

4

How to evaluate the programme

How to evaluate the programme

4.1 Planning the evaluation	121
4.1.1 Aims of evaluation	122
4.1.2 Types of evaluation	122
4.2 Choosing the evaluation methods	123
4.2.1 Study types for formative and process evaluations	124
4.2.2 Study types for impact and outcome evaluations	124
4.2.3 Choosing the performance indicators	128
4.2.4 Conducting an economic evaluation of a programme. .	129
4.2.5 Determining sample size	130
4.3 Dissemination and feedback	132
Summary	135
References	136

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 drinking and driving programme. It is divided into three key sections:

- **4.1 Planning the evaluation:** This important initial stage involves collecting data, in a baseline study, to assess the situation before going on to develop the programme. Based on the information collected, it is then necessary to define the aims of an evaluation, and to consider the different types of evaluation methods that could be used for your evaluation.
- **4.2 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.
- **4.3 Dissemination and feedback:** This section describes how to feed the result of an evaluation back into the planning and implementation stages, as well as ways that the results of an evaluation can be shared with different interested parties.

4.1 Planning the evaluation

The process of designing and implementing a drinking and driving programme was covered in Module 3. 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). Baseline measures need to be collected before the intervention is put in place so that change in such measures over time may be gauged.

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

4.1.1 Aims of evaluation

The aims of the programme will determine how best to carry out the evaluation. The evaluation will focus on assessing the extent to which the programme objectives have been met, and may have one or more aims. For example, an evaluation of a drinking and driving legislation and increased enforcement programme may primarily be aimed at determining whether drinking and driving rates have reduced as a result of the programme. However, 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, but note that a well-designed, simple evaluation can be as powerful as a more complex and costly one.

4.1.2 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.

Process evaluation

Rather than measuring change in outcomes, this aspect of 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 reduce drinking and driving, a process evaluation may ask these sorts of questions:

- Have the campaign products (posters, billboard, 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?
- If the intervention involves enforcement of drinking and driving 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 what are known as “formative”. That is, the enquiries carried out are designed to provide information to guide programme improvement (1). For example, it may be considered important to determine whether the TV adverts shown as part of a drinking and driving programme are appropriate – do they adequately address the issue, does the behaviour portrayed reflect drinking customs in the region where the adverts will be seen?

Impact assessment

This will determine whether the programme has brought about a change. The impact, or programme effect, refers to a change in the target population that has been brought about by the programme – that is, a change that would not have occurred if the programme had not happened (1). For example, if the programme involved airing television advertisements on drinking and driving, the impact assessment might examine whether people who had seen the advertisements believe that there is a good chance that they will be stopped and breathalysed by the police if they drink and drive. Unlike a process evaluation, this would tend to take place at the end of a programme, as the focus would be on the outcome.

Outcome evaluation

This is where the outcomes are measured to see if the programme was successful. Are less people now drinking and driving than before? Have road crashes involving alcohol been reduced? Are fewer injured drivers/riders admitted to hospital with high BAC levels? 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.

The US NHTSA (2) has developed a guide to evaluating road safety projects. This provides a high-level overview of the steps involved in the evaluation of road safety programmes, from defining the problem to reporting results.

4.2 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 formative and process evaluations, e.g. focus groups, short-answer or open-ended questionnaires.

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 methods used will depend on the aim and the budget for the evaluation.

4.2.1 Study types for formative and process evaluations

Qualitative studies

Qualitative research 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 (3, 4). For example, a question in a formative evaluation of a media campaign aimed at reducing drinking and driving 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.

Researchers in Ghana evaluated the effectiveness of televised road safety messages on speeding and drink-driving (5). 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 about the behaviour the advertisements were telling viewers to take. Opportunities for strengthening the messages included using other media, increasing the number of languages and stressing the change in behaviour being recommended.

4.2.2 Study types for impact and outcome evaluations

There is a well-defined hierarchy of study designs for examining the effectiveness of interventions (see Table 4.1). These range from randomised control trials, which provide a high level of evidence, to uncontrolled before–after studies which provide weak evidence about the effectiveness of an intervention.

