A COMMON APPROACH TO UNDERSTANDING STRENGTHS AND LIMITATIONS OF DIFFERENT COST BENEFIT ANALYSIS TECHNIQUES

Tanya, Robinson (nee Smith)
Iain, Knight
TRL
United Kingdom
Paper Number 09-0395

ABSTRACT

Policy makers require evidence of the costs and benefits of a safety measure to inform their views in policy decisions. These analyses are often required in a short period of time with limited research budgets. Increasingly, the measures considered are advanced control systems intended to help drivers to avoid a collision. It is inherently difficult to accurately assess the casualty effects of such systems and this, combined with resource constraints, often results in a wide range of conflicting predictions based on different assumptions, simplifications and analytical techniques. Substantial variation in the presentation of results can make it difficult for researchers to directly compare different studies. In turn, this makes it difficult for policy makers to be confident of the right approach. As a result, studies of very different levels of reliability are often given equal weight in policy debates, risking the possibility of less than optimal implementation of new safety features.

This paper describes the development of a methodology intended to allow a preliminary assessment of the potential benefits of advanced safety systems to be undertaken in a consistent and objective manner. An initial methodology was developed, based on literature and expert opinion, and then tested and refined by applying it to an assessment of existing studies of advanced braking systems for motorcycles.

The research was, therefore, limited to a relatively narrow scope. However, the potential for the method to be expanded in future was explored to assess the possibility of providing a generic methodology to provide guidance for policy makers and researchers alike regarding the:

- Scientific confidence required from a new study or implied by existing analyses;
- Suitability of different analysis techniques for the measure being assessed; and
- Consistent presentation of results to aid subsequent comparison of different studies.

INTRODUCTION

The need for evidence-based policy decisions has developed during recent years. There are a number of techniques that are available to determine the impact of different policy options or proposed legislative changes. The extent of the evidence provided for regulatory change is often directly related to the proposal under consideration.

This paper discusses the issues surrounding the generation of evidence for regulatory change. A number of examples of recent benefit studies are used to highlight the issues. Smith *et al* (2008) developed a methodology for assessing the benefits of active safety systems for Powered-two wheelers (PTWs). Information from this study which was funded by the UK Department for Tranpsort is used as a starting point to stimulate and inform further debate and the future improvement of the research and policy making community's efforts to improve the evidence on which decisions are based.

IDENTIFICATION OF THE PROBLEM

Historically, safety improvements have been developed quite slowly. For example, the European Frontal Impact Directive came into force in the mid-nineties but was supported by research and development going back to the 1970s/80s. The rapid development of new active safety systems coupled with the fact that safety has become a strong selling point for manufacturers and consumers means that governments are often under pressure to regulate much more quickly. For example, the very first collision mitigation braking systems (CMBS) came on the market only a few years ago and currently are only available as options on a small number of high end passenger cars and one or two truck models, but it is proposed that they are mandatory on all new heavy vehicles by 2013. This means that impact assessments for the regulations must be completed at a time when little, if any, accident data for vehicles equipped with such systems exists. This situation is further complicated when assessing the benefits of primary safety (accident avoidance) technologies. Unlike secondary safety (severity reduction) measures, success means that there is no accident and

therefore no accident data with which to compare the outcome before and after implementing the measure. These factors combine to result in considerable variation in the quality/depth of the analyses produced.

However, each level of analysis has its place with respect to the change in legislation being proposed. For example, well supported proposals, such as the introduction of Electronic Stability Control (ESC) for Heavy Goods Vehicles (HGVs), may require only minimal evidence to be successful. However, more controversial proposals such as those for Brake Assistance Systems (BAS), Daytime Running Lights (DRL) and Advanced Emergency Braking Systems (AEBS) have been supported with a wide variety of studies of differing scientific quality/depth. The scientific quality of benefit assessments is influenced by a number of factors, such as the availability of data upon which to base the assessment and suitable information about the effect of the proposal.

Many benefit assessments consider the potential effect of each safety measure on its own. However, in many cases, a number of different measures could influence the same groups of casualties. One example of this is the proposal of the European Commission to mandate Automated Emergency Braking Systems (AEBS). The systems have an automated braking function, the benefits of which can be predicted using existing accident data. However, it is anticipated that the production systems will include functions such as Adaptive Cruise Control, Forward Collision Warning and pre-impact adaptive restraint systems, which will not be mandatory. These types of system are already fitted to some vehicles and will be fitted to more vehicles than AEBS. Therefore the fleet penetration of such systems will be ahead of AEBS, thus reducing the benefits of the AEBS function itself; a factor not accounted for in the benefits study undertaken. There are also potential effects on completely separate systems such as anti-whiplash seats, because AEBS will influence the frequency/severity of rear impacts.

DEVELOPMENT OF A METHODOLOGY

Smith *et al* (2008) developed a generic methodology to evaluate the casualty benefits of advanced safety systems for PTWs. The research was funded by the UK Department for Transport and the objectives of the methodology were to:

- Identify the most suitable and cost effective method of providing evidence of a safety benefit for a range of motorcycle safety systems.
- Include provision to estimate the potential for accident avoidance or injury mitigation using accident statistics or in-depth accident data by identifying causation factors and then assessing the likely impact of advanced safety systems for relevant accidents.

In addition to achieving these two objectives, it should also be possible to use the methodology developed to appraise critically research that has already been completed. Although these objectives are specific to assessing advanced safety systems for PTWs, the principles of the methodology can be applied to all safety measures across all types of vehicle. The following section describes the methodology developed by Smith *et al* (2008). The methodology consists of three main steps and an overview is shown in Figure 1.

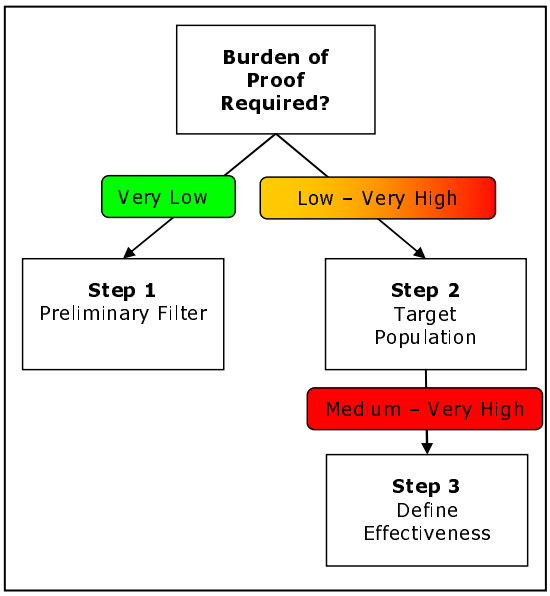

Figure 1. Overview of methodology.

