MONASH University
Accident Research Centre

THE DEVELOPMENT OF A FRAMEWORK FOR ROAD SAFETY STRATEGY EVALUATION

Prepared for

Queensland Transport

by

MONASH UNIVERSITY
ACCIDENT RESEARCH CENTRE

May, 2009

Report No. Final Draft

Contents

EXECUTIVE SUMMARY

Queensland Transport commissioned Monash University Accident Research Centre (MUARC) to develop an evaluation framework that would monitor and report on the outcomes of the 2004-2011 Queensland Road Safety Strategy. It was intended that the previous strategy covering the period 1993-2003 and revised in 1998 be used to trial the framework.

The project consisted of the following work tasks:

1. A review of Australian and international literature on best-practice approaches for evaluating road safety strategies and action plans.
2. A review of the current methods of providing evaluations and analyses within Government.
3. The development of an evaluation framework that could be applied to the 2004-2011 Queensland Road Safety Strategy and associated action plans.
4. A 'test-run' of the evaluation framework on the previous 1993-2003 Queensland Road Safety Strategy (revised in 1998).

LITERATURE REVIEW

A review of Australian and international literature on best-practice approaches for evaluating road safety strategies and action plans recommended that the GOSPA model be used as a basis to define the elements of the evaluation framework. The GOSPA framework for evaluation was developed by MUARC for application in Western Australia. This method relates evaluation elements to a strategic planning framework (Cameron, 1999). The structure of the GOSPA framework is as follows:

<u>G</u>oal	General (idealistic) statement of the Program's overall goal
<u>O</u>bjectives	Specific (pragmatic) statements of the Program's measurable objectives to reach the Goal
<u>S</u>trategies	General (idealistic) strategies to achieve each Objective
<u>P</u>rograms/<u>P</u>lans	Specific (pragmatic) programs/plans contributing to each strategy, with measurable activity levels and outputs
<u>A</u>ctions	Actions undertaken in each program.

GOSPA EVALUATION FRAMEWORK FOR THE QUEENSLAND ROAD SAFETY STRATEGY

The GOSPA framework defines a pyramid of increasing detail in defining the elements of a road safety strategy. The top of the pyramid defines the broad goals for which the strategy is aiming (Goals) whilst the next level down gives specific measurable targets (Objectives) against which the goals can be assessed. The Strategies area of the framework typically defines the target areas on which the road safety strategy will focus to achieve its goals along with local objectives within each target area that will jointly contribute to achieving the global objectives. Finally the Programs and Actions areas contain the specific details on the type of activities to be carried out in each target area and the amount of effort that will be applied to each activity.

An evaluation framework for the Queensland road safety strategy has been developed in this study that mirrors the pyramid structure defined under the GOSPA framework. It is designed to assess the progress of the strategy against the pre-determined goals and objectives at various levels of detail through a multi-tiered modelling approach. Each tier of modelling focuses on a specific level of disaggregation of the strategy elements corresponding to particular levels of the GOSPA framework pyramid. These are described broadly as follows:

- **The Global or First Tier Assessment Model**: This level of assessment aims to measure the ongoing performance of the road safety strategy in achieving the broad **G**oals and **O**bjectives set out for the strategy as a whole. Specifically it aims to measure whether the strategy has reduced overall road trauma levels, as defined by measures specified in the strategy Objectives, from that expected based on pre strategy implementation trends.
- **The Second Tier Asssessment Model:** This level of assessment has aims similar to the First Tier models except here assessment is aimed at each individual target strata defined by the **S**trategies of the Queensland Road Safety Strategy and corresponding Program areas. Assessment at this level will aim to articulate in what particular areas the strategy is working and to what degree.
- **The Third Tier Assessment Model:** This level of assessment aims to relate the trends observed in each of the target strata defined in the Second Tier models to explicit measures of road safety program for major activities defined in the **P**rogram and **A**ction elements of the strategy.

The following table depicts how the GOSPA framework relates to the three-tiered models and fits in with the test-running of the framework on the earlier strategy.

Summary of the link between the GOSPA framework and the three-tier modelling approach

GOSPA FRAMEWORK COMPONENTS		THREE-TIER MODELLING
Component	**Definition**	
Goal	Overall goal of strategy (i.e. to prevent road trauma through safe road use, safe roads and safe vehicles)	Global assessment model (top-tier model) to measure effect on road trauma of the Strategy overall (includes an intervention model)
Objective	Objectives to reach goal (e.g. to achieve a reduction in the fatality rate to under 5.6 deaths per 100,000 people)	
Strategies	General strategies to achieve objectives given in the Qld action plans and road safety strategy	Second-tier modelling of specific strata targeted by the strategies in the action plans (e.g. crashes occurring during high alcohol times of the week)
Programs	Specific programs relating to target group outcomes	Third-tier modelling of the individual program elements of the strategy (e.g RSIP evaluation)
Actions	Actions undertaken in each program	

A fourth tier of the evaluation framework is also included which describes the imperative of undertaking specific targeted evaluations of major road safety programs implemented or enhanced under the road safety strategy. Such evaluations will give the most rigorous scientific assessment of the effectiveness of the programs evaluated and contribute to understanding of the mechanisms of effectiveness. Furthermore, results of the specific evaluations are critical in assisting the formulation of the Tier 3 models by identifying the most relevant measures of program activity that predict the outcomes being measures.

The statistical methodology used in tiers one and two of the three-tiered evaluation framework was based on structural time series (or ‘state-space’) modelling theory. Structural time series models are a general class of time series model that allow the most ready selection of parsimonious model structures to reflect the inter-correlated nature of time series data. The also allow the easy inclusion of covariates into the model structure and provide simple an accurate forward forecasts of trends. Forecasts from the pre-strategy time series data were used to estimate expected post intervention trends in the absence of strategy implementation. Actual road trauma outcomes were them plotted against the forecasts to assess strategy effectiveness either globally or within particular target strata. Observed post strategy implementation data below the forecast indicated the strategy was successful in reducing road trauma from levels expected had the strategy not been implemented. Performance can be assessed month by month by simply plotting actual road trauma levels against the trends forecast in the absence of the strategy. Additional time series models were also fitted to both the pre and post strategy implementation data to formally assess the impact of the strategy in reducing road trauma by including post strategy intervention terms in the model.

To demonstrate how the evaluation framework works in practice, a test-run of the framework was trialled on the previous 1993-2003 Queensland Road Safety Strategy. Sufficient crash data prior to 1993 was not available to support the modelling process so it was decided to trial the evaluation framework on assessing the outcomes of the revision to the strategy implemented from 1998. Detailed description of each tier of the evaluation framework follows along with an assessment of its practical application to the previous Queensland road safety strategy. 'State-space' modelling techniques were used to fit models to the Queensland fatal and injury crashes for the period April 1991-December 2004. Explanatory factors were also included in the models to improve fit and predictive power as well as give insight into how changes in these factors might affect forecast trends. Factors included were the monthly unemployment rate and monthly fuel sales for Queensland.

DEATILED DESCRIPTION OF THE FRAMEWORK ELEMENTS AND TEST-RUN OF THE EVALUATION FRAMEWORK ON THE 1993-2003 STRATEGY

To demonstrate how the evaluation framework could be applied to a Strategy, a test-run of the framework was trialled on the previous 1993-2003 Queensland Road Safety Strategy, revised in 1998.

GLOBAL ASSESSMENT MODEL

This global assessment model aimed to measure the effect on road trauma of the Queensland Road Strategy overall. Road trauma was quantified in terms of fatal crash risk, serious casualty crash risk and all injury crash risk.

The global models that were estimated demonstrated how road trauma levels can be modelled each month prior to the introduction of the Strategy and then used to forecast the levels of road trauma that would have been expected to have occurred after 1997 had no strategy been in place (based on past trends). Confidence limits were also placed on the estimates. The actual road trauma trends that occurred in Queensland after the introduction of the (revised) 1993-2003 Strategy were also plotted. Plotting the actual trends against those forecast in the absence of the Strategy was used as a means of monitoring overall strategy performance over time.

The demonstration of the evaluation framework appeared to have worked best for the first few years of the projection (i.e. 1997 to late 2000). However in the last year of the prediction period (i.e. late 2000 to 2001) assessment of the actual values against the forecast was difficult due to the typically wide confidence limits on the forecast. This suggests that there may be a need to re-calibrate the forecasts at regular intervals, say every 2 years to correspond with the development of each new action plan, based on a longer data period. The evaluation strategy would then measure the incremental improvement in global outcomes related to each action plan period.

The time-series models that investigated road trauma trends pre and post the Queensland Road Safety strategy against a forecast trend post strategy in the absence of the strategy are similar in philosophy to the control-chart methodology used by a number of agencies in the past, including Queensland Transport, to monitor road safety strategy performance. The advantage of the methodology for the global assessment model used in this study is that it employs much more sophisticated and robust statistical methodology yet is still amenable to use by those without statistical training once established.

The global assessment model was also used at various time periods after implementation of the 1993-2003 Strategy to formally evaluate the overall performance of the Strategy to that point in time based on the key outcome measures specified in the Strategy's goals and outcomes. The pre and post implementation data to the time available were modelled using state-space techniques with annual intervention terms being included at the time of Strategy implementation. The intervention term parameters then represent the effect of the strategy on the outcome measure which can be tested formally for statistical significance. The intervention terms can be modified accordingly to reflect increasing effects of the broad Strategy over time which might be expected if components of the Strategy are introduced in a staggered manner over time or take some time to become fully effective. Application of the global assessment model in this manner would require high level trained statistical expertise.

For the 'trial-run', at the global assessment level, the performance of the 1993-2003 strategy was measured using an intervention type forecasting model. As a demonstration of how the intervention analysis works, the fatal crash risk global assessment model was used at various time periods after the Strategy was implemented (i.e. annual step functions at the end of each year for the post-period 1996-2003) to formally evaluate the overall performance of the Strategy to that point in time. The intervention model showed indications of annual reductions in fatal crash risk in December 1998-December 2001, perhaps associated with the implementation of the Queensland Road Safety Strategy.

The results of the intervention analysis suggest that the 'test-run' of the evaluation framework on the earlier Strategy captured some of the performance of the Strategy by finding an effect (a reduction in fatal crash risk) during 1998-2001.

SECOND-TIER MODELS

The second tier model relates to the objectives defined by the strategies and programs of the 2004-2011 Queensland Road Safety Strategy and the 2004-2005 associated Action Plan. This model considered the effects on road trauma for specific strata defined by the road user groups or situations (e.g. high alcohol crashes; speed-related crashes) at which the strategies and programs are targeted.

For each of these strata, a specific analysis model equivalent in structure to that defined by the global assessment model above was estimated for each of the key outcomes being measured. Like the global assessment models, the second tier models were formulated at time of implementation to forecast road trauma outcomes in each stratum of interest had the Strategy implementation not taken place. Actual post implementation road trauma trends were then compared to those forecast to assess Strategy effectiveness.

The modelling results showed that for most of the second-tier models the actual road trauma trends were less than those forecast had the 1993-2003 Strategy not taken place– particularly during the first few years of projection.

In many instances the observed data were outside the confidence limits of the forecast of pre strategy trends showing that the strategy had significantly improved road trauma outcomes from what would have been expected had the strategy not been implemented. In the later months of the prediction period the confidence limits on the pre strategy trend projection typically become very wide meaning statistical comparison of the observed data with the forecast becomes problematic. Again, this is particularly the case in the period more than 2 years after strategy implementation and further suggests that comparisons with the forecast should be limited to 2 years of forecast data. Accuracy of the initial forecast data will rely somewhat on the length of pre-strategy data available for analysis with longer time periods giving more accurate forecasts. The exact time period at which forecasts need to be re-estimated will depend on how much prior data the initial forecast is based. Similar comments on forecasting accuracy are also relevant to the global assessment level models.

THIRD AND FOURTH-TIER MODELS

The third tier models attempted to model the individual program elements and actions (for example the effects of speed camera operations) of the Queensland Road Safety Strategy and Action Plans on crash outcomes.

The third tier modelling strategy is an extension of the second-tier model in that it typically targets the same strata defined in tier 2. However, instead of modelling historic trends through general level, slope and seasonal terms, the model includes specific measures of road safety program effort under different activity areas as model covariates. In this way, the model makes estimates of the effects of individual initiatives (where there is sufficient data for the estimate to be reliable) by establishing the relationship between measurable road safety program effort and the key strategy outcome measures and relating the real variation in program effort to the reduction in road trauma observed. The results from the third-tier modelling process give specific estimates of the relative contributions of each of the major program elements in the road safety strategy to achieving the measured outcomes.

Application of the Tier 3 modelling approach to the previous road safety strategy period in Queensland has been demonstrated previously in evaluating the effects of the Road Safety Initiatives Package (RSIP) implemented as part of the strategy. The demonstration application of the Tier 3 approach on the RSIP example clearly shows the efficacy of the methodology in being able to identify a range of specific major road safety program activities that are significantly associated with observed crash outcomes. Furthermore, the Tier 3 modelling approach successfully identified the relative crash reductions associated with each, thereby giving evidence as to the relative effectiveness of each program component on producing the outcomes observed.

A fourth tier of evaluation recommended for the Queensland Road Safety Strategy is *specific* evaluation of major program components. Demonstration of the evaluation of specific programs (such as those listed above) is outside the scope of the evaluation framework being described and is unnecessary given the range of excellent examples related to the previous Queensland strategy that already exist. It is essential that specific programs in the current and future Queensland road safety strategies are independently evaluated in order to understand the value of individual programs, particularly those programs that are expensive in terms of implementation or enforcement, but require a high level of statistical and analytical expertise to evaluate properly.

PRACTICAL USE OF THE EVALUATION FRAMEWORK

The global assessment and second-tier models developed can be used as predictive models to monitor the road safety performance of the Strategy or action plans. The predictive models should be re-calibrated every two to three years, perhaps to link in with each two-yearly action plans that are released by Queensland Transport. It should be noted that the predictions should not be made too far into the future – other underlying factors may influence the trends and it is difficult to predict what changes will occur in these factors (e.g. Unemployment, fuel sales, alcohol consumption) during the prediction period.

The evaluation framework developed as part of this study and its links with the three-tiered modelling approach has not been used previously for Strategy evaluation. To the best of the authors' knowledge this is a new way to monitor road safety strategy performance. Trained professionals with statistical modelling experience should initially develop the three-tiered models. The models can then be re-calibrated every two years to fit in with the two-yearly action plans.

1. INTRODUCTION

Queensland Transport commissioned Monash University Accident Research Centre to develop an evaluation framework that would accurately monitor and report on the outcomes of the Queensland Road Safety Strategy 2004-2011. It was intended that the previous strategy for 1993-2003 be used to trial the framework.

The objectives of the project were to:

- Identify a best-practice method for evaluating the road safety strategies and action plans, and

- Develop an evaluation framework based on crash and enforcement data, community perceptions and needs, Government and departmental priorities, and perhaps other relevant information indicating outcomes of the strategy.

The project encompassed the following four work tasks:

5. A review of Australian and international literature on best-practice approaches for evaluating road safety strategies and action plans.

6. A review of the current methods of providing evaluations and analyses within Government.

7. Development of an evaluation framework that could be applied to the Queensland Road Safety Strategy 2004-2011 and/or associated action plans.

8. A ‘test-run’ of the framework on the Queensland Road Safety Strategy (revised) 1993-2003.

In this draft report, works tasks 1 and 2 are described in chapter 2, whilst chapters 3 and 4 relate to the issues associated with the development of the evaluation framework. Chapter 5 gives a demonstration of how the evaluation framework was trialled on the previous 1993-2003 Queensland Road Safety Strategy that was revised in 1998.

1.1 BACKGROUND

Monash University Accident Research Centre (MUARC) has had an extensive history of involvement in evaluations of road safety strategies, both at the program level and at the level of individual initiatives.

Perhaps the most extensive of these program evaluations has been MUARC’s evaluation of Victoria’s road safety strategy launched in September 1989 involving the introduction of “booze buses” for RBT, extensive use of covert speed cameras, high-profile graphic road safety television advertising, and reduction in freeway speed limits (Cameron et al 1994, Newstead et al 1995, 1998). This was followed by progressive evaluations of New Zealand’s Supplementary Road Safety Package of enforcement and advertising initiatives launched in 1995/96 (Vulcan and Cameron 1996, Cameron and Vulcan 1996, 1998, Cameron et al 2002). These research programs included process evaluations of the strategy components as well as outcome evaluations of each strategy. This research complemented numerous individual evaluations of each initiative in the jurisdictions.

More recently, MUARC has undertaken a comprehensive evaluation of Queensland's Road Safety Initiatives Package (RSIP) during 2002-2003, which is understood to form part of Queensland's 1993-2003 Road Safety Strategy. This research has demonstrated the feasibility of evaluating road safety strategies in Queensland at the program level (Newstead et al 2004). It provided estimates of crash savings for the RSIP as a whole as well as individually for many of the key initiatives in the RSIP. MUARC has also provided technical reviews of unpublished work by the Queensland Government Statistician's Office which modelled the collective and individual contributions of the earlier initiatives in the 1993-2003 Road Safety Strategy up to 1999 (viz. State Highway Patrols, RBT, Random Road Watch and bicycle helmet legislation).

Many of the road safety initiatives in Queensland's 1993-2003 Strategy have been individually evaluated by MUARC. These include Random Road Watch (Newstead et al 2001), mobile speed cameras (Newstead and Cameron 2003, Newstead 2004, 2005, 2006), and the 50 km/h default urban speed limits in South East Queensland (Hoareau et al, 2007) and regional Queensland (Hosking et al, 2007). In addition, MUARC has undertaken a technical review of the "control chart" evaluation method used within Queensland Transport for rapid monitoring of the effects of initiatives (Newstead 2000). This background experience of evaluation of Queensland's initiatives, while peripheral to the strategic evaluation framework required in the proposed project, will provide a useful basis for testing the sensitivity of the framework in the 'test-run'.

Perhaps most relevant to the framework development is a method of evaluation developed by MUARC in Western Australia. This method relates evaluation to a strategic planning framework (Cameron 1999; see Appendix A). Strategic plans, typified by Queensland's Road Safety Strategy, are often structured according to the GOSPA framework:

Goal	General (idealistic) statement of the Program's overall goal
Objectives	Specific (pragmatic) statements of the Program's measurable objectives to reach the Goal
Strategies	General (idealistic) strategies to achieve each Objective
Programs/Plans	Specific (pragmatic) programs/plans contributing to each strategy, with measurable activity levels and outputs
Actions	Actions undertaken in each program

The GOSPA framework allows an evaluation structure to be defined. Programs are not considered to be implemented unless actions are taken, strategies are not achieved unless planned programs are implemented, objectives are not met unless the strategic directions of programs are correct, and goals are not achieved unless the targets of objectives are met. Through this strategic framework, the key pathways through which each action, program and strategy contributes to the overall goal becomes apparent. Measurable criteria at each level provide the basis for assessment that real change has occurred, and potentially developing linkages to the overall goal through modelling of the linkages.

While the Cameron (1999) research paper given in Appendix A was written in the context of a specific enforcement initiative to expand Western Australia's speed camera operations, its key role in funding other initiatives in their Integrated Road Safety Program led to an evaluation framework which embraced the full road safety strategy (at least conceptually). Hence this approach was suggested as being potentially relevant to evaluation of Queensland's Road Safety Strategy.

The following chapter describes stages 1 and 2 of the project namely the review of Australian and international best-practice for evaluating road safety strategies, and the review of current methods within Government.

2. REVIEW OF LITERATURE – BEST PRACTICE FOR EVALUATING ROAD SAFETY STRATEGIES AND ACTION PLANS

2.1 BACKGROUND

Road safety strategies play a major role in reducing death rates and injury rates resulting from road trauma. Such strategies consist of a range of individual road safety programs, each contributing a varying level of success towards reducing crash rate/frequency and crash severity.

To gain an understanding of how successful a road safety strategy has been is not as simple as looking at the numbers of deaths and injuries to see whether they have increased or decreased over a certain period. Any number of explanatory variables may have an impact on the outcomes of a road safety strategy. Changes in the economy, increased population and increased motorisation all have the capacity to affect the level of success achieved in reaching road safety targets (Johnson, 2001). Evaluation plays a vital role in assessing which aspects of a road safety strategy work and which aspects need to be reviewed. Such a tool can help reallocate funds and resources towards programs which will have the greatest impact on reducing deaths and injuries.

The process of road safety evaluation is not always simple and it is unlikely that the evaluation methodology used for one strategy would be effective if used for another state or country's strategy. Factors such as population, urban/rural distribution, socio-economic status, vehicle ownership, vehicle use, climate, infrastructure and government could vary across states/countries thus rendering any single strategy and its evaluation invalid for other regions.

The more commonly reported method of road safety evaluation to date is individual program evaluation. Examples of this throughout Australia include programs related to speed cameras, random breath testing, moving mode radar, accident blackspot treatments, 50 km/h speed limits and Random Road Watch. With information on the effectiveness of each individual program, estimates can be made of the number of lives and injuries saved/prevented as a result of each program. However, there may be difficulty in measuring the interaction between individual programs and their respective contribution to the overall strategy goals. For example, a reduction in speed-related crashes could be influenced by speed cameras, 50 km/h speed limits and mass media advertising, yet the extent to which each of these programs contributed to the outcome is often much less clear.

Road safety strategies implemented across Australian jurisdictions have had varying degrees of evaluation undertaken to assess what effect the various programs and the strategy as a whole had on road safety. Whilst there were examples of individual program evaluations (as mentioned above), few were found to evaluate the entire strategy with consideration not only to each of the programs, but to the many factors influencing road safety outcomes. Much of the work is unpublished and often takes the form of reports to Government departments or related agencies.

2.2 METHODS OF EVALUATION

To develop an evaluation framework, it must be clear what the objectives and targets are for the whole strategy and each sub-program. Furthermore these objectives must be measurable, and stated clearly, preferably in quantitative terms before the evaluation can be undertaken (Cameron, 1999).

Two methods of evaluation were identified:

i. The GOSPA model, and
ii. The PIARC procedure.

2.2.1 The GOSPA Model

The GOSPA model has been employed successfully by Monash University Accident Research Centre (MUARC) to develop a framework for several different evaluations. Such a model allows evaluation at the level of each action, program or strategy, and the respective contributions of each to the achievement of the overall goal(Cameron, 1999). The structure of the model is as follows:

Goal	General (idealistic) statement of the Program's overall goal
Objectives	Specific (pragmatic) statements of the Program's measurable objectives to reach the Goal
Strategies	General (idealistic) strategies to achieve each Objective
Strategic Objectives	Measurable indicators of extent to which each Strategy has been achieved
Programs/Plans	Specific (pragmatic) programs/plans contributing to each strategy, with measurable activity levels and outputs
Actions	Actions undertaken in each plan
Action Targets	Measurable indicators of extent to which each Action has been achieved

Evaluation can at this stage be classified in three categories - process evaluation, outcome evaluation and impact evaluation. Process evaluation assesses whether the appropriate actions were taken and whether sub-programs were implemented as planned. The process evaluation has parallels with normal management procedures associated with implementing a complex project, but should go beyond normal requirements to record regularly and systematically the key measures of activity of major sub-programs. This will be valuable in attempting to separate the contributions of the individual sub-programs to meeting the overall objectives. It is unable, however, to assess whether a successfully implemented plan has actually achieved the objectives of the program.

Outcome evaluation assesses whether the total program produces real change at each level. Whether the strategies are achieved does not guarantee that the objectives will be met, but if the strategies are well chosen, it could be expected that this is highly likely. The nature of road safety programs is such that it is usually possible to obtain information on the extent to which the strategies have been achieved well before it is clear that the program objectives are being met. For this reason, the emphasis should be placed on clearly articulating the strategies for the overall program, as well as defining the strategic objectives. Much of the evaluation outcome in the evaluation plan will focus on the achievement of the strategic and program objectives.

An important part of outcome evaluation is an assessment of the impact on the criteria defined in the objectives (usually crash-, injury- or behaviour-related criteria), known as impact evaluation.

2.2.2 The PIARC Procedure

The World Road Association (PIARC) incorporated a four-step procedure for a systematic framework of evaluation (Haaland and Odeck 1998). These steps were:

i. defining gains and losses according to some set of objectives

ii. listing advantages and disadvantages

iii. measuring gains and losses in some unit or different units

iv. decisions are made on the basis of (i) to (iii) by using explicit weights of importance (eg monetary values) or implying such weights ex-post

The PIARC study identified diversity within the process of project evaluation, often selected dependent upon the nature of the road safety issue. Whilst basing its methodologies on Benefit Cost Analysis (BCA - where all benefit and cost items can be assigned monetary values), it also allowed for a broader view with BCA on one end of the spectrum and Multi Criteria Analysis (MCA - where information is not restricted by monetary values) on the other. Many countries, Australia included, have not uniformly accepted any single methodology.

2.3 REVIEW OF AUSTRALIAN AND INTERNATIONAL LITERATURE ON BEST-PRACTICE APPROACHES FOR EVALUATING ROAD SAFETY STRATEGIES AND ACTION PLANS

Key road safety organisations in Queensland, Victoria and Western Australia jurisdictions were contacted to obtain information on their current processes for evaluating and monitoring the implementation and outcomes of their road safety strategies. It was expected that each organization would have processes for monitoring outcomes at a general level (macro-level evaluation) and some would also undertake scientific evaluations of individual initiatives (micro-level evaluation). Monitoring of implementation levels to provide process evaluations was also expected to be common. However, global-level evaluations of strategies or packages of initiatives, aimed at measuring their effect collectively and individually, were found to be relatively rare.

The same information was sought from the more progressive road safety countries with a history of monitoring and evaluation (U.K., Netherlands, Sweden, Norway, Finland, New Zealand). Outputs from the SUNflower project, which compared the road safety strategies and outcomes in Sweden, U.K. and the Netherlands, provided key information. Other pan-European projects (e.g. ESCAPE, etc.) also provided information.

2.3.1 Road Safety Strategy Evaluations in Australia

2.3.1.1 Queensland

The previous Queensland road safety strategy (1993-2003) highlighted five principles behind the management of road safety in the state. Assessment of these five principles and the actions which arose from them was addressed by means of a control chart technique (Leggett 1999). The comprehensive assessment of crash problem categories resulted in a ranking by size of the problem, and formed the columns of the control chart (Figure 2) A range of options to address these problems were formulated and ranked in order of priority (based on cost-effectiveness and other factors), forming the rows of the control chart. The resulting risk-response diagram provided a summary of the action plan and illustrated which actions had a direct or indirect impact on crash risk / crash severity.

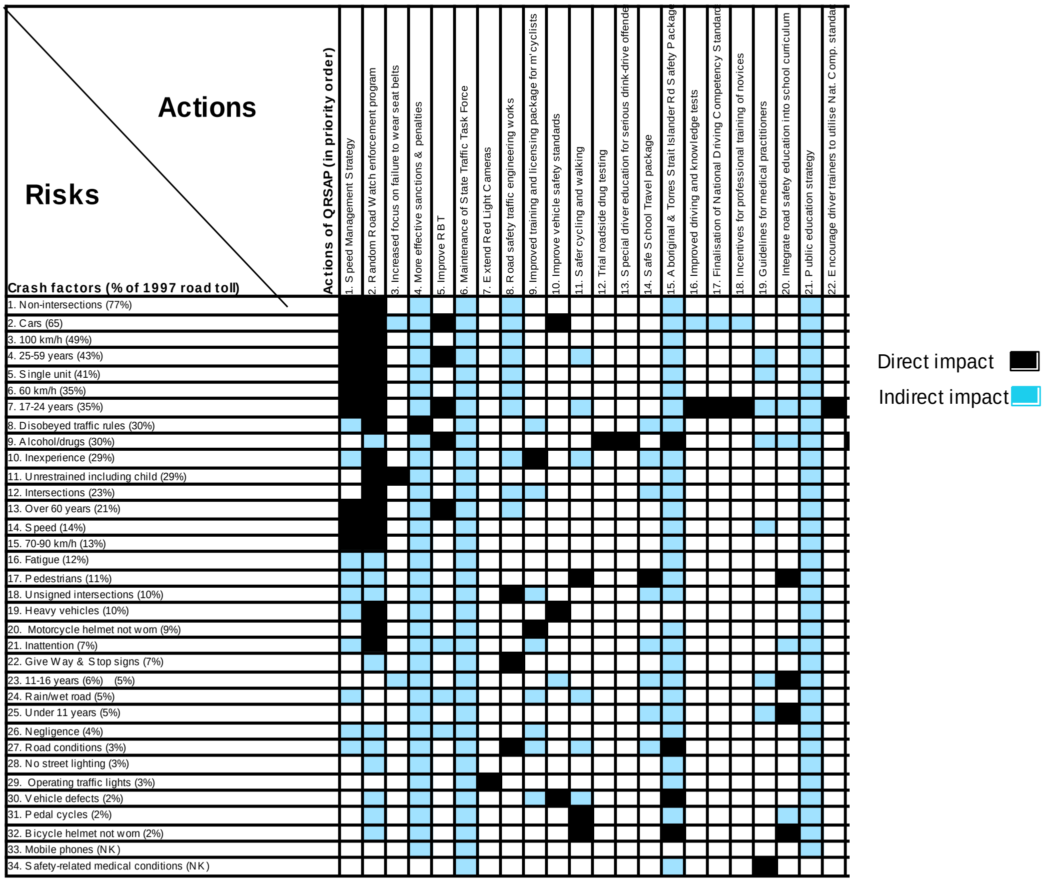


Fig. 2: Queensland Road Safety Action Plan: Risk/response diagram: risks in priority order by responses in priority order

Regular monitoring and evaluation was achievable in Queensland because of the timely supply of information. Electronic data entry from every police station meant that basic crash information from anywhere in the state was usually available within a matter of days (Leggett 1999). This data, fed into an information management system, could then be used to predict road safety outcomes through an econometric model. The rapid turnaround of information into the system and predictions from the model enabled adjustments to be made if necessary, thus saving time and resources.

Independent external evaluations could be used to confirm the estimates from the management information system. It was deemed potentially feasible that the estimated impacts of the individual programs could be aggregated to give an overall estimate of the effectiveness of the Queensland road safety strategy. Results showed that the aggregation of countermeasures method could realistically only account for 80% of the crash savings identified using the macro model which estimated the overall impact of the strategy (Leggett 1999). Furthermore, the aggregation method could not take into account any positive crossover effect one program may have had on another.

2.3.1.2 Victoria

Victoria's "Safety First" road safety strategy for 1995-2000 may not have resulted in a reduction in absolute numbers of fatalities or serious injuries in the period, yet the strategy was still considered successful. This was because absolute numbers are less important than exposure to risk in terms of road safety (Johnston, 2001). Two measures exist for exposure to risk: 1) death/injury per head of population (measure of risk to an individual); and 2) death/injury per unit travel (measure of risk within the road transport system). The decision over which of these measures to use for evaluation, once again will usually be dependant upon the stakeholders and the objectives of the various road safety programs within the strategy.

A number of explanatory variables need to be identified as having an impact on the outcomes of the road safety strategy. Changes in the economy, increased population and motorisation all have the capacity to affect the level of success achieved in reaching targets, and need to be considered in the evaluation(Johnson, 2001).

Economy - Economic factors may certainly impact on the outcomes of road safety strategies and need to be considered in the evaluation model. When the economy is strong, there is an increase in discretionary travel (especially recreational travel), unemployment falls (further increasing travel) and alcohol sales increase (thus increasing high risk travel). This means more vehicles on the road at any given time point, and a higher risk of crashing due to increased alcohol consumption. The opposite is true in times of economic recession (less discretionary travel meaning less vehicles on the road and lower alcohol sales decreasing the risk of crashing), and may well contribute to the success of meeting targets in a road safety strategy, should the economy fall into recession as it did in the late 1980s in Australia. Different measures for economic strength at the State level include unemployment, Gross State Product and Final State Demand.

Population - Population growth ensures that there will be more licensed drivers and other road users (eg. pedestrians, cyclists) exposed to traffic risk. Results of the Victorian evaluation indicated that the reduction in the levels of risk (as a result of the initiatives within the "Safety First" road safety strategy) was offset by the increase in total exposure to risk, thus resulting in a stable number of deaths and serious injuries.

Motorisation - The increased level of motorisation needs to be considered regarding exposure to risk within a road transport system. For example, whilst the Victorian population is projected to grow by 3.9% from 2001-2006, the number of registered vehicles is expected to grow by 12%. Compounding this is the ageing of the road user population, a group with an increased risk of crash per road unit and an increased risk of serious injury given that a crash occurs. The overall effect of motorisation on road crash statistics however was not expected to change greatly, as these increases would be offset by the expected reduction in the proportion of road users under the age of 25 years, the highest risk group of road user.

2.3.1.3 Western Australia

One of the few documented reports located which evaluated the effects of each program on a road safety strategy as a whole was for Western Australia's "Arriving Safely 2003-2007" strategy. The first annual report for the Road Safety Council of Western Australia (RSCWA) was completed in late 2004, effectively a progress report after year 1 of the strategy. Some of the key points regarding data quality, data collation and developing an evaluation model were noted.

To undertake a meaningful evaluation of a road safety strategy, the RSCWA needed to ascertain how best to collate the currently available data. It was necessary to discuss with stakeholders any issues regarding the provision of data within their scope. Such stakeholders included the state police (crash and enforcement data) and a range of government departments. Issues to discuss included the current and predicted frequency of data, the geographical disaggregation of the data, the physical format and transfer of available data and the most appropriate methods for collating data that was not currently available.

Following the collation of data, a model could be developed. The model included all the performance indicators (PIs) identified as important measures of the strategy (RSC 2004). Each program within the strategy had its own set of PIs. Performance was measured in three categories, in order of importance - outcome measures (eg. crash rates), intermediate measures (eg. behaviours/attitudes) and process measures (eg. enforcement intensities). Some PIs were simple counts or rates; others were more complex such as levels of enforcement and advertising. Different measures for some performance indicators needed to be assessed (usually dependent upon stakeholders). Additionally, some overall PIs were identified to gauge performance relative to overall targets.

Performance reports were also an essential part of the evaluation, as they enabled PIs to be compared at different time points of the strategy as well as compared with baseline data. Evaluation of the "Arriving Safely 2003-2007" strategy was planned to consist of an annual report and quarterly performance reports throughout the life of the strategy. This constant monitoring of performance of the strategy would enable identification of problems that required immediate attention. Consultation with stakeholders guided the nature, frequency and form of performance reporting. One issue to be considered was whether each stakeholder should receive the same level of detail in the performance report, as certain stakeholders may not want sensitive information released to any external parties.

2.3.2 Individual Road Safety Program Evaluations in Australia

2.3.2.1 Queensland Speed Camera Program

The Queensland speed camera program was implemented in 1997 and underwent constant monitoring which enabled the focus to be modified dependent upon progress results. Newstead and Cameron (2003) completed an evaluation on crash effects as a result of the program. The broad aim was to establish the effect of the speed camera program on crash frequency in Queensland. Other aims were to investigate the differential effects of the program by crash severity as well as over time, and to establish the mechanisms of program effectiveness through relating estimated crash changes to program operational measures. Outcome measures (i.e. crash effects) were the only performance measures of interest. Dependent upon stakeholders, as was the case with this study, it may not be necessary to identify intermediate or process measures (behavioural changes, attitude changes and enforcement activity). The rationale behind this may be that as long as crash reductions are happening, it doesn't matter whether behaviours or attitudes have changed.

2.3.2.2 Queensland Random Road Watch Program

The evaluation of the Queensland Random Road Watch (RRW)(Newstead et al, 1999) program consisted of a number of stages. Crash frequencies were compared pre- and post-RRW intervention. The statistical model was able to estimate the effect of RRW by police region and urbanisation, as well as region wide and state wide. The homogeneity of the effect of RRW within a group or region was tested, as was the homogeneity of crash data from areas and times outside the influence of RRW. A pseudo experimental design which could potentially control for factors other than the treatment was chosen. Such factors include socio-economic effects and the effects of other road safety programs occurring throughout the duration of this study period. The final stage was to convert the estimated crash reduction into cost benefits to the community.

2.3.2.3 Western Australia Black Spot Program

Another individual road safety initiative to be evaluated was the Black Spot program in Western Australia 2000-2002 (Meuleners et al. 2005). Sources of data were Main Roads WA and the Western Australian Road Injury Database. Crash data included crash date, crash severity, local government area of crash and specific crash location. One assumption of this study was that it was more appropriate for a black spot program to include all crash types rather than isolating specific crash types which may potentially miss some benefits or detrimental effects of a treatment. The outcome measures of interest in this study were overall crash rates, number of casualty crashes only and cost of crashes (based on the Bureau of Transport Economics road crash severity costs for Australia in 1996). The statistical analysis involved a generalised estimating equation (GEE) Poisson regression model, which was used to allow for the inherent correlation of longitudinal data.

