}$
$VAR\{\theta\}$	Variance of the random variable θ

Terms

Bayesian statistics—statistical method of analysis which bases statistical inference on a number of philosophical underpinnings that differ in principle from frequentist or classical statistical thought. First, this method incorporates knowledge from history or other sites. In other words, prior knowledge is formally incorporated to obtain the “best” estimation. Second, the method considers the likelihood of certain types of events as part of the analysis process. Third, it uses Bayes’ theorem to translate probabilistic statements into degrees of belief (e.g., the belief that we are more certain about something than others) instead of the classical confidence interval interpretation.

before-after study—the evaluation of implemented safety treatments, accomplished by comparing frequency or severity of crashes before and after implementation. There are several different types of before-after studies. These studies often develop CMFs for a particular treatment or group of treatments. Also known as *BA studies*.

comparison group—a group of sites, used in before-and-after studies, which are untreated but are similar in nature to the treated sites. The comparison group is used to control for changes in crash frequency not influenced by the treatment.

control group—a set of sites randomly selected to not receive safety improvements.

cost-effectiveness - a type of economic criteria for assessing a potential implementation of a countermeasure or design to reduce crashes. This term is generally expressed in terms of the dollars spent per reduction of crash frequency or crash severity.

countermeasure—a roadway-based strategy intended to reduce the crash frequency or severity, or both at a site.

crash—a set of events not under human control that results in injury or property damage due to the collision of at least one motorized vehicle and may involve collision with another motorized vehicle, a bicyclist, a pedestrian or an object.

crash estimation—any methodology used to forecast or predict the crash frequency of an existing roadway for existing conditions during a past period or future period; an existing roadway for alternative conditions during a past or future period; a new roadway for given conditions for a future period.

crash modification factor/function - an index of how much crash experience is expected to change following a modification in design or traffic control. CMF is the ratio between the number of crashes per unit of time expected after a modification or measure is implemented and the number of crashes per unit of time estimated if the change does not take place.

crash reduction factor (CRF)—the percentage crash reduction that might be expected after implementing a modification in design or traffic control. The CRF is equivalent to (1 – CMF).

crash severity—the level of injury or property damage due to a crash, commonly divided into categories based on the KABCO scale.

dependent variable—in a function given as $Y = f(X1, \ldots, Xn)$, it is customary to refer to $X1,\ldots, Xn$ as independent or explanatory variables, and Y as the dependent or response variable. In each crash frequency prediction procedure, the dependent variable estimated in the base model is the annual crash frequency for a roadway segment or intersection.

efficiency assessment tools - a systematic assessment of the improvement in road safety that can be realised by means of various road safety measures.

Empirical Bayes (EB) methodology—method used to combine observed crash frequency data for a given site with predicted crash frequency data from many similar sites to estimate its expected crash frequency.

entrance ramp—a ramp that allows traffic to enter a freeway.

expected average crash frequency—the estimate of long-term expected average crash frequency of a site, facility, or network under a given set of geometric conditions and traffic volumes (AADT) in a given period of years. In the Empiracal Bayes (EB) methodology, this frequency is calculated from observed crash frequency at the site and predicted crash frequency at the site based on crash frequency estimates at other similar sites.

expected crashes—an estimate of long-range average number of crashes per year for a particular type of roadway or intersection.

experimental studies—studies where sites are randomly assigned to a treatment or control group and the differences in crash experience can then be attributed to a treatment or control group.

ex-post evaluations - before-after ex-post evaluations examine how safety has changed following the implementation of some action; cross-section ex-post evaluations examine how the safety of units where action X was implemented differs from the safety of other units where action Y was taken.

external validity - the possibility of generalising the findings of a study (or set of studies) to other contexts than those in which the study was made. In this study, the term “transferability” is used interchangeably with external validity.

geometric condition—the spatial characteristics of a facility, including grade, horizontal curvature, the number and width of lanes, and lane use.

holistic approach—a multidisciplinary approach to the reduction of crashes and injury severity.

human factors—the application of knowledge from human sciences, such as human psychology, physiology, and

independent variables—a variable which is used to explain (predict) the change in the value of another variable.

meta analysis—a statistical technique that combines the independent estimates of crash reduction effectiveness from separate studies into one estimate by weighing each individual estimate according to its variance.

motor vehicle crash—any incident in which bodily injury or damage to property is sustained as a result of the movement of a motor vehicle, or of its load while the motor vehicle is in motion. Also referred to as a *motor vehicle crash*.

observational studies—often used to evaluate safety performance. There are two forms of observational studies: before-after studies and cross-sectional studies.

predicted average crash frequency—the estimate of long-term average crash frequency which is forecast to occur at a site using a predictive model found in Part C of the HSM. The predictive models in the HSM involve the use of regression models, known as Safety Performance Functions, in combination with Crash Modification Factors and calibration factors to adjust the model to site-specific and local conditions.

publication bias – the tendency not to publish studies whose findings are not statistically significant, counterintuitive or difficult to interpret.