The methodology starts by identifying the burden of proof that the evidence must satisfy, for example:

- Policy makers are sometimes confronted with a large number of proposals for a huge range of potential new safety measures. In this type of situation it is considered useful to have an initial filter to help identify which measures warrant further investigation. It is not necessary to have rigorous proof of the exact effects, merely a broad indication of the potential. This sort of requirement is considered to represent a very low burden of proof.
- By contrast, if a major new safety regulation is planned, that is likely to have a high cost, carries a risk of adverse effects on other policy areas (e.g. GHG emissions) and/or is likely to encounter significant opposition, then very rigorous supporting analysis that accurately and incontrovertibly demonstrates the effects

might be required. This would be considered to require a very high burden of proof.

Step 1 – preliminary filter

Step 1 of the methodology is the definition of a preliminary filter that can be applied to accident data. The primary objective of this step is to define groups of accidents, against which an initial evaluation of the potential benefit of a safety measure can be assessed. Additionally, a secondary objective is to allow the most frequent or most severe groups of accidents to be identified in order to inform the development of new safety systems.

Step 1 can be used for quick stand-alone comparisons for a range of potential safety measures. It could also be used to quickly assess how relevant proposals from one country are to the vehicle and accident population in another country, or as a quick reference to assess the maximum potential benefit of a new safety system. For example, such a filter was developed as part of a review of heavy vehicle safety priorities (Smith *et al*, 2007). During a subsequent policy debate about the possible extension of the scope of UNECE R66 to double deck buses and minibuses, this filter was used, in a matter of minutes, to identify that in the UK large bus/coach occupant casualties (i.e. including those in single deck vehicles already included in R66) in rollover accidents were the 157th most important casualty group involving heavy vehicles (out of a total 244 groups) with an annual casualty valuation of £1.8m. Thus, extending the scope of R66 to double deck vehicles was considered unlikely to be cost effective in the UK unless the measure could be implemented very cheaply. A similarly quick analysis found that extension to minibuses had much greater potential.

In order to carry out step one, a definition of the system specification is required. This should set out the functional requirements of the system under consideration, allowing the casualty groups that could be affected to be identified. When setting up a preliminary filter, there are three main considerations:

<u>What is an appropriate data set?</u> - It is recommended that the data set is a national sample, or is known to be representative of the national sample (evidence of how the data set represents the national population should be presented). The data should cover a period of at least one year, ideally an average of a number of years and be as up to date as possible.

<u>How should accidents be grouped in the filter?</u> - The grouping of accidents can be influenced by the vehicle type to which the safety system is to be fitted, as well as the type(s) of system under consideration. The following aspects should be considered and any limitations of the grouping should be noted.

- The grouping should allow comparison of accident types and be independent from the detailed functionality of the safety systems.
- The grouping should be appropriate to the systems being reviewed. It should allow differentiation between different systems where possible (e.g. a braking system could influence a small proportion of a large number of groups, whereas a cornering stability control system might influence only one or two groups).
- The groups should be mutually exclusive to avoid double counting where multiple groups are affected by a system.
- All casualties within the accidents should be included if possible, i.e. casualties in the vehicle to which the system is to be fitted, casualties in the opponent vehicle (1st impact) and any other casualties in the accident (including pedestrians).

<u>How will the groups be compared?</u> - The accident groups can be compared using a number of different measures that reflect the frequency and/or severity of the casualties (e.g. number of casualties, number of fatalities, monetary valuation associated with the prevention of casualties etc.).

The output from this step is an estimate of the maximum potential benefit of the system. The estimate will be the sum of the casualty groups that can potentially be affected. Although a relatively crude assessment, the preliminary filter will identify if there are 10s or 1,000s of casualties that could be addressed by the system.

Additionally, the preliminary filter will produce groups of casualties. This enables a reference tool for policy makers and researchers to identify where resources should be targeted. For example, in 2005 there were no fatally injured riders of PTWs with engine capacity less than 50cc in collision with a minibus, and only three seriously injured. In comparison, there were 101 fatally and 818 seriously injured PTW riders where a PTW with engine capacity more than 500cc was the only vehicle involved.

Step 2 – target population

Step 2 of the methodology is intended to identify more accurately the accidents that could be affected by the system under consideration (defined as the target population). The target population is specific to the safety system and should be as accurate as possible including causation factors where required.

The term "target population" can be used in a variety of ways, for example:

- The number of casualties that could be prevented by a system that is 100% effective in each of the accident situations it is intended to influence e.g. works in all weather conditions, at all speeds and accounts for driver behaviours etc.; or
- Casualties within a group of accidents that could potentially be influenced by the measure e.g. head-on collisions, rear-end shunts etc.

This measure is one that is often used differently in different studies. For example, the number of detailed data fields (e.g. impact location, speed, driver behaviour factors etc.) that are used to identify it can vary considerably between studies, often leading to misunderstandings and difficulties for policy makers comparing the results of different studies. There is, therefore, a need for a more common understanding of what is meant by the term. For the purpose of this methodology, the target population is defined as the number of casualties that could be prevented by a system that is 100% effective in each of the accident situations it is intended to influence. For example, for a forward collision warning system it would be all casualties where the impact location was the front of the vehicle, the vehicle was moving forward prior to impact and the driver/rider was considered to have been inattentive. This number can also be expressed as a percentage of all accidents.

To carry out step 2 of the methodology a detailed specification of the system and appropriate accident data are required. There are five aspects to be considered in this step:

In what situations is the system intended to be of influence? - In order to define the target population it is necessary to understand how the system operates and the situations where it is intended to be of influence. There should be a written description included in the write-up of the analysis.