2.3.2.4 Western Australia 50 km/h Speed Limit Program

As stated previously, evaluations need to have objectives which are clearly stated and measurable. For example, the WA default 50km/h speed limit program evaluation (Hoareau and Newstead 2004) had the objectives of:

i) estimating the net effect of the program on crash frequency and severity of injury for all crashes combined and for selected target groups.

ii) investigating the effect of the program on vehicular speeds by means of speed monitoring surveys.

iii) assessing the level of community support for the initiative over time by means of attitudinal surveys.

The first of these objectives was interested in outcome measures (crash frequency/severity), whilst the final two objectives were intermediate measures (behaviours/attitudes). This evaluation also intended to examine speed enforcement data (process measures), however there was insufficient data available to conduct a meaningful analysis. The evaluation design for this type of study was a quasi-experimental design, incorporating a control group characteristic of the treatment group, but not exposed to the intervention. Data sources used were crash data, speed monitoring data, community attitude surveys (supplemented by focus groups) and speed enforcement data.

2.3.2.5 Victorian road safety evaluation studies

MUARC has completed eighteen studies involving the evaluation of a range of road safety initiatives in Victoria. Whilst too numerous to mention individually, MUARC report no. 177 (Cameron and Newstead, 2000) listed twelve of these reports (to September 1999) and provided a brief description of the statistical framework used in each evaluation. A further six reports have been completed since then (available on the MUARC website), most of which use a similar method of evaluation to the studies mentioned above. A quasi-experimental evaluation (time-series or before-after comparison) of the impact of these initiatives was applied for most of these studies. For some, a multiple regression time-series analysis was the preferred methodology to link crashes with road safety initiative measures and other factors.

The use of before-and-after studies such as the above Black Spot and RRW evaluations raises the issue of which confounding variables should be controlled for. Often the variables that are controlled for are those which have been shown to confound the results of previous studies (Elvik 2000). Adding to the problem, the lack of appropriate data often determines which confounding variables are controlled. Some of the commonly regarded confounding variables include regression-to-the-mean, long term trends affecting crash rates or injured road users, changes in traffic volume, and any other events introduced at the same time as the road safety measure. These shall be discussed in the next section on international evaluation methodology.

These studies all provide evidence that there is no single best method for evaluating road safety programs. All methods require a number of assumptions to be made, and the nature of the program and vision of the stakeholders will certainly mould the way in which the evaluation is designed.

2.3.3 International Experiences in Evaluation

Perhaps the key issue when evaluating a road safety strategy is "How do we know to what extent programs within the strategy had a significant effect on reducing road crashes?" Internationally, as in Australia, many attempts have been made to evaluate the effect of individual road safety programs, most of which try to quantify the effects. However Wong et al (2004) proposed a qualitative assessment methodology using a cluster analysis and autoregression analysis allowing evaluation of the overall effects of the road safety strategy and the effects and significance of each individual program. The cluster analysis involved placing each road safety program into categories according to the broad nature of their road safety measure (publicity, education, legislation and enforcement). Each cluster was identified after consideration of strategic goals, measure characteristics and target population, and was given a description. The autoregression analysis included casualty rates for 12 dependent variables; four road user types (driver, passenger, motorcyclist and pedestrian) for each of three levels of crash severity (fatal, serious and slight). These casualty rates were in log scale to allow for temporal variations. Autocorrelation was taken into account in the regression analysis, and a multiplicative autoregression model based on the rate of change of the dependent variables over time was considered most appropriate. This type of evaluation did have the limitation of only qualitatively assessing whether certain road safety strategies might be effective in reducing crash rates and fatalities. There is no way of quantifying the effect of the strategies, and any results that show an indication of an effect warrant in-depth quantitative analysis.

A Cochrane Collaboration review of the effect of red-light cameras highlighted two methodological issues to consider when evaluating road safety programs (Aeron-Thomas and Hess 2005). The first of these was regression to the mean, whereby the location for an intervention usually has an increased frequency of crashes to begin with and thus would be expected to display a decrease regardless of whether an intervention was implemented or not. One method, which could be used to overcome regression to the mean, is to collect data for the pre-intervention period for five years. This would help to reduce the likelihood of obtaining unusually high crash rates in the one or two years pre-intervention. The second issue was a spillover effect, which occurs when behaviour is modified both at intervention sites and non-intervention sites possibly due to a media campaign or increased general awareness of the intervention. Results of statistical analyses need to be adjusted to account for these and perhaps other influencing factors, in order to have an accurate evaluation.

2.3.3.1 ESCAPE project

"Escape" is a pan-European project which aims to identify important areas of traffic and driver non-compliant behaviour and to assess the potential of enforcement tools. Much of the project is targeted towards speed and alcohol related crashes. Evaluation is used to ensure that time, money and resources aren't used unnecessarily on ineffective programs. The areas looked at in this report were selection of Performance Indicators (PIs), research design, data collection, sample design and data analysis (Gelau et al. 2000).

Selection of PIs - These could be categorised as indicators of enforcement activity, attitudes and behaviours, and crash statistics.

- Enforcement activity for speeding can be measured as automatic enforcement (number of cameras, number of camera working hours), or conventional enforcement (number of instruments, number of hours instruments were used, number of hours personnel worked). For alcohol, it could include number of drivers stopped, number of breath tests, number of hours personnel worked, and number and duration of large-scale controls.

- Driver's attitudes must be collected in appropriate surveys. The indicators should include the probability of discovery in case of infringement, perceived control intensity, attitude towards infringement, and acceptance of enforcement activities.

- Driver's behaviours can be measured either by surveys or observations, depending on the road safety issue of interest. For speeding, observations are often most appropriate and can be supplemented with self-report surveys, whereas for drink-driving the proportion of drivers exceeding the alcohol limit (at roadside breath tests) and self-reported degree of compliance would be most appropriate.

- Crash indicators include absolute crash numbers, degree of crash risk, ratios and proportion numbers.

Research design - True experimental designs are not feasible in road safety, hence quasi-experimental design is recommended. The variants of this method are with/without comparison, pre-/post-comparison, pre-/post-comparison with control group(s), and pre-/post- with/without comparison. Other designs include time-series analysis and investigations with multiple and repetitious application of the measure.

Data collection - Absolute numbers of crashes and crash rates should be available for collection. Behaviour and attitude data can be collected either by roadside surveys, telephone interviews or household surveys.

Sample design - The selection of a test or control area should take into account traffic, population, workplace structure as well as traffic density and road type. Size of the area and time of investigation should be sufficient to allow for at least 150 crashes both pre- and post-time interval. Simple random sampling should be observed for telephone surveys, whilst time-intervals should be proportional to traffic density in roadside surveys.

Data analysis - The likelihood-ratio-test is appropriate for testing statistical significance of absolute frequencies. For attitudes and behaviours, a two-way ANOVA with group type and investigation period as the independent variables should be used. With the observational data, more complex logit-models should be used to examine differences between proportions.

2.3.3.2 SUNflower project

SUNflower was a comparative study of road safety performance in Sweden, the United Kingdom and the Netherlands (SUN countries). The road safety policies in these three countries have seen them achieve the lowest number of road deaths per 100,000 population in the world. The aim of this project was to determine what makes the policies and programs in the SUN countries, particularly effective in dealing with the traffic safety problem. Secondary to this, a further aim was to identify policy improvements which may produce casualty reductions, both in SUN and other European countries.

The relative benchmarks selected for comparison, not only with each other but also for other European countries, were:

1. the nature and content of road safety plans and actions
2. PIs
3. final outcomes of road traffic fatalities in comparable terms (i.e. rates by exposure)

Ideally, each road safety initiative would be monitored and shown how it contributed to the reductions in crash/injury/death rates. Whilst the effects of many programs are monitored and documented, there are still many smaller initiatives that cannot be measured and may collectively change attitudes and behaviours significantly.

2.3.3.3 Finland

The Finnish Road Safety Programme 2005, published in 1999, described the Finnish National Road Administration's (Finnra) measures and objectives for road safety improvement (Finnish National Road Administration 1999). Rather than being a selection of specific actions, it consisted of a series of non-specific activities under a number of principles and objectives aimed at developing a methodology for an action plan. Following on from this document was the Road Safety Programme 2001-2005, which set out the measures (approximately 50) with which the achievement of targets could be met. It also identified a number of factors which were expected to impact on road safety outcomes, such as the ageing population, economic growth, individualism and social fragmentation, and technology. Annual summary reports of the progression of the Programme were planned, as was the systematic analysis of the various measures within the Programme.

2.3.3.4 Ireland

A review of the Irish road safety strategy identified that speeding, restraint non-use and drink driving were the three areas that could most improve their standard of road safety (Wegman 2002). Parts of the evaluation were difficult, if not impossible, because of the lack of clarity in whether action plans had been set up, whether they had measurable goals, whether they were implemented and also whether implementation had led to a change in PIs. Other factors which made it difficult to evaluate was the choice of 1997 as the reference year (baseline), where in fact it would be more accurate to obtain an average over three to five years. Furthermore, the strategy was to apply for four years (1998-2002), which is considered too short a time frame to achieve any meaningful reductions in road trauma.

2.4 REVIEW OF CURRENT METHODS OF PROVIDING EVALUATIONS AND ANALYSES WITHIN GOVERNMENT

2.4.1 Queensland

Unpublished resources give further insight into how current and previous road safety strategies and action plans have been evaluated. The previous "Queensland Road Safety Strategy 1993-2003" achieved many of its targets well before the estimated completion date. As a result, in 1999 a revised strategy was proposed along with a series of actions for immediate implementation. In assessing the current road safety climate, all crash types were ranked by size of involvement. Each crash type was investigated, largest first, and a range of options to address each problem was identified. These options were then assessed for and prioritised by their potential cost-effectiveness and other factors guided by the Strategy's principles (Leggett 1999). Following this process, a series of actions were selected and the "Queensland Road Safety Action Plan 1998-1999" highlighting the lead agencies for each action area, was formulated.

With an action plan implemented, the focus then shifted onto monitoring and evaluation. The types of feedback available varied across the different actions - some data was available within days (such as Police entered electronic data), but most was usually collected periodically (monthly, quarterly, biannually or annually). Crash information was fed into a model which assessed both expected trends and actual progress of each individual program and their outcomes.

Previous evaluations in Queensland have made use of a control-chart technique, and studied the effects of individual countermeasures by using a comparison group (Strategy Branch 1994). It was believed that this one method could be applied to any number of road safety interventions. Data was analysed over a six-year period prior to the road safety treatment. Linear regression analysis of that trend was used to predict the likely value of the subject/comparison ratio in future years given the absence of the treatment. This value was then compared with the observed subject/comparison ratio under the condition of the treatment. The observed ratio divided by the expected ratio and subtracted from one, was the point estimate of the effect size. A chi-square test was used for direct comparison of observed and expected counts. Dollar costs per crash were then assigned for a crash social cost analysis. The benefit-cost estimate was determined by the aggregate value of crash decreases divided by the cost of providing the treatment. Theoretically, the benefits of an individual countermeasure were in general nett of the effects of other countermeasures, hence the aggregation of the effects of all countermeasures could estimate the aggregate effect on road trauma of the program in general.

This control-chart methodology was assessed by Newstead (2000) to determine how robust the model was when compared with the log-linear analysis method (widely accepted for use in medical research in the analysis of case-control trials and cohort studies). Theoretical comparison of the two methods showed that the control chart method would be just as robust as the log-linear method provided the rates of change in both the control and treatment sites were roughly equivalent in the log space.

Two simulation studies were considered to assess the relative performance of the control chart and log-linear methods - one included a Poisson random error component to each simulated data cell, the other did not. Results showed that estimates of program effects were accurate using the control chart method until a difference of 15 percentage points was realised. When the growth rate difference between treatment and control series was greater than 15%, the accuracy of the control chart methodology was considerably less (Newstead, 2000).

The final test for accuracy was to use real crash data to estimate how often growth rate differences of greater than 15% occurred in practice. Crash data from the Queensland speed camera program was analysed using both methods. Results showed that both methods produced highly consistent estimates of crash effects whether using monthly or annual data (Newstead, 2000). These estimates were extremely close when aggregated annual data was used, more so than when monthly data was used. This suggests that the maximum level of data aggregation should be used when using the control chart method to ensure the highest level of estimation accuracy.

A further resource available for evaluation is the use of independent external evaluations. Recent examples of this in Queensland include the micro-level evaluations of the Random Road Watch program (Newstead et al, 1999) and the Speed Camera Program (Newstead and Cameron, 2003), and the global-level evaluation of the Road Safety Initiatives Package (RSIP) (Newstead et al, 2004). Evaluation of the RSIP in itself offers an indication of some of the processes required for road safety strategy evaluation. The road safety issues the RSIP targeted were drink-driving, speeding, fatigue and restraint use. Measures to achieve the goal included increased speed camera hours, increased hours of police enforcement of the four targeted areas, increased mass media publicity and increased hours of police educative activities. The evaluation examined the crash effects and economic impact of the program as a whole and also for the individual program elements. The methodology involved analysis of crash data, speed surveys, and attitude and behavioural surveys. Some of the outcome measures used were crash severity, date, time of day, police region and speed limit, speed camera hours of operation, sites used, RBT operations, seat belt offences detected, mobile phone offences detected and monthly awareness levels of mass media television advertising.

The RSIP evaluation is a good example of a global-level evaluation. A global-level evaluation has the capacity to include the impacts of initiatives which cannot be evaluated separately using individual program evaluation methods. Possible reasons why the impacts cannot be evaluated separately include:

- relatively low expected impact,
- relatively small target groups of crashes and
- the unavailability of a viable comparison group to take into account the influence of factors other than the RSIP or Queensland Road Safety Strategy.

The research in Newstead et al (2004) was able to provide estimates of crash savings for the RSIP as a whole as well as individually for many of the key initiatives in the RSIP. Hence the RSIP evaluation has demonstrated the potential to evaluate the initiatives in a road safety strategy both collectively and individually.

2.4.2 New Zealand

The current New Zealand "Road Safety Strategy 2010" is very much target driven, and guided by the principle of "safety at reasonable cost". The strategy targets risk factors across the four key contributors to road injury - exposure to risk, crash frequency, crash severity and severity of post-crash injury. Jeanne Breen Consulting (2004) undertook the first review of the strategy (Breen 2004). One limitation was the short time frame between initiating the actions within the strategy and completing the review, meaning significant evaluation of many initiatives was not possible. The review did however evaluate a selection of final outcomes and intermediate outcomes. The final outcomes were social costs of crashes, deaths and hospitalisations. These outcomes generally only needed comparison with the target or estimate, and were often represented as frequency per billion vehicle kilometres, per 100,000 people, or per 10,000 vehicles.

The intermediate outcomes were speed, alcohol consumption, restraint use and regional outcomes. Speed was detected in New Zealand by Hawk Radar, Lasers and speed cameras. Performance indicators included hours of speed camera operation, mean speed and 85th percentile for both open and urban roads, rural speed percentage over 110km/hr, offence notices issued and speed camera infringement notices. Alcohol-related performance indicators were number and percent of driver deaths with excess alcohol, number of compulsory breath tests (CBTs) and mobile breath tests (MBTs) over 12-month periods, offence notices issued, and hours to be delivered. Safety belt and child restraint use rates were reported, along with police targets for hours to be delivered and offence notices issued. Regional outcomes were shown for deaths plus hospitalisations, deaths plus hospitalisations over 1 day, and deaths plus hospitalisations over 3 days, and were compared with the national picture.

A series of working papers were developed to complement the strategy. Working paper 6 (LTSA 2000) described the methods used for predicting and costing road safety outcomes. Similar to the Queensland experience, the issue of how to combine the potential hundreds of individual results from a number of interventions on a number of traffic categories with various other factors was discussed. Within outcome categories, the effects of interventions were combined multiplicatively because they affect the same group of crashes. This was important to ensure that no prevented crash outcome was included twice in the target outcome. When interventions applied to different crash types (between categories), there was no chance of overlap hence the effects between groups were combined additively.

Experience from New Zealand in the last two decades shows the benefit of systematic monitoring and evaluation. The National Road Safety Plan (NRSP), developed in 1991, set targets based upon figures from 1990. These targets for deaths, injuries and hospitalisations were achieved easily by 1994. As a result, the 1995 NRSP set more challenging targets for the same outcomes to be achieved in the six years to 2001. Early indications suggested that these new targets would not be met without substantial changes to the road safety approach. In an effort to assist reaching these targets, a Supplementary Road Safety Package (SRSP) was implemented in the second half of 1995.

Evaluation of the SRSP took place after the first five years (1995-2000). The SRSP is an example of the benefit of systematic monitoring and evaluation to reassess where the strategy is heading, not only if it is effective or ineffective, but also in case targets are reached earlier than anticipated. The first part of the evaluation sought to identify how successful each of the initiatives were by key program area (drink driving, speeding and restraint use) in reducing road trauma. As the intent was to assess the overall effect of the SRSP, it was not necessary to disaggregate the data by type of intervention (Cameron 2002).

Along with the conventional sources of data (crash data, police data, hospital data), a number of different surveys have been used in New Zealand to evaluate drivers' attitudes and behaviours towards various road safety issues (Cameron 2002). Roadside alcohol surveys have been conducted annually since 1995, and have expanded from 9,800 drivers to over 20,000 drivers annually, to become nationally representative. An on-going speed monitoring system in urban and rural areas was expanded by almost 100%, to 70 nationally representative sites in 1995, capturing 34,213 vehicles. A further expansion to 170 sites saw approximately 60,000 vehicles captured in each of the 1996 and 1997 Winter driving surveys. Annual surveys have also been conducted since 1992 for occupant restraint use, and since 1990 for bicycle helmet use. The LTSA implemented an annual survey of public attitudes towards road safety in 1995, covering issues such as drink-driving, speeding, safety belt non-use and enforcement. These surveys were conducted face-to-face with approximately 1,640 people aged 15 and over in 14 regions of New Zealand, and were considered nationally representative. One final survey was the quarterly surveys of public recall and reactions to the road safety television advertisements.

2.4.3 Western Australia

A number of unpublished documents describe the evaluation and monitoring of the 2003-2007 Western Australian road safety strategy "Arriving Safely". One report identified a range of performance indicators including outcome measures, intermediate measures and process measures, and gave the rationale behind the selection of each one (Haworth and Vulcan 2000). Some of the sources and examples of data collation included the Police (crash statistics, RBT operation, speed camera/radar operation), Office of Road Safety (road safety surveys - observational and community), Main Roads (crash statistics, speed data), Australian Bureau of Statistics (population demographics) Local Government (expenditure on road safety), Hospital admissions (injury & severity) and Transport Licensing (new car sales and safety features).

Two types of models may be deemed appropriate for performance monitoring and evaluation (Haworth and Vulcan 2000). The first of these is a hierarchical model with an overall indicator being the "bottom line". Lower level indicators seek to explain the changes in the overall indicator. This model is best suited for communication to those with little technical understanding. The second model is more of a "synthetic" model, where the focus is not on a single overall indicator, rather there are a series of component indicators which evaluate the various components of the strategy. An example of this is multivariate log-linear regression, which relates measure of road safety and economic effects with general trend and monthly variation, to the observed road trauma series via a regression equation (Newstead et al. 1995; Newstead et al. 1998). A further example is the Poisson regression model, seen in the evaluation of the Queensland RSIP program (Newstead et al, 2004). This "synthetic" model appears to be more robust and better suited to rigorous evaluation, however it may be difficult to interpret for those with limited technical understanding.

2.5 CONCLUSION

Development of a framework for evaluating a road safety strategy is a consultative process. The objectives of the strategy and the various agendas of stakeholders often guide the manner in which data is recorded, collected and analysed. Some of the issues that need to be considered in the analysis are regression to the mean, long-term trends, economic growth/decline, population growth, and increased motorisation. Evaluation can take the form of process evaluation, outcome evaluation and impact evaluation. Selection of the appropriate performance indicators (and their units of measurement) is critical to the evaluation.

Evaluation needs to be undertaken at a micro level (to assess the effect of any individual road safety initiative) as well as at a macro level (to assess the general effect of a group of initiatives) to ensure that the effect of all road safety initiatives within the strategy is measured both collectively and individually. There is also a need for global-level evaluation in which the individual effects of many (if not all) of the initiatives are modelled collectively, in order to assess the effects of those initiatives which cannot be subjected to micro-level evaluation for a variety of reasons.

2.6 RECOMMENDATION

Based on the literature review of the existing practices detailed in this report it is recommended that the most appropriate evaluation framework to apply to the 2004-2011 Queensland Road Safety Strategy will be that outlined in section 3.1; i.e. the GOSPA model.

The GOSPA model relates evaluation to a strategic planning framework. The Queensland Road Safety Strategies for both 1993-20003 and for 2004-2011 have structures similar to that described by the GOSPA model. To a large extent, the 1993-2003 and 2004-2011 strategies and action plans are structured in terms of an overall goal, objectives, strategies, programs/plans, actions and targets (at each level).

The GOSPA model allows an evaluation structure to be defined. Programs are not considered to be implemented unless actions are taken, strategies are not achieved unless planned programs are implemented, objectives are not met unless the strategic directions of programs are correct, and goals are not achieved unless the targets of objectives are met. Through this strategic framework, the key pathways through which each action, program and strategy contributes to the overall goal becomes apparent. Measurable criteria at each level provide the basis for assessment that real change has occurred, and potentially developing linkages to the overall goal through modelling of the linkages.

Whilst this method of evaluation was developed for a specific enforcement initiative to expand Western Australia’s speed camera operations, its key role in funding other initiatives in their Integrated Road Safety Program led to an evaluation framework which embraced the full road safety strategy (at least conceptually). Hence this approach is suggested as being potentially relevant to evaluation of Queensland’s Road Safety Strategy.

3. DATA SOURCES AVAILABLE FOR EVALUATION

To develop a reliable evaluation model to estimate the expected levels of road trauma there is a need to have a good understanding of the ways in which certain factors affect trends in road crashes. These factors may include road safety initiatives, Police enforcement, and economic, social and environmental factors. When these factors are taken into account a more reliable model estimating the expected levels of road trauma can be developed. Once an explanatory model has been developed, a comparison of the actual crash levels with the expected levels can be made.

To complete the third and fourth work tasks of the project it was necessary to obtain relevant road safety program data, socio-economic data and other data from the appropriate agencies. The following sections give a list of the data requested and obtained from various sources.

3.1 ROAD SAFETY PROGRAMS

3.1.1 Police Traffic Enforcement Data

Road Safety Program data relating to on-road traffic enforcement operations including speed camera data, seat-belt and mobile phone offence data and drink-driving data was received from the Queensland Police. MUARC requested that the data supplied should cover the period January 1993 (or as close to this date as possible) to the most recent data available. The data supplied covered the period January 1993 to March 2006. MUARC placed another request to Queensland Police for further enforcement data going back to 1993. However not all of this data could be provided.

The Police traffic enforcement data received is listed as follows:

- Number of **Random Breath Tests** (RBTs) conducted per month for each police region and for each type of Test (e.g., Booze Bus, Other Stationary, Mobile, Incident/Infringement) for the period January 1997 to March 2006.

- Number of **reported drink-driving offences** detected per month for each police region for the period January 1997 to March 2006.

- Number of **drink driving infringement notices issued** per month for each police region for the period January 1997 to December 2003.

- Number of speeding offences detected by **Mobile Radar** per month for each police region for the period January 1997 to March 2006.

- Number of speeding offences detected by **LIDAR (laser)** per month for each police region for the period January 1997 to March 2006.

- Number of hours of **Mobile radar operations** conducted per month for each police region for the period January 1997 to March 2006.

- Number of hours of **Laser operations** conducted per month for each police region for the period January 1997 to March 2006.

- Total number of hours of all types of traffic enforcement per month for each police region for the period January 1997 to March 2006.

- Number of seatbelt and mobile phone offences detected per month for each police region for the period January 1997 to December 2003 (Note: incomplete data received for 2004-2005).

- Number of seat belt and mobile phone infringement notices issued per month for each police region for the period January 1997 to December 2003.

3.1.2 Speed Camera Operations

The following speed camera activity data per month and per region for the period May 1997 to March 2006 was received by MUARC:

- Number of speed camera hours
- Number of speed camera detections
- Number of speed camera deployments
- Number of speed camera sites available
- Number of sites visited in compliance with scheduler
- Percentage of sites in compliance with scheduler.
- Number of speed camera traffic infringement notices (TINs) issued.

3.1.3 Seat belt and Mobile Phone Infringement Data

In addition, the following

- Number of seatbelt and mobile phone offences detected per month for each police region for the period January 1997 to December 2003 (Note: incomplete data received for 2004-2005).

- Number of seat belt and mobile phone infringement notices issued per month for each police region for the period January 1997 to December 2003.

3.1.4 Road Safety Advertising Data

In addition to Police enforcement data, mass-media publicity that accompanied the enforcement was considered as a potential explanatory factor. The publicity can cover a variety of themes, namely speed, drink-driving, fatigue, seat belts and concentration. The road safety publicity data is often quantified in terms of television exposure, measured weekly using Target Audience Ratings Points (TARPs). The advertising awareness of television viewers can then be measured by a function of TARPs called “Adstock” (Broadbent, 1979) which represents the audience’s retained awareness of current and past levels of advertising.

MUARC had previously received TARPs data covering various road safety themes (e.g. fatigue, seatbelts, fatal four, speed, drink-driving) for an earlier evaluation conducted for Queensland Transport. This road safety advertising data covered the period 1998 to 2002.

For this current project data was requested in a similar format to that provided earlier for the periods 1993-1997 and 2003-2004. However it was not readily available.

3.2 SOCIO-ECONOMIC DATA

A variety of other factors (apart from enforcement data and supporting publicity) can be used to provide links with road trauma. These factors, primarily economic and social factors, include labour force data, population data and fuel sales.

The socio-economic data requested and obtained from various agencies in Queensland included population, fuel sales and labour force data as follows:

3.2.1 Labour force data

A large number of variables exist which may reflect the state of the economy, including real average weekly earnings, unemployment rates, number of persons employed in the work force, building approvals and consumer sentiment index.

Most economic data is provided on an annual or quarterly basis, and either at a national or state level. For the proposed analysis in this study, data at the monthly level was requested and obtained as follows:

Monthly time series data by region for the period September 1992-December 2005 consisting of:

- Employed persons (part-time, full-time);
- Unemployed persons (full, part-time and total);
- Those not in the labour force;
- Participation rates and unemployment rates.

3.2.2 Fuel Sales

Total fuel sales can be used as a proxy for vehicle travel in studies that require some measure of exposure to risk. For this study, monthly fuel sales by fuel type for Queensland were obtained covering the period January 1986 to September 2005. Regional fuel sales data was not available.

3.2.3 Population Data

Population numbers can be considered as a measure of exposure to risk. The following population data was received from the Office of Economic and Statistical Research, Queensland Treasury:

- monthly civilian population data by Region for the period September 1992 to December 2005;

- estimated annual resident population by region, district and Police Division for the period 1991 to 2005.

3.3 OTHER DATA

Other data deemed suitable for evaluation included data obtained from speed surveys, as well as attitudinal survey data.

3.3.1 Speed Survey Data

MUARC requested speed survey data, to compliment existing data sets already held for November 2002, March 2003 and August 2003.

Queensland Transport (MainRoads) provided MUARC with monthly speed survey data (i.e. mean speeds and percentiles) by site, speed limit, LGA and statistical division for the following periods.

- March 2004
- August 2004
- March 2005
- August 2005
- March 2006.

3.3.2 Attitudinal Survey Data

The following WAVE and RSPAT reports were received by Muarc. These reports measure road users' attitudes to a variety of road safety issues in Queensland.

Wave 1	Aug 1997
Wave 2	Feb 1998
Wave 3	Sept 1998
Wave 4	April 1999
Wave 5	Nov/Dec 1999
Wave 6	Aug/Sept 2000
Wave 10	Dec 2003
Wave 7	
Wave 8	
Wave 9	
Wave 10	
Wave 11	May 2004
Wave 12	Feb 2005.

RSPAT 1	May 2001
RSPAT	Feb 2002
RSPAT speed	2002
RSPAT cycle & ped	2002
RSPAT	2003
RSPAT cycle final	2003
RSPAT	2004
RSPAT final	2005.

3.4 CRASH DATA

Police reported crash data for crashes occurring on Queensland roads was received covering the period January 1991 to December 2004.

The crash data files included the following variables:

- crash date;
- crash severity;
- time of crash;
- Police Division;
- Region,
- DCA and
- Speed limit.

4. DEVELOPMENT OF AN EVALUATION FRAMEWORK FOR THE QUEENSLAND ROAD SAFETY STRATEGY 2004-2011 AND ASSOCIATED ACTION PLANS

4.1 The GOSPA framework and proposed Methodology

Based on the review of the existing practices described in Chapter 2 it was suggested that a modification of the evaluation framework outlined by Cameron (1999) be considered for the Queensland Road Safety Strategy. This framework envisaged that the Queensland Road Safety Strategy would be structured in terms of an overall goal, objectives, strategies, programs/plans, actions and targets (at each level), as outlined above. To some extent, the 1993-2003 and 2004-2011 strategies and action plans already match that structure, perhaps with different terminology.

Ideally the aim of the evaluation framework should be to provide information on the extent to which:

- each action contributes to the relevant program meeting its objectives,
- each program contributes to the relevant strategy meeting its objectives, and
- each strategy contributes to the overall goal and its target.

The staged way in which actions contribute to programs, programs contribute to strategies, and strategies contribute to the overall goal are illustrated in Figures 1 and 2 in the Cameron (1999) paper given in Appendix A. Such staging is an inherent part of the evaluation framework outlined here.

It may be possible to show directly how a larger, more powerful action or program contributes individually to the overall goal, but many actions and programs address a relatively small issue and such a direct effect cannot be expected to be seen conclusively. That is why it is important to consider each action, program and strategy within a consolidated framework, and to develop evaluation criteria which are relevant and feasible for their proper assessment. The process of developing feasible evaluation criteria is outlined below. However, it should be noted that it may not be feasible to evaluate the Road Safety Strategy at all levels.

4.2 PROCESS OF DEVELOPING FEASIBLE EVALUATION CRITERIA

The GOSPA framework has been employed successfully by Monash University Accident Research Centre (MUARC) to develop a framework for several different evaluations. Such a model allows evaluation at the level of each action, program or strategy, and the respective contributions of each to the achievement of the overall goal (Cameron, 1999). The structure of the model is as follows:

Goal	General (idealistic) statement of the Program's overall goal
Objectives	Specific (pragmatic) statements of the Program's measurable objectives to reach the Goal
Strategies	General (idealistic) strategies to achieve each Objective

Strategic Objectives Measurable indicators of extent to which each Strategy has been achieved

Programs/Plans Specific (pragmatic) programs/plans contributing to each strategy, with measurable activity levels and outputs

Actions Actions undertaken in each plan

Action Targets Measurable indicators of extent to which each Action has been achieved

This method relates evaluation to a strategic planning framework. Strategic plans, typified by Queensland's Road Safety Strategy, are often structured according to the GOSPA frameworkl.

In the following sub-sections the Queensland Road Safety Strategy (and action plan) was formulated in terms of the GOSPA framework above where possible stating the Goals, Objectives, Strategies, Programs and Actions.

4.2.1 Goals and Objectives of the Strategy

The main vision of the 2004-2011 Queensland Road Safety Strategy is to prevent road trauma through safe road use, safe roads and safe vehicles. The Strategy's overall goals are to minimize crash severity and to reduce the long-term consequence of injuries. This general, idealistic statement defines the **G**oal of the GOSPA frameworkl in reference to the Queensland Road Safety Strategy.

The key targets of the strategy are:

- to achieve a reduction in the number of fatalities per 100,000 population to under 5.6 deaths per 100,000 people by the year 2011, and

- to achieve a reversal in the increasing trend in hospitalisation casualties and the hospitalisation rate.

The above targets define the Strategy's measurable **O**bjectives to reach the Goals of a minimization in crash severity and a reduction in the long-term consequence of injuries. Hence, these targets give the **O**bjectives of the Gospa frameworkl in relation to the Strategy.

4.2.2 Strategies, Programs and Actions of the Queensland Road Safety Strategy

The GOSPA framework then needs to formulate how the 2004-2011 Queensland Road Safety Strategy will achieve these objectives or targets. This will be achieved by defining the general strategies, specific programs and actions of the Strategy as follows:

Strategies**:**

To achieve the objectives of fewer than 5.6 deaths per 100,000 population and a reversal of the increasing trend in the hospitalisation rate the general (idealistic) strategies include a focus on the key issues; key at-risk behaviours and key at-risk road user groups relevant to Queensland road safety. The general **S**trategies to achieve each **O**bjective will be through:

- Safe attitudes and behaviours and optimal health outcomes in the event of a crash;
- Safe roads, safe road environments and safe management of traffic;

- Safe vehicles that reduce injury severity and maximise the chance of avoiding a crash
- A community that values road safety as a priority.

According to the 2004-2011 Queensland Road Safety Strategy the broad key target road user groups and behaviours needed to be targeted for evaluation are:

- Alcohol and drug-driving (which accounts for about 30% fatal crashes)
- Speed
- Fatigue
- Young adult inexperienced drivers
- Older road users
- Fatal and serious crashes occurring in rural Queensland
- Pedestrians
- Motorcycle riders
- Indigenous road users.

To assist in achieving the above target the Queensland Government has implemented four two-yearly action plans. The first of these was the Queensland Road Safety Action Plan 2004-2005, and consisted of new or modified existing initiatives to be implemented during 2004-2005. The action plan presents activities that directly address the key outcomes of the strategy.

The specific target groups and target areas as defined in the 2004-2005 Action Plan include:

- Roads with poor crash records
- Local roads
- At-risk road users (e.g. bicyclists on major roads; pedestrians and bicyclists on local roads)
- Drink-drivers
- Intoxicated pedestrians (emphasis on rural and indigenous communities)
- Fatigued drivers
- Speed-related crashes and speeding drivers
- Unrestrained occupants
- Crashes in rural and remote areas of Queensland
- Young drivers
- Unlicensed drivers and riders
- Older drivers
- Fleet-related crashes.

<u>**P**rograms and **A**ctions:</u>

Each action, program and strategy within the Queensland road safety Action Plan was reviewed to develop measurable criteria. These criteria included the extent of action taken, change in knowledge, attitude, behavioural intention or actual behaviour, and reduction in crash frequency and injury severity. It should be noted that for some low level actions and programs it was not feasible to associate them individually with specific behaviour changes or road trauma reductions. Other higher level actions, programs and strategies were associated with specific behaviour changes and crash reductions, but because of the relatively small target groups they address, it was not feasible to conclusively evaluate their impact on these criteria within time and cost constraints. It has already been recorded that Queensland's more powerful road safety initiatives addressing substantial target groups during 1993-2003 have individually (and collectively) been shown conclusively to reduce road trauma criteria. It is expected that statistically significant effects will continue to be demonstrable for the more powerful new initiatives during 2004-2011.

It should be noted that for some actions and programs it was not feasible or cost-effective to measure their effect on perceptual, behavioural or road trauma criteria, and that the evaluation criteria are limited to examining whether the action was taken or the program implemented (i.e. process evaluation only).

The specific programs contributing to the 2004-2005 Queensland Road Safety Action Plan are given in Table A (in appendix B). Where possible measurable activity levels and outputs have been stated.

This Table lists only actions given in the 2004-2005 Action Plan that were either

1. 'New Key Initiatives" or

2. "Modifications to existing key initiatives".

The 2004-2005 Queensland Road Safety Action Plan also listed other action types, that fall under the broad categories of i.e. Actions that set a platform for future key initiatives and "Actions that continue to deliver proven best-practice enforcement, education and engineering initiatives". However these are not included in the table because they are either future actions or continuing practices which do not as yet have feasible evaluation criteria attached to them.

The actions undertaken in each program are also given, including what was achieved. Where applicable the target crash groups are given. Behavioural and attitudinal changes are also given where appropriate.

4.2.3 The Three-Level Model Approach

The modelling process used to test-run the GOSPA framework was based on a three-level model approach. This consisted of:

The Global Assessment model or First Tier Model: This model relates to the broad **G**oals and **O**bjectives of the 2004-2011 Queensland Road Safety Strategy.

