5

How to evaluate the programme

Monitoring and evaluation of any programme or intervention is vital to determine whether it works, to help refine programme delivery and to provide evidence for continuing support of the programme. Evaluation will not only provide feedback on the effectiveness of a programme but will also help to determine whether the programme is appropriate for the target population, whether there are any problems with its implementation and support and whether there are any ongoing concerns that need to be resolved as the programme is implemented.

This module describes the process of developing and conducting an evaluation of a seat-belt programme. It is divided into three key sections:

5.1 Planning the evaluation. This important initial section describes how to define the aims of an evaluation, and details the various forms evaluation may take (process, impact and outcome).

5.2 Choosing the evaluation methods. Once the type of evaluation has been determined, there are different methods that can be applied to carry out an evaluation. This section describes the different study types possible, explaining the advantages and disadvantages of each type of method. It outlines the types of performance indicators that can be used to measure the success of a programme. This section also briefly describes how to conduct an economic evaluation, and provides guidance on calculating sample size.

5.3 Dissemination and feedback. This section describes how to feed the result of an evaluation back into the planning and implementation stages, and ways that the results of an evaluation can be shared with different interested parties.

5.1 Planning the evaluation

The process of designing and implementing a seat-belt or child restraint programme was covered in Module 4. Work carried out prior to implementation should ensure that the programme is clearly defined and that it is implemented in a consistent and standardized way. It is far easier to evaluate the impact of a complete, well-planned and executed programme than one that is implemented in an inconsistent way.

It is essential that the evaluation framework is developed and implemented alongside the proposed programme. Thus, this work would be carried out by the working group as they develop the action plan for the programme (see Module 3). For most forms of evaluation, baseline measures need to be collected **before** the intervention is put in place so that change in such measures over time may be gauged. It is therefore important that the evaluation framework is well established early on.

The type of evaluation to be conducted will depend on a number of factors. These include the aims of the evaluation itself and the objectives of the programme being evaluated. The type of methodology chosen may also depend on resource constraints.

5.1.1 Aims of evaluation

Establishing the aims of the evaluation will help to determine how best to carry out the evaluation. The evaluation may have one or more aims. For example, an evaluation of seat-belt legislation and an increased enforcement programme may primarily be aimed at determining whether seat-belt wearing rates have gone up as a result of the programme. Secondary aims may include determining whether the enforcement has increased, whether training of police is effective and whether the programme is acceptable to the stakeholders. The evaluation in this case needs to be multifaceted.

The breadth of an evaluation will always be limited by the resources available, and a well-designed simple evaluation can be as powerful as a more complex one.

If the objective is to increase the provision of seat-belts fitted in vehicles, stakeholders may wish to assess:

- the number and type of vehicles with seat-belts installed;
- the number and type of vehicle manufacturers fitting seat-belts as standard;
- the number and type of imported vehicles sold in the country with seat-belts installed;
- seat-belt sales for retrospective fitting;
- the number and type of vehicles retrospectively fitted at approved test centres;
- seat-belt standards and legislation (implemented and enforced).

If evaluation of a specific discipline is required, for example publicity, then questions must be asked that are specific to that element of the campaign. For instance, if a seat-belt cinema advertising campaign has been implemented aimed at vehicle users aged 17–24 years, it is important that evaluation of success for this element is not clouded by data regarding all vehicle occupants. Therefore stakeholders may wish to assess using the following data:

- baseline wearing rates for 17–24-year-olds;
- the number of 17–24-year-olds attending the cinema where advertising is shown;
- surveys of 17–24-year-olds' opinions and awareness prior to viewing advertising;
- surveys of 17–24-year-olds' opinions and awareness after viewing advertising;
- national and local wearing rates for 17–24-year-olds;
- crash and injury severity data for 17–24-year-olds.

5.1.2 Types of evaluation

Evaluation may take several forms, and one or more may be appropriate, depending on the aims of the specific programme to be evaluated (*1*).

Process evaluation

Rather than measuring change in outcomes, process evaluation examines whether the programme was carried out as planned. This involves creating a list of indicators that need to be measured, depending on the aims of the programme. The results will help to identify the strengths and weaknesses of the programme, and where improvements may be made.

For example, in a media campaign designed to increase voluntary use of seat-belts and child restraints, a process evaluation may ask these sorts of questions:

- Have the campaign products (posters, billboards, radio and television spots) been pre-tested?
- How often were the campaign advertisements run?
- How many people saw them?
- Was the target group being reached?
- Are high-quality child restraints available in local shops?

If the intervention involves enforcement of seat-belt legislation:

- Is there noticeable enforcement by police?
- Are the police supportive of the campaign?
- Is the penalty sufficient to change behaviour?
- Are people able to circumvent the process (for example, using bribery)?