What are the relevant types of accident and vehicle for the system being assessed? - The definition of each accident type and relevant vehicles should follow these guidelines:

- The accident types that could be influenced by the safety system should be identified in as detailed manner as possible for the data source being used. The definition should include criteria that will allow the accidents relevant to the specific system to be identified. For example, head-on collisions can have a number of different causative factors (inattention of one or more of the drivers involved, impairment of the rider/driver etc.). It is recommended that the accident type is defined by the impact configuration (where appropriate) as well as at least one causation factor such as rider/driver behaviour (where appropriate). There may be multiple types of accident that could be influenced.
- It is often appropriate to define the target population in relation to the vehicle type to which the system is to be fitted (e.g. HGV, passenger car, PTW etc.). The composition of the vehicle fleet can be very different when comparing different countries. Sometimes it may be appropriate to define the target population for a sub-set of one vehicle type. For example, when considering ABS for PTWs, the target population can be separated by engine capacity, PTW less than 50cc and PTW greater than 50cc because small urban mopeds are involved in different types of accidents to larger, more powerful motorcycles.

What information is available to estimate the target population? - The target population can be estimated based on different sources:

- Accident data will allow the most flexibility in defining the target population (within the constraints of the data sample being used). This is the preferred method for defining target population.
- Existing scientific literature and benefit studies can also be used but the definition of the target population is likely to vary between different studies and if no studies are available for the required country, the answer could be misleading, particularly where patterns of use vary considerably between different countries, as is the case for PTW accidents.

How can these relevant accidents be identified in the accident data? -

- Does a national data sample have a sufficient level of detail? Is causation data and pre-impact information available to identify the relevant accidents in the national data sample?
- If it is not possible to identify the relevant accidents using a national data sample, is there an in-depth study available that has appropriate detail and represents the national sample appropriately (at the level of detail required)? If so, the use of a more detailed accident database should be considered. However, it is necessary to identify the limitations of such an approach. One of the most important limitations will be related to the representativeness of the data sample. Any assumptions must be reported, for example if the representativeness is not known at the level of detail required (e.g. rider behaviour factors) but is known at a high level (types of casualties and vehicles involved), it can be reasonably

assumed that the rider behaviour data is also representative. However, such assumptions must be stated clearly in the report of the analysis.

- What factors could influence the target population? The target population can be defined in a number of ways. This could lead to the inclusion or exclusion of accidents where certain factors were involved. For example, should impaired drivers (e.g. intoxicated through alcohol or drugs) be included in the target population? In general, accidents should only be excluded from the target population on the basis of this type of factor if it is clear that there is no chance that the measure under consideration will affect them. This will help to allow consistency in study approaches using different data sets, for example, in different countries. However, in some circumstances it will be appropriate to restrict the target population in this way and wherever this occurs the restriction should be stated and the calculation of effectiveness that will be applied in step 3 should be modified accordingly.
- Are there any limitations with the criteria that have been used to define the target population? Some data recorded in databases have inherent limitations. For example, the information required may frequently be unknown, or some may rely on subjective assessments.

Have the correct accidents been identified? - It may be possible that the criteria used to select a specific group of accidents could unintentionally return some non-relevant accidents. The analysis should be accompanied with some indication of confidence in the query that has been used. If the data source has written descriptions of the accidents then these could be used. However if there are no written descriptions then an alternative method should be considered, for example cross-referencing to another database that does have written descriptions.

The output from Step 2 is the target population for the specific system that is being assessed. The target population is a group of accidents that are relevant to the system under consideration. This is the maximum potential benefit for the system, i.e. if it were 100% effective; target population is equal to the expected benefit. In reality, most systems are not 100% effective at preventing the collisions/casualties for which they are designed and thus, step 3 is required to more accurately quantify the expected benefits.

Where possible, the target population should be expressed for each casualty severity as a proportion of all casualties of that severity. However, in some cases it is not possible to identify all casualties of a particular severity. For example, official statistics for the EU-27 provide the number of fatalities and the number of all accidents but not the number of serious and slight casualties. Therefore, the target population should also be shown as a proportion of all accidents (of all severities) within the sample. This will assist direct comparisons across different countries. However, when using the target population as a proportion of all accidents, care should be taken when translating results from one country to the accident numbers from another country because of variations in the definitions used for the casualty severities. A table showing how that data should be presented is shown in the example below. Figures that may not be readily available in all countries/regions are identified by an asterisk.

Table 1.
Example presentation of target population data

	Fatal	Serious	Slight	Total
Target Population (number of casualties)	123	467	1252	1842
Total number of GB casualties (by severity)	3512	24571*	256830*	284913*
Total number of GB accidents	-	-	-	197856
Target population (% of GB casualties by severity)	3.50%	1.90%*	0.48%*	0.65%*
Target Population (% of all GB accidents)	0.06%	0.24%	0.63%	0.93%

Step 3 – effectiveness

Step 3 of the methodology is intended to refine the benefit estimate that was defined in Step 2, that is, to translate the analysis from the maximum possible benefit (target population) to a realistic likely benefit. The main objective of this step is to determine how effective the system will be for preventing the casualties/accidents that make up the target population. There are a number of different methods for determining/identifying the effectiveness of the system and this step is intended to help identify the most appropriate method for the quality of estimate/burden of proof required. It is possible to define the effectiveness of the system under consideration without defining the target population in Step 2. The inputs into Step 3 of the

methodology can depend on the approach taken, but can include:

- Accident data;
- Literature;
- System specification;
- Quality requirements;
- Test/trial results.