The Second Tier Model: This model gives the Objectives related to the strata defined by the Strategies and Programs of the Queensland Road Safety Strategy and/or Action Plan.

The Third Tier Model: The model gives the **P**rogram and **A**ction elements relating specifically to target group outcomes (e.g. the RSIP evaluation).

Greater detail of the modelling process is given in the next sections.

4.3 MODELLING OF KEY OUTCOMES: OVERALL AND BY KEY STRATA

This section discusses the theoretical foundation for undertaking modelling of data to evaluate the performance of a road safety strategy. Concepts discussed include risk, exposure, key outcomes and strata.

4.3.1 Broad Issues in Modelling: Risk and Exposure

4.3.1.1 Relationship between crashes, risk and exposure

When considering any effect on safety, a well-established conceptual framework is generally used where the total number of crashes are seen as the product of two elements (Haight, 1986), risk (the probability of a crash per unit of exposure) and exposure (a measure of situations that potentially lead to crashes: for drivers, often equated to distance driven):

Number of crashes = Risk x Exposure (1)

As can be seen from equation (1), the number of crashes can decrease because risk has decreased, or because exposure has decreased. Often road safety strategies need to be evaluated under the complex situation where exposure is increasing (people are driving more) but risk has decreased (due to the successful implementation of a road safety program). This example highlights the need to have some measure related to exposure available in the evaluation of a road safety program as a successful and effective program may not necessarily reduce the absolute number of crashes in the face of increasing exposure levels. Where a particular sort of exposure carries an unusually high risk (e.g., driving after drinking), a high priority needs to be given to measuring such exposure. Changes in these types of exposures can have a marked effect on crash levels.

Crash risk estimates are determined by dividing the number of crashes per unit time by the exposure per corresponding unit of time, often defined as the ‘opportunity to have a crash’. Exposure can be measured as the distance travelled on the road, the number of trips of travel, the number of driver licence holders, the number of registered vehicles, fuel consumption or the number of people in the population. The particular measure chosen for use in the evaluation framework will be dependent on how the goals of the road safety strategy are formulated. Crash risk estimates given in Chapter 5 have used the appropriate number of people in the population during each time period as the exposure measure, e.g.

Risk of a crash = (number of crashes) / (population).

This measure of risk is consistent with how the goals of the Queensland road safety strategy have been formulated in terms of crashes per 100,000 population.

4.3.1.2 Exposure measurement

Often, direct measures of exposure are not available and suitable proxies are used, such as the economic activity of the jurisdiction as indicated by unemployment rates (Cameron, 1997). A more direct measure that can be used is the quantity of fuel sold, reflecting the amount of driving being undertaken. The advantage of an economic measure over the more direct measure of fuel sales is that under good economic circumstances, certain types of higher-risk driving may increase (Cameron, 1997), including all discretionary travel, and particularly driving at night to attend social events, where drinking and driving may be more common. Another useful exposure measure, because of its ability to measure changes in levels of particularly high-risk driving, is sales of alcohol, indicating potential levels of drinking and driving.

A change in the amount of exposure is very common, as has been experienced in Australia by gradual increases in traffic over the last few decades. Changes in types of exposure are also fairly common. One such example is a movement away from car driving to motorcycle riding (which is inherently more risky than car driving).

As mentioned above, any evaluation of the effects of a road safety program or intervention needs to account for changes in risk or exposure that *would have happened anyway*, even if the program or intervention had not been initiated. Cameron (1997) provides a number of examples of such situations leading to gradual change in crash or casualty rates:

- Increased proportion of total travel on motorways
- Changes in travel mode mix away from unsafe modes
- Changes in the population age distribution, especially the decrease in the proportion of teenagers and young adults
- Changes in driver licensing rates, including the decrease in numbers of new licences and novice drivers, and the increased rate of licensing among women
- Other general road and transport infrastructure changes.

4.3.1.3 Identifying changes in risk

A successfully implemented effective road safety initiative will usually be associated with a reduction in risk. The objective of statistical models to evaluate road safety programs is to identify and quantify this reduction in risk associated with the program. A program may have an abrupt effect, where there is a sudden change in the road safety environment due to the program (see *Figure 4.1*, modelling the effects of the British seatbelt law), or there may be a more gradual change, reflecting the amount of effort (enforcement hours, publicity, etc.) being devoted over time.

Cameron (1997) provides the following guidelines:

> "If a program involves new legislation, then it may be appropriate for the extended models to use a term representing a change in level of the crashes (either long-term or short-term). If the program involves enforcement and/or mass media publicity, then appropriate terms would be monthly measures of the intensity of the enforcement (eg. hours of operation or number of offences detected) or the publicity (eg. total television rating points or number of radio broadcasts). This type of data on each new road safety program will need to be recorded systematically in the future (if it is not already) in order to make the explanatory models ultimately feasible, so far as representing the relative contributions of the programs."

4.3.2 Concept of the Key Outcome Model

At the highest level of evaluation, the performance of the strategy can be measured using an intervention model that measures change in aggregate in trauma levels for the situation where the road safety strategy is implemented as a package compared to a situation modelled in which the strategy is not implemented at all.

4.3.2.1 An example of types of casualty savings estimated by a model

Estimates of the overall crash effects of the Queensland RSIP resulting from fitting regression models are given in *Table 4.1*, from (Newstead et al, 2004). This shows how an overall program may be evaluated at several different severity levels of road trauma. Note that the confidence intervals show the range of values that the reduction is likely to have taken and the statistical significance provides a guide of how likely it is that the measured crash reduction was zero (a small value indicates that the true value of the reduction was very unlikely to be zero).

Table 4.1**:* ***Estimated Total Crash Reductions Attributable to the RSIP (from Newstead et al, 2004)

Crash Severity	Estimated Crash Reduction	Upper 95% Confidence Limit	Lower 95% Confidence limit	Statistical Significance
Fatal + Hospital	13.12%	6.09%	19.62%	0.0004
Medically Treated	14.20%	7.91%	20.06%	<.0001
Other Injury + Non-Injury	4.34%	-0.36%	8.83%	0.0693
All Crashes	8.80%	5.52%	11.96%	<.0001

NB: Negative percentage crash reduction estimates indicate an estimated percentage crash increase.

4.3.2.2 An example of a State Space model used to model casualty rates

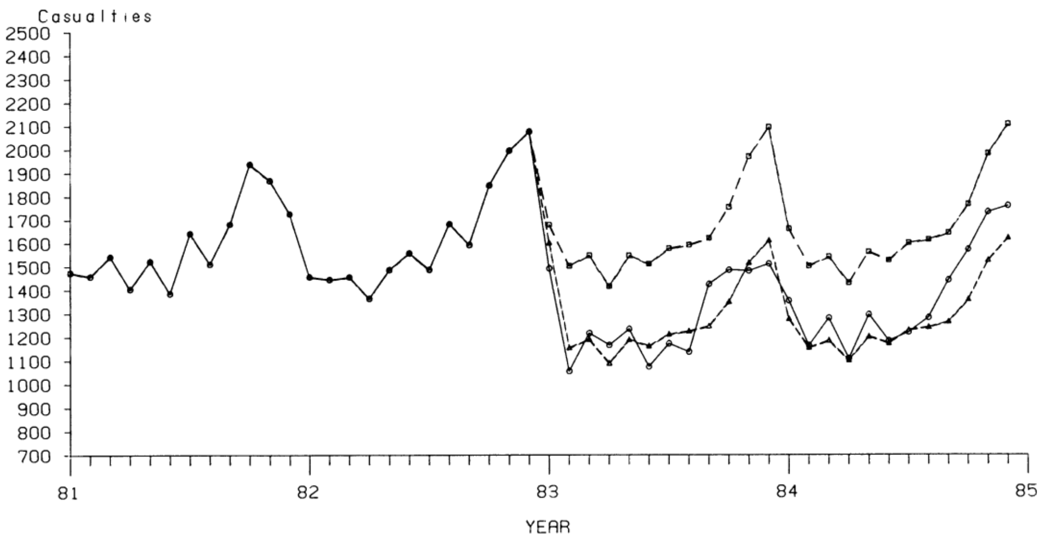


Figure 4.1: From (Harvey and Durbin, 1986): Car drivers killed and seriously injured, o, actual values of series; Δ, predictions with intervention effect included; □, predictions with intervention effect removed (top line). The seatbelt law was introduced at the beginning of 1983.

Figure 4.1, from one of the first papers to utilise Structural time series (State Space) modelling methods (Harvey and Durbin, 1986), shows UK data with predictions of casualties for 1983 and 1984 modelled under two circumstances: seatbelt law enacted and seatbelt law not enacted. These were obtained using observations of the explanatory variables for 1983 and 1984 but not using observations of the car drivers killed and seriously injured (KSI) series itself for 1983 and 1984. As can be seen from this figure, the bottom two lines, representing the actual numbers killed and seriously injured as well as the number predicted by the model are quite close to one another, showing that the model represents reality well. Explanatory variables included in the model were petrol prices and traffic density to explain variations in exposure levels over time. At the beginning of 1983, a sudden change in risk can be seen corresponding to the increased rate of seatbelt wearing.

4.3.2.3 Estimating effects on casualties and crashes

An estimate of casualties saved can be made from a model such as that presented in *Figure 4.1* by summing over the relevant time periods the total number of casualties predicted for the situation where the intervention is not implemented (the top line of *Figure 4.1* from 1983 onwards) compared to the predicted numbers affected by the intervention (the lower dotted line of *Figure 4.1*). When estimating the effect on crashes and casualties of an entire strategy, a model can be fitted to predict the numbers of crashes and casualties for the situation where no strategy was implemented, which may include several measures of the type evaluated in *Figure 4.1*. This can then be compared to the prediction made given the strategy was implemented. Such a model can then produce estimates such as shown in *Table 4.1* for the RSIP (Newstead et al, 2004).

4.3.2.4 Identifying outcome measures

The above example uses the number of drivers killed and seriously injured as the outcome measure. Other outcome measures may be specified in a road safety target, such as "drivers killed per capita". By including a denominator for number of crashes or injuries (per unit of population in this example), the target allows for growth in population that might otherwise lead to higher casualty rates, even in the face of effective measures to reduce road crash and injury risk.

Table 4.1, from Newstead et al. (2004), estimates the effect of the RSIP package of road safety measures in terms of changes in crash and trauma levels (the outcome measures) of different severities. As can be seen from this table, the largest effect of the RSIP was on medically treated injury crashes, compared with a fairly modest reduction in minor and non-injury crashes. This is a not unusual pattern for a successful road safety program that targets aspects such as excessive speed and alcohol consumption by drivers: a reduction in speed can reduce the forces involved in a crash so that an injury that would have been more serious becomes minor. When assessing a strategy or individual programs, a further outcome measure that may be worthwhile modelling is the total number of fatal and serious injury crashes as a proportion of the total number of crashes of all severities. This accounts more specifically for the phenomenon of reduced injury severity just discussed.

4.3.3 Overall model versus stratum specific models

Several programs that form part of the Queensland 2004-2005 Action Plan are intended to have most effect on specific strata. For example, the plan to provide at least 170,000-190,000 police officer hours towards random breath testing over each year of the Action Plan is likely to affect mainly crashes and casualties during high alcohol times. Another program element of the 2004-2005 strategy involves trialling the re-introduction of L-plates on all vehicles driven or ridden by holders of learner licenses. This program element is targeted at learner drivers, who are largely young drivers, and can potentially have an effect on both young driver crashes and a comparatively small effect on other crashes and casualties involving young drivers.

4.3.3.1 Relation between stratum specific effects and overall effects

This subsection discusses how stratum specific effects (the second level in Figure 4.2) can be combined and how they relate to overall estimates of strategy performance (the top level of Figure 4.2). These are important issues when the performance of the strategy is assessed using more than one model. Models fitted at different levels are necessary to provide different insights into the performance of the strategy.

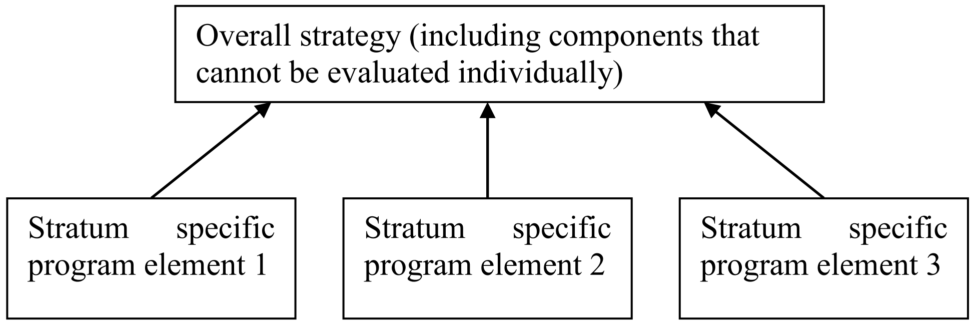


Figure 4.2: Diagram showing program elements that feed into the overall strategy. A program element may be "to provide at least 170,000-190,000 police officer hours annually towards random breath testing". Program elements target particular road safety issues. The effects of the strategy overall include the effects of the individual program elements able to be assessed, as well as the effects of program elements that cannot be readily assessed individually.

When program elements overlap in their effects, there are questions about whether these effects are experienced multiplicatively or additively. For example, does a 5% reduction in high alcohol hour crashes and a 10% reduction in young driver crashes suggest a 15% (=5% + 10%) reduction in young driver high alcohol hours crashes? In general, changes in risk levels are considered to operate multiplicatively such that a change in risk has a *proportional* effect on the number of crashes (Newstead et al, 1995). Using this example, a 14.5% reduction in young driver high alcohol hours crashes is the multiplicative effect of the 5% and the 10% reduction. i.e. (1 – (1-.05)*(1-.10)) = [1 – (.95*.9)]= [1-.855]*100 = 14.5%.

Newstead et al., (2004) evaluated the Queensland RSIP using models fitted at two levels:

(i) An overall effect was estimated by an "Intervention model" that measured the overall effect of the RSIP on crashes. This model measured the RSIP effect after adjusting for the effects of changes in socio-economic factors and road safety programs other than the RSIP on crash outcomes as well as accounting for seasonality and long term trends in the crash data.

(ii) The second model, the "Program Component Effects model", replaced the RSIP intervention term with explicit measures of RSIP program components. This model aimed to measure the relative effects of the RSIP component activities on crash outcomes.

A significant proportion of the total crash savings estimated by the Intervention model (see top level of Figure 4.2) attributable to the RSIP program was unexplained by the aggregate effects of the individual program components found to have a statistically significant effect estimated by the Program Component Effects model (see second level of Figure 4.2). This suggests there were one or more RSIP components, other than those identified, that led to substantial crash savings, which could include those for which explicit measures were available but were not significant in the analysis model. There were also likely effects due to unmeasured program effort such as publicity generated through program launches or enforcement blitzes. This is a good example of the additional insights provided by fitting models at different levels. The overall effect is best measured at a high level, which nevertheless fails to show the relative contributions of the programs. Lower level models measure the effects of individual programs and associated effort, but fail to account for the full effects of the strategy.

4.3.4 Defining the Key Strata from the Current and Previous Strategies

Key strata are identified according to the way that the current and previous strategies are targeted. For example, the program element "to provide at least 170,000-190,000 police officer hours towards random breath testing over each year of the Action Plan" of the 2004-2005 strategy is likely to affect mainly crashes and casualties during high alcohol times. The strategy thus identifies high alcohol crashes as an important stratum that should then be used in the analytical framework of the evaluation. An evaluation approach that uses time series also needs to account for the effects of previous and existing road safety measures and strategies in order to estimate the effects of a new strategy over and above the effects of an older strategy. Thus, strata identified in previous strategies also need to be accounted for in the analytical framework.

4.4 DEFINING INPUT VARIABLES FROM PROGRAMS

Explanatory models need to relate program activity to observed crash outcomes through the statistical modelling process. To achieve this, it is necessary to have measures of activity associated with the programs, including data to indicate when the program's effects would first be expected to influence crash rates and any measures of intensity (e.g., police hours of enforcement) that would lead to larger effects on crash rates during periods of greater intensity. The modelling methods are described in greater detail in Section 5. Table 4.2 shows some examples of programs with the appropriate input variables and strata for evaluating the effects of the programs via statistical models. A limitation for defining strata optimally is data quantity and availability. Clearly, only strata for which data can be obtained are able to be used. Also, the cross-classification of data by numerous strata can result in some sparsely populated cells, which can be difficult for the analysis methods to handle.

Table 4.2: Examples of programs with appropriate input variables to be used in explanatory models; also key strata for categorising outcome measures

Examples of programs	Input variables	Key strata
Implement best practice planning and design of cycling and pedestrian facilities for local roads.	Number of facilities with these features per time period	Pedestrian and cyclist injuries; local roads
Provide at least 170,000-190,000 police officer hours towards random breath testing over each year of the Action Plan.	Number of tests done per time period; number of hours of enforcement done per time period	High alcohol hours crashes; crashes specific to region and hours of enforcement specific to region
Advertising campaign to promote the dangers of driving tired, or without due care and attention.	Measures of publicity intensity (e.g. TARPS) per time period; advertising air time per time period NOTE: TARPs do not measure awareness, Adstock does	Region level crashes if exposure to advertising different per region
Undertake key enforcement activities such as the delivery of at least 43,800 hours of speed camera activity	Enforcement hours per time period	Crashes specific to region and hours of enforcement specific to region; low alcohol hours crashes
Introduce new penalties and sanctions for non-use of restraints.	Date when penalties were first publicised; hours of dedicated enforcement per time period	Crashes specific to region and hours of enforcement specific to region
Trial the re-introduction of L-plates on all vehicles driven or ridden by holders of learner licenses.	Date of introduction of measure	Young or learner driver crashes

4.5 EVALUATION OF SPECIFIC PROGRAMS

Evaluation of specific programs is outside the scope of the evaluation framework being described. However, it is essential that specific programs are evaluated thoroughly for the following reasons:

- Although the effects of specific programs are included in the estimates of the higher level modelling described above, the effects specifically attributable to a program cannot be evaluated separately.
- The higher level approach is not effective in providing scientific evidence of program effectiveness. There needs to be detailed assessment of the

implementation of the program to accompany the analysis. More effective implementation in one region than another, for example, should lead to larger crash effects.

- Major programs need specific evaluations, e.g. Random Road Watch (Newstead et al, 1999b), the Queensland Speed Camera Program (Newstead and Cameron, 2003, Newstead, 2004, 2005, 2006), the Queensland 50 km/h evaluation (Hoareau et al, 2007; Hosking et al, 2007). See Table 3. These evaluations included assessment of the implementation, as mentioned above. In the light of international experience, changes in implementation were recommended that could improve the effectiveness.

- The results of these specific program evaluations can be fed into the higher level models to generate more accurate results for the evaluation of the strategy as a whole.

These evaluations are carried out in much more detail than the levels of modelling described above as:

- the effects of the specific program are estimated;

- the analysis separates out the effects of the specific program from the effects of other programs; (However if no control or comparison group is available this may not be possible).

- a process evaluation (an assessment of how well the program was implemented) is integral to the evaluation of the program;

- there are greater demands on data, in particular data relating to intensity of effort (e.g., enforcement hours specific to the program) or intermediate measures (e.g., for a program intended to reduce speeds, measures of vehicle speeds);

- usually, a control area or crash type needs to be identified that is unaffected by the specific program that can be contrasted with the treatment area or crash type; (However if no control or comparison group is available this may not be possible).

- a cost-benefit analysis is desirable as major programs involve major effort (cost), which needs to be justified;

- analysis is made of "lessons learnt" that may improve the implementation or effectiveness of the program in the future or for other jurisdictions that may wish to adopt the program, as mentioned above.

Table 4.2: Examples of specific program evaluations in Queensland

Program	Reference	Design	Data requirements	Estimated reduction
An evaluation of the 50km/h default speed limit in regional Queensland	(Hosking et al, 2005)	Time series analysis of intervention (50km/h areas) with control areas (60-70km/h)	Crash data; Speed survey data; separate analysis for young drivers, older drivers, pedestrians	13.5% for all crashes; 19.3% for fatal, serious and medical attention crashes combined
Evaluation of crash effects of the Queensland speed camera program	(Newstead and Cameron, 2003)	Time series analysis of intervention (areas within 6km of the camera site) with control areas (areas further than 6km)	Crash data inside and outside speed camera zones; number of speed camera zones, sites, site density and hours of operation by police region	32% reduction in fatal crashes, a 26% reduction in fatal to medically treated crashes combined and a 21% reduction in all reported casualty crashes
Evaluation of the Queensland Random Road Watch Program	(Newstead et al, 1999b)	Analysis of treatment and control site crash rates before and after intervention, stratified into rural/urban and region	Operational details of program; crashes by strata defined and by severity level	11% crash reduction in aggregate; 13% for serious injury crashes; 9% for damage only crashes

Table 4.3 shows examples of evaluations of specific programs. In each of these cases, the design of the study and the analysis arose out of an understanding of the way each program was implemented (a process evaluation) and used appropriately identified control areas and outcome criteria. It is this attention to the detail of each program, together with the additional demands on relevant data, that make specific evaluations necessary. In each of these cases, cost-effectiveness analyses were also undertaken, showing that the resources dedicated to each program were more than justified by the benefits in terms of crash and injury reductions.

This section describes the modelling methods proposed to evaluate the road safety strategy and the levels of expertise required to fit the models and monitor the results.

4.6 Modelling methods TO EVALUATE THE STRATEGY

4.6.1 Three-tiered approach to strategy evaluation

The evaluation of the road safety effects of initiatives due to the implementation of the road safety strategy can be considered in three tiers:

1. **Global level evaluation** of the strategy overall.
2. Effects for **specific road user groups** or situations (e.g., pedestrians; young drivers; high alcohol crashes)
3. Evaluation of **individual program elements** (e.g., effects of speed cameras)

Figure 4.3 shows how the three levels relate. Individual programs potentially affect specific road user groups (typically high-risk groups identified by the strategy), but also have effects on road safety generally. Both the individual programs and effects of initiatives targeted to specific groups feed into the achievement of the overall target. Each of these three levels needs to be evaluated to adequately assess the performance of the road safety strategy.

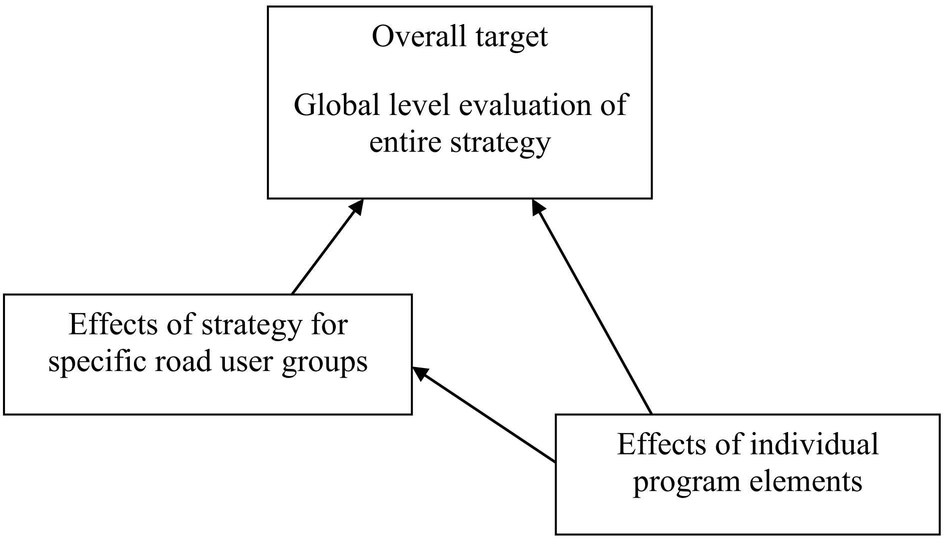


Figure 4.3: Diagram of three-tiered approach to strategy evaluation

4.6.2 Structural time series (State Space) models

Time series models with stochastic trends are a more flexible family of models than econometric or traditional 'fixed-effects' models in that they allow certain parameters to have stochastic (random) variation from one period of observation to the next. Stochastic trend models are often defined, using the 'state-space or structural time series' model. The name comes from the representation of the model in 'state-space' form, consisting of:

- *A measurement* equation, representing the observed 'state' of the dependent variables as a function of the explanatory factors; one part of the explanatory factors being unobserved model components, representing level, trend and seasonality effects in the dependent variable itself; the other part, factors representing interventions and other explanatory variables, and

- *Transition (or state)* equations, representing the model structure, describing the unobserved model components at time *t* as a function of their values on time *t-1*.

The essential feature of time series models, represented by the transition equation, is that they use the fact that the state of a developing process (in this case the number of crashes) at time *t*, is (partly) dependent on the state of the process at time *t-1*.

Thus structural time series models retain the strength of econometric models in that they allow a wide variety of functional forms and factors to be used to explain variations in road trauma series. They avoid the rigidity of ARIMA models and allow factors other than time to be included as explanatory terms. Indeed, structural time series models can be thought of as a generalisation of ARIMA type modelling with ARIMA models representing a specific form of the structural time series model.

Historical Background

Structural time series models were probably first used in a road safety context by Harvey & Durbin (1986), who used the-then new approach to evaluate the effects of the British seat belt wearing legislation on road casualties. They modelled monthly casualties from 1969 to 1984 using a state equation including level, trend, seasonality and an intervention effect due to the law.

Scuffham (1998) developed state-space models for quarterly variations in fatal crashes and fatal crash rates per kilometre travelled in New Zealand from 1970 to 1994. His models included terms representing level, trend, seasonality, interventions (speed limit increase, oil crises, seat belt law) and a range of socio-economic variables (e.g. unemployment rate, income per capita, beer consumption per capita, and alcohol and petrol tax rates). The most satisfactory models were those obtained for the fatal crash rate series rather than the frequency of fatal crashes.

The Dutch SWOV Institute for Road Safety Research has experimented with the application of structural time series models to data on fatalities and hospital admissions (two sources: police reports and hospital records) of car occupants aged 25 to 49 at the quarterly level (Bijleveld & Oppe, 1998). The models appear to be very adaptive to short-term changes in the casualty series and hence appear suitable for short-term forecasting (one to two years). Structural time series models have the advantage that it is possible to decompose the model into components representing the separate contributions of each term in the model. This attribute is valuable if the models are established for explanatory purposes.

In summary the key features of structural time series modelling techniques that make them potentially more powerful for the purpose of developing models of crashes are:

(i) The parameters of the 'structural time series' models can have stochastic variation that can represent real variation over time. The 'fixed-effects' models traditionally used by MUARC to date may suffer by their implicit need to average the estimation of the model parameters over the full period of the data.

(ii) The structural time series models are capable of representing time-series data, which is also cross-sectional in nature, in one unified model structure.

Structural time series models (with random effects) are thus considered to produce better estimators of time series data than traditional 'fixed-effects' models.

Form of a structural time series or state-space model

The 'state-space' model as applied consists of the following components, which can be used to explain monthly trends in crashes:

- a level term, representing the true number of crashes after removing seasonal and random variations from the data;
- a slope term, which tells us how much the level is changing;
- time-dependent components (i.e. seasonal and/or trend terms, accounting for variation in crashes due to month-specific effects);
- explanatory terms, which may be variables representing levels or quantities (e.g. enforcement variables, socio-economic factors
- an intervention component where necessary (e.g. an impulse function and/or a step function)
- a noise(error) term, representing random fluctuations around the true level.

A general form of this type of model may be written as:

$$y_t = trend_t + slope_t + seasonal_t + explantory_t + observation\_error_t .$$

In this general form, we have not specified how the unobserved components will evolve over time. A popular way of introducing time variation is to allow these components to evolve as random walks:

$$trend_t = trend_{t-1} + slope_{t-1} + trend\_error_t ,$$
$$slope_t = slope_{t-1} + slope\_error_t ,$$
$$seasonal_t = seasonal_{t-k} + seasonal\_error_t ,$$

where k=12 for monthly data or k=4 for quarterly data. This model, with trend, slope and seasonal terms is referred to as the Basic Structural Model (BSM), and is applicable to a wide range of data. The Kalman filter is used to calculate the error terms and produce maximum likelihood estimates of the unobserved components. The likelihood function is maximised over the variances of the error terms, so that for the BSM, we must estimate 4 different parameters (assuming no explanatory variables are used) – variances of observation_error, trend_error, slope_error and seasonal_error. By allowing for random variation in the trend and seasonal components, a structural time series model is very flexible.

The above equation gives the traditional form of a state-space model. A recent trend model used in work form the OECD by Moauro and Savio (2001), is a particular expression of the general state-space modelling framework further validating its use in this context.

4.6.3 Evaluation of programs and interventions using structural time series models

Instead of modelling the road trauma time series data explicitly as a function of road safety variables, the critical underlying components of a historical trauma time series can be represented and identified accurately. Features of the identified underlying estimated components can then be related broadly to the introduction and operation of key road safety programs. The simple model described above provides a framework for answering the following questions:

- How do current crash levels compare with those of the past? Examining the level term yields insights.
- Do we expect to see increasing, decreasing or steady crash levels based on trends in the data prior to the strategy being implemented? Examine the evolution of the slope term over time.
- How confident are we of our forecasts? The magnitude of the irregular term tells us how much error we can expect in our forecasts as represented by confidence limits on the forecast values.
- How effective have past programs been at reducing road trauma? Intervention analysis can be used to test for changes in the level, slope or seasonal terms, which correspond to certain programs.

4.6.4 The role of explanatory variables in evaluating the Queensland road safety strategy 1993-2003

The following subsection identifies important aspects of the road safety strategy 1993-2003 and variables available for evaluating these aspects in structural time series models. In order to relate program activity to observed crash outcomes through the statistical modelling process, it is necessary to have measures of the activity associated with the program. The activities and measures of activity or intensity are described below. Measures of activity were collected under a number of broad program component areas with the data chosen largely dictated by what was reliably collected by the relevant authorities. Also described below are the outcome data used (crashes of different severities) and stratification used.

4.6.4.1 Description of Socio-Economic Factors as covariates

Changes in socio-economic factors are known to have effects on observed road trauma. It is necessary to include measures of socio-economic effects in the statistical models to accurately describe trends in the crash data driven by factors other than the strategy being evaluated. Measures of several socioeconomic factors can be included in the statistical models on a monthly and regional basis. These are: population size, unemployment rate, and fuel sales. Each of these measures is known to reflect differences in total exposure to crash risk, each in a subtly different way:

- Population Numbers, stratified into South East QLD and rest of QLD (continuous, all strata)
- Unemployment Rate stratified into South East QLD and rest of QLD (continuous, all strata)

- Fuel Sales in whole of QLD (continuous, all strata)

4.6.4.2 Data Stratification

The crash data analysed can be prepared as monthly crash counts from January 1993 to December 2004. To enable specific road safety measures to be related to the crash population towards which they are targeted, the crash data can be stratified for analysis by three factors. Three categorical variables can be used in the analysis data to define the strata to which the monthly crash count applied. The strata are classified as follows:

- Police region (defining the eight Police regions),
- Alcohol hours (low or high indicating the prevalence of drink-driving offences) If possible will use a Queensland version of the high and low alcohol hours, otherwise will use the Victoria definitions.
- Road type.

The three way categorisation defined 32 (8x2x2) data strata for analysis.

4.6.4.3 Outcome Crash Data

The following crash severities can be investigated in the models.

- Fatal + Hospital Admission crashes
- Fatal + Hospital Admission + Medically Treated crashes
- All Crashes
- Hospital Admission Crashes
- Medically Treated Crashes

It is not possible to reliably analyse fatal crash counts on their own given the likely very small average monthly fatality count when divided across the 32 defined strata (e.g. if there were 400 fatalities per year, you would get: 400 / 12 months / 32 strata = 1 per stratum per month). It should be noted that the results would not be reliable if fatal crash counts were analysed on their own.

4.6.5 Global-level evaluation

At the highest level, a global-level model evaluates the overall effects on crashes of the major initiatives in the strategy (Newstead et al, 1998; Newstead et al, 1995). Such a model will include variations in major socio-economic factors across time and regions of the State. Hence these influences, which are potentially contaminating influences on the apparent effects of the initiatives, will be removed. The effects of initiatives addressing relatively small target groups or relatively short periods can also be accounted for in this global-level model, allowing their potential evaluation for the first time in some cases. The results will include a global assessment of the impact of the 1993-2003 Strategy on crashes at each severity level.

4.6.6 Evaluation of program elements

This model makes estimates of the effects of individual initiatives (where there is sufficient data for the estimate to be reliable). The results will also indicate the type of initiative whose impact on crashes cannot be assessed in this way (due to relatively small target group and/or duration of operation), but whose aggregated effects are accounted for in the Level 1 model.

As noted in describing the stratification of the crash data for analysis, the motivation for the data stratification is to be able to relate certain measures used in the statistical models to those crash sub-populations to which they most directly relate. The measures considered only relating to certain crash strata and the strata to which they relate are as follows:

- Number of RBTs: High Alcohol Hour crashes
- Moving mode radar hours: Crashes on rural roads
- LIDAR speed detector hours: crashes on urban roads during Low Alcohol Hours
- Seat-belt and mobile phone penalties: Crashes on urban roads during Low Alcohol Hours
- Road safety publicity Adstock with drink-driving theme: High Alcohol Hour crashes
- South-east Queensland 50km/h local street speed limit: Police regions in south-east Queensland
- Regional Queensland 50km/h local street speed limit: Police regions outside of south-east Queensland

4.6.6.1 Description of road safety activities to be evaluated

The component areas and the specific component activities within each area are summarised as follows.

1. **Speed Camera Activity**: Six measures of speed camera activity that have been found to be key predictors of crash outcomes in the full formal evaluation of the Queensland speed camera program (Newstead and Cameron, 2003) are proposed to be used in this evaluation:
    - Total number of speed camera operation hours per month by police region
    - Number of active sites available for use by police region
    - Hours of operation per active camera site available for use by police region (derived from the above measures)
    - Percentage of sites visited as expected according to randomised speed camera operations schedule by police region
    - Monthly rate of increase in active camera sites by police region
    - Monthly rate of increase in speed camera operation hours by police region

2. **On-road (non speed camera) Police Enforcement**:
    - On road speed enforcement effort by Moving Mode Radar (MR) and Laser Speed Detection (LIDAR) operational hours

- Random Breath Testing (RBT) operations (the number of monthly random breath tests[1] conducted from booze buses and other stationary vehicles)
- Seat belt offences detected (monthly detected seat belt offences, available monthly by police region apart from 3 months in 2000 where the data was unavailable)
- Mobile phone offences detected (monthly detected mobile phone offences, available monthly by police region apart from 3 months in 2000 where the data was unavailable)

3. **Mass Media Publicity**:
    - Monthly awareness levels (AdStock[2]) of mass media television advertising with the following themes; speed, fatigue, seat belts, and drink driving.
    - Monthly awareness levels (AdStock) of mass media television advertising on all themes
4. **Change in crash reporting levels** from October 2000, associated with changes in the rules for making injury compensation claims following a motor vehicle crash.

5. **Other measures**, including legislation and penalty changes:

    - Introduction of the default 50km/h local street speed limit in south-east Queensland in June 1999 (enforcement amnesty period from March to May 1999)
    - The Holiday Period Road Safety Trial from December 2001 to end of January 2002
    - Introduction of the regional 50km/h local road speed limit from May 2003 (enforcement amnesty period from February to April 2003)
    - the increase in speeding penalties from April 2003
    - the increase in penalties for use of hand-held mobile phones while driving in December 2003.

Each of the above factors can be represented in the statistical model as a step functions, each step function defined as a binary variable with two levels: either "off" prior to the introduction of the initiatives, or "on" after the initiative was first introduced.

---

[1] The number of tests conducted should be used rather than the number of offences detected as RBT is considered to be effective in reducing crashes primarily through creating the perception of a high probability of offence detection through testing of a large proportion of the driving population.

[2] AdStock is a measure of retained awareness following exposure to advertising and is a function of the measure Target Audience Ratings Points (TARPs). Advertising was scheduled separately for South East Queensland and the rest of Queensland, so AdStock can be calculated separately for each of these regions and related to the relevant police regions in the statistical models.

4.6.7 Expertise required to fit models and monitor the results

Some of the models can be set up by a statistician experienced with the methods in a way that can then be used by non-statisticians to monitor the performance of the strategy in an ongoing way, identifying significant changes in crashes or injury rates that threaten the ability to meet the targets set by the strategy. Such monitoring can identify when further effort or intervention is required to keep the targets within reach.