Table 4.1 Study types and their advantages and disadvantages

	Formative and process evaluation	Impact and outcome evaluation	Pros and cons
QUALITATIVE			
Focus groups/ in-depth interviews	formative – process	– outcome	<ul style="list-style-type: none"> – Can provide information on why intervention may or may not have worked – Cheap – Sample (participants) are not random sample – Results cannot be generalised
QUANTITATIVE			
Randomised controlled trials		<ul style="list-style-type: none"> – impact – outcome 	<ul style="list-style-type: none"> – Most rigorous evidence – Expensive – Randomisation not always feasible
Controlled before–after study		<ul style="list-style-type: none"> – impact – outcome 	<ul style="list-style-type: none"> – Most practical design – Must have comparable control group
Interrupted time series design		<ul style="list-style-type: none"> – impact – outcome 	<ul style="list-style-type: none"> – Practical design if sufficient numbers of events and accurate surveillance systems in place
Before–after study (no control group)		<ul style="list-style-type: none"> – impact – outcome 	<ul style="list-style-type: none"> – Cheap – Low level of evidence

Further detail about study types is available in references 6 and 7. There is also a useful online glossary of epidemiological terms at www.cochrane.org/resources/glossary.htm

Randomised control trial (RCT)

The gold standard of evaluation, the randomised control trial will provide the highest quality level of evidence that an intervention or programme is successful. A RCT design means that individuals or groups of individuals (e.g. a school, or village, known as a cluster randomised trial) are randomly allocated either to 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 group. However, although RCT designs should always be considered when evaluating effectiveness of an intervention, they require significant resources and may be difficult to conduct with a limited budget. It may also be difficult to identify a non-intervention group when some aspects – for example changes in the national law, or national television broadcasts – apply to all regions. There may also be ethical considerations in randomising 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 a randomised controlled trial on the effectiveness of reducing drinking and driving as part of the programme. There is overwhelming evidence from many studies that clearly demonstrates that reducing the number of drunk drivers is effective in reducing crashes and the injuries and fatalities that result (see Module 1).

Quasi-experimental designs

These study designs, while not as rigorous as randomised trials, if well conducted may also be used to establish the effectiveness of an intervention. That is, using the information collected on trends of the indicators measured, these studies allow conclusions to be drawn as to whether or not the intervention (the programme) is associated with change in the outcome.

Controlled before-after study

This is often the most practical design for programme evaluation. Randomisation is not always feasible, for example where some areas have already adopted an intervention or some of the changes apply nationwide. The controlled before–after study design involves observing the outcome of interest (e.g. drinking and driving 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 as a result of the programme are taken into account.

BOX 4.1: **Controlled before-after study, United States**

“A controlled before-after study was undertaken in the United States to determine whether reductions in alcohol-related fatal crashes following adoption of 0.08% legal BAC limits were independent of general regional trends. The first five states that lowered legal BAC limits to 0.08% were paired with five nearby states that retained at 0.10% legal standard.

The study found that states adopting 0.08% laws experienced 16% and 18% declines in the proportion of fatal crashes involving fatally injured drivers whose BAC levels were 0.08% or higher and 0.15% or higher. The study concluded that if all states adopted 0.08% legal BAC limits, at least 500–600 fewer fatal crashes would occur annually.”

Source: (8)

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 and the possibility that other factors occurring simultaneously to the programme actually led to the observed effect. However, 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.

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 weak 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 a research team 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.

BOX 4.2: Compulsory breath test programme, New Zealand

Compulsory breath testing (CBT) was introduced in New Zealand on 1 April 1993, following nine years of random breath testing (RBT). Passive alcohol detectors were used for the initial screening. At the same time, the legal blood alcohol limit was lowered from 0.08 for drivers under 20 years of age to 0.03. CBT checkpoints are essentially car-based operations at which all drivers stopped are supposed to be tested. Paid publicity was largely limited to the first three months.