In order to determine the effectiveness of the system in the most appropriate manner, the following aspects require consideration:

What burden of proof is required? - The burden of proof required should be classified on a scale from very low to very high. Step 3 is typically only required when the burden of proof is medium or higher, for example, proposals for voluntary or mandatory fitment. Figure 2 summarises how to determine the most appropriate method. Additional guidelines are provided below:

What is most appropriate assessment method for the information available? - The selection of the method to be used will be based on the burden of proof required, the availability of the system being assessed, constraints on cost and time and the availability of accident data and literature. The main types of method that can be used for determining/estimating the effectiveness of the system are:

- **Predictive studies** examine accidents where vehicles were not equipped with the specific feature under consideration and make calculations and/or judgements to assess whether the accident would have been avoided or mitigated if the safety feature had been present. There are a number of different methods that can be used when carrying out a predictive study. The most appropriate method will again be influenced by the burden of proof required and budgetary/time constraints:
- A parameter based predictive study is the most straightforward, and it is likely to be appropriate for a medium burden of proof. This type of study is an extension of the target population exercise described in Step 2 and involves interrogating an accident database to identify in more detail the casualties where a system is likely to be effective. If a forward collision warning system was assessed, the target population might be all front to rear shunt collisions where the vehicle to which the system is to be fitted approached from the rear. The effectiveness calculation might further restrict the target population to exclude accidents where the driver of the vehicle of interest was impaired, accidents on a bend, or those that occurred in severe weather conditions where the system was known not to function well. The quality of this type of analysis will depend upon the detail, accuracy and representativeness of the data source used, the available definitions of system functional performance and any assumptions made to overcome limitations in the data.
- Case by case analysis involves the detailed review, reconstruction and prediction of effects in a range of individual accidents. The predictions can be made in a number of ways:
    - An assessment of the effectiveness of the safety system can be made for each accident case identified based on the information available and engineering judgments. Again the quality of this assessment can be influenced by the source of data. If the cases have been reconstructed based on the sequence of events, the evidence left at the scene (e.g. tyre marks) and mathematical calculations (e.g. police accident reconstruction) there is more information available than what may be available from a limited number of database fields. The method can be made less subjective by providing guidelines that define when the system is expected to be effective. Weighting of the assessment with estimates of the probability of effectiveness (e.g. definitely, probably or maybe) can also reduce the subjectivity of the assessment. This method is likely to be appropriate for a high burden of proof (e.g. proposal for mandatory regulation of a moderately costly system).
    - Mathematical modelling can be used on a case by case basis and is less subjective than the method described above. This method involves creating a computer model of an accident and simulating the outcome with the fitting of the safety system. Such an approach has the advantage of being fully objective but is more complex and time consuming and, because it is firmly rule based, can miss some more subtle factors that influence outcomes. This method is likely to be appropriate when a high burden of proof is required.

A limitation of both techniques above is that it is difficult to rigorously include driver behaviour factors associated with the new system in the assessment of its effectiveness. Particularly for primary

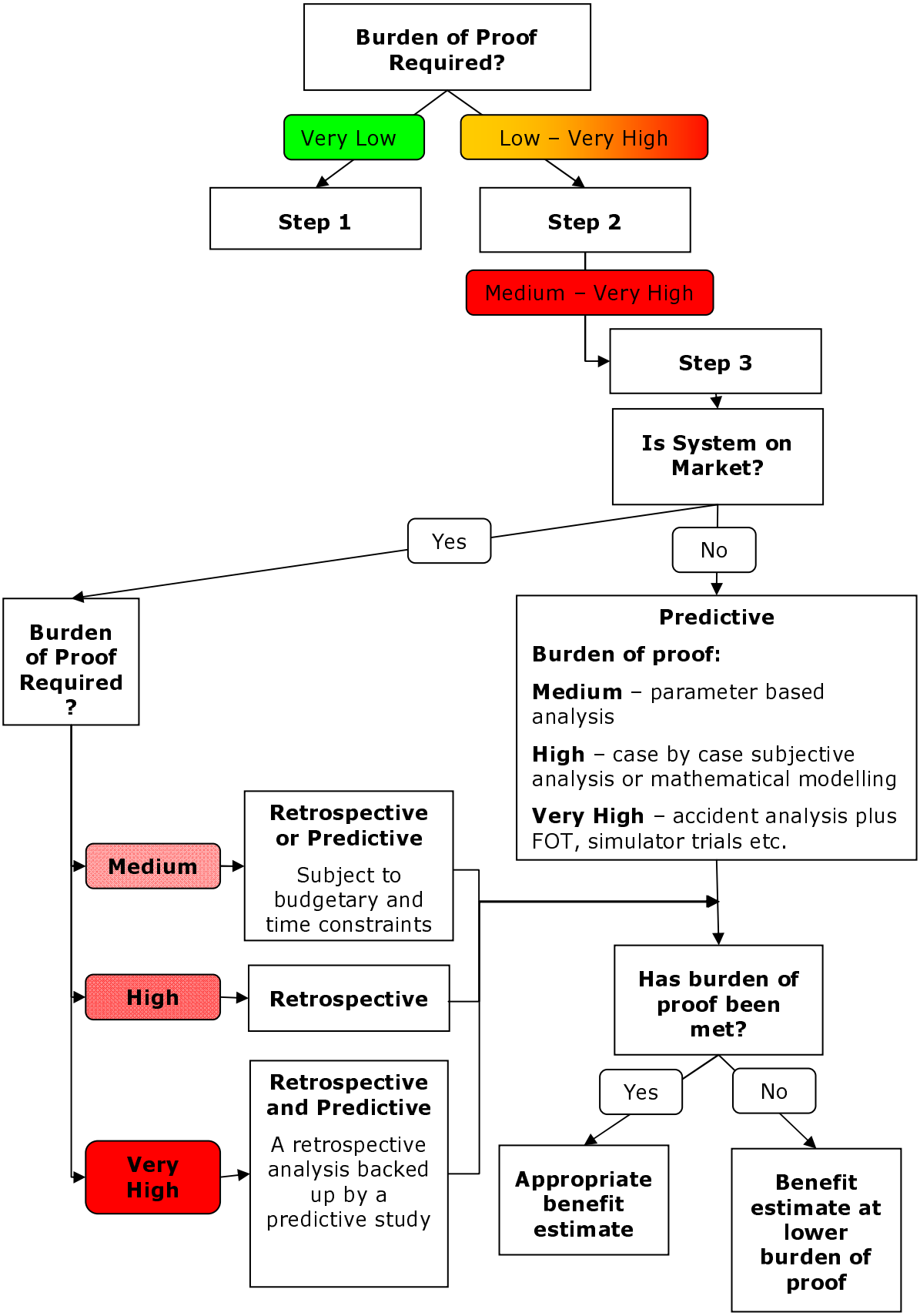

Figure 2. Identification of appropriate method for determining effectiveness.

safety systems, this means that it can be easy for critics to argue that the results are not valid because the system would induce a behavioural change that would reduce or eliminate the predicted benefits. Where the highest burden of proof is required this limitation can be overcome through the use of physical trials involving ordinary drivers as subjects. These can take the form of simulator trials, track trials or field operational trials. This method can allow for human factors issues to be combined with the accident data assessment, however the reliability of the data is dependent on the assumptions made and the experimental methods used.