4.7 SPECIFIC PROGRAM EVALUATIONS

Evaluation of specific programs (such as those listed in Table 3) is outside the scope of the evaluation framework being described. Such evaluations are essential to understanding the value of individual programs, particularly those programs that are expensive in terms of implementation or enforcement, but require a high level of statistical and analytical expertise to evaluate properly.

4.8 SUMMARY of framework development

The framework proposed to evaluate the Queensland 2004-2011 Road Safety Strategy follows a three-tiered modelling approach, and where possible has been structured to fit the GOSPA framework method.

The three-level modelling approach including what each model contains and proposes to measure is summarised below.

4.8.1 The Global Assessment or Top Tier Model

The Global Assessment Model relates to the broad goals and objectives of the 2004-2011 Queensland Road Safety Strategy, namely to achieve a reduction in the Queensland road fatality rate to under 5.6 deaths per 100,000 people by the year 2011, and to achieve a reversal in the increasing trend in hospitalisation casualties and the hospitalisation rate.

At the highest level of evaluation, the performance of the strategy can be measured using an intervention type forecasting model that measures change in aggregate in road trauma levels for the situation where the road safety strategy is implemented as a package compared to a situation modelled in which the strategy is not implemented at all.

The global assessment model evaluates the overall effects on crashes of the major initiatives in the strategy (Newstead et al, 1998; Newstead et al, 1995). Such a model will include variations in major socio-economic factors across time and regions of the State. Hence these influences, which are potentially contaminating influences on the apparent effects of the initiatives, will be removed. The effects of initiatives addressing relatively small target groups or relatively short time periods can also be accounted for in this global-level model, allowing their potential evaluation for the first time in some cases.

This global assessment model aims to measure the effect on road trauma of the Queensland Road Strategy and associated Action Plan overall. Road trauma will be quantified in terms of fatalities, serious injuries, fatality rates and serious injury rates.

The global assessment model can be formulated in two different ways depending on the time frame in which the evaluation model is being formulated relative to the implementation of the strategy. In the case of formulating the evaluation model at the commencement of the strategy, a time series model (based on state space modelling techniques) is estimated that models road trauma levels each month before the introduction of the Queensland Road Safety Strategy (i.e. Pre 2004) and forecasts the levels of road trauma that would have been expected (together with confidence limits on the estimates) to have occurred after 2004 had no strategy been in place (based on the past trends). Against the forecasts from the resulting model can be plotted the actual road trauma trends in Queensland that occurred after the 2004-2011 Queensland Road Safety Strategy was introduced. Plotting of the actual trends against those forecast in the absence of the strategy can be easily achieved by Queensland Transport staff without statistical training as a means of monitoring overall strategy performance over time. Figure 4.4 depicts this graphically.

The time-series model that will investigate road trauma trends pre and post the Queensland Road Safety strategy against a forecast trend post strategy in the absence of the strategy are similar in philosophy to the control-chart methodology used by a number of agencies in the past, including Queensland Transport, to monitor road safety strategy performance. The advantage of the proposed methodology for the global assessment model here is that it employs much more sophisticated and robust statistical methodology yet is still amenable to use by those without statistical training once established.

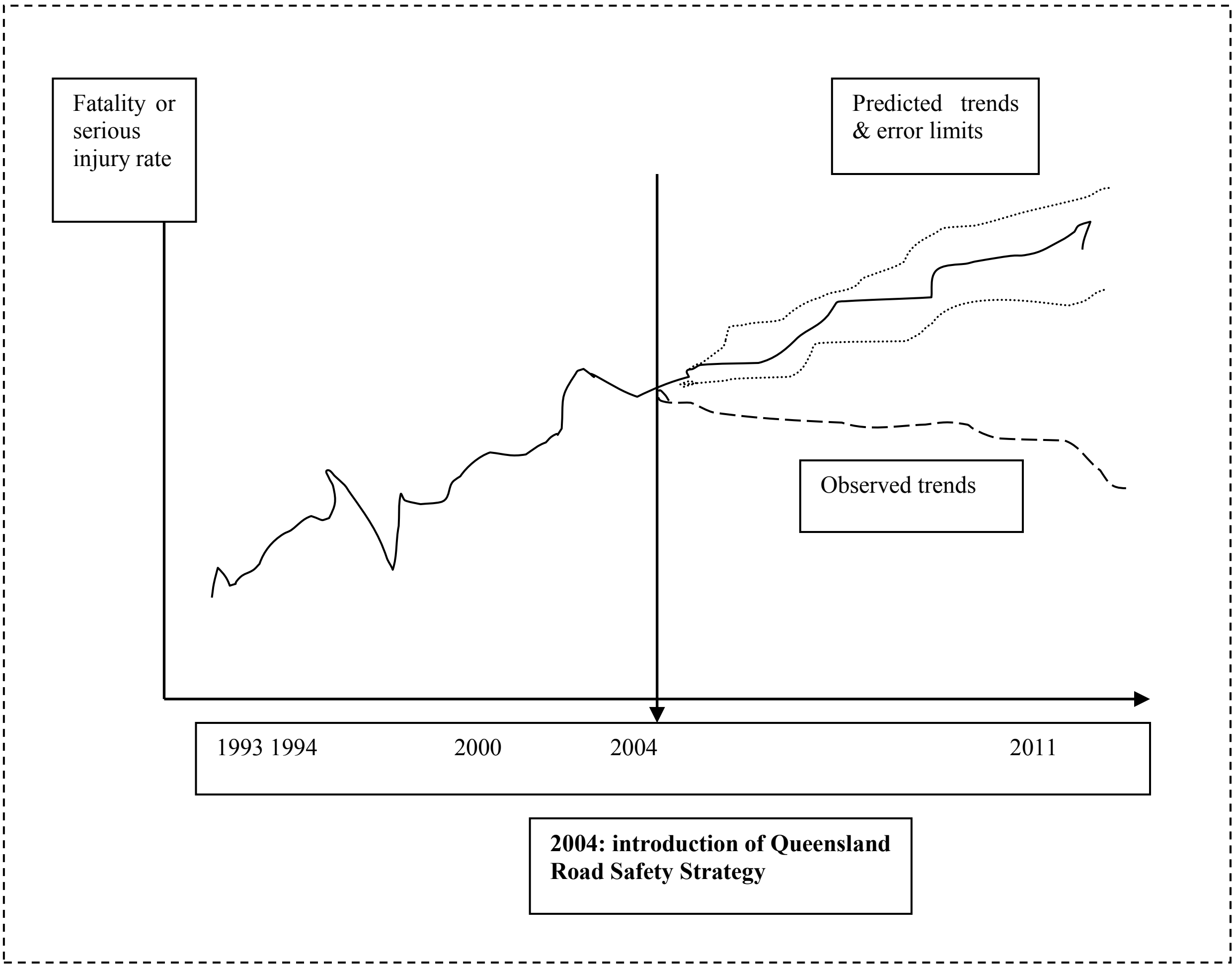


Figure 4.4: Model of observed and predicted road trauma trends pre and post the introduction of the Queensland Road Safety Strategy

The global assessment model can also be used at various time periods after strategy implementation to formally evaluate the overall performance of the strategy to that point in time based on the key outcome measures specified in the Strategy's goals and outcomes. Here the pre and post implementation data to the time available are modelled using the state space techniques with an intervention term being included at the time of strategy implementation. The intervention term parameter then represents the effect of the strategy on the outcome measure which can be tested formally for statistical significance. The intervention term can be modified accordingly to reflect increasing effects of the broad strategy over time which might be expected if components of the strategy are introduced in a staggered manner over time or take some time to become fully effective. Application of the global assessment model in this manner would require high level trained statistical expertise.

4.8.2 4.8.2The Second Tier Model

The second tier model relates to the objectives defined by the strategies and programs of the 2004-2011 Queensland Road Safety Strategy and the 2004-2005 associated Action Plan. This model considers the effects on road trauma for specific strata defined by the road user groups or situations (eg. pedestrians; young drivers; high alcohol crashes) at which the strategies and programs are targeted.

Based on the 2004-2011 Queensland Road Safety Strategy and the 2004-2005 Queensland Road Safety Action Plan the broad key road user groups and behaviours/situations that need to be targeted in the second-tiered model include:

- Alcohol and drug-driving (modelled during high alcohol times of the week)
- Speed-related crashes and speeding drivers
- Fatigued drivers
- Young adult inexperienced drivers
- Older drivers
- Fatal and serious crashes in rural Queensland
- Pedestrians, including intoxicated pedestrians
- Unrestrained vehicle occupants
- Unlicensed drivers and riders
- Motorcycle riders
- Indigenous road users
- Roads with poor crash records
- Fleet-related crashes.

For each of these strata, a specific analysis model equivalent in structure to that defined by the global assessment model above would be estimated for each of the key outcomes being measured. Like the global assessment models, the second tier models can be formulated at time of implementation to forecast road trauma outcomes in each stratum of interest had the strategy implementation not taken place. Actual post implementation road trauma trends are then compared to those forecast to assess strategy effectiveness. Intervention models can also be estimated at time points after strategy implementation to formally assess the statistical significance of outcome changes related to the strategy for each strata defined above. Expertise required for each approach is the same as for the global assessment model.

4.8.3 The Third Tier Model

The third tier model will attempt to model the individual program elements and actions of the Queensland Road Safety Strategy and Action Plans if possible (for example; the effects of speed camera operations).

The third tier modelling strategy is an extension of the tier 2 model in that it will typically target the same strata defined in tier 2. However, instead of modelling historic trends through general level, slope and seasonal terms, the model will include specific measures of road safety program effort under different activity areas as model covariates. In this way, the model makes estimates of the effects of individual initiatives (where there is sufficient data for the estimate to be reliable) by establishing the relationship between measurable road safety program effort and the key strategy outcome measures and relating the real variation in program effort to the reduction in road trauma observed. Results from the tier 3 modelling process will give specific estimates of the relative contributions of each of the major program elements in the road safety strategy to achieving the measured outcomes.

Poisson or negative binomial regression models will be used in this third tier modelling approach. These will be fitted to the outcome data series at some point after program implementation when sufficient post strategy and program element implementation experience has been accumulated to allow for successful modelling outcomes. The tier 3 modelling process must be carried out by someone with high level statistical expertise.

There will be some types of initiatives whose impact on crashes cannot be assessed using the tier 3 modelling approach due to relatively small target group and/or duration of operation or because program element input cannot be measured in a meaningful way. In some instances these program will be represented as local interventions. Where this is not possible, the aggregate effects of such programs will be assessed through comparing the third tier modelling outcomes with the aggregated effects accounted for in the global assessment model.

Types of road safety activities that can be included in the tier 3 evaluation models include:

Speed camera activity:

Six measures of speed camera activity that have been found to be key predictors of crash outcomes in the full formal evaluation of the Queensland speed camera program (Newstead and Cameron, 2003) and could be used as inputs into the tier 3 models. They are:

- Total number of speed camera operation hours per month by police region
- Number of active sites available for use by police region
- Hours of operation per active camera site available for use by police region (derived from the above measures)
- Percentage of sites visited as expected according to randomised speed camera operations schedule by police region
- Monthly rate of increase in active camera sites by police region
- Monthly rate of increase in speed camera operation hours by police region

On-road (non speed camera) Police Enforcement

- On road speed enforcement effort by Moving Mode Radar (MR) and Laser Speed Detection (LIDAR) - operational hours or offences issued
- Random Breath Testing (RBT) operations (the number of monthly random breath tests[3] conducted from booze buses and other stationary vehicles)
- Seat belt offences detected (monthly detected seat belt offences)

[3] The number of tests conducted should be used rather than the number of offences detected as RBT is considered to be effective in reducing crashes primarily through creating the perception of a high probability of offence detection through testing of a large proportion of the driving population.

- Mobile phone offences detected (monthly detected mobile phone offences)

Mass Media Publicity

- Monthly awareness levels (AdStock[4]) of mass media television advertising with the following themes; speed, fatigue, seat belts, and drink driving.
- Monthly awareness levels (AdStock) of mass media television advertising on all themes

Change in crash reporting levels from October 2000, associated with changes in the rules for making injury compensation claims following a motor vehicle crash.

Other measures, including legislation and penalty changes. Previous examples include:

- Introduction of the default 50km/h local street speed limit in south-east Queensland in June 1999 (enforcement amnesty period from March to May 1999)
- The Holiday Period Road Safety Trial from December 2001 to end of January 2002
- Introduction of the regional 50km/h local road speed limit from May 2003 (enforcement amnesty period from February to April 2003)
- the increase in speeding penalties from April 2003
- the increase in penalties for use of hand-held mobile phones while driving in December 2003.

4.8.4 Specific Evaluation

A final tier of evaluation recommended for the Queensland Road Safety Strategy is specific evaluation of major program components. The tier 3 models described above measure the general association between measures of specific road safety program activities in a multivariate setting. However, for large complex road safety program elements, specific evaluation of major elements is generally needed for two reasons. First, only specific evaluation can establish the cause and effect relationship between road safety program element implementation and road trauma outcomes with a sufficient degree of scientific rigour. Second, specific evaluation is often needed to establish the measure of road safety program operation that is best related to the outcomes achieved which in turn is fed into the tier 3 models as a key input. For example, specific evaluation of the

[4] AdStock is a measure of retained awareness following exposure to advertising and is a function of the measure Target Audience Ratings Points (TARPs). Advertising was scheduled separately for South East Queensland and the rest of Queensland, so AdStock can be calculated separately for each of these regions and related to the relevant police regions in the statistical models.

Queensland speed camera program identified the 6 key measures of program activity listed in the previous section that best predicted crash outcomes.

For these reasons, it is recommended that Queensland Transport continue to commission specific evaluations of key road safety programs implemented as part of the broader road safety strategy. In the past this has included such programs as:

- the mobile speed camera program
- Random Road Watch
- 50km/h default urban speed limits in South East and the rest of Queensland.

4.8.5 Relationship between GOSPA Framework and Three-Tiered Modelling

The following table (Table 4.4) summarises the links between the GOSPA framework and the three-tier modelling method and how they relate to the Queensland Road Safety Strategy and Action Plans.

Table 4.4: Summary of the link between the GOSPA framework and the three-tier modelling approach

GOSPA FRAMEWORK COMPONENTS		THREE-TIER MODELLING
Component	**Definition**	
Goal	Overall goal of strategy (i.e. to prevent road trauma through safe road use, safe roads and safe vehicles)	Global assessment model (top-tier model) to measure effect on road trauma of the Strategy overall (includes an intervention model)
Objective	Objectives to reach goal (e.g. to achieve a reduction in the fatality rate to under 5.6 deaths per 100,000 people)	
Strategies	General strategies to achieve objectives given in the Qld action plans and road safety strategy	Second-tier modelling of specific strata targeted by the strategies in the action plans (e.g. crashes occurring during high alcohol times of the week)
Programs	Specific programs relating to target group outcomes	Third-tier modelling of the individual program elements of the strategy (e.g RSIP evaluation)
Actions	Actions undertaken in each program	

The next chapter of the report will include a demonstration of the three-tier modelling approach on the 1993-2003 (revised) Queensland Road Safety Strategy.

5. TEST RUN OF THE FRAMEWORK ON THE 1993-2003 QUEENSLAND ROAD SAFETY STRATEGY

5.1 Overview of 1993-2003 Queensland Road Safety Strategy

The 1993-2003 Queensland Road Safety Strategy was launched in April 1993. It included an outline of the vision of road safety for 2003, and also outlined the objectives and principles to achieve that vision. The strategy contained more than 120 specific actions of which 85% had been implemented by 1998.

The target of the strategy was for actions to lead to a 30% reduction in fatal crashes relative to the trend by 2003. This target was met in 1998 with an estimated 8.07 fatalities per 100,000 population recorded. This was a 42% reduction on the 1992 fatality rate of 13.7 deaths per 100,000 population. In addition there were improvements for most road user groups and crash categories.

Were these reductions due to the Strategy? During 1993-2003 as part of the Queensland Road Safety Strategy a number of initiatives were introduced including:

- Random Road Watch (1991-1993)
- Bicycle helmet laws (1991-1993)
- Audible linemarking (1991-1993)
- Driver reviver program (1991-1993)
- Speed management strategy (1997).

Because the target of a 30% reduction was met by 1998, a revised Strategy was launched in 1999. The main target of this revised Strategy was to achieve a 20% reduction in the annual fatality rate by 2003 compared with the 1998 figure.

To test-run the framework described in the previous chapter, structural time series (also known as 'state-space') models were fitted to the crashes. 'State-space' modelling techniques were used to fit models to the Queensland fatal and injury crashes. All variables (both crashes and explanatory factors) were transformed into natural logarithms.

5.2 FIRST TIER or global assessement MODELS

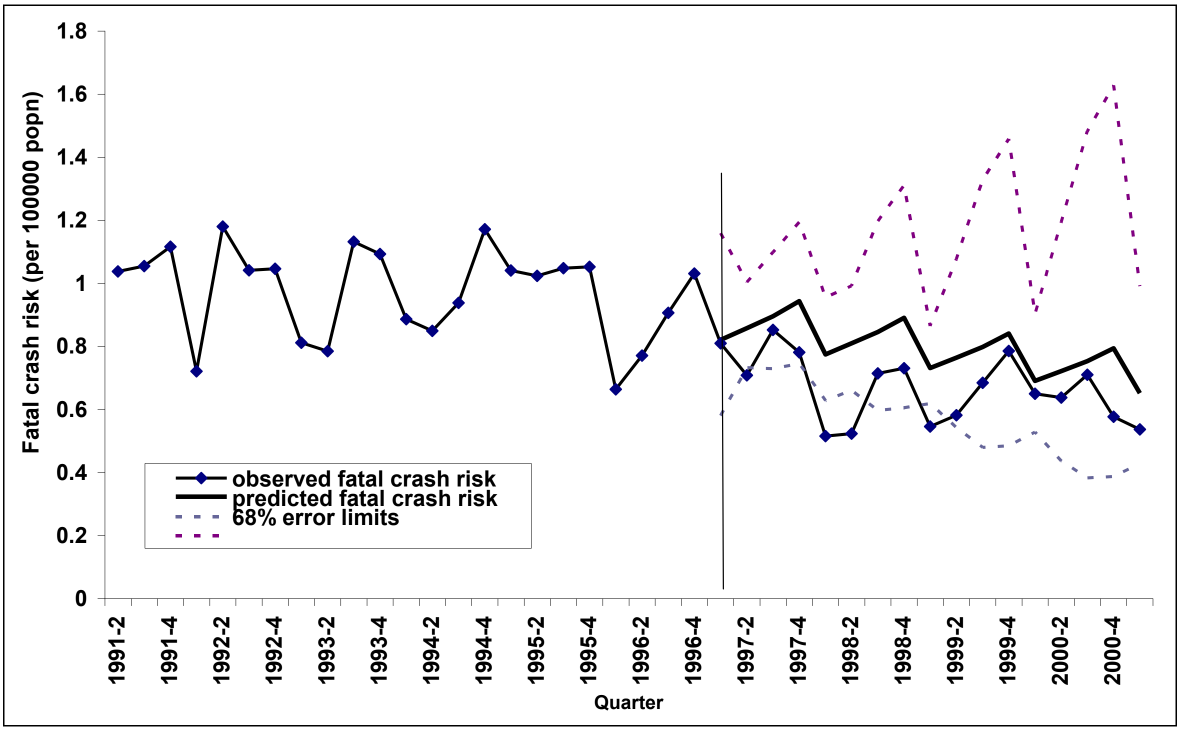
Queensland crash data for the period April 1991-December 2004 was considered in the modelling process. Road trauma was measured as a crash risk per 100,000 population, with fuel sales and unemployment rate included as covariates. Crash risks at various severity levels were considered, e.g. fatal crash risk, serious casualty crash risk, casualty crash risk and all crash risk. Crash risk was modelled for the 'before' period April 1991 to December 1996 – using quarterly crash risk figures for fatal crashes and monthly crash risk figures for the other crash severities. Crash risk trends were then predicted from 1997 to 2001. Confidence limits (68%) were also estimated for these predicted crash risks. The models included covariates monthly (or quarterly) unemployment rate and the monthly (or quarterly) fuel sales. For the state-space models the covariates as well as the crash risks were transformed into natural logarithms before being modelled.

Sixty-eight percent confidence limits were chosen to be consistent with control chart analysis techniques (Montgomery, 1991). Often, when using control chart theory to investigate if time series predictions represent real trends or chance fluctuations, 68% error bounds are typically used. Since the global assessment models are essentially defining a control chart type methodology where we are seeking to establish whether the road safety strategy has driven road trauma 'out of control' (i.e. away from the path it was previously heading), it seemed appropriate to apply control chart based confidence limits.

5.2.1 Fatal crash risk

The first global assessment model estimated was the fatal crash risk model for all of Queensland. Fatal crashes per 100,000 population were modelled per quarter for the period April 1991 to December 1996. The state-space model included the quarterly unemployment rate and quarterly fuel sales as covariates. From the fitted model, the predicted fatal crash risk trends were then estimated for the period January 1997 to December 2001, together with 68% confidence limits on each estimate. The fatal crash risks are shown in Figure 5.1, together with what was actually observed in the period of the strategy 1997-2001. As can be seen in this graph, the observed fatal crash risk during the period of the revised strategy were somewhat lower than that predicted from the model.

Figure 5.1: Queensland Fatal crash risk per 100,000 population per quarter, April 1991-March 2001



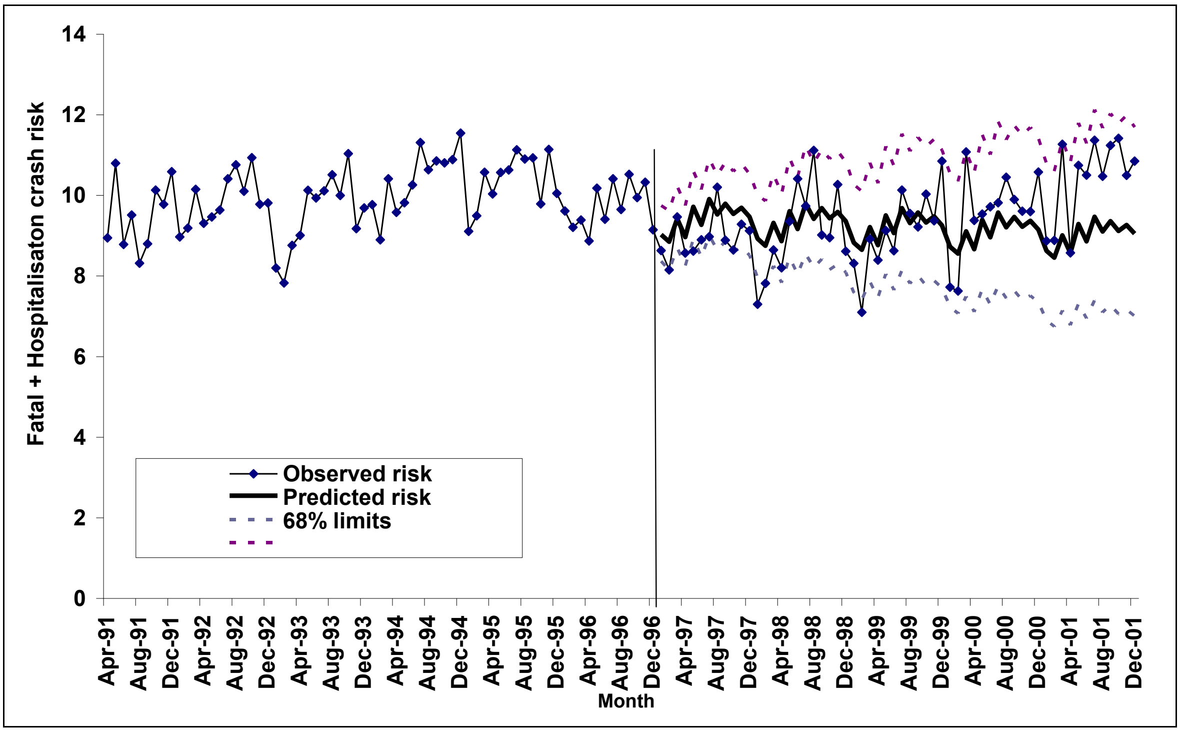
5.2.2 Serious Casualty Crash risk

Two definitions of serious casualty crashes were used namely those that resulted in i) a fatality or hospitalisation and those that resulted in ii) a fatality, hospitalisation or medically treated injury.

i) Fatal & Hospitalisation Crash Risk

Fatal plus hospitalisation crashes per 100,000 population were modelled per month for the period April 1991 to December 1996 to give an estimate of the serious casualty crash risk in the before period of the revised Strategy. The monthly serious casualty crash risk was then predicted from the fitted model from January 1997 to December 2001. Confidence limits (68%) were placed on each crash risk estimate. Figure 5.2 depicts the fatal and hospitalisation crash risk. Whilst the predicted serious casualty crash risk appears quite stable during 1997-2001, there was some variation in the observed risk with an apparent increasing trend in 2001 greater than the predicted crash risk.

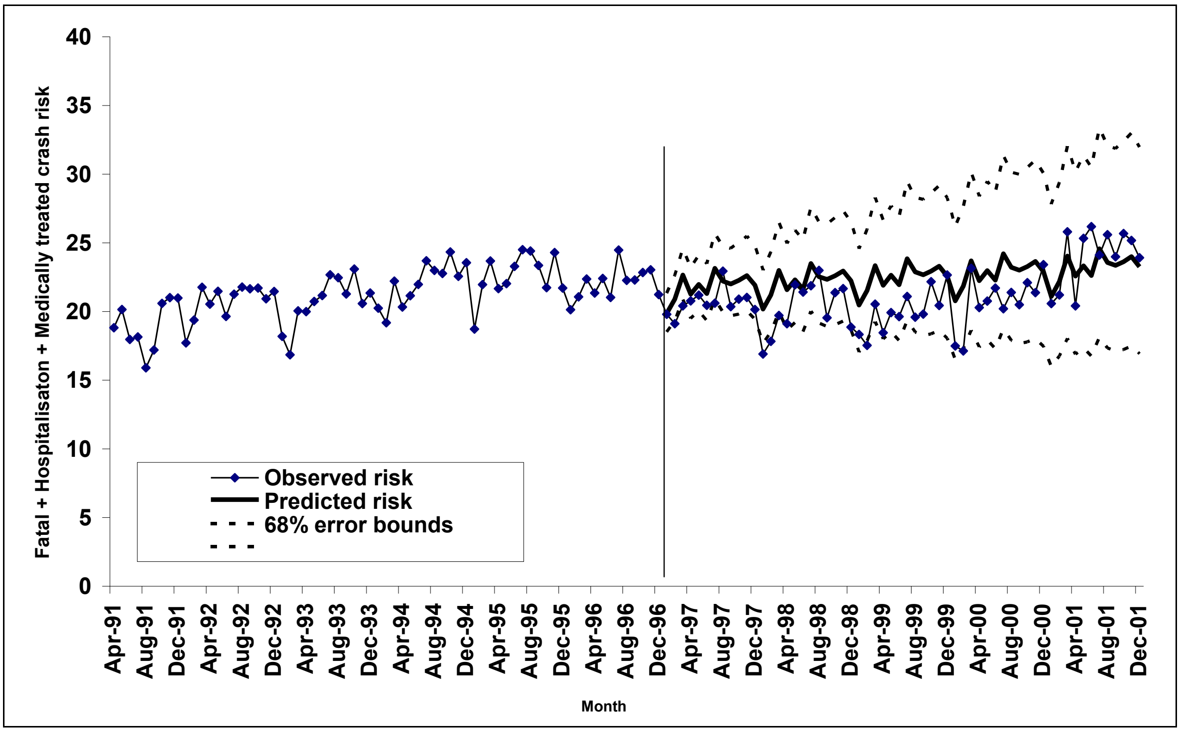
Figure 5.2: Queensland 'Fatal & Hospitalisation' crash risk per 100,000 population per month, April 1991-December 2001



ii) Fatal & Hospitalisation & Medically treated injury Crash Risk

Figure 5.3 gives the analogous chart to Figure 5.2 however the serious casualty crash risk was determined from crashes that resulted in a fatality, a hospitalisation or a medically treated injury. Again there is an increasing trend in the observed serious casualty crash risk compared with that predicted from the model.

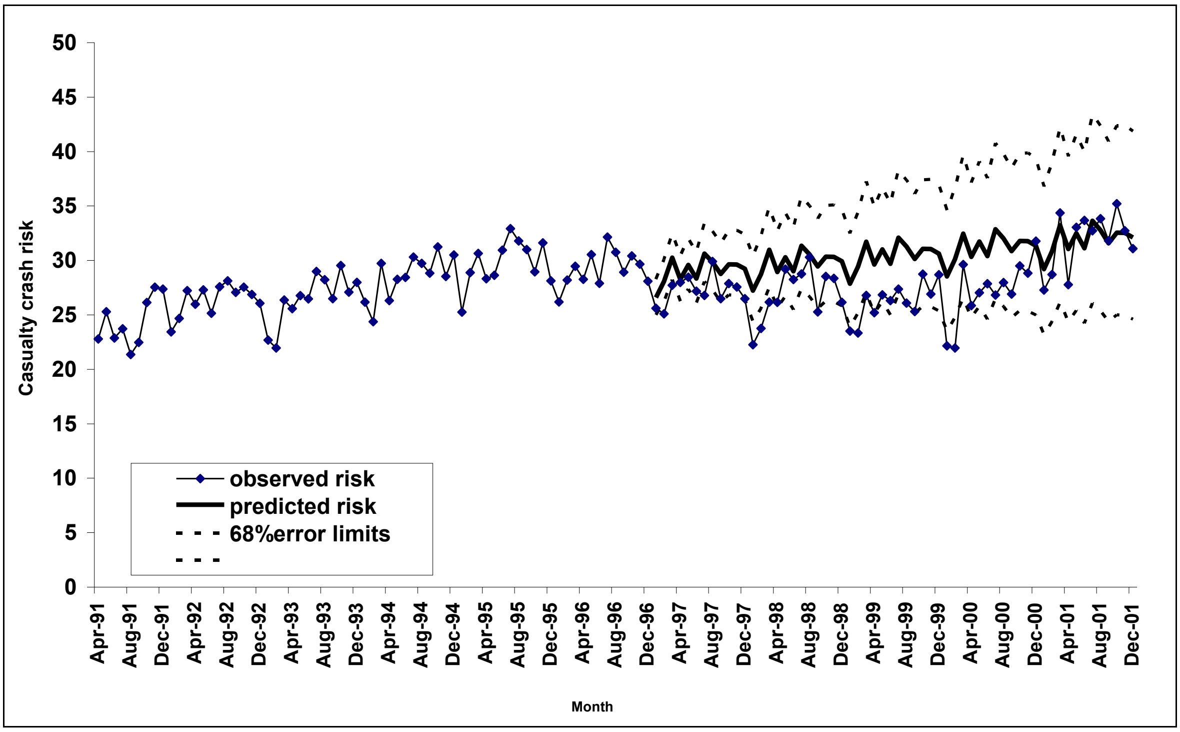
Figure 5.3: Queensland 'Fatal & Hospitalisation & Medically Treated' crash risk per 100,000 population per month, April 1991-December 2001



5.2.3 Casualty Crash Risk

A casualty crash was defined to be one that resulted in either a fatality, a hospitalisation, a medically treated injury or a minor injury. The monthly casualty crash risk per 100,000 population was modelled for the period April 1991-December 2001, and predicted from the fitted model from January 1997 to December 2001 (Figure 5.4). There was an increasing trend in the observed casualty crash risk from 2000 onwards, however the observed casualty crash risk was generally less than that predicted.

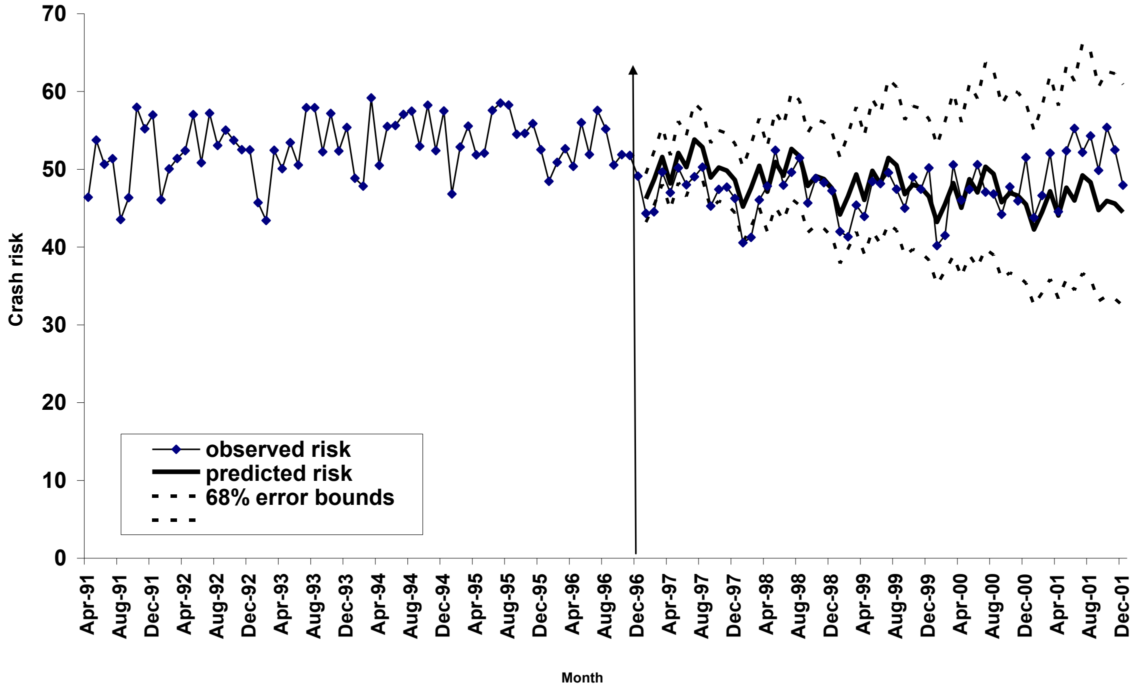
Figure 5.4: Queensland Casualty crash risk per 100,000 population per month, April 1991-December 2001



5.2.4 All Crash risk

Property damage crashes were added to crashes that resulted in a fatality or injury to estimate the 'all crash' risk. The monthly crash risk per 100,000 population was modelled for the period April 1991 to December 1996. The crash risk from the fitted model was predicted from January 1997 to December 2001. (Figure 5.5). From 1997 to 2000 the observed 'all crash' risk was generally smaller than that predicted from the model.

Figure 5.5: Queensland Crash risk per 100,000 population per month, April 1991-December 2001



5.2.5 Intervention Model

The Queensland fatal crash risk model was fitted per quarter from April 1991 to December 2001. As in the previous models the quarterly unemployment rate and quarterly fuel sales were included as covariates.

Intervention variables are dummy variables used to take into account outlying observations or structural breaks. Structural breaks indicate where the predicted crashes were substantially different from the observed crashes in that month or quarter, representing the effects of the road safety strategy in improving road trauma outcomes. To account for a structural break, a step function is included in the model.

For this model annual step intervention functions were fitted in December for the years 1996 to 2003 (Figure 5.6). The actual fatal crash risk series is shown by the blue line in Figure 5.6. The second series (shown in red) shows the trend in the series (represented by the stochastic level) as well as adding in the effects of the explanatory variables (reg) and the yearly interventions.