Process evaluations are "formative"; that is, the enquiries carried out are designed to provide information to guide programme improvement (*2*). For example, it may be considered important to determine whether the television advertisements shown as part of a seat-belt programme are appropriate – do they adequately address the issue, do they give the right information about how to choose child restraints?

Impact assessment

An impact assessment will determine whether the advertisements have brought about a change. The impact, or programme effect, refers to whether the programme has made an impact on the target population (*2*). For example, if the seat-belt programme involved airing television advertisements on seat-belt use, the impact assessment might examine whether people who had seen the advertisements believed that there was a good chance that they would be fined by the police if they did not wear a seat-belt.

CASE STUDY: **Evaluation of rear seat-belt campaign, United Kingdom**

A new campaign to raise awareness of the need to wear seat-belts in the rear of cars was launched by the Department for Transport in July 1998. Research among the target audience had indicated that most were not aware that an unbelted rear seat passenger could kill the driver in the event of a crash. The campaign focused on this fact (*3*).

Monthly tracking research monitored the impact of the campaign in terms of awareness, understanding and attitudes to particular road safety behaviours. Observational, evidential and accident data were also collected to give a longer-term measure of campaign impact.

Independent survey work commissioned by the department before and after the campaign indicated significant attitudinal changes. For example, there was a substantial increase in the numbers recognizing that, in a crash, a rear seat passenger not wearing a seat-belt could injure or kill the driver or front seat passenger. Research among 2000 adults showed significant shifts in awareness of the seat-belt wearing message. Recognition that an unbelted rear seat passenger might kill the driver in the event of a crash rose from 33% to 53% among 16–24-year-olds.

The week-long campaign ran in July 1998 and was effective not only in raising awareness but also in affecting seat-belt wearing behaviour. It is almost certain that many drivers who took note of the advertising were asking passengers to belt up.

Observational research is carried out each April and October by the Transport Research Laboratory. In April 1998, three months before the campaign, the adult wearing rate in the rear was 48%. In October it was 54%. The advertising campaign has been run over occasional week-long bursts since then. The increasing understanding of what can happen has been accompanied by an improving trend in seat-belt wearing rates. In October 2006 the rate was 65%, following a new campaign aimed at young men (see also United Kingdom case study in section 4.6.7).

Outcome evaluation

An outcome evaluation measures outcomes to see if the programme was successful. Are more people now wearing seat-belts than before? Have deaths and serious injuries been reduced? Are more children restrained in child restraints suitable for their age and size? Measuring a change in outcomes is probably the most common form of evaluation as it provides information as to whether the programme or intervention has actually made a difference (*4*).

The case study on the China Seat-Belt Intervention Project describes the evaluation of the seat-belt campaign in Guangzhou, China. This evaluation included a process evaluation assessing the changes in police enforcement practice, and an outcome evaluation in terms of the effect on seat-belt wearing rates. It also estimated the cost–effectiveness of the campaign.

CASE STUDY: **Seat-Belt Intervention Project, China**

In China, injuries are the leading cause of death from age 1 to 44 years, accounting for approximately 750 000 deaths and 3.5 million hospitalizations each year. Much of the injury-related mortality and morbidity is due to road traffic injury, perhaps expectedly, given that motor vehicle production has tripled since the early 1990s. In the large urban centres such as Beijing, Shanghai and Guangzhou, where the burden of road traffic injuries is growing rapidly, there is an urgent need to implement proven road safety interventions. In the absence of widespread implementation of road safety interventions, the epidemic of road traffic injuries will worsen. The China Seat-Belt Intervention Project was an initiative proposed by the representative of WHO in China and developed by the George Institute for International Health. It involved cooperation between the Chinese Government, international organizations and industry. It was officially launched in Guangzhou in mid-2005. In this one province, almost 10 000 road traffic crashes took place during 2004, resulting in 1800 deaths and almost 12 000 people injuries.

In 1993, the Ministry of Public Security in China instituted a regulation requiring all vehicles less than 6 m in length to have seat-belts fitted for the front seats. Despite the availability of seat-belts in almost all cars (excluding trucks) in China and laws on the use of seat-belts, their use has not reached levels achieved in countries such as Australia, where seat-belt use is as high as 90%.

Aims

The main purpose was to implement a comprehensive intervention over 12 months focused on increasing seat-belt use in Guangzhou, with the following specific aims:

- to increase the rate of correct seat-belt use by 20%;
- to estimate the cost–effectiveness of the intervention;
- to build capacity in road traffic injury prevention in Guangzhou.