- **Retrospective studies** treat the feature under investigation as a risk factor and use statistical methods to compare the relative risk of accidents in real world accident data where vehicles can be identified that both do and do not have the safety feature fitted. Where such an approach is possible, it has the most potential for providing a rigorous and defendable outcome because it seeks to objectively measure the actual effect on real vehicles in service with real drivers, thus accounting for many of the factors that can confound predictive studies. The size of the sample will have a strong affect on whether statistically significant conclusions can be drawn and the analytical design, particularly the control of confounding factors (e.g. systematic biases such as age of driver etc.), will strongly affect the quality of the results.

Is the system on the market? - Whether the system is on the market, or available for trials will influence the type of analysis that can be completed.

- No – If the system is not on the market, or at least not in significant numbers, then the estimation of effectiveness is restricted to a predictive study.
- Yes – If the system is on the market then either a retrospective study, a predictive study, or both can be carried out depending on the burden of proof required, analytical design factors and budgetary constraints for completing the analysis.

What sources of information are available for determining the effectiveness? -

- **Literature**, which could include the findings from a range of studies that have already been carried out which could have determined the effectiveness of the system under consideration. The findings from other studies should be reviewed critically and any assumptions made should be identified in order to determine if the effectiveness quoted is appropriate for the target population. The use of multiple sources is recommended, identifying where there is agreement or differences between studies. It may be necessary to define a range of effectiveness if there is no consensus in the literature and the logic used to define the range should be reported. Where sufficient detail exists a formal meta-analysis can be undertaken. This essentially involves calculating a statistical weighted mean of the effects identified by the previous studies. However, this can require substantial time and effort and requires the data in the literature to be well reported in considerable detail.
- **Specific research studies** can be used as a substitute for accident data and can include field operational trials or questionnaire surveys to compare the accident involvement of equipped and unequipped vehicles and estimate the relative change in risk for equipped vehicles.
- **Accident data** can be used to allow either predictive or retrospective studies. The data sources used will be influenced by the burden of proof required, the type of analysis and also the function that the system is intended to achieve. For example a parameter based predictive analysis for assessing the benefits of improved helmets is likely to require a different source of data to a case-by-case predictive analysis of an advanced braking system. Retrospective analyses have different requirements again, and are typically based on national accident data.
- **Vehicle equipment data** can be used to identify whether the specific vehicles recorded in the accident data are fitted with a specific safety system. This type of information is an essential pre-requisite of retrospective analyses.
- **Exposure data**, or the use of an induced exposure technique, is required to allow the probability of an accident occurring to be determined when carrying out a retrospective analysis.

The output from Step 3 is an estimate of system effectiveness that is relevant to the target population that was defined in Step 2. This can then be applied to the target population to estimate the casualty benefits for the safety system.
The estimated benefits should be clearly expressed as a percentage of the target population (so it can be seen how effective the system is at addressing the intended group of accidents) and as a percentage of all accidents. In particular, the latter measure is important for comparison with other studies and for context for the predicted casualty benefit.

The estimated casualty benefits can be combined with vehicle registration data, casualty valuation information and details of the costs of the system to produce a full cost benefit analysis. Defining procedures or guidelines for the generation of cost benefit analyses was beyond the scope of this project. However, it is possible to define a preliminary assessment on the basis of the calculation of a break-even cost for the system. This is calculated by multiplying the number of casualties by their casualty prevention value and dividing by the number of new registrations expected each year. This represents the maximum cost that can be associated with fitting the system to a single vehicle that could still produce a benefit-to-cost ratio to equal one. If the actual costs of the system are likely to be substantially below this break-even cost then it is likely that the system would prove to have a positive benefit to cost ratio (greater than one) if a full analysis was undertaken. If the cost is substantially greater it is likely to have a negative ratio (less than one). Where the actual costs are relatively close to the break-even cost, the simplifications inherent in this method mean that the outcome remains uncertain.

APPLICATION OF THE METHODOLOGY

Smith *et al* (2008) applied the proposed methodology to assess the potential benefits of advanced braking systems for PTWs. The methodology was applied for three systems:

- Anti-lock braking systems (ABS);
- Combined braking systems (CBS); and
- Brake assist systems (BAS).

This paper describes the information collated for the assessment of ABS as an example of how the methodology can be applied. The assessment of the potential benefits of ABS was restricted (by available budget) to the use of existing information only, i.e. mainly literature supplemented by limited analysis of existing accident data.

Step 1 – preliminary filter

The preliminary filter is intended to be used to identify casualty groups that could potentially be affected by the technology under consideration. However, ABS can influence a broad range of casualty groups to varying extents. It is, therefore, only possible to generate a coarse estimate of the target population for braking systems using this tool. However, analysis by the type of PTW involved could provide an insight into where to target the technology. Additionally, it is possible to logically assess the types of accident where advanced braking systems are more likely to have an influence, for example accidents at junctions where the PTW is travelling ahead and single vehicle accidents involving loss of control on a bend, and then quantify the number of casualties occurring in these "more likely" accident types. Such an assessment will be imperfect because not all of these casualties will be influenced by the technology and there will also be other casualty groups that have been excluded but may be influenced. However, it could give a closer indication than considering "all" accidents only.

A preliminary filter was developed based on a three year sample of national road accident data (STATS19). The PTW casualties were grouped by the type of PTW being ridden (i.e. <50cc, 50-125cc, 125-500cc and >500cc) and the number of vehicles involved in the accident (single vehicle vs multi-vehicle). Further categorisation is based on criteria such as: where the accident occurred (at a junction or not), whether there was loss of control, the first point of impact on the motorcycle and what manoeuvre the PTW was making. Table 2 shows an example of the data that can be obtained by using the preliminary filter.