Table 5.1 gives the estimated coefficients for each annual intervention together with their significance probability, as well as the estimates of the coefficients of the explanatory variables, unemployment and fuel sales. The estimated coefficients can be interpreted in the same way as regression coefficients. For example the coefficient of ln(unemployment) is -0.1956, indicating that a 1% increase in unemployment leads to a reduction in the fatal crash risk of approximately -0.2 with an associated significance of 0.1554. Similarly the coefficient for the December 1998 intervention was -0.2693, indicating a reduction in the fatal crash risk after introduction of that intervention (with a two-sided significance probability of 0.09615)

Figure 5.6: Queensland fatal crash risk per quarter April 1991-December 2006 with yearly interventions

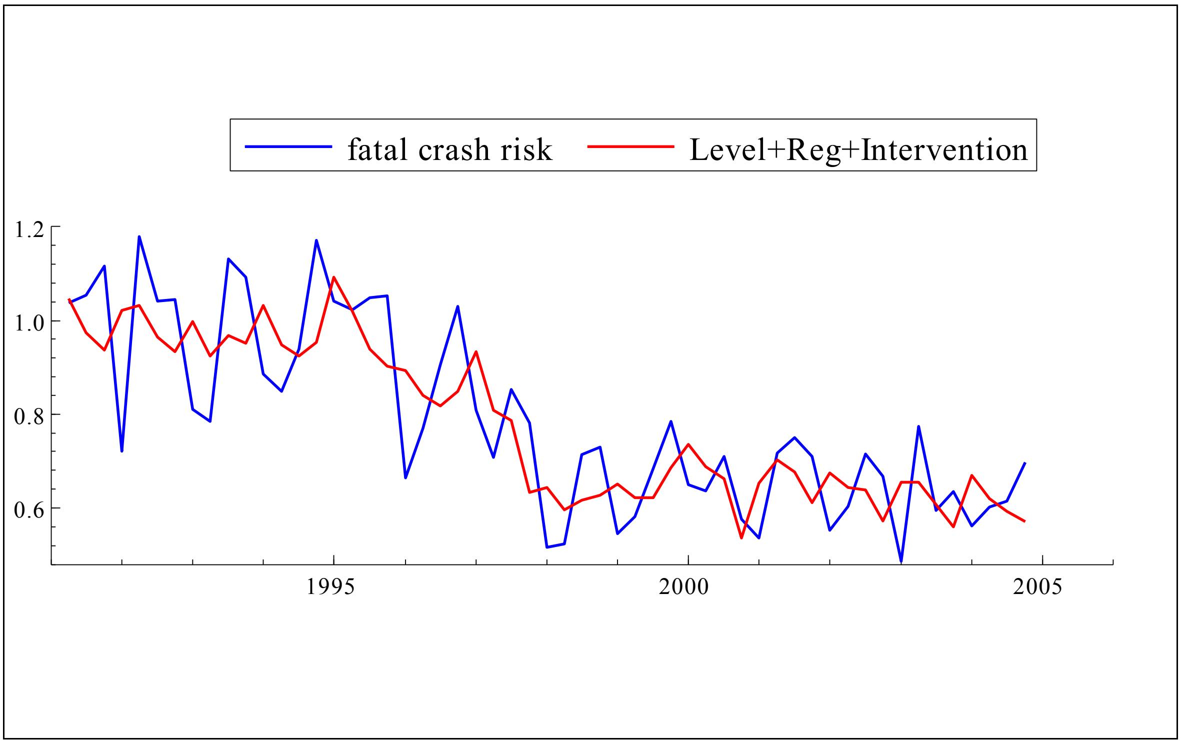


Table 5.1: Estimated coefficients for each yearly intervention in the fatal crash risk model.

	Coefficient	**Standard error**	**Significance probability**
ln(mgL fuel sales)	0.0437	0.1361	[0.74970]
ln(unemployment)	-0.1956	0.1351	[0.15544]
Intervention			
Level break 1996	0.0347	0.1416	[0.80754]
Level break 1997	0.0508	0.1419	[0.72214]
Level break 1998	-0.2693	0.1580	[0.09615]
Level break 1999	-0.1321	0.1429	[0.36077]
Level break 2000	-0.1339	0.1438	[0.35730]
Level break 2001	-0.0616	0.1439	[0.67068]
Level break 2002	0.0629	0.9140	[0.94546]
Level break 2003	0.6870	0.4129	[0.10395]

5.2.6 Summary of First-Tier Modelling

The evaluation models formulated in this section have demonstrated how road trauma levels can be modelled each month prior to the introduction of the Queensland Road Safety Strategy and then used to forecast the levels of road trauma that would have been expected to have occurred after 1997 had no strategy been in place (based on the past trends). The estimated forecasts also had confidence limits placed on them. The actual road trauma trends that occurred in Queensland after the introduction of the (revised) 1992-2003 Strategy were also plotted. Plotting the actual trends against those forecast in the absence of the Strategy can be used as a means of monitoring overall strategy performance over time.

The demonstration of the evaluation framework appeared to have worked quite well for the first few years of the projection (i.e. 1997 to late 2000), as shown in Figures 5.1 to 5.6. However in the last year of the prediction period (i.e. late 2000 to 2001) the process seemed to be somewhat ‘out-of-control’. This suggests that there may be a need to re-calibrate the forecasts from then onwards. In addition, the confidence limits placed on the forecasts are relatively wide indicating the predictions have substantial variation.

Section 5.3.5 gave a demonstration of how an intervention can be formally measured. The fatal crash risk global assessment model was used at various time periods after the Strategy was implemented (i.e. annual step functions at the end of each year for the post-period 1996-2003) to formally evaluate the overall performance of the strategy to that point in time.

The results in Table 5.1 (i.e. the high-lighted negative coefficients) indicate that there has been a period of sustained reduction from about 1998 to 2001 before levelling off again after that. Although the step intervention functions fitted to the model were not statistically significant they are indicative of a reduction in the fatal crash risk. A more definitive result may occur if an intervention model was fitted that also included injury crashes as well as fatal crashes.

5.3 SECOND-TIER MODELS

The key road user groups or behaviours targeted in the revised Strategy are given in Table 5.2. Where possible (depending on data availability) these road user groups/behaviours were considered for the second-tier modelling analysis.

Table 5.2: Key road user groups/behaviours and associated strata for the second-tier models

Key target group/behaviour	Stratum
DRINK-DRIVING	
1. Drink-drivers	High alcohol hour serious casualty crashes in Qld
2. Drink-drivers in rural regions	High alcohol hour serious casualty crashes in rural regions (i.e. regional Queensland)
3. Drink-drivers in metropolitan regions	High alcohol hour serious casualty crashes in metropolitan regions (i.e. South East Queensland)
4. Young drink-drivers	High alcohol hour serious casualty crashes involving drivers aged 17-29 years
SPEED	
Speeding drivers	Casualty crashes by speed zone and by road type (highways, local roads) during low alcohol times of the week
Speed by region (metro and rural)	Casualty crashes by speed zone by region (metro/rural) during low alcohol times of the week
Speeding around schools and residential areas	Casualty crashes by road type and by speed zone (50km/h, 40 km/h) in S.E. Qld and other regions
VULNERABLE ROAD USERS	
Rural road users	Casualty crashes in rural regions
Older road users	Casualty crashes involving drivers/pedestrians aged over 60 years
Young/inexperienced drivers	Casualty crashes involving drivers aged 17-29 years

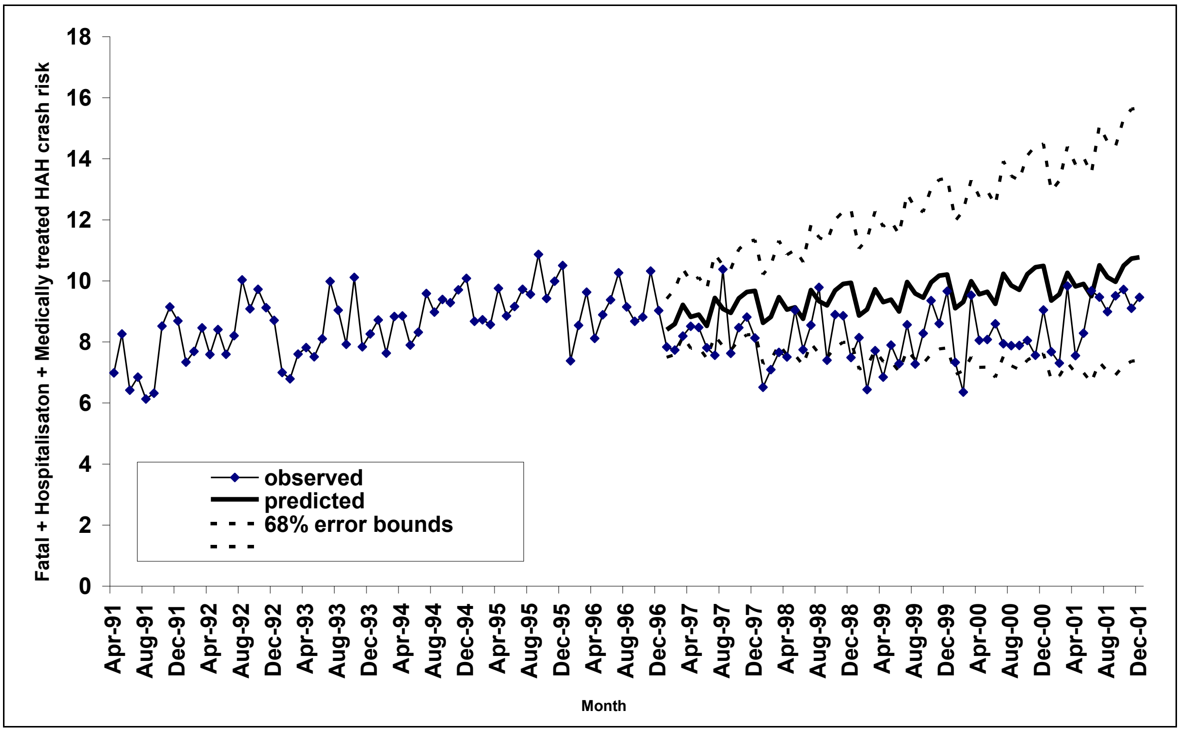
For each target group, state-space models were fitted to the relevant monthly crash risk (i.e. number of crashes per 100,000 population) for the period April 1991-December 1996. The monthly unemployment rate and monthly fuel sales were included as co-variates in the models. The future crash risk trends as predicted from the fitted models were compared with the crash risks that actually occurred during January 1997-December 2001. In addition 68% error limits were estimated for each prediction.

5.3.1 Drink-driving

Using the definitions of alcohol times given by Harrison (1990) the serious casualty crashes occurring in Queensland were categorised into those that occurred during high alcohol hours of the week (HAH) and those that occurred during low alcohol hours (LAH) of the week. Harrison's definitions were based on Victorian crashes and were used as an approximation of Queensland alcohol times for this study. Serious casualty crashes were defined to be those that resulted in either a fatality, a hospitalisation or a medically treated injury.

Figure 5.7 gives the high alcohol hour serious casualty crash risk for Queensland during April 1991 to December 2001. This graph depicts the observed crash risk for April 1991-December 2001 and that predicted from the second-tier model from January 1997-December 2001. Generally the observed HAH serious casualty crash risk was less than that predicted from the model during 1997 to 2001, suggesting that during the period of the revised Strategy there may have been a reduction in alcohol-related crashes.

Figure 5.7: Queensland High Alcohol Hour Serious Casualty Crash Risk per 100,000 population per month, April 1991-December 2001



5.4.1.1 Drink-driving by Region

From the Queensland database of Police-reported crashes, region was defined to be either South-east Queensland or Regional Queensland. State-space models were then fitted to the high alcohol hour serious casualty crash risk for each region for the period September 1992-December 1996. Monthly fuel sales and monthly unemployment rate were included as socio-economic explanatory factors in the models. These models were used to estimate the predicted crash risk for January 1997-December 2001. It should be noted that the estimation pre-period commenced at September 1992 rather than at April 1991 because monthly population and unemployment data were not readily available by region prior to September 1992.

Figures 5.8 and 5.9 show the HAH serious casualty crash risk models for South East Queensland and for regional areas of Queensland respectively.

Figure 5.8: Serious Casualty Crash Risk per month during High Alcohol Times of the week for South-East Queensland: September 1992 to December 2001

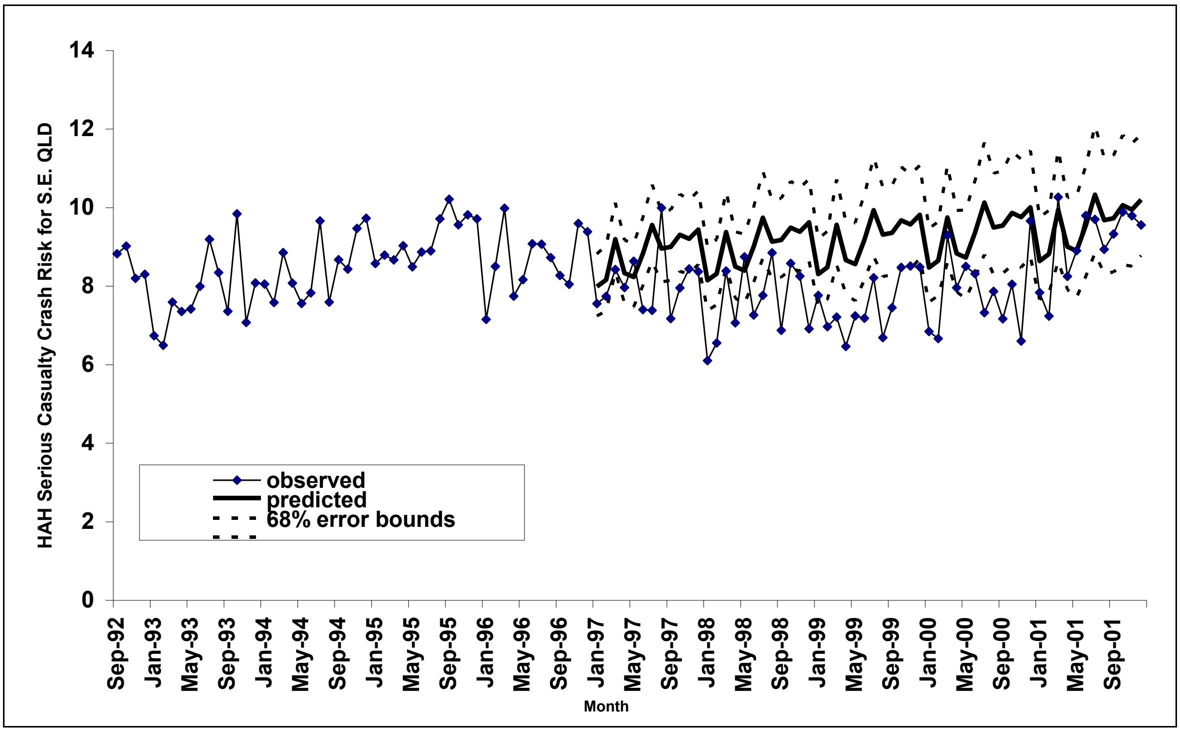
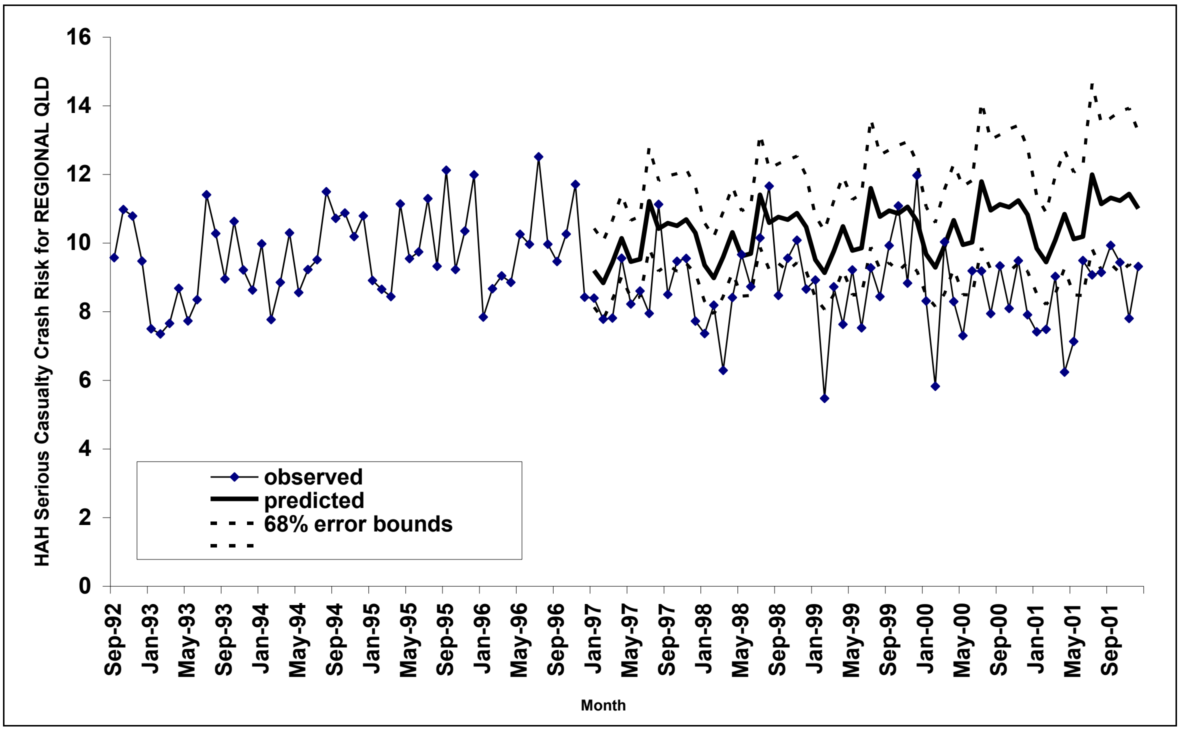


Figure 5.9: Serious Casualty Crash Risk during High Alcohol Times of the Week in REGIONAL Queensland: September 1992 to December 2001



For both regions the observed HAH serious casualty crash risk was generally less than that predicted from the state-space models - similar to the overall Queensland trends.

5.4.1.2. Young Drink Drivers

The crash data supplied to MUARC for this project was not person-based hence crashes involving young drivers aged 17-29 years could not be extracted. Consequently the second-tier crash risk models involving young drink-drivers were not attempted as part of this study.

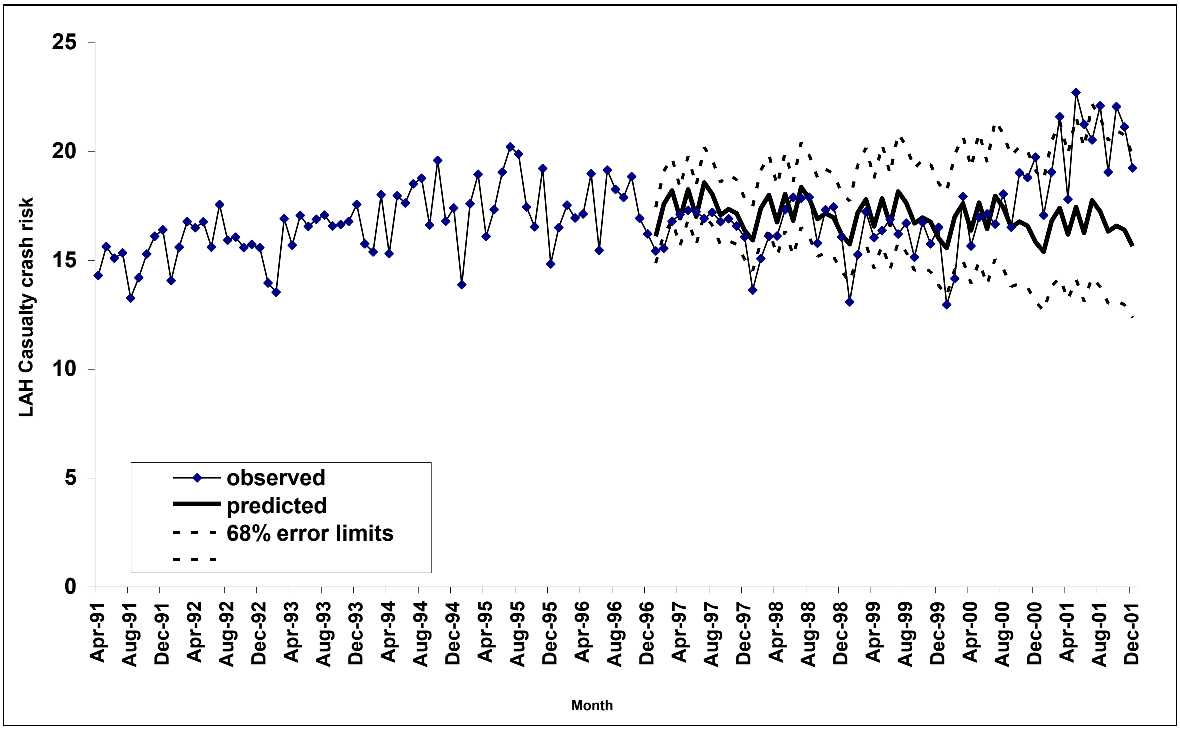
For future evaluations it is recommended that the crash database contains person-based information such as the age and sex of each person involved in a crash.

5.3.2 Speed

5.3.2.1 All Speed Zones

Casualty crashes occurring during low alcohol hours (LAH) of the week per 100,000 population were modelled for the period April 1991-December 1996. A casualty crash was defined as one that resulted in either a fatality or a hospitalisation or a medically treated injury or a minor injury. For the first model all speed zones in all of Queensland were considered (Figure 5.10). The monthly LAH casualty crash risk was then predicted from the fitted model from January 1997 to December 2001, together with 68% confidence limits. As seen in Figure 5.10 the observed and predicted crash risks were similar up until 2000, with the observed crash rate increasing during 2000-2001.

Figure 5.10: Queensland LAH Casualty Crash Risk per 100,000 population per month, April 1991-December 2001 – all speed zones

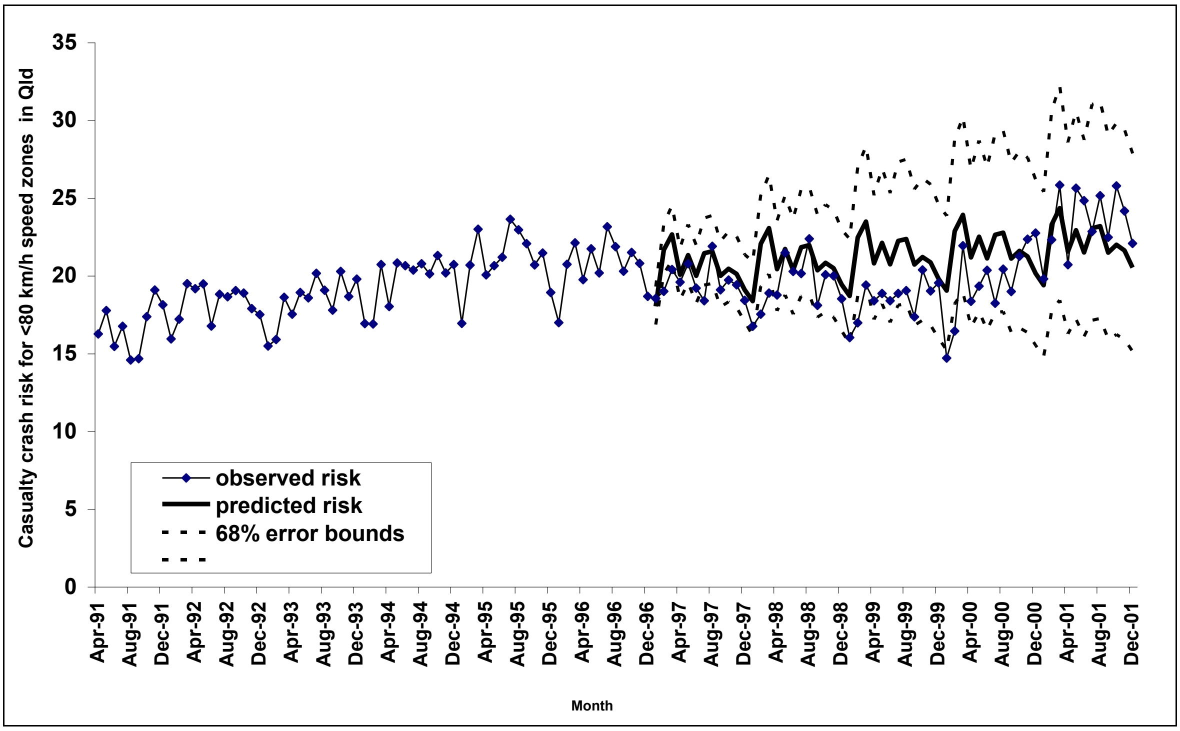


Since the type of road was not given in the Queensland database of Police-reported crash a proxy for speed-related crashes used for this study was crashes that occurred in speed zones under 80 km/h and those that occurred in zones of 80 km/h and above.

5.3.2.2 Speed zones below 80 km/h

The casualty crash risk per month for speed zones below 80 km/h in all of Queensland was estimated for the period April 1991-December 1996 using state-space model techniques. The monthly unemployment rate and monthly fuel sales were included in the model as covariates. This estimated model was then used to predict the crash risk that would have been expected to occur during January 1997-December 2001. In addition, confidence limits (68%) were placed on each estimate. Figure 5.11 shows the predicted casualty crash risk as well as the actual casualty crash risk that was observed in speed zones <80 km/h during April 1991-December 2001. Generally during January 1997-December 2000 the predicted casualty crash risk in speed zones <80 km/h was lower than that observed. However in 2001 there was an increasing trend in the observed casualty crash risk compared with that predicted.

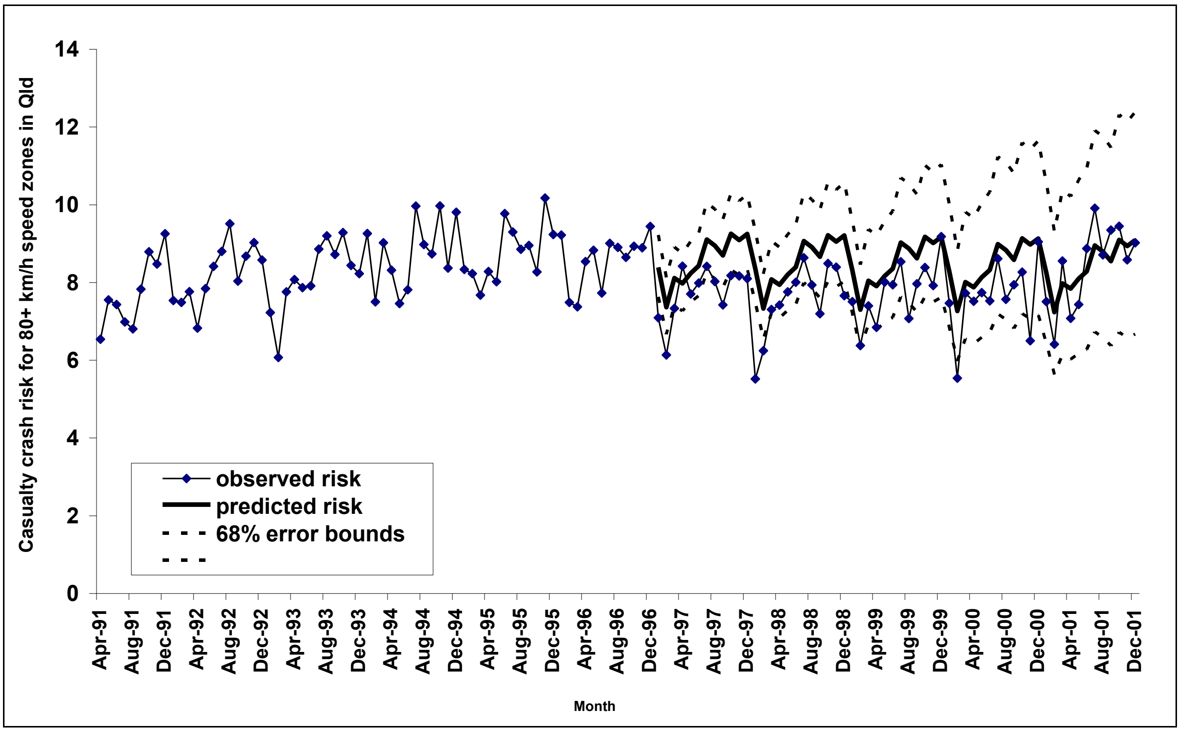
Figure 5.11: Casualty Crash Risk per month for <80 km/h speed zones in Queensland: April 1991-December 2001



5.3.2.3 Speed zones 80 km/h and above

Figure 5.12 gives the casualty crash risk for crashes occurring in speed zones of 80 km/h or greater that was observed in Queensland during April 1991-December 2001. In addition this chart shows the corresponding casualty crash risk predicted from the state-space model for the period January 1997-December 2001. Except for the later half of 2001 the observed crash risk was generally less than that predicted from the model for crashes occurring in speed zones of at least 80 km/h.

Figure 5.12: Casualty Crash Risk per month for 80+ km/h speed zones in Queensland: April 1991-December 2001



5.3.2.4 Speed zone by region

The casualty crash risk was estimated for crashes occurring in speed zones under 80 km/h and those in speed zones of at least 80 km/h separately for South East Queensland and for Regional Queensland. The appropriate monthly unemployment rate and monthly fuel sales were included in the models as explanatory factors. The period of estimation was from September 1992 to December 1996 because the unemployment rate and the monthly population estimates by region were not available prior to September 1992.

South East Queensland

For South East Queensland the casualty crash risk estimates for speed zones <80 km/h and speed zones of at least 80 km/h are shown in Figures 5.13 and 5.14, respectively.

The observed casualty crash risks for speed zones under 80 km/h in South East Queensland were generally less than those predicted by the model from January 1997 to December 2000 – after that there appeared to be an increasing trend.

For speed zones of 80 km/h or more in South East Queensland the observed casualty crash risk was relatively stable up until October 2000. From late 2000 and throughout 2001, an increasing trend was apparent.

Figure 5.13: Casualty Crash Risk for <80 km/h speed zones in South East Queensland: September 1992-December 2001

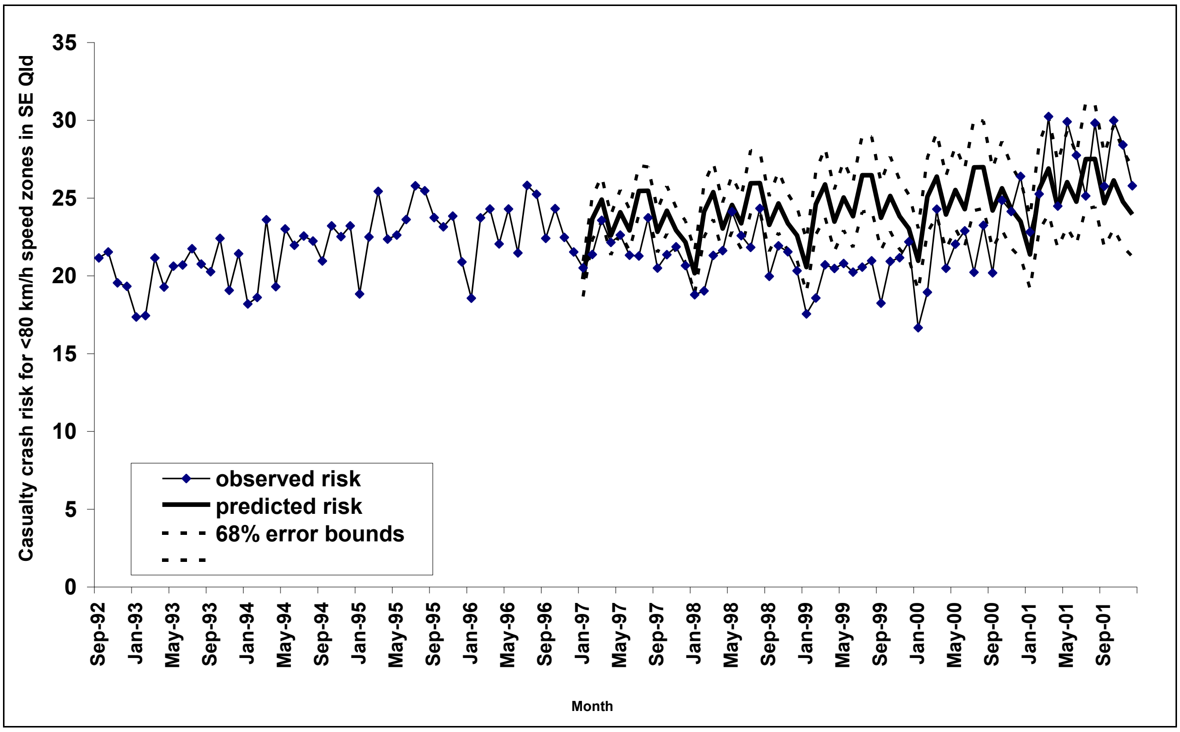
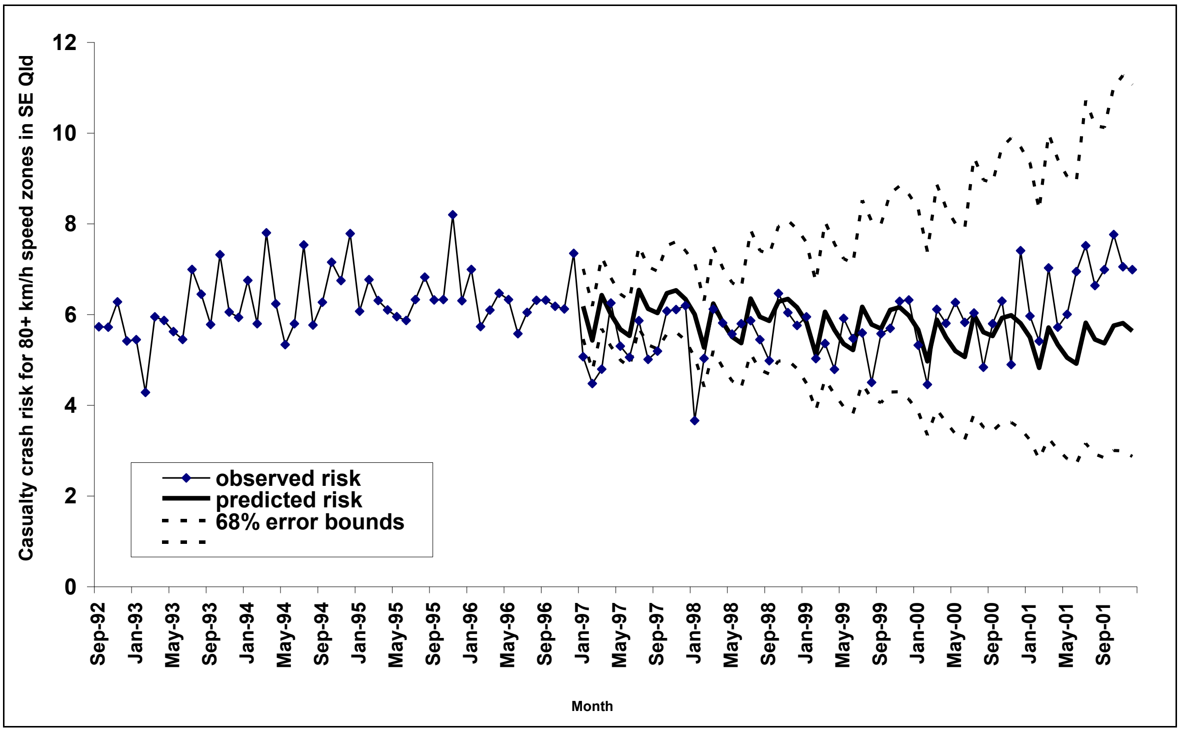


Figure 5.14: Casualty Crash Risk for 80+ km/h speed zones in South East Queensland: September 1992-December 2001



Regional Queensland

For Regional Queensland the casualty crash risk estimates for speed zones <80 km/h and speed zones of at least 80 km/h are shown in Figures 5.15 and 5.16, respectively.