Assessing the effects of the CBT program in New Zealand was difficult because of the lack of a suitable control or comparison group and because, preceding the introduction of CBT, there were large generalised and local changes, such as:

- economic conditions;
- the gradual reduction in the amount of alcohol drunk per person;
- the removal of the number and time-of-day conditions on liquor licences;
- the amalgamation of Ministry of Transport traffic officers with the New Zealand Police;
- the introduction of speed cameras.

It was not possible to conclude that CBT had a positive effect on reducing crashes involving drinking and driving (over RBT). Further work will be required to determine this.

More information: www.druglibrary.org/schaffer/Misc/driving/s29p3.htm

4.2.3 Choosing the performance indicators

Performance indicators (or outcome measures) are a measure of how successful the programme has been. They should relate directly to the objectives of the programme. 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 impact of alcohol on the human body and its consequent effects in increasing crash risk have been robustly established (see Module 1). There is no need to replicate these findings in a large scale and expensive piece of experimental research. What is less well known in many countries is the extent of drinking and driving, and whether there is a high proportion of victims from crashes where alcohol was a factor.

It is possible to use routinely collected BAC data to establish injury and death rates. However, the efficiency with which such rates can be calculated depends on whether BAC levels are collected systematically for road crash victims, and the accuracy of local surveillance. If there is a uniform capture, coding and reporting system already set up in hospitals and/or health departments there may be aggregated data available on the proportion of crash victims where alcohol is involved. Similarly, crash and/or death data may be routinely collected from police or transport authorities – some of whom may record alcohol presence as a factor.

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

Drinking and driving rates

An appropriate performance indicator is the proportion of riders and drivers who have consumed alcohol, and their BAC levels. To obtain this requires police commitment to stop and breathalyse riders/drivers on a systematic basis. It is also desirable that all riders/drivers involved in crashes be breathalysed as a standard procedure. The trends in observed BAC levels from such tests offer a good basis for tracking changes in the extent to which alcohol is a factor in road crashes.

Calculating rates

Comparing changes in absolute numbers in injury and death outcomes, or in riders/drivers at a certain BAC level, before and after a programme is of limited value, as absolute numbers may change because of an increase or decrease in the numbers of riders and drivers, registered or otherwise, and the numbers of breath or blood tests carried out. It is therefore important that rates be calculated. Denominators may

include number of drivers, registered vehicles, or kilometres travelled. For example, for injury outcomes, a rate may be the number of drink-driving injuries per licensed vehicle or licensed driver/rider, or number of drink-driving injuries per 100 000 km travelled. For alcohol intoxication rates, the appropriate rate would be the proportion of drunk drivers/riders over the total number for which alcohol was measured.

4.2.4 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 \$y on programme B? To address this sort of question, it is apparent therefore that a comparison of two or more options is needed (often 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 (*o*). 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 (e.g. 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 e.g. a programme may result in reduced hospitalisations 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 the total cost of programmes alongside a defined outcome to produce a “cost-effectiveness ratio” (e.g. 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 (*io*). One modification to conventional cost effectiveness analysis is cost-utility analysis which is based on an outcome measure, Quality Adjusted Life Year (QALY), that incorporates change in survival and quality of life and thereby enables a wider set of interventions to be legitimately compared than would be possible with CEA.

Another form of economic evaluation, often used to evaluate transport sector investment, is **cost-benefit analysis** (CBA) which seeks to evaluate interventions in terms of total costs and total benefits – both dimensions being valued in monetary terms (e.g. 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.

Choosing the appropriate type of economic analysis for the needs of the particular programme will depend on resources available (both economic and human), and the aims of the evaluation. Taking quality of life into account is a powerful measure for evaluations of road crashes where lifelong disability resulting from serious injury may be an outcome.

4.2.5 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 road crashes are relatively rare events and a study using serious injury or death as an outcome would involve a large sample size. Measuring drinking and driving 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 randomised 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 randomised trials see reference 11.

Sample size calculators are freely available on the internet¹, but it is wise to consult a statistician regarding such estimates, particularly where cluster randomised trials or random and/or stratified samples are necessary.

Statistical analysis

For quantitative study designs data will require statistical analysis. For more advice on how to go about this refer to reference 6, or see the relevant lectures in the basic methods and injury sections at www.pitt.edu/~superi.