Ideally the target population for all three braking systems would be any accident where the vehicle braked. Unfortunately, this cannot be identified from the available data. Therefore, the only rigorously acceptable target population is all casualties. The preliminary filter has been used to compare the relative size of the different casualty groups. It can be seen that the greatest benefits would appear to lie with larger motorcycles, simply because of their greater involvement.

Table 2.
Examples of casualty groups where ABS is more likely to have an influence

	Junction accidents - PTW going ahead		Single vehicle loss of control on bend	
	Casualties	% of all PTW accidents by severity*	Casualties	% of all PTW accidents by severity*
Fatal	108	18.5%	46	7.9%
Serious	1238	20.7%	296	4.9%
Slight	3218	18.6%	320	1.8%
KSI	1347	20.5%	342	5.2%
Total	4565	19.5%	662	2.8%

**Note: percentage values are not a direct output from the preliminary filter but are calculated using only the data generated by the preliminary filter.*

A more subjective approach can be used to try to get a more realistic target population. An upper estimate was based on excluding accidents where the PTW was waiting or turning (where it is logical to assume braking might be less relevant). A lower estimate was derived from considering only accidents with loss of control on a bend and junction accidents (where logically, braking is likely to occur frequently). The estimated target population using this approach is:

- 154 to 572 PTW fatalities;
- 1534 to 5624 serious PTW casualties; and
- 3538 to 15323 slight PTW casualties.

Step 2 – target population

If a new full-scale analysis of GB accidents was being carried out, then the target population would be identified as defined previously. Data from a detailed in-depth study which is representative of the national statistics (STATS19 data) would be used to identify the proportion of PTWs that braked prior to impact and this proportion would be applied to the casualty numbers recorded in STATS19 to obtain a sound estimate of the target population nationally. However, the scope of the research was restricted to analysis based on existing accident data and literature. So, although a definition for the target population is provided in the previous section, based on the methodology defined in this report, the analyses reported in existing literature may have structured their findings differently. Table 3 summarises the literature and data relating to target populations for ABS that was identified.

Step 3 - effectiveness

If a new analysis of the potential benefit of ABS was to be undertaken in accordance with the methodology defined earlier, to meet a high burden of proof then the programme of work could involve:

- Detailed definition of the performance characteristics of the system;
- Predictive analyses, based on case by case review and reconstruction of on-the-spot and/or fatal cases to assess the influence of each system with extrapolation of results to the national statistics (STATS 19) for an estimate of national benefits.
- Human factor experiments on the test track to assess rider response to the system and identify any behavioural risks;
- Identification of makes and model of PTW fitted with ABS (which is possible based on manufacturers literature but is labour intensive);
- Retrospective statistical analysis of the relative accident involvement of PTWs with and without the system.

However, the scope of this research was limited to a review of existing literature. Table 4 summarises the findings from this review with respect to the effectiveness of ABS.

Table 3.
Summary of target populations for ABS as defined in the literature

Target Population	Source
All cases in which it can be conclusively proven that braking took place in the pre-crash phase (45% of all accidents)	Gwehenberger *et al* (2006) Allianz centre of technology
All in-depth data collected from Hurt (1981) and MAIDS (2004) studies, i.e. a sample substitute for all motorcycle accidents	Kebschull and Zellner (2007) Dynamic Research for IMMA
All motorcycle accidents	McCarthy and Chinn (1998 & 1999) TRL
All motorcycle accidents involving downfall* prior to first impact	Baum *et al* (2007) University of Cologne
Collisions between motorcycles and cars that involved braking (65% of all accidents)	Sporner and Kramlich (2000) cited in Vavryn and Winklebauer, 2005 and Gwehenberger *et al*, 2006
All accidents in sample	Teoh (2008) Insurance Institute for Highway Safety USA.
All accidents in sample where at least one wheel has locked prior to impact or loss of control (34% of all fatal PTW accidents)	McCarthy *et al* (2008) PISa

** downfall accidents are when the rider becomes detached from the PTW before the first impact.*

To allow a detailed assessment of the benefit of fitting ABS, the effectiveness of the system for each accident severity is required. The literature review did not identify the effectiveness for PTW fatalities that could be applied to the target population. Therefore, the effectiveness of ABS used to identify the proportion of fatal casualties that can be prevented, was based upon the information found from a review of the PISa Fatal

Accident Database. The development and analysis of the database was reported in McCarthy *et al* (2008). However, the target population used was not consistent with the other research identified and therefore the data was re-analysed. From this additional review of fatal accidents, an effectiveness of between 8.8% and 35.7% was estimated, with a best estimate of effectiveness of 18%.

For serious casualties, the effectiveness used is based upon the estimates outlined in Gwehenberger *et al* (2006) and McCarthy and Chinn (1999). Although Gwehenberger *et al* (2006) include accidents of all severities, the sample is most representative in relation to serious casualties and states an effectiveness range of between 8% and 17%. McCarthy and Chinn (1999) state an effectiveness of 3% for fatal and serious casualties, however the estimate is likely to be dominated by the effectiveness for serious casualties because the sample included only a relatively small number of fatalities. Therefore 3% was selected as an approximate lower boundary for the effectiveness for serious casualties

There was no research that specifically identified the effectiveness of ABS for slight casualties. However, Sporner (2000, cited in Gwehenberger *et al*, 2006) stated that ABS is effective in 10% of PTW accidents of all severity levels. In comparison, Kebschull and Zellner carried out a comprehensive study resulting in an overall effectiveness of between 1% and 3%. However, there were limitations associated with both studies as described below.

Sporner *et al* (2000), cited in Vavryn and Winklebauer (2006), undertook a study of 610 in-depth accident reports. Vavryn and Winklebauer (2006) stated that Sporner *et al*'s findings were that on average:
"Approximately 55% of the motorcycle accidents could be avoided or at least positively influenced by ABS".
Multiple papers by Sporner *et al* (2000, 2002 and 2004) are cited by Gwehenberger *et al* (2006). Gwehenberger *et al* (2006) stated that Sporner *et al*'s findings were that:
"approx. 10% of motorbike accidents involving bodily injuries can be avoided or at least positively influenced through ABS".