Figure 5.15: Casualty Crash Risk for <80 km/h speed zones in Regional Queensland: September 1992-December 2001

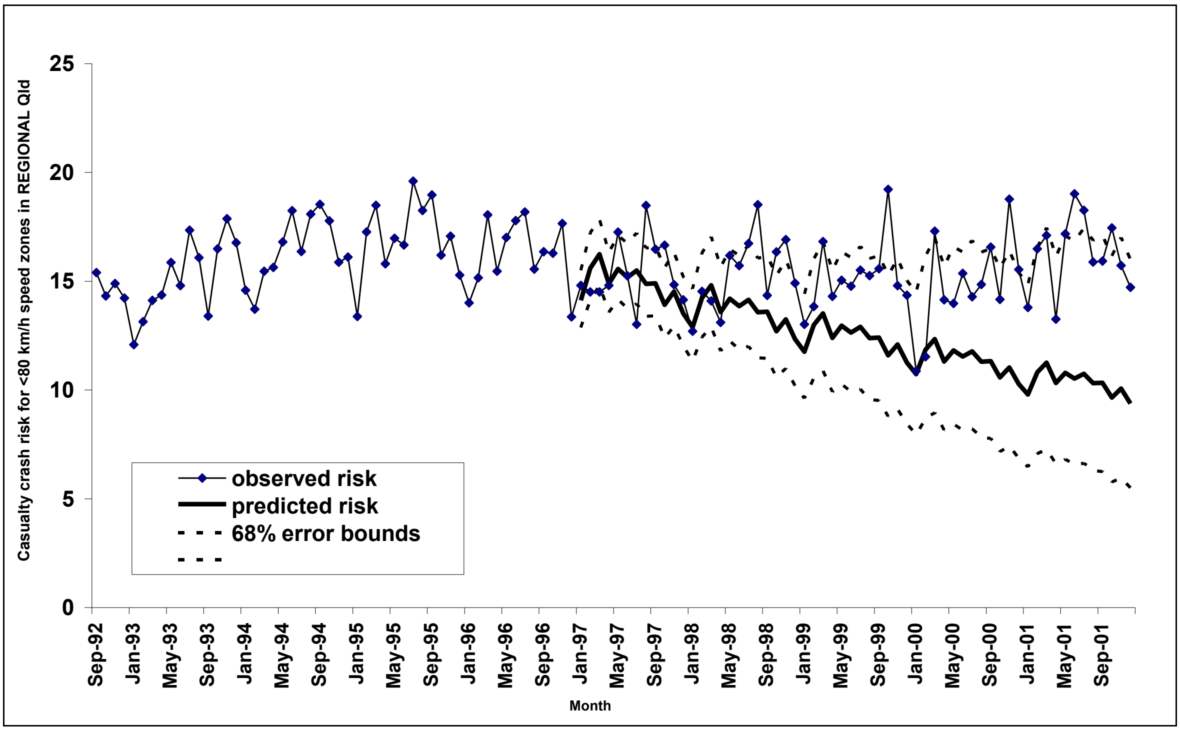
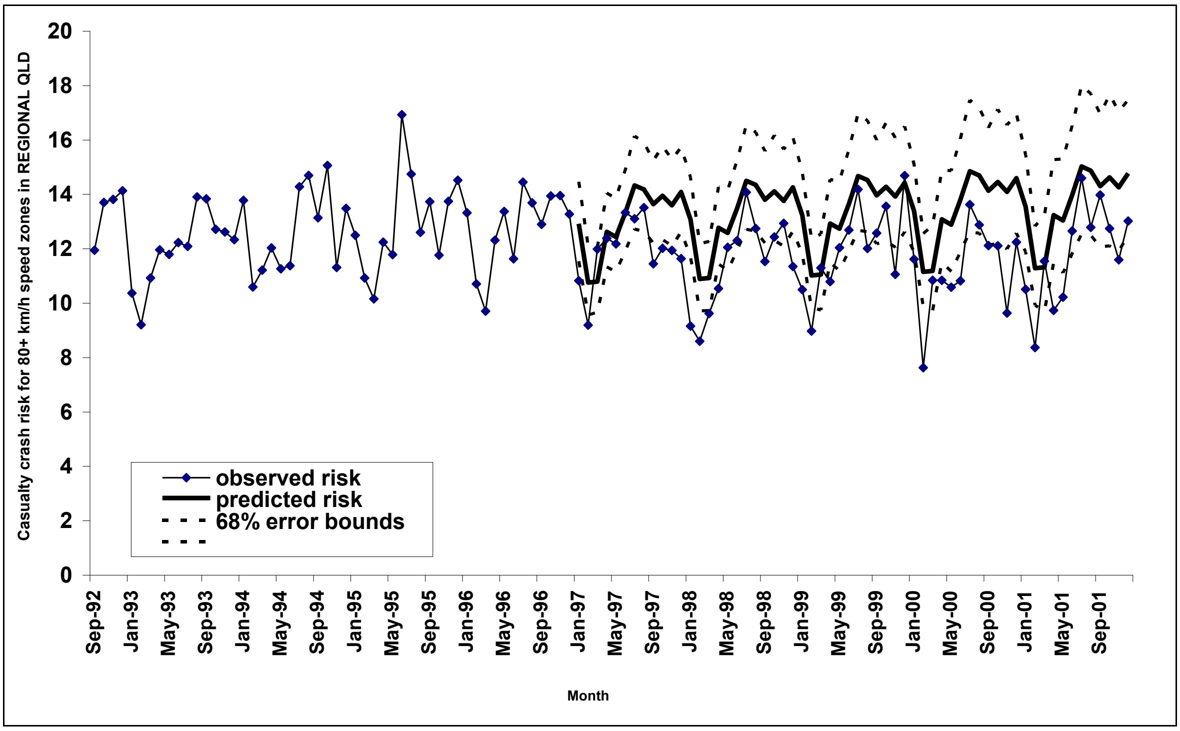


Figure 5.16: Casualty Crash Risk for 80+ km/h speed zones in Regional Queensland: September 1992-December 2001



For speed zones under 80 km/h in regional areas of Queensland the predicted crash risk decreased sharply during January 1997 to December 2001, whilst the observed crash risk remained relatively constant (Figure 5.15).

The observed casualty crash risk in speed zones of at least 80 km/h in regional Queensland was generally less than the corresponding predicted crash risk for the period January 1997-December 2001 (Figure 5.16).

5.3.3 Summary of Second-Tier Modelling

Like the first-tier global assessment models, the second-tier models in the above sections were formulated at the time of implementation to forecast road trauma outcomes in each stratum of interest (e.g. serious casualty crashes during high alcohol times of the week) had the strategy not take place. Actual post implementation road trauma trends were then compared to those forecast to assess strategy effectiveness.

For most of the second-tier models the actual road trauma trends were less than those forecast had the strategy not taken place– particularly during the first few years of projection. In many instances the observed data were outside the confidence limits of the forecast of pre strategy trends showing that the strategy had significantly improved road trauma outcomes from what would have been expected had the strategy not been implemented. In the later months of the prediction period the confidence limits on the pre strategy trend projection typically become very wide meaning statistical comparison of the observed data with the forecast becomes problematic. This is particularly the case in the period more than 2 years after strategy implementation and suggests that comparisons with the forecast should be limited to 2 years of forecast data. In practice, this means that the forecast will need to be re-estimated periodically, for example every 2 years for the examples considered, against which to compare the ongoing success of the strategy. Accuracy of the initial forecast data will rely somewhat on the length of pre-strategy data available for analysis with longer time periods giving more accurate forecasts. The time period at which forecasts need to be re-estimated will depend on how much prior data the initial forecast is based. Similar comments on forecasting accuracy are also relevant to the global assessment level models.

5.4 The Tier 3 Model: EVALUATION OF SPECIFIC STRATEGY COMPONENT EFFECTS

As described, the third tier models in the framework for evaluation of the Queensland road safety strategy aim to model the effects of individual program elements and actions of the Queensland Road Safety Strategy and Action Plans on crash outcomes. The third tier modelling strategy is an extension of the Tier 2 model in that it will typically target the same strata defined in Tier 2 although it could also be applied as an extension to the Tier 1 global assessment model. Instead of modelling historic trends through general level, slope and seasonal terms, the Tier 3 models will include specific measures of road safety program effort under different activity areas as model covariates. In this way, the model estimates the effects of individual initiatives (where there is sufficient data for the estimate to be reliable) by establishing the relationship between measurable road safety program effort and the key strategy outcome measures and relating the real variation in program effort to the reduction in road trauma observed. Results from the Tier 3 modelling process will give specific estimates of the relative contributions of each of the major program elements in the road safety strategy to achieving the measured outcomes.

5.4.1 An Example of Tier 3 Type Modelling: The RSIP Evaluation

Application of the Tier 3 modelling approach to the previous road safety strategy period in Queensland has been demonstrated previously in evaluating the effects of the Road Safety Initiatives Package (RSIP) implemented as part of the strategy which is described in detail below. Undertaking Tier 3 analysis such as in the RSIP example given is a complex process requiring high level statistical expertise to design and execute the analysis successfully. Careful detailed planning is required to successfully construct such models along with a reasonable a-priori knowledge of the road safety programs considered in the analysis in terms of the most appropriate measures of the key operational elements leading to their success. This latter information is typically gleaned from specific formal program evaluations like those described in the Tier 4 models of Section 6.6. Because of these specific requirements and the substantial effort require to achieve these it was beyond the scope of this study to build a successful Tier 3 model from scratch, hence the decision to use the RSIP example. Furthermore, the RSIP study represented a best practice example of Tier 3 modelling that clearly demonstrated all the required elements of the process leading to a proven successful outcome. Formulating a new Tier 3 model specifically for this study would most likely not have improved on the RSIP example to demonstrate the techniques required for this level of modelling and the outcomes that can potentially be achieved.

To address the Queensland road toll, Queensland Transport (QT) and the Queensland Police Service (QPS) developed the RSIP. The RSIP was a continuation of the Holiday Period Road Safety Enforcement and Education Campaign (HPRSEEC) that was trialled between 13th December 2001 and 8th February 2002, and re-implemented from 13th December 2002 to 27th April 2003. The RSIP commenced on 28th April 2003 and continued into 2004.

The RSIP involved the following measures to target the road toll:

1. Increase in the hours of operation of speed cameras, from five to eight hours per camera per day.
2. Increase in the hours of on-road Police enforcement to target the “Fatal Four” behaviours – drink driving, speeding, fatigue, and non-seat belt wearing:
   2.1. To target drink driving: increase in RBTs by buses and other stationary or mobile vehicles.
   2.2. To target speeding: increase in the hours of operation of moving mode radar (MMR) and LIDAR speed detectors.
   2.3. To target non-seat belt wearing and fatigue: increase in Police hours of “Stop and inspect” operations for seat belts, mobile phones and fatigue.
3. Increase of mass-media publicity (planned TARPs) to target the “Fatal Four”:
4. Increase in hours of Police educative activities: For example, educating motorists who have been pulled over about the Fatal Four and publicising analysis of crashes on the previous day.

The RSIP package was evaluated for Queensland Transport by MUARC during 2004 with the results of the evaluation made public in late 2007 (Newstead et al, 2007). The aim of the evaluation was first to examine the effects of the RSIP on crash outcomes in Queensland for the program as a whole as well as for specific program elements including estimation the economic worth of the program. As such, the first part of the evaluation assessing the intervention effects of the RSIP on crashes across Queensland as a whole essentially describes post hoc application of the Tier 1 global assessment modelling approach which statistically evaluated the effect of a program or strategy on crashes over the state as a whole based on observed after implementation data. The second component of the RSIP evaluation describes an application of the Tier 3 evaluation framework modelling approach to the RSIP period of the previous road safety strategy. It related measures of road safety program activity under the RSIP elements to observed crash outcomes over the RSIP period.

This illustration of the application of the Tier 1 and 3 evaluation framework components does not apply to the previous strategy as a whole or even the modified strategy from 1997 onward. However, the aim of this section is to demonstrate the efficacy of the evaluation framework when applied to real world data. On this basis, the RSIP program can be thought of as representing a mini road safety strategy commencing in December 2002 and running to the end of January 2004.

5.4.1.1 Analysis Design and Supporting Data

The evaluation of RSIP crash effects sought to identify relationships between measures of the RSIP and crash outcomes in terms of a time trend in observed monthly crashes through statistical regression modelling. The success of such an approach relies on the ability to effectively represent the majority of factors other than the RSIP that have influenced observed crash counts over an extended time period in order to be able to measure the pure effects of the RSIP. To do this, it was necessary to have accurate measures of the other influential factors and to model the crash data for a period sufficiently long to allow accurate associations between the available measures and the crash outcomes to be firmly established. This required crash trends to be modelled over a time period including the RSIP implementation period but also for a significant time period before the introduction of the RSIP. The basic premise of the modelling approach was to accurately represent crash trends in the pre-RSIP period by the non-RSIP factors included in the regression model and then measure the perturbation from the pre-implementation crash trends once the RSIP program was in place. The perturbation is then inferred to represent the effect of the program on crashes.

A two stage approach to the regression modelling was used. The first models the perturbation on the crash series attributed to a program effect as a single global effect (the Tier 1 model) measuring the overall crash effects of the RSIP on crashes across the whole of Queensland. The second measures the perturbation as a function of measures of key RSIP component activity measures, aiming to measure the crash effects of each of the key RSIP activities (the Tier 3 model).The evaluation of the crash effects covers the time frame December 2002 to January 2004 inclusive. The period from January 1998 to November 2002 was used as the pre RSIP intervention period from which general trends in the crash data were estimated. The length of this period was dictated largely by the available data on key factors influencing the observed crash trends.

Since elements of the RSIP were targeted specifically at crashes occurring in certain regions or at certain times of the day and because Queensland is such a geographically and socially diverse state, it was considered likely that the key factors considered in this study influence road trauma trends differently in different parts of the state. To accommodate this in the crash analysis, the monthly crash data series was stratified for modelling. The first level of stratification was by the 8 Queensland Police Regions. Within each Police Region the monthly crash frequencies were then further stratified into two times of week (High versus Low Alcohol Hours as defined by Harrison, 1990) and two road environments (urban, < 80km/h, versus rural, >= 80km/h) defining 32 crash strata (8x2x2). For each of the 32 strata, the monthly crash counts from January 1998 to January 2004 were assembled for analysis. Other specific applications of the Tier 3 level of evaluation may not need to stratify the crash population in the way described here, particularly if the crash stratum being modelled is relatively homogeneous. In instances where relatively broad strata are defined, such as in the example here, it may be technically useful to further stratify the data for analysis to improve the fit of the model as well as the interpretation of the model outcomes.

Crash data modelled in evaluating the RSIP covered the period January 1997 to April 2004. It included all reported crashes in Queensland over that period with each unit record in the data representing a reported crash. Each record in the crash data contained information on the crash severity, date, time of day, police region and speed limit. These fields allowed crashes to be allocated to each of the 32 analysis strata defined above. Injury outcome in the data is classified into one of five levels based on the injury level of the most seriously injured person involved, being fatal, serious injury (requiring hospital admission), medically treated injury, other injury and non-injury. The data was collapsed into 3 crash severity levels for analysis: fatal and serious injury crashes, medically treated injury crashes and other injury and non-injury crashes.

Measures of RSIP specific road safety program activity were collected under a number of broad program component areas with the data chosen for use largely dictated by what was reliably collected by the relevant authorities. The component areas and the specific component activities along with variable name abbreviations within each area are summarised as follows.

1. Speed Camera Activity: hours of operation, number of sites used and compliance with randomised schedular
2. On-road (non speed camera) Police Enforcement: Moving Mode Radar (MMR) and Laser Speed Detection (LIDAR) operations, Random Breath Testing (RBT) operations, seat belt offences detected and mobile phone offences detected.
3. Mass Media Publicity: Monthly awareness levels of mass media television advertising (AdStock) with the following themes; speed, fatigue, seat belts, and drink driving (see Broadbent, 1979, 1984 for the definition of AdStock)

A number of other road-safety related initiatives were introduced during the RSIP evaluation period, but which were not the central focus of this evaluation. In addition, there was also a change in crash reporting levels in Queensland in late 2000 associated with changes in the rules for making injury compensation claims following a motor vehicle crash. The factors included were as follows.

- Introduction of the default 50km/h local street speed limit in south-east Queensland in June 1999 (enforcement amnesty period from March to May 1999)
- The Holiday Period Road Safety Trial from December 2001 to end of January 2002

- Introduction of the regional 50km/h local street speed limit from May 2003 (enforcement amnesty period from February to April 2003)
- the increase in speeding penalties from April 2003
- the increase in penalties for use of hand-held mobile phones while driving from December 2003
- the change in crash reporting levels from October 2000

Each of the above factors was represented in the statistical model as a step function. Whilst these were not the focus of the RSIP evaluation, in a general application of the Tier 3 evaluation modelling approach, these factors would also be of interest.

In applying the Tier 3 modelling strategy in the context of the RSIP evaluation, it was necessary to include measures of socio-economic effects in the statistical models to accurately describe trends in the crash data driven by factors other than the RSIP. Measures of several socioeconomic factors were included in the statistical models on a monthly and regional basis. These were: population size, unemployment rate, and fuel sales. In general, it would be necessary to include such factors in any application of the Tier 3 strategy to describe latent trends in the data and hence ensure unbiased estimates of the effects of the specific road safety program efforts being assessed.

As noted in describing the stratification of the crash data for analysis of RSIP crash effects, the motivation for the data stratification was to be able to relate certain measures used in the statistical models to those crash sub-populations to which they most directly relate. The measures considered only relating to certain crash strata and the strata to which they relate are as follows: Number of RBTs - High Alcohol Hour crashes, Moving mode radar hours - Crashes on rural roads, LIDAR speed detector hours - crashes on urban roads during Low Alcohol Hours, Seat-belt and mobile phone penalties - Crashes on urban roads during Low Alcohol Hours, Road safety publicity AdStock with drink-driving theme -High Alcohol Hour crashes, South-east Queensland 50km/h local street speed limit - Police regions in south-east Queensland, Regional Queensland 50km/h local street speed limit - Police regions outside of south-east Queensland.

5.4.1.2 Analysis Methods

In order to build statistical regression models that are robust and easily interpreted the first step in the statistical modelling process was to test for co-linearity between the regression input (independent) variables. The presence of potential co-linearity between independent variables in the regression models has been investigated through analysing the correlations between the variables. Variables having a raw correlation co-efficient higher than 0.5 were considered to have a co-linearity high enough to be of concern to the analysis interpretation. In order to eliminate the co-linearity problem, one of each pair of highly correlated input variables was removed from the regression equation. The remaining variable could be considered to represent the effect of both itself as well as the removed highly correlated variable.

The analysis model used in the RSIP evaluation to relate measures of RSIP effort and other factors to observed monthly crash counts was a Poisson Regression model which has been widely used to model crash count data (Nicholson, A.J. 1985a, b, Maher and Summersgill, 1996, Newstead, Cameron & Leggett, 2001). An advantage of the Poisson regression analysis model was the ability to conveniently measure the average RSIP crash effects over all 32 analysis strata through application of single regression model to the entire data stratified data. The overall purpose of the Poisson regression analysis models is to measure the level and statistical significance of association between RSIP program measures and observed crash outcomes. As an alternative, the Tier 3 evaluation could also use a structural time series model as demonstrated in the previous sections. Structural time series methodology is not so straight forward to apply in an integrated manner to a stratified analysis however in application to a single stratum in the Tier 3 context it may produce superior results.

In describing the output from the statistical modelling in the RSIP evaluation example, a shorthand notation for each variable included in the models has been used. A summary of the shorthand notations as well as an indication of the structure and treatment of the variable in the models is given in Table 5.3.

Table .5.3: *Summary of modelling variable shorthand names and structures*

Variable Grouping	Variable Description	Shorthand Name	Structure in Modelling Process
Crash data	Monthly fatal and serious injury crash counts	FATHOSP	Count variable
	Monthly medically treated injury crash counts	MEDICAL	Count variable
	Monthly other and non injury crash counts	MINPROP	Count variable
Stratification Variables	Police Region	REGION	Categorical (8 level)
	Urban or Rural speed zone	METRUR	Categorical (1=Urban, 2=Rural)
	Alcohol Hours	ALCHOUR	Categorical (1=LAH, 2=HAH)
	Month of Crash	MONTH	Numeric (1-12)
	Year of Crash	YEAR	Numeric(1998-2004)
Socio-Economic Variables	Population	POP	Continuous Numeric
	Unemployment	UNEMPLOY	Continuous Numeric
	Fuel Sales	FUELSALES	Continuous Numeric
Other Road Safety Initiatives	S.E. QLD 50km/h local street speed limit	SEQLD50	Categorical (1=Post, 2=Pre)
	Initial Holiday Period Road Safety Trial	HOLTRIAL	Categorical (1=During, 2=Not During)
	Change in Crash Reporting Levels	CRASHREP	Categorical (1=Post, 2=Pre)
	Regional QLD 50km/h local street speed limit	REGQLD50	Categorical (1=Post, 2=Pre)
	Change in Speeding Penalties	SPDPEN	Categorical (1=Post, 2=Pre)
	Increase in Mobile Phone Use Penalties	MOBPEN	Categorical (1=Post, 2=Pre)
RSIP Measures	Drink-Driving AdStock	AD_DRINK	Continuous Numeric
	Speed AdStock	AD_SPEED	Continuous Numeric
	Belt Use AdStock	AD_BELT	Continuous Numeric
	Fatigue AdStock	AD_FATIGUE	Continuous Numeric
	Total AdStock	AD_ALL	Continuous Numeric
	Number of Random Breath Tests	BT	Continuous Numeric
	Hours of LIDAR Enforcement	LASHRS	Continuous Numeric
	Hours of Moving Mode Radar Enforcement	MOBHRS	Continuous Numeric
	Mobile Phone Offences	MOB	Continuous Numeric
	Seat Belt offences	BELT	Continuous Numeric
	Number of Hours of LIDAR and MR enforcement combined	NOCAMHRS	Continuous Numeric
	Hours of Speed Camera use	CAMHRS	Continuous Numeric
	Number of Active Speed Camera Sites	ACTSITE	Continuous Numeric
	Hours of Speed Camera Enforcement per Active Site	HRSACT	Continuous Numeric
Non-RSIP Controlled Speed Camera Variables	Compliance with Randomised Schedular	COMPLY	Continuous Numeric
	Monthly Difference in Active Site Numbers	ACTDIFF	Continuous Numeric
	Monthly Difference in Hours of Operation	HRSDIFF	Continuous Numeric

The Tier 1 (global assessment) intervention model for the RSIP example was specified to include all relevant factors aside from those associated with elements of the RSIP. By not using the RSIP program element measures, the model was designed to measure the total effect of the RSIP over its implementation period, without regard to the mechanisms producing the effect. The specification of the model is given by Equation 5.1.

$$\ln(crash_{asrm}) = \alpha_{asr} + \beta_r^{Pop} Pop + \beta_r^{Unemploy} Unemploy + \beta_{mr}^{Month} Month_m + \beta_r^{year} Year$$
$$+\beta^{SEQLD50} I_{SEQLD50} + \beta^{REGQLD50} I_{REGQLD50} + \beta^{HOLTRIAL} I_{HOLTRIAL} \quad \text{...Equation 5.1}$$
$$+\beta^{SPDPEN} I_{SPDPEN} + \beta^{MOBPEN} I_{MOBPEN} + \beta^{CRASHREP} I_{CRASHREP}$$
$$+\beta^{RSIP} I_{RSIP}$$

In Equation 5.1, a is the index for alcohol hours, s is the index for urban or rural speed zone, r is the index for region, m is the index for month, α and β are parameters of the model and I is an indicator function being 1 if the program/factor is active and 0 otherwise. The RSIP intervention term is the final term in Equation 5.1. It is defined as 1 for the RSIP period and 0 before the RSIP program was implemented and is included as a categorical variable in the analysis model. It is the key term for assessing the intervention effect of the RSIP on crash outcomes.

The long term trend and seasonal terms were included in the intervention model to better describe the data trends in the pre RSIP period. This was necessary because the majority of the RSIP program components were also operational in the period before the RSIP was introduced, albeit at a much lower level of effort or exposure. The trend and seasonal terms were included in the model to represent the potential variation in the monthly pre-program crash counts resulting from the operation of these measures at their pre RSIP levels. To assess the robustness of the estimate of the RSIP intervention parameter in Equation 5.1, a number of different forms of the model were fitted. These included allowing for separate seasonal and yearly trend terms for each of the 32 crash strata rather than just by region and similar modifications for effects of population, unemployment and other road safety programs. The modifications made essentially no difference to the estimated RSIP effect showing the result was robust to fine detail of model specification. Separate analysis models were fitted to the monthly data series of each of the three crash severity levels considered in the evaluation. All models were fitted using the SAS statistical software.

Assessment of the association between specific measure of RSIP program activities and observed monthly crash outcomes (the Tier 3 model) was made through a modification of the Poisson regression model described in Equation 5.1.

$$\ln(crash_{asrm}) = \alpha_{asr} + \beta_r^{Pop} Pop + \beta_r^{Unemploy} Unemploy +$$
$$+\beta^{SEQLD50} I_{SEQLD50} + \beta^{REGQLD50} I_{REGQLD50} + \beta^{HOLTRIAL} I_{HOLTRIAL}$$
$$+\beta^{SPDPEN} I_{SPDPEN} + \beta^{MOBPEN} I_{MOBPEN} + \beta^{CRASHREP} I_{CRASHREP} \quad \text{...Equation 5.2}$$
$$+\beta^{COMPLY} COMPLY + \beta^{ACDIFF} ACDIFF + \beta^{HRSDIFF} HRSDIFF$$
$$+\beta^{AD\_ALL} AD\_ALL + \beta^{BT} BT + \beta^{CAMHRS} CAMHRS$$
$$+\beta^{HRSACT} HRSACT + \beta^{NOCAMHRS} NOCAMHRS + \beta^{BELT} BELT$$

Interpretation of the components of Equation 5.2 is the same as for Equation 5.1. As before, separate models were fitted to the monthly crash counts for each crash severity level considered using the SAS software.

Estimates of percentage crash change attributable to a factor included in the Poisson regression were derived directly from the model parameter associated with that factor via Equation 5.3.

$$\%Change = (1 - \exp(\beta)) \times 100\% \quad \text{...Equation 5.3}$$

For a binary categorical variable, such as the RSIP intervention variable in the intervention model, Equation 5.3 gives the net percentage crash change due to the introduction of the program. For the parameter estimates associated with continuous measures, Equation 5.3

gives the estimated percentage crash change associated with a unit increase in the continuous measure. To calculate the percentage crash change associated with a change in a continuous variable from level a_0 to level a_1, Equation 5.4 was is applied.

$$\%Change = (1 - \exp(\beta(a_1 - a_0)) \times 100\% \quad \text{...Equation 5.4}$$

Equation 5.4 was used in assessing the effect of RSIP project components on crash outcomes during the RSIP period. For this assessment, a_0 was assigned as the average of the RSIP program component in the period before the RSIP was implemented (January 1998 to November 2002) whilst a_1 was the average of the RSIP program component after RSIP implementation.

Statistical confidence limits can be placed on the estimated percentage changes by using the parameter standard error to estimate confidence limits on the parameter and then transforming the confidence limits into percentage changes using Equation 5.3 or 5.4. Estimates of absolute monthly crash savings attributed to the RSIP or its components have been derived from the percentage crash savings through multiplying the estimated percentage crash savings by the average monthly crash count in the pre RSIP implementation period (January 1998 to November 2002). Annual crash savings or crash savings for the entire 14 months post RSIP implementation period were then derived by multiplying the average monthly savings by 12 or 14 respectively. Confidence limits have been estimated for the absolute crash savings by converting the confidence limits on the estimated percentage crash savings into absolute crash savings in the same way. If desired, the crash savings can also be translated into measures of economic worth for each program element considered in the models.

The methods of converting the model parameters into estimates of crash savings for each program element considered is specific to the Poisson model structure and linear form of the model predictor variables used in this particular application of the Tier 3 evaluation framework. Analogous methods could be derived for other forms of Tier 3 evaluation framework models.

5.4.1.3 Tier 3 Test Results

In order to build statistical regression models that were robust and could be easily interpreted, it was first necessary to test for co-linearity between the regression input (independent) variables. The presence of potential co-linearity between independent variables in the regression models has been investigated through analysing the Pearson correlations between the variables over the months for which data was available. Variable pairs having a raw correlation co-efficient higher than 0.5 were considered to have a co-linearity high enough to be of concern to the analysis interpretation. Variable pairs with high correlations were as follows with the level of correlation indicated: Fuel sales with unemployment rate (-0.500); Laser enforcement hours with seat belt offences (0.912), mobile phone offences (0.755); Mobile phone offences with seat belt offences (0.555); All AdStock with belt AdStock (0.704), drink driving AdStock (0.637), fatigue AdStock (0.737), speeding AdStock (0.807); Fatigue AdStock with belt AdStock (0.583); Speed camera hours with fuel sales (0.540); Active speed camera sites with speed camera hours (0.727), fuel sales (0.539), hours per active speed camera site (-0.585)

To overcome the model co-linearity problems, a number of variables were excluded from the analysis. Because of the high correlation, the effect of the excluded variables is represented by the correlated variable included in the analysis. The following variables were excluded, whilst the variable representing the effect of each excluded variable is indicated in brackets: Fuel sales (in favour of unemployment rate), All individual AdStock themes (in favour of total AdStock), Active speed camera sites (in favour of speed camera hours), Mobile phone offences (in favour of seat belt offences). Similar pre-conditioning of independent variables would be necessary in any application of the Tier 3 level modelling approach to ensure the analysis outcomes are readily interpretable.

Total Program (Tier 1 Global Assessment Model) Effects

Estimates of the overall crash effects of the RSIP resulting from post hoc application of the Tier 1 or Global evaluation model estimated using the regression model described in Equation 5.1 are given in Table 5.4. Given along with the estimated crash reduction attributable to the program are 95% confidence limits on the estimate as well as the statistical significance of the estimate. Low statistical significance values indicate the crash effect is unlikely to have arisen through chance variation in the data.

Table 5.4: ***Estimated Total Crash Reductions Attributable to the RSIP***

Crash Severity	Estimated Crash Reduction	Upper 95% Confidence Limit	Lower 95% Confidence limit	Statistical Significance
Fatal + Hospital	13.12%	6.09%	19.62%	0.0004
Medically Treated	14.20%	7.91%	20.06%	<.0001
Other Injury + Non-Injury	4.34%	-0.36%	8.83%	0.0693
All Crashes	8.80%	5.52%	11.96%	<.0001

NB: Negative percentage crash reduction estimates indicate an estimated percentage crash increase.

Table 5.4 shows that the estimated overall reduction in fatal and hospitalisation crashes attributable to the introduction of the RSIP was 13.12%. The estimate was highly statistically significant. A similar estimate of program effectiveness was also obtained for medically treated crashes. In contrast, the estimate for the more minor crash severity levels, other injury and non-injury crashes, was a crash reduction of only 4.34% which was marginally statistically significant. This result indicates that the RSIP was much more effective in reducing higher severity crashes than lower severity crashes. The estimated reduction in all crashes due to the program was a highly statistical significant 8.8%.

Tier 3 Models: Effects of Specific RSIP Program Elements

This analysis aimed to identify those program components principally responsible for the intervention effect measured in the previous analysis. Given the co-linearity problems between input variables had been largely solved by the process of pre conditioning described above, it was considered unnecessary to undertake a model building process. Instead, a single model described by Equation 5.2 was fitted including all the factors remaining after the process of removing co-linearity. As there were no variable co-linearity problems, leaving non-significant factors in the model had little bearing on the coefficient estimates or significant levels of the other factors in the model. Model parameter estimates for the RSIP program elements are summarised in Table 5.5 for each of the models by crash severity.

Table 5.5: ***Parameter Estimates of RSIP Component Crash Effects***

Crash Severity Level	RSIP Program elements	Regression Co-efficient	Standard Error	Statistical Significance
Fatal + Hospital	AD_ALL	-0.000012	0.000003	0.0003
	BT	0.000000	0.000002	0.7477
	CAMHRS	-0.000204	0.000101	0.0429
	HRSACT	0.000156	0.005524	0.9774
	NOCAMHRS	0.000023	0.000021	0.2731
	BELT	-0.000236	0.000141	0.0937
Medical	AD_ALL	-0.000008	0.000003	0.0168
	BT	-0.000003	0.000002	0.023
	CAMHRS	-0.000074	0.000092	0.4212
	HRSACT	0.000103	0.004759	0.9827
	NOCAMHRS	-0.000014	0.000019	0.4686
	BELT	-0.000441	0.000112	<.0001
Other Injury and Non-Injury	AD_ALL	-0.000011	0.000002	<.0001
	BT	-0.000001	0.000001	0.1932
	CAMHRS	-0.000018	0.000062	0.7714
	HRSACT	-0.002599	0.003178	0.4135
	NOCAMHRS	0.000027	0.000012	0.0316
	BELT	-0.000263	0.000079	0.0009
All Crashes	AD_ALL	-0.000011	0.000002	<.0001
	BT	-0.000001	0.000001	0.0481
	CAMHRS	-0.000071	0.000046	0.1218
	HRSACT	-0.001553	0.00238	0.5142
	NOCAMHRS	0.000017	0.000009	0.0747
	BELT	-0.000317	0.000059	<.0001

A number of RSIP program elements were significantly associated with crash outcome in each of the models considered by crash severity level. RSIP program elements of total AdStock (AD_ALL), speed camera hours (CAMHRS) and number of seat belt offences detected (BELT) were estimated to be statistically significantly associated with crash outcomes when considering the more severe crash levels (fatal and hospitalisation). The negative coefficient of the parameter estimate also shows that an increase in these measures was associated with a decrease in observed crash numbers. For the lower crash severity levels (medically treated and below), total AdStock and number of seat belt offences were again significant along with the number of breath tests carried out (BT). It should be recalled that the high correlation between mobile phone and seat belt offences means that the BELT factor in the model is representing the effect of both these factors. It should be noted that a lack of measured association between a program component measure and crash outcome does not imply that the component is ineffective. It is still possible that program components not found to be significantly related to crash outcomes here had a significant and enduring effect on crash outcomes at the time of their initial introduction. It is also possible that the relationship between their variation and crash outcomes is more subtle than can be measured using the analysis design and techniques employed here.

Effects of the significant RSIP program components were translated into percentage crash effects following implementation of the RSIP using Equations 5.3 and 5.4 in conjunction with average program component effort both before and after the implementation of the RSIP. Program effort levels are shown in Table 5.6 along with the estimated percentage crash reductions derived from the regression parameter and change in average monthly program component effort level. The 95% confidence limits on the estimated percentage crash reductions are also shown.

Table 5.6: *Estimates of Percentage RSIP Component Crash Effects*

Crash Severity Level	Significant RSIP Component	Average Pre RSIP Level	Average Post RSIP Level	Estimated Crash Reduction	Lower 95% CL	Upper 95% CL
Fatal + Hospital	AD_ALL	3126	939	-2.66%	-1.35%	-3.99%
	CAMHRS	417	856	8.58%	0.27%	16.20%
	BELT	306	390	1.95%	-0.34%	4.19%
Medical	AD_ALL	3126	939	-1.76%	-0.46%	-3.08%
	BT	20333	24390	1.21%	-0.37%	2.77%
	BELT	306	390	3.62%	1.83%	5.37%
Other Injury & Non-Injury	AD_ALL	3126	939	-2.43%	-1.56%	-3.32%
	BELT	306	390	2.17%	0.90%	3.43%
All Crashes	AD_ALL	3126	939	-2.43%	-1.56%	-3.32%
	BT	20333	24390	0.40%	-0.39%	1.19%
	BELT	306	390	2.62%	1.67%	3.55%

NB: Negative percentage crash reduction estimates indicate an estimated percentage crash increase.

Of the RSIP component measures that were significantly associated with crash outcomes, the increase in speed camera hours was estimated to be associated with the largest crash savings. Table 5.6 shows that an estimated doubling of the speed camera hours was associated with a reduction in fatal and hospitalisation crashes of around 9 percent. This represents over half the total effect of the RSIP on these crash severity levels estimated in the intervention analysis. Increased detection of mobile phone and seat belt offences under the RSIP was estimated to have been associated with a 2-3 percent crash reduction. Increases in random breath testing were estimated to have reduced crashes by only around 1 percent during the RSIP period. If desired, the estimates of the overall percentage crash reduction attributable to the RSIP could be readily converted to absolute crash savings by multiplying the percentage crash effect estimate by the average monthly crash frequency in the pre-RSIP implementation evaluation period.

Tallying the estimated crash reductions associated with each significant RSIP component leads to a total crash reduction associated with these components which is far less than the overall program crash effects measured in the intervention analysis. This implies that there were other components of the RSIP not identified amongst those above that were also effective in reducing crashes. These could include components for which explicit measures were available and included in the model but which did not show significant association with monthly crash counts. If this was the case, the contributing program components must have produced crash effects in a way that was inconsistent with the relationship assumed by the analysis model structure. Further specific evaluation of each program component would be needed to identify such effects. It is also possible that there have been other effects from the RSIP for which no explicit measure was available but which have produced crash savings. One such example is general community awareness of the RSIP generated through media interest in the general implementation of the program. Queensland Police Service indicated this is highly likely, as their officers aimed to get significant media coverage of events such as enforcement blitzes. Unlike scheduled road safety advertising, there is no direct measure of audience exposure to the media coverage generated. Consequently, the evaluation process cannot assess the effect of this unscheduled publicity.

Such limitations as observed here are likely to manifest in many such applications of this Tier 3 modelling approach. Despite this, the demonstration application of the Tier 3 approach on the RSIP example clearly shows the efficacy of the methodology in being able to identify a range of specific major road safety program activities that are significantly associated with observed crash outcomes. Furthermore, the Tier 3 modelling approach successfully identified the relative crash reductions associated with each, thereby giving evidence as to the relative effectiveness of each program component on producing the outcomes observed.