Methods

The evaluation of the intervention was undertaken by adopting a comparison group pre-test/post-test design. Two sites were selected, one that received the intervention (Guangzhou City) and another that acted as a comparison site (Nanning City). Prior to the implementation of the intervention, extensive baseline (pre-test) measures of the key outcomes were undertaken. These measures were repeated 12 months post-intervention (October 2006).

An economic evaluation was undertaken to determine the cost–effectiveness of the intervention. The reason for conducting an economic analysis was to ensure that injury prevention resources were being allocated to programmes that represented value for money.

Status and results

In order to explore the barriers to seat-belt use amongst drivers and front seat passengers, a number of focus groups were hosted in Guangzhou City. A survey into seat-belt availability and seat-belt quality was conducted in November 2004.

A successful launch of the intervention was held on 25 April 2005 at Times Square in Guangzhou City. Representatives from the Ministry of Public Security, the Ministry of Health and provincial government, as well as representatives from WHO and the project funder, BP China, attended the launch.

From 20 to 24 June 2005, 50 senior traffic police officers received skills-based training and were educated on strategies for enforcement by Mr Ray Shuey, the former Assistant Commissioner of Police in Australia. Training for all police officers was conducted across all the branches of Guangzhou Traffic Command and Control Centre (TCCC) between August and October 2005, with 1125 traffic police (82.6%) in Guangzhou trained in enhanced enforcement practices targeting the enforcement of seat-belt use.

Throughout the month of September 2005, the Division for Safety and Education of TCCC conducted a comprehensive educational programme for taxi companies in Guangzhou, aiming to raise the prevalence of seat-belt use among taxi drivers.

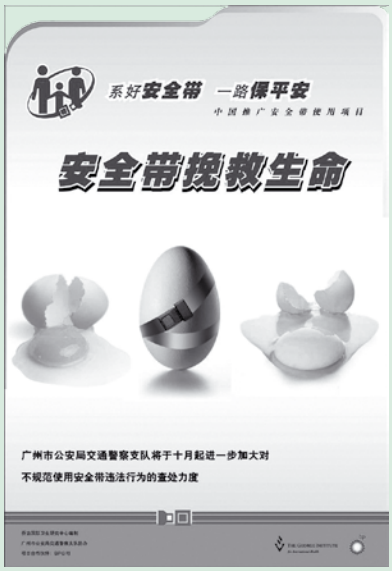
Five intensive law enforcement activities aimed at promoting seat-belt use, linked with a social marketing campaign, were conducted between October 2005 and August 2006. Each traffic police branch in Guangzhou established two checkpoints within its precincts. Target locations included locations

Continues...

Continued from previous page

with high crash risk, the train station, the airport and tollgates. Traffic police officers who had attended the train-the-trainer course were appointed to oversee the enforcement. Normal routine enforcement was still applied when specific campaigns were not in operation. During the study period, several approaches included in the training were implemented by TCCC to encourage traffic police officers to continue enforcing the use of seat-belts and penalizing infringers.

As described above, comprehensive social marketing campaigns linked with intensive enforcement were implemented. Related information was published on the traffic management web site; stickers and brochures were distributed among drivers; posters were displayed in a number of parking lots and taxi company premises; bus advertisements delivering the key message "Seat-belts save lives" were displayed on 15 buses along three bus routes that ran through the city centre; the intervention advertising was released in two major newspapers, featured on Guangzhou CATV six to seven times per day and broadcast on Guangzhou Radio traffic channel (FM 106.1) 18 times per day. The highly popular television commercial for the intervention, developed by collaborators and the Chinese Export Commodities Fair Advertising Co. Ltd, won a bronze prize at the 13th Guangdong Advertising Collection in 2006.



The post-intervention observations were collected in late 2006 and statistical analysis of results was completed in early 2007. The results showed that following the 12-month intervention period, the prevalence of seat-belt use increased significantly, from a prevalence of 50% (range 30–62%) to a prevalence of 62% (range 60–67%) in the intervention city. In contrast, the prevalence of seat-belt use declined, significantly, in the comparison city; an overall difference between the increased prevalence in the intervention city and decline in the comparison city of 18%. Importantly, the prevalence of seat-belt use increased significantly ($P < 0.01$) from pre- to post-intervention in the intervention city across all factors, namely seat-belt use by male or female drivers, both drivers and front seat passengers, and across road types and vehicle types. The greatest increase in the prevalence of seat-belt use was among drivers and passengers of taxis in the intervention city. The prevalence of seat-belt use among taxi drivers increased by 21% ($P < 0.01$) (from pre-test 30%, range 10–58%, to post-test 51%, range 45–60%). In contrast, the prevalence of seat-belt use significantly declined ($P < 0.01$) in the comparison city over the observational period across the same factors.