There appears to be some discrepancy between these two interpretations of an estimate of effectiveness from a single source. The only immediately apparent difference in the citations is the effectiveness estimate and the fact that Gwehenberger *et al* reference their effectiveness as a proportion of PTW accidents involving bodily injury whereas Vavryn and Winklebauer's citation does not mention injury severity so could refer to a specific severity level. However, it has not been possible to locate an English language version of the original paper to clarify the exact findings. The estimate of 10% is most likely to be applicable to the target population that has been defined for this study for the following reasons:

- The effectiveness of 55% was written as though it may be the effectiveness for a different target population.
- Sporner was one of the authors of the Gwehenberger *et al* (2006) paper and would be expected to ensure that his previous research was cited correctly.

Kebschull and Zellner (2007) found a relatively low effectiveness compared with other studies. A large percentage of the 900 European accidents investigated contained accidents which involved either no braking or braking with no loss of control, which was assumed in their investigation to be sub-limit braking. A large proportion of the accidents that involved over braking also involved an emergency steering action. In general, PTW ABS does not allow the PTW to maintain stability while braking heavily in a curve/swerve. This was reflected in the ABS model used in this study, which was not capable of maintaining stability in a swerve when braking was severe enough to activate the ABS. This was a predictive study that used computer simulation to predict how the outcome of real accidents involving PTWs without ABS would have been changed if the vehicle had been fitted with ABS. This approach would result in evidence that has a high burden of proof according to the methodology defined earlier in this paper. However, the assumption that ABS would have no influence in any accident where braking occurred without loss of control contradicts several other studies which suggests that ABS gives the rider more confidence and in turn, results in higher maximum achievable deceleration. Therefore, the method used in the analysis may tend to under-estimate the benefits.

Because of the limitations with both studies, it was not clear which effectiveness was most appropriate and therefore a weighted average from these two studies has been used for the best estimate. Based on the mid range value from Kebschell and Zellner of 2% and a quality rating of 3 for each study, the best estimate is 6%. This was generated by multiplying the effectiveness by the score (2%x3 and 10%x3), summing (6+30) and dividing by the sum of the effectiveness scores (36/6). The extreme values from the two studies have been used to generate the overall range of effectiveness.

Using the 6% value for all accidents and the best estimates of 18% for fatalities and 10% for serious

Table 4.
Effectiveness of ABS as identified from the literature

Effectiveness	Source	Region	Study type	Sample size
85% of all downfall accidents with downfall before initial impact	Baum *et al* (2007) based on a retrospective study.	Germany	Retrospective	
Approximately 10% of motorbike accidents involving injury can be avoided or positively influenced	Sporner *et al* (2000,2002,2004) cited in Gwehenberger (2006) describe the dangers of braking with conventional braking systems and the avoidance potential through ABS in several studies based on the GDV accident database (insurance claims).	Germany	Predictive	
Avoids 8%-17% of serious motorbike accidents	Gwehenberger *et al* (2006). Results of analysis of 200 serious accidents by Allianz Center of Technology. Extrapolated to Germany would result in around 100 deaths and more than 1,000 serious injuries avoided a year	Germany	Predictive case by case; subjective	200 accidents
Net injury benefit 1%-3% of all casualties	Kebschull and Zellner (2007) conducted a series of computer simulations based on data collected in the MAIDS (2004) and Hurt (1981) studies. Several configurations of ABS were simulated.	USA and Europe	Predictive case by case; computer modelling	1800 accidents
Analysis of Austrian statistics showed that the benefit was comparable to the 55% stated by Sporner *et al* (2000)	Vavryn and Winkelbauer (2005)	Austria and Germany	Predictive	
Increase in braking performance observed of novice and experienced test riders from $5.7ms^{-2}$ to $7.7ms^{-2}$ for novice riders and $6.6ms^{-2}$ to $7.8ms^{-2}$ for experienced riders	Vavryn and Winkelbauer (2005)	Austria	Human factors study	47 novice riders and 134 experienced riders
ABS reduces risk of riders being thrown from the bike. May lead to a reduction in forward collision and off-road crashes.	Bayly *et al* (2006)	Australia	N/A	N/A
3% reduction in fatal and serious casualties	McCarthy and Chinn (1999)	UK	Retrospective	
The effectiveness of ABS is currently under investigation as part of the PISa project. However, the report contains a ranking of various safety systems including ABS, CBS and BAS. Each system was given a score based on the potential influence on the accident outcome, however was not expressed as a percentage of the target population. ABS was given a score of 2.39	McCarthy *et al* (2008), review of GB OTS/COST327 cases for PISa project	UK and Europe	Predictive case by case; subjective	60

casualties, a best estimate effectiveness of 4.2% was calculated for slight casualties. The upper and lower effectiveness values are calculated using the same method. Table 5 shows the estimated benefit of fitting ABS. The target population and effectiveness are shown in the table to allow readers to understand how the benefits have been derived. A best estimate of the effectiveness is shown in the table, accompanied by minimum and maximum effectiveness values.

Table 5.
Estimated benefit of fitting ABS to all PTW

Severity	Target population (All PTW casualties)	Effectiveness (%)	Estimated benefit
Fatal	585	18 (9-36)	105 (52-209)
Serious	5991	10* (3-17)	599 (180-1018)
Slight	17293	4 (0-7)	692 (0-1159)
Total	23870	6 (1-10)	1432 (239-2387)

** This is the mid-point of the range (rounded to nearest integer) and not a best estimate*

DISCUSSION

An increase in the number of safety measures and the rate at which they are coming to market can put an increased burden on the regulatory process. Impact assessments are, therefore, often required before there is sufficient voluntary market penetration to effectively measure the impact on the number and severity of road casualties using a retrospective statistical approach. Literature exists that describes the different types of research methods available (Elvik and Vaa, 2004) or to provide guidelines for assessing benefits (Burgette *et al,* 2008). However, within budgets and timescales available, it is often not possible to follow such guidance. TRL have seen an increase in requests for the assessment of the benefits associated with safety measures, based on existing literature, rather than new research. These are often required in short timeframes and on low budgets, thus limiting the depth of analysis that can be undertaken. This type of study has frequently identified widely varying and conflicting results amongst the existing literature meaning that if scientific rigour is applied, only wide ranges of potential benefits can be produced, which do little to resolve policy debate about the merits of proposals. It also allows stakeholders to select different values from within the quoted range, based upon broad assumptions that may or may not be accurate.