5.5 TIER 4: TARGETTED Evaluation OF SPECIFIC PROGRAMS

A fourth tier of evaluation recommended for the Queensland Road Safety Strategy is specific evaluation of major program components. The Tier 3 models described above measure the general association between measures of specific road safety program activities in a multivariate setting. However, for large and complex road safety program elements, specific evaluation is generally needed for two reasons. First, only specific evaluation can establish the cause and effect relationship between road safety program element implementation and road trauma outcomes with a sufficient degree of scientific rigour. Second, specific evaluation is often needed to establish the measure of road safety program operation that is best related to the outcomes achieved which in turn is fed into the Tier 3 models as a key input. For example, specific evaluation of the Queensland speed camera program identified the 6 key measures of program activity listed in the previous section that best predicted crash outcomes.

For these reasons, it is recommended that Queensland Transport continue to commission specific evaluations of key road safety programs implemented as part of the broader road safety strategy. For the past road safety strategy, specific evaluation was undertaken for a number of key programs with some specific examples undertaken by MUARC detailed in the following.

5.5.1 Random Road Watch

Random Road Watch was a traffic policing program in operation in Queensland during the period of the previous road safety strategy. The difference between Random Road Watch (RRW) and conventional traffic policing is the resource management technique used for scheduling Police enforcement in a manner intended to maximise road safety benefits. Extensively developed from the original model used in the U.S.A., the technique involves dividing each Police jurisdiction into a number of sectors, and the week into a number of time blocks. The sector to be visited and the time at which it is to be visited are assigned randomly with the whole week being enforced. Enforcement involves conspicuous stationing of a marked police vehicle in the chosen sector for the allocated time block to undertake general road safety enforcement duties.

Evaluation of the effects of the Queensland Random Road Watch program on crash frequency using a scientifically rigorous quasi-experimental design showed the program to be effective overall in producing a significant reduction in crashes of all severities in all police regions except Metropolitan South, where reductions were observed but were not generally statistically significant. Crash savings were estimated to be highest for fatal crashes and diminishing with crash severity. Crash effects of the program were estimated to be uniform on fatal crashes across areas of Queensland outside of Brisbane, there was significant variation in program effects on non-fatal crashes between Police regions and urban and rural areas. Further analysis showed that variation in program effectiveness between police regions was strongly related to the crash population covered by the program and less strongly related to the hours of enforcement achieved and offences issued from the program.

Full results of the Queensland Random Road Watch program evaluation are reported in:

Newstead, S., Cameron, M. and Leggett, M. (2001) The crash reduction effectiveness of a network-wide traffic police deployment system *Accident Analysis and Prevention*, 33 (2001), pp393-406.

5.5.2 Mobile Speed Camera Program

The Queensland mobile speed camera program was introduced on 1 May 1997 using an overt mode of operation to maximize general deterrence. It represented a key activity of the previous road safety strategy in Queensland. The cameras were deployed initially in white commercial vans and later in other models of vehicles, marked as a speed camera unit on the side of the vehicle and with a 700mm high sign placed five to ten meters past the vehicle, advising motorists of the camera presence. In the early stages of the program, cameras were deployed to 500 sites, on state controlled roads only where a speed limit review had been completed. Zones were chosen on the basis of crash history and were approved by Traffic Advisory Committees. The sites at which cameras were operated at on a particular day were chosen using a randomised scheduling procedure with some scope for variation. By June 2001, the number of speed camera sites in use had grown to over 2,500 with a proportionate increase in the total number of hours of camera use.

Specific evaluation of the Queensland speed camera program investigated the effects on crashes of the speed camera program in Queensland across the period from its introduction in May 1997 to the end of June 2001. Subsequent updates of the evaluation have investigated the effects through to June 2006 using the same rigorous scientific design and analysis methods employed in the initial evaluation.

Due to the overt nature of the Queensland program, analysis examined crash reductions associated with the program in areas within 6km of speed camera sites that had been used up to the end of the study period. A quasi-experimental study design was employed comparing crash trends in areas within 6km of the camera site to areas further than 6km from camera sites both before and after introduction of the program. Analysis results showed evidence of larger crash reductions nearer to speed camera sites. Significant increases in crash reductions associated with the program were observed over time, reflecting the substantial growth of the program over the study period. As over 70% of reported crashes in Queensland occurred within 2km of a speed camera site, estimated crash effects within this area were considered to be most indicative of program performance. When operating at maximum coverage, the initial Queensland speed camera program was estimated to have produced a reduction in fatal crashes of around 45% in areas within 2km of speed camera sites. Corresponding reductions of 31%, 39% 19% and 21% were estimated for hospitalisation, medically treated, other injury and non-injury crashes respectively. This translates to an annual crash saving in the order of 110 fatal, 1100 hospitalisation, 2200 medically treated, 500 other injury and 1600 non-injury crashes.

In terms of total annual road trauma in Queensland, these savings represent a 32% reduction in fatal crashes, a 26% reduction in fatal to medically treated crashes combined and a 21% reduction in all reported casualty crashes. The benefit cost ratio estimated for the program over the period from its introduction to June 2001 was 47. Comparison of the estimated crash reductions associated with the Queensland speed camera program and program operational measures gave insight into the mechanisms of program effectiveness. Variations in estimated crash reduction over time were strongly related to the size of the overall program as well as the density of enforcement. Periods of program growth were also associated with larger crash reductions beyond that expected from the increasing size of the program alone. Higher levels of true randomness in selection of speed camera sites for operation was also associated with higher levels of crash reduction when comparing differential performance of the program across police regions in Queensland. Finally, higher levels of site coverage per crash were also related to larger crash reductions although this is not reflected at higher severity crashes.

Full results of the initial evaluation are reported in:

Newstead, S., Cameron, M. (2003) *Evaluation of the Crash Effects of the Queensland Speed Camera Program*. Report No. 204, Monash University Accident Research Centre.

Results of further updates of the evaluation can be found on the MUARC web site (www.monash.edu.au/muarc)

5.5.3 50km/h Default Urban Speed Limits.

A 50 km/h default speed limit was introduced in built-up areas in South East Queensland in March 1999. The 50km/h default speed limit was extended to local roads in regional Queensland in February 2003. Following a three-month amnesty period, full enforcement of the 50km/h speed limit commenced in May 2003. The primary aim of both implementations was to reduce the incidence and severity of casualty crashes on the local road network. The effects of both implementations were the subject of specific evaluation with respect to the outcomes on crash frequencies and vehicle speeds.

In South East Queensland, the evaluation found that the implementation was associated with statistically significant average yearly reductions of 88%, 23% and 22% for fatal crashes, all casualty crashes and all reported crashes, respectively. Crash reductions appeared to have increased with time after program implementation for each crash severity level considered. Analysis of speed survey data associated with the program implementation suggested these crash reductions stemmed largely from a reduction in excessive speeding in 50km/h zones rather than large reductions in mean speeds at the affected sites.

Evaluation of effects in areas outside of South East Queensland also found strong evidence of significant crash reductions associated with implementation of the 50km/h default speed limit. Crash reductions were estimated both in the amnesty period immediately after implementation and in the subsequent period during which the new speed limit was fully enforced. In the period of full enforcement, the analyses found statistically significant crash reductions of 13.5% for all crashes reported to police. Greater reductions were estimated in higher severity crashes with a reduction of 24.9% estimated for serious casualty crashes (crashes involving death or hospitalisation) and 19.3% for fatal, hospitalisation and medical attention severity crashes combined. The analyses also found that the 50km/h default speed limit was associated with significant reductions in crashes that involved younger drivers both during the amnesty and full enforcement periods. It also estimated statistically significant reductions in speed related crashes that occurred primarily during the amnesty period. Analysis of speed survey data associated with the program implementation indicates that the reported crash reductions were associated with a reduction in vehicle speeds above 60km/h on roads that became 50km/h subsequent to the default limit introduction, consistent with the patterns observed in SE Queensland.

Full reports on the evaluations can be found in:

Hoareau, E., Newstead, S., Oxley, P. and Cameron, M. (2007) *An Evaluation of 50km/h Speed Limits in South East Queensland*, Report No 264, Monash University Accident Research Centre.

Hosking, S., Newstead, S., Hoareau, E. and Delaney, A. (2007) *An Evaluation of the 50km/h Default Speed Limit in Regional Queensland* Report No 265, Monash University Accident Research Centre.

6. DATA REQUIREMENTS & RECOMMENDATIONS FOR LEVELS 1-3 AND SPECIFIC PROGRAM EVALUATION

6.1 DRINK-DRIVING REQUIREMENTS

On-road alcohol measurements

Driver alcohol measurements should be made at randomly-selected sites during the times of the week when alcohol-involvement is considered to be more prevalent, say between 10 p.m. and 2 a.m. on Friday and Saturday nights

In New Zealand during 2004, 28,898 drivers were stopped at 336 randomly-selected sites and breath-tested between 10pm and 2 am on Friday and Saturday nights during late summer and early autumn. The estimates taken were designed to measure changes occurring from one year to the next by taking measurements at the same sites and at the same times of the night.

A possible design for Queensland, similar to the New Zealand model, is to select 40 urban sites by randomly choosing grid squares on city maps. Potential sites can be chosen within selected squares and the sites visited to check that they are suitable. Sites are deemed suitable if safe parking is available for police vehicles; there are no side roads within visibility of the site where drivers can turn off; the roads chosen are not too minor – although some lower volume roads should be allowed to enter the sample. There also needs to be an agreement with the police to go to these sites at the times and days specified. Operations need to be low profile, as low visibility is possible. A civilian surveyor needs to attend each operation to record start and finish times of the survey; to count the volume of passing traffic, to record or estimate the age and gender of drivers and if possible, the number of passengers and the vehicle type. The surveyor also needs to ensure that survey procedures are followed. Once the surveys have been completed there needs to be some way of downloading the breathalyser data to ensure that it can be matched with the data that has been recorded manually.

Crashed driver alcohol measurements

A further measure that can either supplement, or even possibly replace the above survey, is Police routine breath testing of crash-involved drivers, whether or not they are suspected of drinking. If the rate of testing were high enough (i.e. about 70%-80% or more), then it would be possible to impute drink-driving exposure, given certain assumptions.

Drink-walking measurements

On-site measurements of pedestrian behaviour can be problematic for the following reasons:

- There is a need to ensure surveyor safety when observing pedestrian behaviour at times when there are drinking pedestrians, and
- The ethical obligation to intervene when a pedestrian is intoxicated.

An alternative way of obtaining drink-walking measurements would be to investigate the use of security camera footage to record behaviour, which is coded the following day.

This would depend on the adequacy of existing cameras, or whether there are funds to install new cameras.

The only other practical measurement would be to consider crashes involving a drunk pedestrian – need to be cautious about any coding issues that may arise here. This would be a very imperfect measure, however, due to the limited number of such incidents expected.

Drink-Driving Program Data

It is recommended that the following data relating to drink-driving operations be collected:

- Number of random breath tests conducted per month per region for each type of operation (i.e. Booze bus, mobile patrol car, stationary car)
- Number of drink-driving infringement notices issued per month per region
- Number of hours of RBT hours conducted by Queensland Police per month per region
- For each RBT session the following data should also be collected and recorded electronically in a database:
    - The PBT (preliminary breath test) reading of each driver tested
    - The sex and (estimated) age of each driver
    - The exact location of each session (in terms of a map reference or nearest road reference)
    - The volume of traffic

It should be noted that some of these variables are already collected by Queensland Police, however, they should continue to be collected in a systematic way and in the same way from year to year so that comparisons can be made annually.

6.2 Speed-related measurements and requirements

Speed surveys

In New Zealand, about 125 sites are surveyed annually, each for about 2 hours, measuring speeds of about 30,000 free-running vehicles. It is recommended that similar surveys should be conducted in Queensland however fewer sites would be needed unless many sub-national estimates were required.

Separate surveys would be required per major speed limit type or road type for which estimates are required. For each road type or each speed limit area type, 20 surveys would provide good estimates of speeds. The sites should be chosen so that they are uniformly not impeding vehicle speeds via road features or surroundings. An attempt should be made to measure free speeds, where the speed adopted by the driver is not directly affected by that of the vehicle in front. All speeds (not just free speeds) could be measured, but these are less sensitive to changes in driver choice of speeds and are subject to variation solely due to changes in congestion, which is not interesting from the perspective of this study.

Speeds can be measured using automated devices that also measure vehicle headways, or by manual measurements using laser devices, for example. It is desirable to classify vehicles by type, i.e. car, truck, motorcycle.

Measurements should be unobtrusive and should be taken for at least two hours per site. Each site should be surveyed at the same time of year, on the same day of the week and at the same time of the day (unless 24-hour data are collected) for comparative measurements from year to year.

Police speed enforcement data

It is recommended that the following speed enforcement continue to be collected and collated systematically and in the same format annually for future evaluations of Road Safety Strategies**:**

- Number of hours of speed camera operations per month per Police region
- Number of speed camera traffic infringement notices issued per month per region
- Number of speed camera detections and speed camera deployments
- Number of offences detected through mobile radar operations per month per region
- Number of Police hours of mobile radar operation per month per region
- Number of offences detected through laser operations per month per region
- Number of Police hours of mobile radar operations per month per region.

6.3 Pedestrian and Bicyclist Measurements

Improvement of bicycling and pedestrian facilities, as indicated in the 2004-2011 Road Safety Strategy, may increase the use of these facilities by pedestrians and cyclists. This means that some measure of exposure is desirable, both to measure the increased use and to control for this increased use when measuring crash rates. To reduce the cost of this measure, it may be desirable to reduce the scope of the measure to groups most likely to be affected. For example, surveys in the form of roadside counts, could focus on children going to and returning from school. A measure of bicycle helmet use would also be a logical output from such a survey. It would also be desirable to restrict the scope to urban sites; to a particular time of year (e.g. late summer); to fine weather (i.e. not raining) and even to a particular day of the week (as patterns may vary from day to day in a pattern). These restrictions are necessary so that consistent results can be achieved from one year to the next.

In New Zealand an annual national survey of helmet use by bicyclists of all ages is carried out during March and April. About 5,700 bicyclists are observed at 58 sites throughout New Zealand. Each site is surveyed for two hours, typically between 8 a.m. and 9 a.m. and between 3 p.m. and 4 p.m. on normal school weekdays. It is recommended that this survey method be used for Queensland, however fewer sites would be required than the New Zealand survey, because New Zealand publishes estimates at sub-national levels. Thus, a sample of about 20 sites may be adequate, surveyed for two hours per site. Data gathered would include estimated age group; gender of bicyclist/pedestrian; whether the pedestrian crossed the road; whether the cyclist was wearing a helmet, and a five minute count of passing traffic (during which no other observations are made).

6.4 FATIGUE MEASUREMENTS

It would be possible to measure both attitudes and behaviour relating to fatigue based on self-report from drivers stopped at remote service stations, for example. Fatigue-related behaviour (e.g. time since last break/sleep, amount of driving today, etc.) should be recorded before any attitudes are recorded in order to avoid sensitising responses. In addition, before and after surveys should be conducted to measure any change associated with a particular intervention.

6.5 CHILD RESTRAINTS

The annual national survey of child restraint use in New Zealand by children under 5 years is carried out in September and October. Children in more than 5000 cars are observed at 110 sites throughout New Zealand. Sites are surveyed during school hours in the school term. Results are weighted to reflect the population under five in each local area. Restraint use is observed (the vehicles are not stopped) generally in entrances to supermarket car parks.

A similar design would be feasible for Queensland. Once again a smaller sample of about 20 sites would be adequate if only state-wide estimates are required. The same sites should be used from year to year.

6.6 YOUNG DRIVERS

As young learner drivers were part of the focus of the 2002-2004 Action Plan, a survey would logically focus on this group by observing L-plate use at institutions such as secondary schools, universities and other tertiary institutions. A number of such institutions should be sampled and L-plates observed in student car parks. The required sample sizes may be difficult to determine. However, a purposive sample of five of each of these institutions at which all cars parked in student car parks are observed during term time mid-morning should provide an adequate sample.

6.7 ROAD SAFETY ADVERTISING DATA

Road safety advertising data was provided to MUARC for an earlier evaluation conducted for Queensland Transport. This road safety advertising data covered the period 1998 to 2002, and consisted of data covering various road safety themes such as fatigue, seatbelts, fatal four, speed and drink-driving.

For this current project data was requested in a similar format to that provided earlier for the periods 1993-1997 and 2003-2004. However it was not readily available.

It is therefore recommended that weekly road safety advertising data (i.e. Target Audience Ratings Points or TARPs) for each road safety campaign be collected by QT regularly (i.e. per quarter) and stored electronically in Excel format. If possible, the TARPs data should be collected for advertisements that were screened in both metropolitan and rural regions of Queensland. In addition the data should be collected by the theme of the road safety advertisement (i.e. drink-driving, speeding, fatigue, inexperienced drivers, motorcycle safety, seat-belt wearing etc.), and by the style of the advertisement (e.g. emotive, enforcement, educational).

Other studies have found a link between road safety advertising and road trauma outcomes, so it is important that this data is collected regularly and systematically for future evaluations of road safety programs.

6.8 OTHER POLICE ENFORCEMENT DATA

In addition to data relating to drink-driving and speeding operations the following Police enforcement data concerning other driver offences should be collected annually:

- Number of seatbelt-wearing related offences detected per month for each police Region
- Number of mobile phone offences detected per month for each police Region
- Number of drug-related driving offences detected per month for each police Region
- Number of hours of drug-driving operations conducted per month per Region.

6.9 CRASH DATA REQUIREMENTS

In addition to the crash data variables already recorded in the Queensland database of Police-reported crashes the following information should also be collected and collated:

- The Injury status of each person involved in a crash.
- Classification of the type of road on which the crash occurred.

There should also be an update of the crash database from 2006 onwards, and crash data files should be provided at both a crash level and person-injury level.

6.10 SOCIal and ECONOMIC DATA

6.10.1 Unemployment and Employment Rates

A number of variables exist which may reflect the state of the economy. The data on these variables is of differing periodicities. Much economic data is collected at the national level and on an annual or quarterly basis however this data is unlikely to give an accurate representation of the state of the economy in Queensland. A problem with either annual or quarterly data is that it is of a longer periodicity than the variables to be modelled. It is therefore recommended that data on any economic indicators should be collected at a monthly level and at a regional level, if available. The monthly economic data that needs to be collected for future evaluations of road safety strategies or programs in Queensland should include the unemployment rate (%) for S.E. Queensland and for Regional Queensland as well as the number of persons employed in the workforce in both regions.

6.10.2 Population

Population numbers can be considered as a measure of exposure to risk, and were utilised in this study to estimate crash risk in the modelling process. It is therefore recommended that monthly population estimates for each region of Queensland be collected. If monthly population figures are not available, then quarterly or annual estimated population numbers should be interpolated to monthly data. This interpolation technique smoothes the series and makes it a weaker factor to use in modelling, as most of its variability has been smoothed out of the series. Nevertheless population has shown a relationship with crash frequencies.

A better measure of exposure would be the number of driver or motorcycle rider licences on issue in Queensland.

An even better measure of exposure would be an estimate of vehicle travel obtained from on-road travel/exposure surveys, conducted per annum in Queensland. The conceptual population would be each of the kilometres of out-of-house surface travel by persons in Queensland, whether it is by foot, bicycle, motorcycle, motor vehicle of public transport, on each of the days of the year, and for each of the years in future.

6.10.3 Fuel sales

Total fuel sales is often used as a proxy for vehicle travel in studies of crash data which require some measure of exposure to risk. Fridstrom et al. (1995), for example, use total gasoline sales in their study of factors influencing crashes in four Scandinavian countries.

For Queensland it is recommended that fuel sales continue to be collected on a monthly basis for future evaluations.

6.10.4 Alcohol sales

Experience with modelling road trauma trends in Victoria has shown the levels of alcohol sales in that state to have a strong positive relationship with the number of serious casualty crashes occurring during high alcohol hour of the week (Cameron et. al, 1993; Thoresen et al, 1992). The measure of alcohol sales used was the monthly value of retail alcohol trade in Victoria in dollars deflated by the tobacco and alcohol component of the Consumer Price Index for the Melbourne capital city area. This gave a monthly series of retail alcohol sales. It is recommended that analogous monthly retail alcohol trade figure be obtained and collated for Queensland.

6.10.5 Environmental factors

A number of past studies on factors influencing road crash numbers have found strong links between environmental factors and crash numbers. Fridstrom et. al. (1995) in a study of factors affecting crash frequencies in four Scandinavian countries found significant associations between crash numbers and rainfall, snowfall and sunshine hours. Whilst snowfall is not a relevant factor in the context of Queensland road trauma, rainfall and sunshine hours should be considered as potential explanatory factors to consider in crash modelling.

Environmental data for Queensland should be collected and this data should include:

- The monthly total rainfall per weather district/station in Queensland;

- The average number of sunshine hours per day by month for each weather district in Queensland.

Once this data is obtained, it can be aggregated across individual weather stations to give monthly average rainfall and sunshine hours for South East Queensland and for regional Queensland.

7. SUMMARY

This project developed an evaluation framework to monitor and report on the outcomes of the 2004- 2011 Queensland Road Safety Strategy and its associated two-yearly action plans. The previous strategy, namely, the 1993-2003 Queensland Road Safety Strategy was used to test-run the evaluation framework.

7.1 LITERATURE REVIEW

A review of Australian and international literature on best-practice approaches for evaluating road safety strategies and action plans recommended that the GOSPA model be used as a basis to define the elements of the evaluation framework.

The GOSPA framework for evaluation was developed by MUARC for application in Western Australia. This method relates evaluation elements to a strategic planning framework (Cameron, 1999). The structure of the GOSPA framework is as follows:

***G*oal**	General (idealistic) statement of the Program's overall goal
***O*bjectives**	Specific (pragmatic) statements of the Program's measurable objectives to reach the Goal
***S*trategies**	General (idealistic) strategies to achieve each Objective
***P*rograms/Plans**	Specific (pragmatic) programs/plans contributing to each strategy, with measurable activity levels and outputs
***A*ctions**	Actions undertaken in each program.

Based on the literature review of existing practices, an adaptation of the GOSPA evaluation framework was considered as the most useful basis for establishing an evaluation framework for the Queensland Road Safety Strategy. This framework reflected that the Queensland Road Safety Strategy, like most other road safety strategies in Australia, is structured in terms of an overall goal, objectives, strategies, programs and actions.

7.2 STRUCTURE OF EVALUATION FRAMEWORK

The 2004-2011 Queensland Road Safety Strategy and associated action plan for 2004-2005 were expressed to fit the GOSPA Framework as follows in Table 7.1:

Table 7.1: Queensland Road Safety Strategy and Action Plans structured as the GOSPA Framework

GOSPA Framework Components	***2004-2011 Queensland Road Safety Strategy and 2004-2005 Action Plan***
Goal	To minimise crash severity and reduce long-term consequences of injuries
Objectives	To achieve a reduction in the number of fatalities to less than 5.6 deaths per 100,000 people by the year 2011, and to achieve a reversal in the increasing trend in hospitalisation casualties and the hospitalisation rate.
Strategies	• Safe attitudes and behaviours and optimal health outcomes in the event of a crash; safe roads, safe road environments and safe management of traffic. • To target broad key road user groups and behaviours, e.g: Drink-drivers; speed-related crashes and speeding drivers; fatigued drivers; young adult inexperienced drivers; older road users; intoxicated pedestrians; unrestrained occupants; unlicensed drivers and riders.
Programs	e.g.: Police enforcement programs targeted at drink-driving and speeding; public education campaigns targeted at fatigue and speed; RACQ'S 'Years Ahead' program targeted at older drivers (a full list of programs is given in Table A in Appendix B of the report)
Actions	E.g: increase in hours of random breath test operations and speed-related enforcement operations; implement countermeasures to address drink-walking; trial the use of automatic number plate recognition technology targeted at unlicensed drivers (a full list of actions is given in Table A in Appendix B of the report).

7.3 THE PROPOSED EVALUATION FRAMEWORK: A THREE-TIERED MODELLING APPROACH

The GOSPA framework defines a pyramid of increasing detail in defining the elements of a road safety strategy. The top of the pyramid defines the broad goals for which the strategy is aiming (Goals) whilst the next level down gives specific measurable targets (Objectives) against which the goals can be assessed. The Strategies area of the framework typically defines the target areas on which the road safety strategy will focus to achieve its goals along with local objectives within each target area that will jointly contribute to achieving the global objectives. Finally the Programs and Actions areas contain the specific details on the type of activities to be carried out in each target area and the amount of effort that will be applied to each activity.

An evaluation framework for the Queensland road safety strategy has been developed in this study that mirrors the pyramid structure defined under the GOSPA framework. It is designed to assess the progress of the strategy against the pre-determined goals and objectives at various levels of detail through a multi-tiered modelling approach. Each tier of modelling focuses on a specific level of disaggregation of the strategy elements corresponding to particular levels of the GOSPA framework pyramid. These are described broadly as follows:

- **The Global or First Tier Assessment Model**: This level of assessment aims to measure the ongoing performance of the road safety strategy in achieving the broad **G**oals and **O**bjectives set out for the strategy as a whole. Specifically it aims to measure whether the strategy has reduced overall road trauma levels, as defined by measures specified in the strategy Objectives, from that expected based on pre strategy implementation trends.
- **The Second Tier Asssessment Model:** This level of assessment has aims similar to the First Tier models except here assessment is aimed at each individual target strata defined by the **S**trategies of the Queensland Road Safety Strategy and corresponding Program areas. Assessment at this level will aim to articulate in what particular areas the strategy is working and to what degree.
- **The Third Tier Assessment Model:** This level of assessment aims to relate the trends observed in each of the target strata defined in the Second Tier models to explicit measures of road safety program for major activities defined in the **P**rogram and **A**ction elements of the strategy.

Table 7.2 depicts how the GOSPA framework relates to the three-tiered models and fits in with the test-running of the framework on the earlier strategy.

Table 7.2: Summary of the link between the GOSPA framework and the three-tier modelling approach

<table>
<tr><th colspan="2">GOSPA FRAMEWORK COMPONENTS</th><th>THREE-TIER MODELLING</th></tr>
<tr><th>Component</th><th>Definition</th><td rowspan="3">Global assessment model (top-tier model) to measure effect on road trauma of the Strategy overall (includes an intervention model)</td></tr>
<tr><td>Goal</td><td>Overall goal of strategy (i.e. to prevent road trauma through safe road use, safe roads and safe vehicles)</td></tr>
<tr><td>Objective</td><td>Objectives to reach goal (e.g. to achieve a reduction in the fatality rate to under 5.6 deaths per 100,000 people)</td></tr>
<tr><td>Strategies</td><td>General strategies to achieve objectives given in the Qld action plans and road safety strategy</td><td>Second-tier modelling of specific strata targeted by the strategies in the action plans (e.g. crashes occurring during high alcohol times of the week)</td></tr>
<tr><td>Programs</td><td>Specific programs relating to target group outcomes</td><td rowspan="2">Third-tier modelling of the individual program elements of the strategy (e.g RSIP evaluation)</td></tr>
<tr><td>Actions</td><td>Actions undertaken in each program</td></tr>
</table>

A fourth tier of the evaluation framework is also included which describes the imperative of undertaking specific targeted evaluations of major road safety programs implemented or enhanced under the road safety strategy. Such evaluations will give the most rigorous scientific assessment of the effectiveness of the programs evaluated and contribute to understanding of the mechanisms of effectiveness. Furthermore, results of the specific evaluations are critical in assisting the formulation of the Tier 3 models by identifying the most relevant measures of program activity that predict the outcomes being measures.

The statistical methodology used in tiers one and two of the three-tiered evaluation framework was based on structural time series (or 'state-space') modelling theory. Structural time series models are a general class of time series model that allow the most ready selection of parsimonious model structures to reflect the inter-correlated nature of time series data. The also allow the easy inclusion of covariates into the model structure and provide simple an accurate forward forecasts of trends. Forecasts from the pre-strategy time series data were used to estimate expected post intervention trends in the absence of strategy implementation. Actual road trauma outcomes were them plotted against the forecasts to assess strategy effectiveness either globally or within particular target strata. Observed post strategy implementation data below the forecast indicated the strategy was successful in reducing road trauma from levels expected had the strategy not been implemented. Performance can be assessed month by month by simply plotting actual road trauma levels against the trends forecast in the absence of the strategy. Additional time series models were also fitted to both the pre and post strategy implementation data to formally assess the impact of the strategy in reducing road trauma by including post strategy intervention terms in the model.

To demonstrate how the evaluation framework works in practice, a test-run of the framework was trialled on the previous 1993-2003 Queensland Road Safety Strategy. Sufficient crash data prior to 1993 was not available to support the modelling process so it was decided to trial the evaluation framework on assessing the outcomes of the revision to the strategy implemented from 1998. Detailed description of each tier of the evaluation framework follows along with an assessment of its practical application to the previous Queensland road safety strategy. 'State-space' modelling techniques were used to fit models to the Queensland fatal and injury crashes for the period April 1991-December 2004. Explanatory factors were also included in the models to improve fit and predictive power as well as give insight into how changes in these factors might affect forecast trends. Factors included were the monthly unemployment rate and monthly fuel sales for Queensland.

7.4 DETAILED DESCRIPTION OF THE FRAMEWORK ELEMENTS AND REVIEW OF THEIR PRACTICAL APPLICATION

7.4.1 Global Assessment Model

The Global Assessment Model measures success in achieving the broad goals and objectives of the Queensland Road Safety Strategy. For the 2004-2011strategy this was to achieve a reduction in the Queensland road fatality rate to under 5.6 deaths per 100,000 people by the year 2011, and to achieve a reversal in the increasing trend in hospitalisation casualties and the hospitalisation rate.

The global assessment model aimed to measure the effect on road trauma of the Queensland Road Strategy overall. Road trauma was quantified in terms of fatal crash risk, serious casualty crash risk and all injury crash risk.

For the 'test-run', the global assessment model was formulated in two different ways as described. In the case of formulating the evaluation model at the commencement of the 1993-2003 Strategy (revised in 1998), a state-space time series model was estimated that modelled road trauma levels each month before the introduction of the (revised) Strategy (i.e. Pre 1997) and forecasts the levels of road trauma that would have been expected (together with confidence limits on the forecast) to have occurred after 1997 had no strategy been in place (based on the past trends). Against the forecasts from the resulting model were plotted the actual road trauma trends in Queensland that occurred after the Strategy was introduced. Plotting of the actual trend against that forecast in the absence of the strategy can be easily achieved by Queensland Transport staff without statistical training as a means of monitoring overall strategy performance over time. Only the initial forecast would need to be estimated by trained statistical practitioners.

The global models that were estimated demonstrated how road trauma levels can be modelled each month prior to the introduction of the Strategy and then used to forecast the levels of road trauma that would have been expected to have occurred after 1997 had no strategy been in place. The actual road trauma trends that occurred in Queensland after the introduction of the (revised) 1993-2003 Strategy were also plotted. Plotting the actual trends against forecasts in the absence of the Strategy was used as a means of monitoring overall strategy performance over time.

The demonstration of the evaluation framework appeared to have worked best for the first few years of the projection (i.e. 1997 to late 2000). However in the last year of the prediction period (i.e. late 2000 to 2001) assessment of the actual values against the forecast was difficult due to the typically wide confidence limits on the forecast. This suggests that there may be a need to re-calibrate the forecasts at regular intervals, say every 2 years to correspond with the development of each new action plan, based on a longer data period. The evaluation strategy would then measure the incremental improvement in global outcomes related to each action plan period.

The time-series models that investigated road trauma trends pre and post the Queensland Road Safety strategy against a forecast trend post strategy in the absence of the strategy are similar in philosophy to the control-chart methodology used by a number of agencies in the past, including Queensland Transport, to monitor road safety strategy performance. The advantage of the methodology for the global assessment model used in this study is that it employs much more sophisticated and robust statistical methodology yet is still amenable to use by those without statistical training once established.

Intervention Model

The global assessment model was also used at various time periods after the implementation of the 1993-2003 Strategy to formally evaluate the overall performance of the Strategy to that point in time based on the key outcome measures specified in the Strategy's goals and outcomes. The pre and post implementation data to the time available were modelled using state-space techniques with annual intervention terms being included at the time of Strategy implementation. The intervention term parameters then represent the effect of the strategy on the outcome measure which can be tested formally for statistical significance. The intervention terms can be modified accordingly to reflect increasing effects of the broad Strategy over time which might be expected if components of the Strategy are introduced in a staggered manner over time or take some time to become fully effective. Application of the global assessment model in this manner would require high level trained statistical expertise.

For the ‘trial-run’, at the global assessment level, the performance of the 1993-2003 strategy was measured using an intervention type forecasting model. As a demonstration of how the intervention analysis works, the fatal crash risk global assessment model was used at various time periods after the Strategy was implemented (i.e. annual step functions at the end of each year for the post-period 1996-2003) to formally evaluate the overall performance of the Strategy to that point in time. The intervention model showed indications of annual reductions in fatal crash risk in December 1998-December 2001, associated with the implementation of the Queensland Road Safety Strategy.

The negative coefficients associated with the annual step functions indicate that there has been a period of sustained reduction from about 1998 to 2001 before levelling off again after that. Although the step intervention functions fitted to the model were not statistically significant at the 5% level due to the limited amount of data when considering only fatalities, they are indicative of a reduction in the fatal crash risk and demonstrate the efficacy of the intervention model. A statistically more definitive result would be achieved if an intervention model was fitted that also included injury crashes as well as fatal crashes.

The results of the intervention analysis suggest that the ‘test-run’ of the evaluation framework on the earlier Strategy captured some of the performance of the Strategy by finding an effect (a reduction in fatal crash risk) during 1998-2001.

7.4.2 Second-Tier Models

Key strata were identified according to the way that the current and previous strategies are targeted. For example, the program element “to provide at least 170,000-190,000 police officer hours towards random breath testing over each year of the Action Plan” of the 2004-2005 Action Plan is likely to affect mainly crashes occurring during high alcohol times. The Action Plan thus identifies high alcohol crashes as an important stratum that should then be used in the analytical framework of the evaluation. An evaluation approach that uses time series also needs to account for the effects of previous and existing road safety measures and strategies in order to estimate the effects of a new strategy over and above the effects of an older strategy. Thus, strata identified in previous strategies also need to be accounted for in the analytical framework.

The second tier model relates to the objectives defined by the strategies and programs of the 2004-2011 Queensland Road Safety Strategy and the 2004-2005 associated Action Plan. This model considered the effects on road trauma for specific strata defined by the road user groups or situations (eg. high alcohol crashes; speed-related crashes) at which the strategies and programs are targeted.

For each of these strata, a specific analysis model equivalent in structure to that defined by the global assessment model above was estimated for each of the key outcomes being measured. Like the global assessment models, the second tier models were formulated at time of implementation to forecast road trauma outcomes in each stratum of interest had the Strategy implementation not taken place. Actual post implementation road trauma trends were then compared to those forecast to assess Strategy effectiveness. Intervention models can also be estimated at time points after Strategy implementation to formally assess the statistical significance of outcome changes related to the Strategy for each strata defined above. Expertise required for each approach is the same as for the global assessment model.

Like the first-tier global assessment models, the second-tier models (as part of the 'test-run') were formulated at the time of implementation to forecast road trauma outcomes in each stratum of interest (e.g. serious casualty crashes during high alcohol times of the week) had the Strategy not take place. Actual post implementation road trauma trends were then compared to those forecast to assess Strategy effectiveness.

The modelling results showed that for most of the second-tier models the actual road trauma trends were less than those forecast had the 1993-2003 Strategy not taken place– particularly during the first few years of projection.

In many instances the observed data were outside the confidence limits of the forecast of pre strategy trends showing that the strategy had significantly improved road trauma outcomes from what would have been expected had the strategy not been implemented. In the later months of the prediction period the confidence limits on the pre strategy trend projection typically become very wide meaning statistical comparison of the observed data with the forecast becomes problematic. Again, this is particularly the case in the period more than 2 years after strategy implementation and further suggests that comparisons with the forecast should be limited to 2 years of forecast data. Accuracy of the initial forecast data will rely somewhat on the length of pre-strategy data available for analysis with longer time periods giving more accurate forecasts. The exact time period at which forecasts need to be re-estimated will depend on how much prior data the initial forecast is based. Similar comments on forecasting accuracy are also relevant to the global assessment level models.