quantitative predictive analysis—methodology used to calculate an expected number of crashes based on the geometric and operational characteristics at the site for one or more of the following: existing conditions, future conditions, or roadway design alternatives.

randomized controlled trial—experiment deliberately designed to answer a research question. Roadways or facilities are randomly assigned to a treatment or control group.

regression analysis—a collective name for statistical methods used to determine the interdependence of variables for the purpose of predicting expected average outcomes. These methods consist of values of a dependent variable and one or more independent variables (explanatory variables).

regression-to-the-mean (RTM)—the tendency for the occurrence of crashes at a particular site to fluctuate up or down, over the long term, and to converge to a long-term average. This tendency introduces regression-to-the-mean bias into crash estimation and analysis, making treatments at sites with extremely high crash frequency appear to be more effective than they truly are.

safety effect – the difference between the number of crashes by severity that one should expect with and without a countermeasure.

safety effectiveness - extent to which predicted or actual safety performance of a countermeasure manifests itself in reduced target crashes.

safety performance function (SPF)—an equation used to estimate or predict the expected average crash frequency per year at a location as a function of traffic volume and in some cases roadway or intersection characteristics (e.g., number of lanes, traffic control, or type of median).

target crashes - a specific type and severity category of crashes that can be affected by a countermeasure. For example, fatal crashes in a region for a countermeasure, all injury crashes on curves, crashes where alcohol was a causal factor for blood alcohol content limits, etc.

transferability – synonymous with external validity. See definition above.

variance- a measure of the average distance between each of a set of data points and their mean value; equal to the sum of the squares of the deviation from the mean value.

CONTRIBUTORS TO THE REPORT

Chair

Patrick Hasson

Federal Highway Administration (United States)

Members of the Working Group

Australia	Peter Cairney ARRB Group	**Germany**	Kai Assing BASt
	Blair Turner ARRB Group	**Greece**	George Yannis NTUA
	Michael Tziotis ARRB Group	**Ireland**	Mike Brosnan Road Safety Authority
Austria	Jennifer Bogner KfV	**Italy**	Francesca La Torre University of Florence
	Thomas Fessl Austrian Institute of Technology	**Japan**	Masahiro Kaneko NILIM
	Christian Stefan KfV		Yoshiharu Namikawa NILIM
	Florian Strohmayer KfV	**Mexico**	Miguel Elizalde Lizarraga FMCA
Canada	Ezra Hauer University of Toronto	**Netherlands**	Wendy Weijermars SWOV
Denmark	Liisa Hakamies-Blomqvist DTU	**Norway**	Rune Elvik TØI
Finland	Harri Peltola VTT	**Spain**	Álvaro Gómez-Méndez DGT/ONSV
France	Laurent Carnis IFSTTAR	**Sweden**	Mats Pettersson Trafikverket
	Sylvain Lassare IFSTTAR	**United Kingdom**	Deirdre O’Reilly DfT
	Jean-Francois Sanchez MEDDTL	**International Transport Forum**	Jari Kauppila

Peer reviewers and other commentators

Ray Krammes, Technical Director for Safety Research, Turner-Fairbank Highway Research Center

Karen Dixon, Professor, Civil Engineering, Oregon State University

David Shinar, Professor, Ben Gurion University

Shalom Hakkert, Professor, Faculty of Civil and Environmental Engineering, Israel Institute of Technology

The Working Group acknowledges the insightful comments provided also by the Joint Transport Research Centre Committee and participants at the International Workshop on Transferability of CMFs, Washington, D.C., 23. January 2011.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Union takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
(75 2012 01 1 P) ISBN 978-92-821-0375-3 – No. 60353 2012

International
Transport Forum

Sharing Road Safety

Each year about 1.3 million people are killed and another 50 million people are injured on roads worldwide. These road crashes cost countries between 1 and 3 percent of their GDP. Many of these crashes can be prevented by effective countermeasures. This report helps identify the most effective safety countermeasures.

Policy makers need to justify expenditure on road safety in terms of effectiveness, competing for the scarce resources available. The risk of making poor decisions and the cost of making better decisions can be reduced by the use of reliable studies on how effective safety measures are, based on Crash Modification Functions (CMFs). This report shows that there is a prospect for significant advances and major cost savings through the transfer of results internationally, allowing for more rapid adoption and dissemination of new life-saving safety measures.

The report serves as a guide to how research results can be shared internationally. It provides checklist for systematic review of road safety studies and a framework for standardising methodology. The report targets the road safety research community but will also find an audience among policy makers at all levels of government. It highlights the value of Crash Modification Functions and the importance of ensuring practicioners use the best CMFs available.

OECD*publishing*
www.oecd.org/publishing

International Transport Forum
2 rue André Pascal
75775 Paris Cedex 16
France
T +33 (0)1 45 24 97 10
F +33 (0)1 45 24 13 22
Email : itf.contact@oecd.org
Web: www.internationaltransportforum.org

9 789282 103753
(75 2012 01 1P1)
ISBN 978-92-821-0375-3

photo credit: © Adam Larsen