Based on the increased prevalence of seat-belt use observed as part of the intervention, the estimated total number of disability-adjusted life years (DALYs) saved as a result of the intervention was 530. Taking account of the cost of implementing the intervention (US$ 221 500), along with the associated cost savings, the incremental cost–effectiveness ratio of the intervention compared with no enhanced programme to increase the prevalence of seat-belt use was 3246 Chinese yuan per DALY saved (or the equivalent of US$ 418 per DALY saved).

The intervention provided the government with the opportunity to build capacity in road safety and, at the same time, provided information on the opportunity to target financial resources in order to reduce the burden of road traffic injury. Importantly, the study provided outcomes such as changes to police enforcement practice and comprehensive road safety communication strategies that are a legacy of the project and will be invaluable to both central and provincial-level governments who are responsible for road safety in China (5).

5.2 Choosing the evaluation methods

The methods used for each type of evaluation will vary. Both qualitative and quantitative methods can be used within the design of an evaluation. Qualitative methods may be employed for the formative or process evaluations, for example focus groups and use of short-answer or open-ended questionnaires (*6*).

Impact and outcome evaluations may be carried out using a variety of quantitative methods. Using an experimental or quasi-experimental design to demonstrate a change (or not) is the most powerful programme evaluation for detecting changes in outcome. The type of methods used will depend on the aim of and the budget for the evaluation.

5.2.1 Study types for formative and process evaluations

Formative evaluation is often in the form of qualitative research, which tends to involve detailed verbal descriptions of characteristics, cases and settings to explain reasons underlying various behavioural patterns. Specific techniques include using focus groups, in-depth interviews or surveys with short answers or open-ended questions (*7, 8*). For example, a question in a formative evaluation of a media campaign aimed at increasing seat-belt use may be whether the television advertisements address the question. Focus groups may be set up to determine whether the audience believes that the message from the television advertisements is appropriate. Feedback will further enhance the development of the advertisement.

CASE STUDY: **Evaluation of televised road safety messages, Ghana**

Researchers in Ghana evaluated the effectiveness of televised road safety messages on speeding and alcohol-impaired driving (9). Focus groups were conducted with 50 commercial drivers and addressed coverage, clarity and appropriateness of messages, including suggestions for improvements. The advertisements reached and were understood by most of the target audience, although some participants were unclear on the behaviour that the advertisements were urging viewers to take. Opportunities for strengthening the messages included using other media, increasing the number of languages used and stressing the change in behaviour being recommended.

5.2.2 Study types for impact and outcome evaluations

There is a well-defined hierarchy of study designs for examining the effectiveness of interventions. These range from randomized controlled trials, which provide a high level of evidence, to uncontrolled before–after studies, which provide very weak evidence about the effectiveness of an intervention.

Randomized controlled trial

The gold standard of evaluation, the randomized controlled trial (RCT), will provide the highest-quality level of evidence of whether an intervention or programme is successful. For the RCT, individuals (or groups of individuals, for example a school or village, in a variant of the RCT known as a cluster randomized trial) are randomly allocated to either receive, or not receive, the programme. As participants (or groups of participants) are randomly assigned to one group or another, other factors that may influence the outcome – measured and unmeasured – are more likely to be balanced between the intervention and non-intervention groups. However, although RCT designs should always be considered when evaluating the effectiveness of an intervention, they do require significant resources and may be difficult to conduct with a limited budget. There may also be ethical considerations in randomizing an intervention with known benefits (that is, in denying an effective intervention to those participants who will be in the non-intervention group).

It is important to note that there is no need to conduct an RCT on the effectiveness of seat-belts themselves as part of the seat-belt programme. There is sufficient evidence from a number of studies that clearly demonstrates that seat-belts are effective at reducing the injuries and fatalities that result during motor vehicle crashes (see Module 1).

NOTE

Examples of randomized controlled trials

Goodson, Buller & Goodson (*10*) carried out an RCT of prenatal safety education in two hospitals in the United States. Reported use of child safety seats was 96% for the intervention group and 78% for the control group.

Stuy (*11*) studied children aged 2–6 years and parents at childcare centres offered educational activities. In the intervention group children observed wearing seat-belts rose from 54% before to 75% after. For the control group the rise in seat-belt wearing rates was from 20% to 30%.

Quasi-experimental designs

Quasi-experimental study designs, while not as rigorous as randomized trials, if well conducted, may also be used to establish the effectiveness of an intervention. Such designs may also involve a "treatment" and a "control" group, but individuals or communities are not randomized into these groups.