The project to develop a methodology to assess the benefits of advanced safety systems for PTWs provided an opportunity to begin to highlight these previous experiences and to consider the wider issues within a more formalised framework for undertaking benefit analyses. Although the application of the methodology was limited to reviewing existing literature and accident data, the methodology itself was developed to include all benefit assessment methods, to assist in identifying the limitations of existing estimates, and also to help identify knowledge gaps.

The methodology provides a framework, in which each method has its place, from a quick look at the casualty groups that can be affected, to full statistical retrospective analyses. It is intended that the methodology will allow policy makers to understand the limitations of the benefit estimates with which they are presented, and also what actions are required to develop the estimates to meet a higher burden of proof, if that is what they deem to be necessary.

The application of the methodology to the estimation of the potential benefits of fitting ABS to PTWs highlighted many of these issues. An estimate was possible but produced a large range of potential benefits because the quality of the estimate was severely limited by the ability to extract appropriate information from the existing literature. Some of the issues identified during the application of the methodology were:

- Variation in the presentation of the data within the studies. It was not always possible to relate the target population or effectiveness to an overall number of accidents/casualties so that they could be applied to the UK accident data.
- Not all assumptions were clearly stated and widely differing assumptions were clearly used in different studies.
- Conflicting results from low effectiveness/high cost to high effectiveness/low cost
- Insufficient detail on context and exposure. For example, papers where an effectiveness was stated for all casualties, but no data was presented about the severity distribution of all casualties so that different severity distributions in different countries could not be accounted for.

Many of the studies that were identified by Smith *et al* (2008) had used appropriate methods to assess benefits. However, there was insufficient information available to directly apply the findings to an alternative source of accident data, i.e. it was

not possible to trace the benefit estimate back to the original source data. Following the methodology described in this paper should lead to a consistent style in which benefit assessments are reported, which in turn will allow wider application of the results in different countries or under different regulatory options.

The methodology that has been developed is appropriate to meet the objectives of the specific research project for which it was intended. However, it could be considered just a starting point for a wider debate about how the scientific community and policy makers could work together to improve the quality, consistency and understanding of casualty benefit assessment. Ideally this would enable more effective implementation of the safety improvements that today's rapid development of advanced active safety systems make possible.

Future developments could include:

- Extending the methodology to include assessment based on regional representation, analytical quality and sample size;
- More detailed guidance on specific analytical techniques (e.g. highlighting known confounding factors that should be accounted for in retrospective statistical studies or the strengths and weaknesses of different ways of accounting for exposure)
- Development of new, improved data sources specifically designed to overcome limitations of existing data with respect to active safety systems
- How to encourage widespread use of a common methodology
- Methods to ensure that the use of a common approach does not compromise the flexibility needed to assess a wide variety of different systems
- A methodology for assessing costs.

CONCLUSIONS

- There are a range of methods that can be used to estimate the benefits of safety measures. None are perfect and each has strengths and weaknesses. However, to the reader, the limitations and assumptions are not always transparent. This can mean conflicting results, extended policy debate and slower implementation of technology.
- A generic methodology has been developed for a specific type of analysis that will assist both researchers and policy makers to identify the most appropriate methods to use and the limitations of each method without unduly limiting the range of analysis that could be undertaken.
- This methodology has the potential to be expanded to the full range of casualty benefit analyses, which if successfully implemented in a wide range of research projects, could substantially improve the overall quality and cost effectiveness of the research and regulatory processes of implementing new technologies.

REFERENCES

Bayly, Regan and Hosking. 2006. "Intelligent Transport Systems and Motorcycle Safety." Monash University Accident Research Centre.

Baum, Westerkamp and Geißler. 2007. "Cost-benefit analysis for ABS of motorcycles." Institute for Transport Economics at the University of Cologne.

Burgett, A, Srinivasan, G and Ranganathan, R. 2008. "A methodology for estimating potential safety benefits for pre-production driver assistance systems". Report number DOT HS 810 945, NHTSA, May 2008.

Elvik, R and Vaa, T. 2004. "The handbook of road safety measures". Elsevier 2004.

Gwehenberger, Schwaben, Sporner and Kubitzki. 2006. " Serious accidents involving motorbikes-analysis of accident structures and effectiveness of ABS." Allianz Zentrum für Technik GmbH, Ismaning.

Kebscull and Zellner. 2007. "Analysis of cost effectiveness of motorcycle antilock brake system using computer simulation." Dynamic Research on behalf of International Motorcycle Manufacturers Association, Geneva.

McCarthy, M.G and Chinn, B.P. 1998. "ABS and its effect on BMW motorcycle accidents." TRL Project Report PR/SE/470/98, TRL Limited, Crowthorne.

McCarthy, M.G. and Chinn B.P. 1999. "ABS and its effect on BMW motorcycle accidents; supplementary report." TRL Project Report PR/SE/470/98, TRL Limited, Crowthorne.

McCarthy, M.G, StClair, V and Halewood, C. 2008. "Powered Two Wheeler Integrated Safety (PISa)- Annual report 2007." Published Project Report PPR351, TRL Limited, Crowthorne.

Page, Foret-Bruno and Cuny. 2005 "Are expected and observed effectiveness of emergency brake assist in preventing road injury accidents consistent." Paper number 05-0268. The 19th

International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV) in Washington D.C, June 6-9, 2005.

Smith T L, Gibson T and McCarthy M. 2008. "Development of a methodology for the evaluation of safety systems for powered two-wheelers – final report." TRL Published Project Report PPR381, December 2008.

Smith T, Broughton J and Knight I. 2007. "Passenger, goods and agricultural vehicle safety – effectiveness of existing measures and ranking of future priorities in the UK." Paper number 07-0452. The 20th International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV) in Lyon, France, June 18-21, 2007.

Teoh. 2008. "Effectiveness of Antilock Braking Systems in Reducing Fatal Motorcycle Crashes." Insurance Institute for Highway Safety.

Vavryn and Winkelbauer. 2005. "Braking Performance of Experienced and Novice Motorcycle Riders- Results of Field Study." Austrian Safety Board, Austria.