1 The statistical package Epi Info™ may be downloaded at www.cdc.gov/epiinfo
A sample size calculator for cluster randomised trials can be found at www.abdn.ac.uk/hsru/epp/cluster.shtml

BOX 4.3: **Estimated savings from measures to reduce drinking and driving in the United States**

A study was undertaken by the *Pacific Institute for Research and Evaluation (PIRE)* on behalf of NHTSA to evaluate the economic costs and benefits of impaired driving prevention measures in the United States with the following results:

Administrative License Revocation: State laws that allow police or driver licensing authorities to revoke a driver's license swiftly and automatically for refusing or failing a BAC test have reduced alcohol-related fatalities in the US by 6.5% on average and saved an estimated \$54,000 per driver sanctioned. Fees paid by offenders to get their driving licenses back typically cover start-up and operating costs.

Zero Tolerance Law: Nationwide laws making it illegal for persons under 21 to drive with a positive BAC have reduced impaired driving fatalities by 4% on average. Per licensed youth driver, these laws cost approximately \$30 and yield net savings of \$700. Medical care cost savings alone exceed the intervention cost. The primary cost is the value of mobility lost by youth who are forced to reduce their drinking or driving.

0.08 BAC Law: Well-publicized laws lowering driver BAC limits to 0.08 have reduced alcohol-related fatalities by an average of 7% in 32 States, the District of Columbia and Puerto Rico. On average, 0.08 laws save an estimated \$40 per licensed driver nationwide.

Minimum Legal Drinking Age (MLDA): To reduce alcohol-related fatal crashes among youth, all 50 States and the District of Columbia have adopted a MLDA of 21. The MLDA of 21 prevents an estimated 700–1,000 traffic deaths annually among

youth targeted. It saves an estimated \$540 per youthful driver.

Intensive Sobriety Checkpoint Program: Intensive enforcement of State BAC limits with highly visible sobriety checkpoints would reduce alcohol-related fatalities by at least 15% and save approximately \$62,000 per checkpoint. Including police resources, costs of travel delay and the value of mobility losses by impaired drivers apprehended and sanctioned, the costs of conducting a checkpoint averages about \$8,800.

Enforcing Serving Intoxicated Patrons Law: Using undercover police officers to enforce the State laws against serving alcohol to intoxicated bar and restaurant patrons would reduce alcohol-related crash fatalities by an estimated 11%. It would cost an estimated \$0.30 per licensed driver and save about \$20 per licensed driver.

Server Training: Generally, 40% to 60% of intoxicated patrons drive after consuming alcohol in bars, clubs or restaurants. A full-day, mandatory, face-to-face server training program with active management support has the potential to reduce nighttime DUI injury crashes by 17%. Implementing such a program costs an estimated \$70 per licensed driver and saves about \$200 in crash costs per licensed driver.

For more information and information on the economic evaluation of other drinking and driving prevention measures see www.nhtsa.dot.gov/PEOPLE/injury/alcohol/impaired_driving_pg2/US.htm

4.3 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 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 publicise the outcomes of the programme, or publishing reports and papers in the scientific literature.

Checklist

- 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.

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 (Box 4.4).

BOX 4.4: Community action on rural drink-driving – the Waikato, New Zealand Rural Drink Drive Pilot Project 1996–1998 Formative Evaluation Report (1998)

The 1996–1998 Waikato Rural Drink Drive Project (WRDDP) was established and funded by the Alcohol Advisory Council (ALAC) as a community action pilot to develop strategies which would support rural communities in reducing problems associated with drinking and driving. The project was set up in what was the wholly rural Te Awamutu Police District. The formative evaluation involved taking account of:

- countermeasures implemented as part of the project
- the types, frequencies and outcomes of project meetings
- aspects of the project that were not working and the apparent reasons for this
- solutions that were employed to address hurdles encountered
- material, financial and personnel resources that were required
- project team members' perceptions of the process and impact of the project
- stakeholders' perceptions of the process and impact of the project
- apparent keys to successful operation of the project.