Controlled before–after study

The controlled before–after study is often the most practical design for programme evaluation. Randomization is not always feasible, for example where some areas have already adopted an intervention. The controlled before–after study design involves observing the outcome of interest (for example seat-belt use rates) before and after the programme, in both the people who receive the programme and those in a control group. The control group should be as similar as possible to the programme group and any important differences between the groups need to be taken into account. Having a control group means that trends that may have been occurring in the population aside from what was happening due to the programme are taken into account (Box 5.1).

BOX 5.1: **Evaluation using a controlled before–after study**

A school-based intervention was developed in 1999 to increase seat-belt use among families living along the Texas-Mexico border, where child use of seat-belts had been found to be low. The Socorro Seatbelt Program sought to increase seat-belt use by changing perceived norms within the community, i.e. making the non-use of seat-belts less socially acceptable. Children in primary schools were shown reasons why family members should wear seat-belts, and reasons why they might not do so. Related activities included poster competitions and role model stories, and seat-belt newsletters were distributed to families. Parents were invited to join their children in a second poster and slogan competition.

Two adjacent communities served as the intervention and control sites. There were five schools in the intervention community and six in the control community. The two communities were similar in terms of demographic characteristics.

Pre-intervention seat-belt use was observed at each primary school during a 13-day period. Post-intervention assessments were conducted exactly one year later. In the intervention community, seat-belt use among children increased from 47% to 57%, whilst in the control community seat-belt use among children showed a non-significant decline from 50% to 47%. Whilst there was no significant change in drivers' seat-belt wearing arriving at the intervention schools, at the control schools drivers' seat-belt wearing declined significantly from 90% to 83% (*12*).

Before–after study (no control group)

The before–after study without a control group is often used to evaluate the impact of a programme, but provides the weakest evidence for the effectiveness of a programme. This design involves measuring the outcome of interest before and after the programme has been run. This study design is simple and may be conducted relatively cheaply, as all that is needed is a sampling frame and research assistants to conduct observations at various sites. However, without a control group the scientific merit of these study types is relatively limited, as it is often difficult to attribute with any certainty the change in outcome to the introduction of the programme.

CASE STUDY: **Before–after study (no control group) to evaluate a seat-belt campaign, Queensland, Australia**

In 1972 it became compulsory for all vehicle occupants to wear seat-belts in Queensland, Australia. Queensland Transport implemented public education campaigns over several years, and in 1997 the Fatal 4 campaign was developed in conjunction with the Queensland Police Service. Research conducted by Queensland Transport in August/September 2000 showed that 10% of people admitted to driving without a seat-belt. The most frequently cited reason was "being on a short trip". The campaign aimed to dispel popular myths:

- Seat-belts are only necessary on long journeys.
- Seat-belts cause injuries.
- It is not necessary to wear a seat-belt in the rear.

The campaign during 1998 to 2000 aimed to encourage motorists and passengers to buckle up by alerting them that police will enforce the wearing of seat-belts and reminding them of the penalties of non-compliance. The campaign included television and radio advertising, billboard and taxi messages, and a brochure aimed at visitors to the Olympics in 2000. Research showed that support for the statement "Since the ad I have made more of an effort to wear my seat-belt" increased significantly from 46% in April 1999 to 58% in August 2000. In 2001, research showed that only 5% of motorists admitted to driving without wearing a seat-belt. Of respondents, 64% claimed that they had seen some or all of the billboard messages, and 50% of motorists believed that the billboard messages strongly encouraged them to belt up (*13*).

Interrupted time series design

It is possible to assess the effect of a programme by using multiple measures of the outcome of interest before and after the programme. There are a number of different variations on this design, some involving control groups. Studies that have used these designs generally use routinely collected measures, such as death rates, as multiple measures are required for appropriate analysis. This study design is, however, subject to time-related challenges to its validity, given the possibility that other factors occurring simultaneously to the programme actually led to the observed effect. Statistical analysis of such data can take into account any such secular trends, meaning that it is possible to establish whether the intervention or programme was responsible for the change in outcome.

Table 5.1 summarizes the characteristics and advantages and disadvantages of the study types discussed in this section. Further detail about study types is available in references *14* and *15*. There is also a useful online glossary of research terms at http://www.cochrane.org/resources/glossary.htm.

Table 5.1 Study types and their advantages and disadvantages

	Formative, process evaluation?	Impact, outcome evaluation?	Advantages, disadvantages
Qualitative			
Focus groups/ in-depth interviews	• Formative • Process	• Outcome	• Can provide information on why intervention may or may not have worked • Cheap • Sample (participants) is not random • Results are not generalizable
Quantitative			
Randomized controlled trial		• Impact • Outcome	• Most rigorous evidence • Expensive • Randomization not always feasible
Controlled before–after study		• Impact • Outcome	• Most practical design • Must have comparable control group
Before–after study (no control group)		• Impact • Outcome	• Cheap • Low level of evidence
Interrupted time series design		• Impact • Outcome	• Practical design if sufficient numbers of events and accurate surveillance systems in place

BOX 5.2: **Effectiveness of programmes promoting use of child restraints**

A systematic review (*16*) identified 16 studies examining the effects of educational campaigns aimed at increasing the use of safety seats and seat-belts for children. A variety of approaches were used, with all studies including an element of professional counselling to encourage use of restraints (seat-belts etc.), supported by a range of media. Some programmes included rewards or coercion such as fines to encourage compliance. There is evidence that the campaigns led to an increase in observed restraint use, although this may not be sustained in the long term. This conclusion is supported by Grossman and Garcia (*17*), who report that the magnitude of positive effects from such programmes diminishes substantially one or more months after the intervention. More intensive programmes appear to produce more positive results. There is some indication that programmes may be less effective for older children.

Two further reviews (*18, 19*) conclude that education alone has a modest effect in increasing use of child restraints. This can be considerably strengthened among younger children where appropriate restraints are provided through a loan, low-cost rental or giveaway programme, alongside education and guidance in their proper use. Incentive and education programmes that reward parents for correctly using child restraints or reward children for using them have been shown to be effective in increasing child restraint use in the short term (1–4 months) (*19*).

5.2.3 Choosing the performance indicators

Performance indicators (or outcome measures) are a measure of how successful the programme has been. Choice of performance indicators will be determined by the aims of the evaluation, the study type used, the resources available and, to a certain extent, the requirements of the funding agency. For instance, government funding agencies may require certain information to ensure support for increased enforcement or for further roll-out of a programme.

Injury and death outcomes

The effectiveness of seat-belts in reducing crash-related injury and death has been well documented in many studies (see Module 1) and there may be no need to replicate these findings in a large-scale (and possibly expensive) piece of experimental research. Although much of this effectiveness research has been conducted in high-income countries (predominantly the United States) there is no reason to believe that seat-belts would be less effective in preventing death and injury in low-income settings.

If it is necessary to calculate death and injury rates it may be possible to use routinely collected data as the basis for these calculations. However, the efficiency with which such rates can be calculated depends on the accuracy of local surveillance. If there is a uniform capture, coding and reporting system already set up in hospitals or health departments there may be aggregated data available on crash-related injury. Otherwise this may need to be abstracted from local data sources. Similarly, motor vehicle crash and death data may be routinely collected from police or transport authorities.

As quality may be variable, completeness and accuracy of these data sources should be carefully checked before use.

Seat-belt wearing rates

Possibly the most useful performance indicator to use in the evaluation of a seat-belt wearing campaign is the proportion of drivers and passengers wearing seat-belts. Observations of vehicle occupants may be made at a number of sites before and after a programme to document whether seat-belt wearing rates have changed.

Calculating rates

Comparing changes in absolute numbers in injury and death outcomes, or in drivers and passengers wearing seat-belts, before and after a programme is not useful, as absolute numbers may change due to an increase or decrease in the numbers of vehicles, registered or otherwise. It is therefore important that rates be calculated. Denominators may include number of vehicles, registered vehicles or kilometres travelled. For example, for injury outcomes a rate may be number of injuries per

licensed drivers, or number of injuries per 100 000 km driven. For seat-belt use, the appropriate rate would be the proportion of belted occupants over total occupants observed. Note that it is preferable to use a population denominator (for example per 100 000 population) rather than the number of vehicles as a denominator. This is because the rapidly increasing use of motor vehicles in many countries may distort the results of an evaluation, if this latter measure is used.

Module 2 includes a detailed section on how to measure seat-belt wearing rates.

5.2.4 Conducting an economic evaluation of a programme

It may also be necessary to conduct an economic evaluation to demonstrate value for money and possible cost savings for government by investing in prevention. Economic evaluation addresses the question of whether one intervention represents a better use of resources than another. In other words, does spending $X on programme A represent a better investment than $X on programme B? To address this question, it is apparent that a comparison of two or more options is needed (sometimes this comparison is with a "do nothing" or "status quo" alternative).

Economic evaluation is based on the comparison of alternatives in terms of their costs and consequences *(20)*. The term "consequences" is used here to represent an outcome of value. There are various forms of economic evaluation that can be conducted, each differing in terms of scope, i.e. the range of variables included in the analysis. Importantly, each form of economic evaluation typically entails a set of starting assumptions; recognition of these is necessary for the policy-maker to make appropriate use of the evidence from such studies.