Based on the evaluation it was concluded that by “pooling research-based information and local and national knowledge, experience and ideas, sound strategies could be initiated and supported,” and that “health promotion provides a useful guide for the range of strategies and approaches. However, local efforts must be supported by regional and national level agencies, through development or maintenance of appropriate policies. Examples are the need to continue resourcing compulsory and mobile breath testing at efficient levels in rural areas, and mass media campaigns to reinforce this activity. This is particularly important when the public perception remains that it is unlikely they will be breath tested on rural roads.”

More information: www.aphru.ac.nz/projects/rural2.htm

BOX 4.5: Evaluation of Checkpoint Demonstration Project, Tennessee, United States

In 1994 Tennessee initiated and evaluated a state-wide sobriety checkpoint programme. The NHTSA, which operates under the US Department of Transportation, funded equipment and the evaluation, while personnel were provided through diversion of existing resources in the highway patrol.

Checkpoints were conducted throughout the state every weekend using four specially equipped vans with generators, lights, cones, signs, video recorders (one to record field sobriety tests outside the van and one to record breath alcohol tests inside the van) and evidential breath testing equipment. Officers also used passive alcohol sensors in flashlights to detect the odour of alcoholic beverages, and standardised field sobriety tests to detect impaired drivers. The number of checkpoints increased from approximately 15 in the preceding year to nearly 900 in the programme year. Nearly 145 000 vehicles passed through these checkpoints. The increased checkpoint activity was publicised extensively through:

- public service announcements (television and radio)
- print media
- outdoor advertising (billboards)



- news coverage
- brochures and other handouts.

Checkpoints were conducted following guidelines contained in the Tennessee Department of Safety General Order pertaining to Sobriety Checkpoints, which requires that at least six troopers and a supervisor staff each checkpoint. On several occasions, typically holidays, the checkpoints were supplemented with enforcement roadblocks, which do not have the same personnel and equipment requirements as sobriety checkpoints.

Evaluation of the program revealed that:

- during the program 773 “driving under the influence” (DUI) arrests and over 8 000 other traffic citations were made
- self-report measures showed increases in exposure to roadblocks (but no substantial change in self-reported drink-driving behaviour)
- nine out of ten people surveyed supported the use of sobriety checkpoints
- there was a 20% reduction in the projected number of fatal drink-driving crashes that would have occurred if there had been no intervention.

More information: www.nhtsa.dot.gov/people/injury/research/ChkTenn/ChkptTN.html

Summary

- Evaluation should be seen as an integral component of any drinking and driving programme. An evaluation plan needs to be determined at the beginning of a programme development, so that a 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.
- The aims of the programme must be reflected in the evaluation plan and the performance indicators selected. This 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 drinking and driving 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 future programmes.

References

1. Rossi PH et al. *Evaluation: a systematic approach*. California, Sage Publications, 2004.
2. *The art of appropriate evaluation: a guide for highway safety program managers*. Washington DC. Department of Transportation Report HS 808894, May 1999. (available at www.nhtsa.dot.gov/people/injury/research/ArtofAppEvWeb/index.htm, accessed 25 January 2007).
3. Britten N. Qualitative research: qualitative interviews in medical research. *British Medical Journal*, 1995, 311:251–253.
4. Kitzinger J. Qualitative research: introducing focus groups. *British Medical Journal*, 1995, 311:299–302.
5. 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.
6. Woodward M. *Epidemiology: study design and data analysis*. 2nd ed. Boca Raton, FL, Chapman and Hall CRC, 2005.
7. Rothman, KJ, Greenland S. *Modern epidemiology*, 2nd ed. Hagerstown, MD, Lippincott-Raven, 1998.
8. Hingson R et al. The case for 0.08% Per Se Laws. *American Journal of Public Health*, 1996, 86, (No. 9):1297–1299.
9. Drummond MF et al. *Methods for the economic evaluation of health care programmes*, Oxford, Oxford University Press, 1997.
10. Tan-Torres T et al., eds. *WHO guide to cost-effectiveness analysis*. Geneva, World Health Organization, 2003.
11. Kerry SM, Bland JM. Statistics notes: sample size in cluster randomisation. *British Medical Journal*, 1998, 316:549.