A common element across all forms of economic evaluation is that they involve measuring costs. Costs usually comprise, at least in part, the direct programme costs – the resources that are used to run the programme (for example, equipment, staff, consumables). However, in principle, other costs may also be relevant, such as those incurred by patients, carers and the wider community. Furthermore, there are downstream costs and cost savings that may enter into consideration; for example, a programme may result in reduced hospitalizations and these savings in resources may be deemed relevant. The type of costs selected generally depends on the perspective taken in the evaluation and the nature of the resource allocation problem being addressed.

Methods used in economic evaluation

The most common form of economic evaluation is **cost–effectiveness analysis** (CEA). This entails placing the total cost of programmes alongside a defined outcome to produce a cost–effectiveness ratio (for example, cost per life saved, cost per life year saved or cost per case prevented). The assumption in CEA is that the objectives of interventions being compared are adequately captured in the measure of

outcome used *(21)*. One modification to conventional cost–effectiveness analysis is cost–utility analysis, which is based on an outcome measure, the quality-adjusted life year (QALY). This incorporates changes in survival and quality of life and thereby enables a wider set of interventions to be legitimately compared than would be possible with CEA. The case study from China in section 5.1.2 shows the results of a cost–effectiveness study.

Another form of economic evaluation is **cost–benefit analysis**, which seeks to evaluate interventions in terms of total costs and total benefits, both dimensions being valued in monetary terms (for example dollars). Therefore if benefits are greater than costs, the decision would be to fund the programme. Valuation of health benefits in this way can be challenging, but one approach would be to elicit from beneficiaries of programmes their maximum willingness to pay for these benefits (i.e. if they had to pay for it in a hypothetical market place). The idea behind this approach is to derive a valuation for an intervention akin to the way in which consumers value goods and services in markets (*22*).

Cost–benefit analyses of seat-belt wearing show that the benefits clearly outweigh the costs, with benefit–cost ratios of between 3 and 8 (*23*).

Choosing the appropriate type of economic analysis for the needs of the particular programme will depend on resources available (both economic and human resources) and the aims of the evaluation.

5.2.5 Determining sample size

For all quantitative study types it is important to have sufficiently large numbers in the study to be sure that if an effect exists it is detectable. The rarer the event, the greater the sample size needs to be in order to detect a difference. Serious injuries from motor vehicle crashes are relatively rare events and a study using serious injury or death as an outcome would involve a large sample size. Measuring seat-belt wearing rates requires a smaller number of participants.

Factors that must be taken into consideration in determining the sample size are the expected size of the effect to be detected, variability in the measures and the prevalence of the variable of interest. For a cluster randomized trial, sample size calculations will also take the size of the cluster and correlation within clusters into account. For further information on sample size calculations for cluster randomized trials see reference *24*.

Sample size calculators are freely available on the Internet, but it is wise to consult a statistician regarding such estimates, particularly where cluster randomized trials or random or stratified samples are necessary.

For quantitative study designs data will require statistical analysis. For more advice on how to go about this refer to reference *7*, or see the relevant lectures in the basic methods and injury sections at http://www.pitt.edu/~super1.

5.3 Dissemination and feedback

Once an evaluation is complete it is important to provide feedback to the stakeholders involved in the programme. Dissemination of the results will help to garner further support for the programme if it is successful, and help others gain support for the introduction of similar programmes. Publicity from dissemination activities may also increase the impact of the programme. If the programme has not been successful it is important to share this with others so that weaknesses or relevant issues are considered in other similar interventions, including whether or not to introduce such interventions.

Dissemination may involve presenting the results at public meetings, using the media to publicize the outcomes of the programme or publishing reports and papers in the scientific literature. Dissemination, translation and diffusion activities are often planned in advance so as to increase the chances for nationwide adoption of the effective interventions (*25*).

5.3.1 Checklist for evaluation process

- Start evaluation process at the beginning of programme implementation.
- Determine aim of evaluation and develop evaluation framework.
- Clearly define target population, place and time.
- Develop and test instruments for data collection, ensuring consistency in training and measurement.
- Collect and analyse data.
- Write and disseminate evaluation report, feeding back into various aspects of programme.

5.3.2 Using evaluation results to feed back into new planning cycle

Consider whether the evaluation demonstrated any tangible benefits – should the programme be continued, or does it require disbanding or modification? Can the existing programme be improved on the basis of the evaluation? Have there been any unexpected side effects of the programme?

The results of the evaluation should be fed back into the planning cycle and the appropriate modifications to the programme made before it is further expanded.

Summary

Evaluation should be seen as an integral component of any seat-belt programme. An evaluation needs to be determined at the beginning of a programme development, such that the plan for data collection for this purpose is built into project implementation. As well as providing information on the effectiveness of a programme, evaluation will help identify if there are any problems in running a programme.

Determining the aims of the evaluation will help to decide how best to carry out the evaluation. There are a number of different methods that can be used to evaluate a seat-belt programme. Each method has various advantages and disadvantages, and the choice of which to use will depend on the aims of the programme and the resources available.

It is important that the results of the evaluation are shared with the appropriate parties, and that they are used in the planning of the programme. Dissemination, translation and diffusion of effective interventions will help increase the chances for nationwide adoption.

References

1. Thompson NJ, McClintock HO. *Demonstrating your program's worth: a primer on evaluation for programs to prevent unintentional injury*. Atlanta, Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, 1998.
2. Rossi PH et al. *Evaluation: a systematic approach*. California, Sage Publications, 2004.
3. THINK! Always wear a seat belt: rear seat belt campaign. Think! Road Safety Campaign, United Kingdom Department for Transport, 2003 (http://www.thinkroadsafety.gov.uk/campaigns/seatbelts/rear.htm, accessed 12 November 2008).
4. Zaza S et al. Methods for conducting systematic reviews of the evidence of effectiveness and economic efficiency of interventions to reduce injuries to motor vehicle occupants. *American Journal of Preventive Medicine*, 2001, 21(4S):23–30.
5. Stevenson M et al. Reducing the burden of road traffic injury: translating high-income country interventions to middle- and low-income countries. *Injury Prevention*, in press 2008.
6. Thompson NJ. Study methods for understanding injury behavior. In: Gielen A, Sleet DA, DiClemente R, eds. *Injury and violence prevention: behavioral science theories, methods and applications*. San Francisco, Jossey-Bass, 2006.
7. Britten N. Qualitative research: qualitative interviews in medical research. *British Medical Journal,* 1995, 311:251–253.
8. Kitzinger J. Qualitative research: introducing focus groups. *British Medical Journal,* 1995, 311:299–302.

9. Blantari J et al. An evaluation of the effectiveness of televised road safety messages in Ghana. *International Journal of Injury Control and Safety Promotion*, 2005, 12:23–29.
10. Goodson J, Buller C, Goodson W. Prenatal child safety education. *Obstetrics and Gynaecology*, 1985, 65:312–315.
11. Dowswell T, Towner E. Social deprivation and the prevention of unintentional injury in childhood: a systematic review. *Health Education Research*, 2002, 17:221–237.
12. Cohn LD et al. A program to increase seat belt use along the Texas-Mexico border. *American Journal of Public Health*, 2002, 12:1918–1920.
13. *Seat belts public education campaign.* Queensland Transport, 2004 (http://www.roadsafety.qld.gov.au/qt/LTASinfo.nsf/ReferenceLookup/SeatbeltsPubEd.pdf/$file/SeatbeltsPubEd.pdf, accessed 12 November 2008).
14. Woodward M. *Epidemiology: study design and data analysis,* 2nd ed. Boca Raton, Chapman and Hall CRC, 2005.
15. Rothman KJ, Greenland S. *Modern epidemiology,* 2nd ed. Hagerstown, Maryland, Lippincott-Raven, 1998.
16. Towner E et al. *What works in preventing unintentional injuries in children and adolescents? An updated systematic review.* London, Health Development Agency, 2001.
17. Grossman G, Garcia D. Effectiveness of health promotion programs to increase motor vehicle occupant restraint use among young children. *American Journal of Preventive Medicine*, 1999, 16(IS):12–22.
18. Di Guiseppi C, Roberts I. Individual-level injury prevention strategies in the clinical setting. *Future Child*, 2000, 10(1):53–82.
19. Dinh-Zarr T et al. Reviews of evidence regarding interventions to increase the use of safety belts. *American Journal of Preventive Medicine*, 2001, 21(4S):48–65.
20. Drummond MF et al. *Methods for the economic evaluation of health care programmes.* Oxford, Oxford University Press, 1997.
21. Tan-Torres T et al, eds. *WHO guide to cost-effectiveness analysis.* Geneva, World Health Organization, 2003.
22. Miller TR. The effectiveness review trials of Hercules and some economic estimates for the stables. *American Journal of Preventive Medicine*, 2001, 21(4S):9–12.
23. Elvik R, Vaa T, eds. *The handbook of road safety measures.* Elsevier, 2004.
24. Kerry SM, Bland JM. Statistics notes: sample size in cluster randomisation. *British Medical Journal,* 1998, 316:549.
25. Sogolow E, Sleet DA, Saul J. Dissemination, implementation and widespread use of injury prevention interventions. In: Doll LS et al., eds. *Handbook of injury and violence prevention*. New York, Springer, 2007.