Intervention analyses for the second-tier models were not attempted as part of this study due to time and budget limitations but are expected to be successful in implementation based on the results of the Tire 1 application.

7.4.3 Third-tier models

The third tier models attempted to model the individual program elements and actions (for example the effects of speed camera operations) of the Queensland Road Safety Strategy and Action Plans on crash outcomes.

The third tier modelling strategy is an extension of the second-tier model in that it typically targets the same strata defined in tier 2. However, instead of modelling historic trends through general level, slope and seasonal terms, the model includes specific measures of road safety program effort under different activity areas as model covariates. In this way, the model makes estimates of the effects of individual initiatives (where there is sufficient data for the estimate to be reliable) by establishing the relationship between measurable road safety program effort and the key strategy outcome measures and relating the real variation in program effort to the reduction in road trauma observed. The results from the third-tier modelling process give specific estimates of the relative contributions of each of the major program elements in the road safety strategy to achieving the measured outcomes.

Application of the Tier 3 modelling approach to the previous road safety strategy period in Queensland has been demonstrated previously in evaluating the effects of the Road Safety

Initiatives Package (RSIP) implemented as part of the strategy. Undertaking Tier 3 analysis such as in the RSIP example given is a complex process requiring high level statistical expertise to design and execute the analysis successfully. Careful detailed planning is required to successfully construct such models along with reasonable a-priori knowledge of the road safety programs considered in the analysis in terms of the most appropriate measures of the key operational elements leading to their success.

This latter information is typically gleaned from specific formal program evaluations. Because of these specific requirements and the substantial effort require to achieve these it was beyond the scope of this study to build a successful Tier 3 model from scratch, hence the decision to use the RSIP example. Furthermore, the RSIP study represented a best practice example of Tier 3 modelling that clearly demonstrated all the required elements of the process leading to a proven successful outcome. Formulating a new Tier 3 model specifically for this study would most likely not have improved on the RSIP example to demonstrate the techniques required for this level of modelling and the outcomes that can potentially be achieved.

The demonstration application of the Tier 3 approach on the RSIP example clearly shows the efficacy of the methodology in being able to identify a range of specific major road safety program activities that are significantly associated with observed crash outcomes. Furthermore, the Tier 3 modelling approach successfully identified the relative crash reductions associated with each, thereby giving evidence as to the relative effectiveness of each program component on producing the outcomes observed.

7.4.4 Fourth-tier models

A fourth tier of evaluation recommended for the Queensland Road Safety Strategy is specific evaluation of major program components. The Tier 3 models described above measure the general association between measures of specific road safety program activities in a multivariate setting. However, for large and complex road safety program elements, specific evaluation is generally needed for two reasons. First, only specific evaluation can establish the cause and effect relationship between road safety program element implementation and road trauma outcomes with a sufficient degree of scientific rigour. Second, specific evaluation is often needed to establish the measure of road safety program operation that is best related to the outcomes achieved which in turn is fed into the Tier 3 models as a key input.

For these reasons, it is recommended that Queensland Transport continue to commission specific evaluations of key road safety programs implemented as part of the broader road safety strategy. For the past road safety strategy, specific evaluation was undertaken for a number of key programs by MUARC, for example

- the mobile speed camera program
- Random Road Watch
- 50km/h default urban speed limits in South East and the rest of Queensland.

Demonstration of the evaluation of specific programs (such as those listed above) is outside the scope of the evaluation framework being described and is unnecessary given the range of excellent examples related to the previous Queensland strategy that already exist. It is essential that specific programs in the current and future Queensland road safety strategies are independently evaluated in order to understand the value of individual programs, particularly those programs that are expensive in terms of implementation or enforcement, but require a high level of statistical and analytical expertise to evaluate properly.

8. REFERENCES

Aeron-Thomas, A.S, and Hess, S. (2005). *Red-light cameras for the prevention of road traffic crashes.* The Cochrane Database of Systematic Reviews 2005; (Issue 2).

Breen, J. (2004). *Review of the Road Safety to 2010 Strategy,* 170/04, National Road Safety Committee.

Broadbent, S. (1979). "One way TV advertisements work". *Journal of the Market Research Society, Vol. 21, No. 3 London.*

Broadbent, S. (1984), Modelling with AdStock. *Journal of the Market Research Society,* Vol. 26, No. 4. London.

Cameron, M. (2002). *An evaluation of the Supplementary Road Safety Package: Review of initiatives and outcomes.* In: An evaluation of the Supplementary Road Safety Package July 1995 to June 2000, Wellington, Land Transport Safety Authority.

Cameron, M. (1999) 1997?. *Methodology for evaluation of the Enhanced Traffic Enforcement Program in Western Australia.* Draft (?) report, Monash University Accident Research Centre.

Cameron, M,, Haworth, N., Oxley, J., Newstead, S. & Le, T. (1993). *Evaluation of the Transport Accident Commission Road Safety Television Advertising.* Report No. 52, Monash University Accident Research Centre, Victoria.

Cameron M, Fitzharris M, Corben B, Jacques N (2002), *Advice to assist the development of an Enhanced Enforcement Program 2002: Part A - Principles of enforcement and site identification*. Conducted for TAC.

Cameron M, Hoareau E (2002) MUARC's evaluation of 50 km/h speed limits in S.E. Queensland, Transport Statistical Steering Committee, Brisbane.

Cameron MH (1997) Review of modelling of road casualties in the Netherlands, Vol Report No. D-97-18. Institute for Road Safety Research SWOV, Leidschendam, The Netherlands.

Cameron MH (1999) A method of evaluation based on a strategic planning framework, Road Safety Research Policing and Education Conference, Canberra pp. 801-08.

Cameron MH, Newstead SV (2000) Response by Monash University Accident Research Centre to 'Re-investigation of the effectiveness of the Victorian Transport Accident Commission's road safety campaigns' (White, Walker, Glonek and Burns, November 2000), Vol Report No. 177. Monash University Accident Research Centre.

Elvik, R. (2000). *The importance of confounding in observational before-and-after studies of road safety measures.* Thirteenth International Cooperation on Theories and Concepts in Traffic Safety Workshop, Corfu.

Finnish National Road Administration (1999). *Road Safety Program 2005*, Finnra, Helsinki.

Fridstrom, L., Ifver, J., Ingebrigtsen, S., Kulmala, R. & Thomsen, L.K. (1995). Measuring the contribution of randomness, exposure, weather and daylight to the variation in road accident counts. *Accident Analysis and Prevention, Volume 27, No. 1, pp 1-20, 1995.*

Gelau, C., Fischer, P. and Pfeiffer, M. (2000). *Methodological guidelines for evaluating Police enforcement projects,* VTT (Finland).

Haaland, K. and Odeck, J. (1998). *Economic evaluation methods for road projects in member countries: Compiled information on existing methodologies*. Final report to C9 (financing and economic evaluation) Committee of PIARC, Norwegian Public Roads Administration, Oslo.

Haight FA (1986) Risk, especially risk of traffic accident., *Accident Analysis and Prevention* 18:pp. 359-366.

Harrison, W.A. (1990) *Update of alcohol times as a surrogate measure of alcohol-involvement in accidents.* Research Note, Monash University Accident Research Centre, Clayton, Victoria.

Harvey AC (1989) Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Harvey AC, Durbin J (1986) The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling, *Journal of the Royal Statistical Society. Series A (General)* 149:pp. 187-227.

Haworth, N. and Vulcan, P. (2000). *Performance indicators and evaluation methodology for the new road safety strategy 2000-2005,* Monash University Accident Research Centre.

Hoareau, E. and Newstead, S. (2004). *An evaluation of the default 50km/h speed limits in Western Australia.* Report no. 230. Monash University Accident Research Centre, Melbourne.

Hoareau, E., Newstead, S., Oxley, P. & Cameron, M. (2007). An evaluation of 50 km/h speed limits in South East Queensland, Report No. 264. Monash University Accident Research Centre, Melbourne.

Hosking S, Newstead S, Hoareau E, Delaney A (2007) An evaluation of the 50km/h default speed limit in regional Queensland. Report No. 265, MUARC, Melbourne.

Johnson C (2001) Towards accountability: Narrowing the gap between NGO priorities and local realities in Thailand. Overseas Development Institute, London. pp. 1-19.

Johnston, I. (2001). *A brief review of Victoria's road safety outcomes and strategies. Report to Department of Treasury and Finance*, Monash University Accident Research Centre, Melbourne.

Leggett, M. (1999). *Results from the use of a statewide strategic framework to reduce injuries from road crashes.* 3rd National Conference on Injury Prevention and Control, Brisbane, Australian Injury Prevention Network.

LTSA (2000) Travel Survey Report 1997/98. Land Transport Safety Authority, Wellington, New Zealand.

LTSA (2000). *Working Paper 6 Safety Directions: Predicting and costing road safety outcomes.* Land Transport Safety Authority, Wellington.

Maher, M.J., and Summersgill, I., (1996) *A Comprehensive Methodology for the Fitting of Predictive Accident Models.* Accident Analysis and Prevention 28(3) 281-296.

Meuleners, L., Hendrie, D., Legge, M. and Cercarelli, L.R. (2005). *An evaluation of the effectiveness of the Black Spot programs in Western Australia,* Report RR 155. Injury Research Centre, University of Western Australia, Crawley.

Moauro, F. and Savio, G. (2001), "Temporal Disaggregation using Multivariate Structural Time Series Models", in www.oecd.org/dataoecd/60/4/21781478.pdf. Organisation for Economic Cooperation and Development.

Montgomery, D.C. (1991) *"Introduction to Statistical Quality Control",* 2nd Edition, John Wiley & Sons**.**

Newstead, S. (2000). *Assessment of the relative performance of the control chart methodology in application to road safety program evaluation*, Monash University Accident Research Centre, Melbourne.

Newstead, S., Cameron, M. and Leggett, M. (2001). The crash reduction effectiveness of a network-wide traffic police deployment system. *Accident Analysis and Prevention*, 33 (2001), pp393-406.

Newstead, S., Bobevski, I., Hosking, S. and Cameron, M. (2007). *Evaluation of the Queensland Road Safety Initiatives Package,* Report No. 272. Monash University Accident Research Centre; Report for Queensland Transport.

Newstead, S. and Cameron, M. (2003). *Evaluation of the crash effects of the Queensland speed camera program,* Report no. 204. Monash University Accident Research Centre, Melbourne.

Newstead, S., Cameron, M., Gantzer, S. and Vulcan, P. (1995). *Modelling of some major factors influencing road trauma trends in Victoria 1989-93,* Report no. 74. Monash University Accident Research Centre, Melbourne.

Newstead, S., Cameron, M. and Leggett, M. (1999b). *Evaluation of the Queensland Random Road Watch program,* Report no. 149. Monash University Accident Research Centre, Melbourne.

Newstead S, Cameron M, Leggett M (1999a) Evaluation of the crash effects of the Queensland Random Road Watch Program, *Road Safety Research Policing and Education Conference,* Canberra pp. 499-508.

Newstead, S., Cameron, M. and Narayan, S. (1998). *Further modelling of some major factors influencing road trauma trends in Victoria: 1990-96*, Report no. 129. Monash University Accident Research Centre, Melbourne.

Nicholson, A.J., (1985a). *The Variability of Accident Counts.* Accident Analysis and Prevention 17(1): 47-56.

Nicholson, A.J., (1985b). *The Randomness of Accident Counts.* Accident Analysis and Prevention 18(3): 193-198.

RSC (2004). *Arriving Safely: Road safety strategy for Western Australia 2003-2007. Summary of performance indicators for 2003,* Monash University Accident Research Centre for Road Safety Council of Western Australia.

Strategy Branch (1994). *Road safety initiatives: Crash outcome evaluation for period ending 30 June 1994,* Report no. 18/94. Road Transport & Safety Division, Queensland Transport, Brisbane.

Thoresen, T., Fry, T., Heiman, L. & Cameron M. (1992). *Linking economic activity, road safety countermeasures and other factors with the Victorian road toll.* Report No. 29, Monash University Accident Research Centre, Victoria.

Wegman, F. (2002). *Review of Ireland's road safety strategy*, Report no. R-2002-27. SWOV Institute for Road Safety Research, Leidschendam.

Wong, S.C., Leung, B.S.Y., Loo, B.P.Y., Hung, W.T, and Lo, H.K. (2004). A qualitative assessment methodology for road safety policy strategies. *Accident Analysis & Prevention 2004; 36(2):281-293.*

APPENDIX A

A METHOD OF EVALUATION BASED ON A STRATEGIC PLANNING FRAMEWORK

Maxwell H. Cameron

Monash University Accident Research Centre

Introduction

The Office of Road Safety in Western Australia sought the development of a methodology to evaluate the Enhanced Traffic Enforcement Program (ETEP). ETEP is an initiative designed to optimise WA Police Service traffic enforcement activities and the effective administration of fine collection. Currently ETEP is focused on reducing speed-related crashes through the improved use of speed and red light cameras. ETEP Phase 2, implementation of which will commence during 1999, will provide the infrastructure and technology to expand the number of WA Police Service speed cameras.

This paper initially outlines a framework on which the short and long term evaluation of ETEP could be based. The evaluation plan is based on a suggested strategic planning framework for ETEP. Because of the key role of ETEP in WA's Integrated Road Safety Program, it is necessary for the framework to go beyond ETEP to cover the strategic plan for the whole Program, at least at a general level. Full details of the evaluation plan derived from this basis, and recommended to the Office of Road Safety, are given in Cameron (1999).

Phases of ETEP

The first phase of ETEP commenced in June 1998. The objectives of Phase 1 were to enhance the current system through the optimisation of the usage of existing speed cameras and to increase the total operational speed camera hours. A related objective was to initiate an increase in the revenue stream from penalty payments for camera-detected offences, one-third of which is received by the Road Trauma Trust Fund. Phase 1 was completed during 1998/99.

ETEP Phase 2 has the objective of developing and implementing a fully integrated electronic system which will link the detection, processing and enforcement of camera-detected offences and have the capacity to handle an increased number of offences. When the system is built, the objective is to expand the number of operational speed cameras to 30 and further increase operational hours. Phase 2 is expected to be completed by December 2000, at which time the 30 cameras will become fully operational and the offences detected will be processed within targeted timeframes.

ETEP Phase 3 will commence at the conclusion of Phase 2. The levels of operational camera hours and offence processing volumes achieved at the end of Phase 2 are expected to be continued during Phase 3 for the foreseeable future (at least until levels of speeding and camera-detected offences decrease due to deterrence processes achieved by the program).

While it is possible that ETEP Phase 2 will achieve further reductions in speeding and speed-related road trauma over and above that which may have been achieved during Phase 1, most of Phase 2 is expected to be a maintenance period during which the new offence processing system is built before the additional cameras are delivered. Thus the most substantial changes in speeding behaviour and associated crashes could be expected to be seen from the commencement of the fully operational phase, Phase 3, rather than during Phase 2. For this reason, this evaluation plan is focused on both Phase 2 and Phase 3 of ETEP, with little distinction between them except to note that the more substantial effects of the program are likely to be found during the latter period. However, the evaluation framework described in this paper is expected to be applicable during each of Phases 2 and 3.

Strategic Planning Framework

Within the Integrated Road Safety Program (IRSP), ETEP appears to have a number of objectives and targets to help the IRSP achieve its goal and objective. One of the objectives of ETEP appears to be to generate a revenue stream which will fund the other activities of the IRSP as well as ETEP itself. Thus ETEP is an integral part of the IRSP, both in terms of the direct contribution which it is expected to make to the reduction of speed-related road trauma and also, because without the revenue to the Road Trauma Trust Fund which it generates, other components of the Program could not be funded.

To develop an evaluation plan for a program it is necessary to have a clear vision of the measurable objectives and targets for the whole program and each of its sub-programs. It is necessary that the objectives of the program be clearly stated, preferably in quantitative terms, before an evaluation of the program against those objectives can be undertaken. For this reason, the objectives and components of the IRSP and ETEP were structured in terms of the GOSPA model of strategic planning, that is:

Goal	General (idealistic) statement of the Program's overall goal
Objectives	Specific (pragmatic) statements of the Program's measurable objectives to reach the Goal
Strategies	General (idealistic) strategies to achieve each Objective
Programs/Plans	Specific (pragmatic) programs/plans contributing to each strategy, with measurable activity levels and outputs
Actions	Actions undertaken in each program

The GOSPA model allows an evaluation structure to be defined. Programs are not considered to be implemented unless actions are taken, strategies are not achieved unless planned programs are implemented, objectives are not met unless the strategic directions of programs are correct, and goals are not achieved unless the targets of objectives are met.

Assessing whether the appropriate actions were taken and whether sub-programs were implemented as planned is known as *process evaluation.* Whether the total program (eg. the IRSP) produces real change at each level is known as *outcome evaluation.* An important part of outcome evaluation is an assessment of the impact on the criteria defined in the objectives (usually crash-, injury- or behaviour-related criteria), known as *impact evaluation.*

The process evaluation has parallels with normal management procedures associated with implementing a complex project, but should go beyond normal requirements to record regularly and systematically the key measures of activity of major sub-programs. This information will be valuable in attempting to separate the contributions of the individual sub-programs to meeting the overall objectives.

The outcome evaluation aims to assess to what extent the strategies have been achieved and the program objectives have been met. Whether the strategies are achieved does not guarantee that the objectives will be met, but if the strategies are well chosen, it could be expected that this is highly likely. The nature of road safety programs is such that it is usually possible to obtain information on the extent to which the strategies have been achieved well before it is clear that the program objectives are being met. It is for this reason that emphasis should be placed on clearly articulating the strategies for the overall program, as well as defining the strategic objectives.

Casting ETEP and the IRSP in the GOSPA framework

A strategic planning framework, using the GOSPA model, for the IRSP and ETEP was developed as shown in Figures 1 and 2. ETEP Phases 1 and 2 have been expanded upon under Strategy 1: “Reduce speeding behaviour” in Figure 2. The existing performance indicators for ETEP Phase 1, and those additional indicators proposed for Phase 2, were suggested as specific Strategic Objectives and Action Targets where they seem appropriate. Targets for some of the actions to be taken in ETEP Phase 2 had not been fully defined.

Details of the programs and actions related to Strategies 2, 3 and 4 are not shown in the Figures. Two of the existing performance indicators for ETEP were suggested as being Strategic Objectives for Strategy 4: “Make community attitudes more positive to ETEP and IRSP”. This was treated as a separate strategy because of its broader role in the IRSP.

Although a framework has been developed in Figures 1 and 2 which potentially embraces the whole of the IRSP, the evaluation plan for ETEP excluded evaluation of whether the objectives of Strategy 3 are achieved and how this contributes to the achievement of the overall objective of the IRSP. It was considered that the programs in Strategy 3, while potentially funded by the revenue stream produced by ETEP, otherwise fell outside the speed-related objectives of ETEP and that their evaluation should be treated separately.

Figure 1 : STRATEGIC FRAMEWORK FOR EVALUATION PLAN

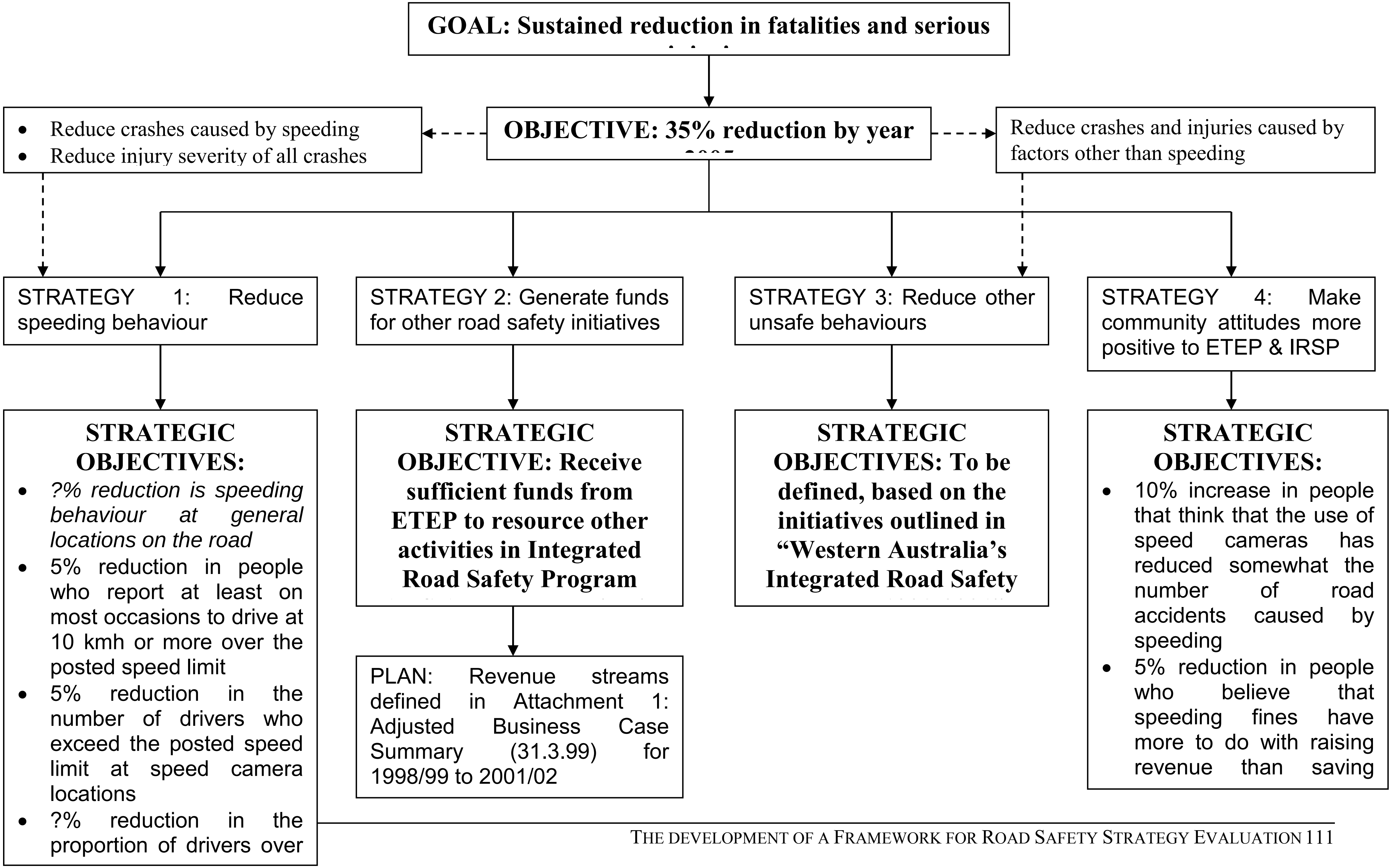


Figure 2 : STRATEGIC FRAMEWORK FOR EVALUATION PLAN (cont.)

STRATEGY 1: Reduce speeding behaviour

8.1

PROGRAM: ETEP Phase 1

PROGRAM OBJECTIVES: To optimise the usage of existing speed cameras (Stage 1) and to increase the operational speed

STAGE 1

ACTION: Optimisation of current speed cameras (and new red light cameras)

TARGET: 25% increase in monthly infringement detection

STAGE 2

ACTION: Increase operational speed camera hours

TARGET: Achievement of benchmark of 56 operational hours per speed camera per week by the 5 optimised central speed cameras

ACTION: Provide additional resources to process infringements

TARGETS:

- 5% increase in the number of prosecutable images

PROGRAM: ETEP Phase 2

PROGRAM OBJECTIVES: To expand the number of operational speed cameras to 30 and to provide offence

ACTION: Increase operational speed cameras to 30

TARGET: Acquisition of additional cameras to bolster operational cameras to 30

ACTION: Optimisation of all current and new speed cameras

TARGETS:

- 10% increase in people who think that in the last 6 months the amount of speed enforcement has increased
- 10% reduction in people who agree you can drive up to 10 kmh over the limit and not risk being booked
- all central speed cameras to maintain the benchmark of 56 operational hours per camera per week
- all regional and metropolitan district speed cameras to maintain the benchmark of 42 operational hours per camera per [illegible]

General Issues in Developing the Evaluation Plan

Focus of the evaluation plan

ETEP can be evaluated at the level of each action, program, or strategy, and their respective contributions towards achievement of the Integrated Road Safety Program goal. In the framework presented in this paper, most of the existing performance indicators and service level targets were assigned as evaluation criteria for Action Targets. In this role, they play a valuable part of the process evaluation of the two ETEP phases, ie. evaluating whether the program was implemented as planned.

However, the process evaluation does not address whether a successfully implemented plan has achieved the objectives of the program. It was decided that the evaluation plan should give emphasis and priority to methods to check whether the strategic objectives of ETEP have been achieved and whether the program has contributed towards the IRSP goal. Compared to the current performance indicators, this puts greater emphasis on behavioural criteria (observed and intended behaviour) and road trauma outcomes related to speed.

The priority focus of the evaluation plan was on the outcomes of the ETEP and IRSP programs. This in turn put a focus on the impact evaluation of the programs and the choice and analysis of appropriate criteria for measuring whether the objectives and goals have been achieved. Following this emphasis, there was a need for the impact evaluation to be designed in such a way that the causes of the changes in the evaluation criteria are clear.

Cause and effect

An important issue in conducting an evaluation is to ensure that a change in the evaluation criterion is due to the action or program under consideration, rather to some other factor. This is less of an issue for criteria so closely linked to specific actions (eg. many of the Action Targets specified in Figure 2) that cause-and-effect is not questionable. In these cases, the simple monitoring of changes in the measured criteria over time is adequate.

However, in the case of criteria for the IRSP overall Objective and the Strategic Objectives, it is possible that any changes may be due to other factors (eg. economic and social factors, changes in traffic patterns, weather fluctuations, and perhaps national road safety initiatives). It is more likely that other factors will distort the apparent effects of ETEP on the criteria for these objectives than they will affect the conclusions from the Action Target criteria. For this reason, the impact evaluation needs to be based on scientific methods and principles. Methods and principles of scientific impact evaluation are described in Cameron (1995). It is not sufficient just to monitor changes in the criteria related to the IRSP and ETEP Objectives.

Impact criteria

The impact evaluation needs to be based on criteria which are appropriate to the objectives of ETEP and IRSP. For example, the strategic planning framework in Figures 1 and 2 suggests that Strategy 1 is to “reduce speeding behaviour”. Hence the principal criterion among the strategic objectives should be one or more measures of actual speeding behaviour on WA’s roads, collected in a way that can be considered representative of all

WA drivers. For this reason it was recommended that an additional strategic objective was required, shown in italics in Figure 1, related to changes in speeding behaviour at general locations on the road.

Two of the existing performance indicators relevant to Strategy 1 were based on speed behaviour measured by speed cameras at camera sites. Because WA's cameras are operated overtly, the measurements are likely to be a biased indicator of actual speeds. Specific measures of actual speeding behaviour, and suggested methods to collect the data, were developed as part of the evaluation plan.

While the overall objective of the Integrated Road Safety Program is a reduction in overall serious road trauma (fatalities and serious injuries), it was considered inappropriate to use all serious road trauma as the criterion in an impact evaluation of ETEP. The direct effect of ETEP (ie. ignoring those parts of the Program aimed at assisting Strategy 3) is aimed at speed-related road trauma.

Drivers exceeding the speed limit or travelling at excessive speeds contribute to serious casualties by causing crashes due to their speed and/or by increasing the injury severity of crashes which occur (not necessarily just those caused by their speed). Thus the criteria for assessing the direct impact of ETEP should be of two types:

1. The number or proportion of crashes in WA caused by speeding
2. A measure of the injury severity of all crashes occurring in WA.

Suitable data sources, definitions, and the reliability and sensitivity of potential criteria that could be used in an impact evaluation of the direct effect of ETEP are outlined in the evaluation plan.

It was recommended that the overall objective of the Integrated Road Safety Program (*"35% reduction in fatalities and serious injuries by the year 2005"*) should be split into objectives related to the direct effect of ETEP, namely:

- Reduce crashes caused by speeding
- Reduce injury severity of all crashes,

and an objective related to the other initiatives in the IRSP, namely:

- Reduce crashes and injuries caused by factors other than speeding.

These expanded program objectives are shown in italics in Figure 1. Appropriate target levels for the extent of the reductions in the two ETEP-related objectives, consistent with the IRSP overall target, were also developed as part of the evaluation plan (Cameron 1999).

Summary and Conclusion

A strategic planning framework for ETEP was developed to provide the basis of the evaluation plan. Because of the key role which ETEP plays in the Integrated Road Safety Program (IRSP), it was necessary to expand the framework to cover the whole IRSP, albeit in less detail for the initiatives not concerned with speed-related road trauma.

The strategic planning framework defined IRSP's goal, objectives, strategies, strategic objectives, programs, program objectives, actions and action targets. The ETEP phases were treated as key programs in this framework. The objectives and action targets were defined as specific measurable criteria.

The priority focus of the evaluation plan was on assessing outcomes related to the higher-level objectives defined in the framework. The impacts on the criteria defined for these objectives were given emphasis, using scientific methods to establish cause-and-effect. The lower-level action targets were considered important for assessing the process of implementation of ETEP. However, the process evaluation should be given lower priority than the outcome evaluation.

It was concluded that structuring a complex road safety program in a strategic planning framework assists the systematic and rational evaluation of the program. The framework makes obvious the key evaluation criteria and their interactions with lower-level measures.

The evaluation plan included five recommendations for short-term initiatives to provide a better basis for conclusive evaluation of ETEP in the longer-term, namely:

1. The impact evaluation criterion representing injury severity in crashes should be a six category classification of injury outcome based on Police crash reports matched with hospital admission records. The matched file should be created on an annual (or more frequent) basis for use in the evaluation of the direct effect of ETEP, and for other monitoring and research purposes.

2. The sensitivity of each of the impact evaluation criteria (number of speed-related crashes, and the injury severity of all reported crashes) should be investigated by examining the monthly trends in each criterion during 1992 to December 1998 and comparing these with the increased speed camera activity levels over the same period.

3. An evaluation of the impacts of speed camera activity on the speed-related road trauma criteria at camera sites should be conducted. This would test the sensitivity of the impact evaluation criteria to the known changes in speeding behaviour at camera sites, and would provide a short-term indication of the benefits of ETEP, albeit limited to road sections in the vicinity of cameras.

4. An annual survey of speeding behaviour at general locations on WA's roads should be established. The survey should cover at least four major classes of road type, chosen on the basis of the most heavily trafficked roads, and aim to collect observations on the free speeds of at least 50,000 vehicles travelling on each class of road.

5. If a program of closely-spaced community education campaigns aimed at a range of road safety issues (including speeding themes supporting ETEP) is established, an on-going series of awareness surveys should be undertaken on a weekly or fortnightly basis. These surveys should measure spontaneous and prompted awareness of the key messages by the target groups (including decision-makers), as well as measuring whether the messages are received positively and considered relevant.

APPENDIX B

TABLE OF PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QUEENSLAND ROAD SAFETY ACTION PLAN

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
Safer roads							
Improve the safety of the road system							
1. Implement the 'Safer Roads Sooner' package which includes the Targeted Road Safety Initiative to improve sites or road sections with poor crash records. (**MR**)	**1**	Y – scheduled finish date 2007/2008;			-identify sites/road sections with poor crash records -select treatments -before/after casualty crash analysis (2002/2003 with 2004/2005)		
2. Map existing enforcement interception sites and explore opportunities to develop new sites, including the potential for shared facilities with rest areas. (**MR* / QPS / QT/ LG**)	**2**	Y- MR to map sites & explore opportunities for new sites –scheduled for completion 2007/2008		i) (increased?) enforcement at new and existing sites ii) (lower?) offence rates at increased enforcement sites –Police behaviour change	- measure crash rates at increased enforcement sites		
3. Improve incident management procedures on highly trafficked stretches of the road network to reduce secondary crashes. (**MR* / LG* / QPS / DES**)	**1**	Y – in Dec 04 modifications made to allow for quick removal of abandoned or stationary vehicles			-measure change in the number of secondary crashes	- measure the change in the severity of secondary crashes	

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
4. Implement a road safety audit policy. (**MR* / LG***)	**1**	Y – districts have been advised to implement policy					
Improve road safety on local roads							
6. Assist local governments to take a proactive approach to road network safety, including the use of crash data and road safety audits. (**QT* / MR* / LG**)	**1**	Y- ARRB has produced a network version of the Road Safety Risk Manager for use by LGs	Educate local govts about proactive approaches to road safety		Reduction in the frequency & severity of crashes on local govt. roads		
Improve the safety of at-risk road users							
7. Implement Main Roads' Cycling Policy and develop guidelines for the safe management of cycling on major roads. (**MR**)	**1**	? – guidelines not started due to lack of resources -implement MR actions identified in State Cycle Strategy	Surveys of cycle use Has there been an increase since cycling policy was implemented? Surveys of other road users re cyclist safety on major roads		Reduction in crash frequency & severity involving cyclists on major roads		
9. Implement best practice planning and design of cycling and pedestrian facilities for local roads. (**QT* / MR / LG**)	**1**	Development to be finalized by Dec 2005	Survey cyclists and pedestrians about their attitudes to	Observational surveys of pedestrians and cyclists on local roads	Reduction in frequency and severity of crashes involving		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N=No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
			safer cycling/ped facilities	before & after implementation	cyclists & pedestrians on local roads (before & after implementation)		
Develop intelligence about making roads safer							
10. Use the learnings from coronial investigations and legal cases involving crashes on the road network. (**MR* / LG**)	1	Y – access to coroners fatal road crash database has been granted			Consider fatal crash trends	Improved survival rate of victims due to quicker response times at crash site; - before/after study – focus on rural roads	
Optimise emergency response times to crashes							
12. Implement enhanced communication, dispatch, and traffic management capability, including rural/remote communications, priority access (e.g., green wave technology) and integrating traffic management and dispatch systems. (**DES*/QPS*/MR***)	1	N – due for completion 2007					
Safer vehicles							
14. Introduce a system for assessing the impact of new vehicle-based technologies (e.g., collision avoidance systems) on road users, and communicate the results to manufacturers and the community. (**QT**)	1	Y – scheduled for completion by Dec2005	Survey community & manufacturers on safety benefits of new vehicle-based technologies				
Influence national standards to make vehicles safer							
17. Advocate a pole crash test for use in ANCAP to test vehicles for their crash worthiness in single vehicle crashes. (**QT* / RACQ***)	1	Y- pole test series on 4WD occurred			Single vehicle crashes – Crashworthin		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
		Oct2004-Mar2005; The new pole test replaced standard side impact test in Nov2005			ess ratings (used car safety ratings) to see if new test provides better side impact protection		
Improve the safety of heavy vehicles							
19. Facilitate industry uptake of under-run protection devices for heavy vehicles. (**QT**)	**1**	Y-expected to be completed Dec2005			Has there been a reduction in under-run heavy vehicle crashes		
Alcohol and drugs							
Improve deterrence of drink driving							
24. Potentially trial an improved intelligence system for deploying all forms of traffic enforcement (including Random Breath Testing) based on crash data and local offense data and knowledge to promote its general deterrent effect. (**QPS* / QT**)	**1**	Expected to be completed end of 05/06; (still at the development phase?)	- Change in no. of RBTs - change in no. over the BAC limit -has new intelligence system changed attitude of drivers to drink-driving; measure this through surveys & RBT changes	- undertake an on-road drink-drive survey	Has there been a reduction in alcohol-related crashes due to improvement in RBT effectiveness?		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
26. Provide at least 170,000-190,000 police officer hours towards random breath testing over each year of the Action Plan. (**QPS**)	**2**	Y – 2004 finalised (222,683 hours of RBT in 2004)	-increase specific & general deterrence effect of RBT in Qld -get #RBTs and #RBT hours per month per region in Qld -link to attitudinal/ advertising surveys	- undertake on road drink-drive surveys	-has there been a reduction in the number & severity of high alcohol hour (HAH) crashes in metro/rural areas		
Reduce the involvement of intoxicated pedestrians in crashes							
28. Pilot local drink walking counter measures in Indigenous and rural communities throughout Queensland. (**CARRS-Q**)	**1**	Y ? – scheduled to commence June 2005	-promote responsible drinking & anti-drink walking strategies through local communities - undertake attitudinal surveys	-observational surveys of pedestrians during HAH times and at specific (drinking) locations in indigenous/ rural QLD communities	- consider alcohol-related pedestrian crashes in rural Qld		
29. Implement counter measures (e.g., education, infrastructure) to address drink walking at tertiary education institutions. (**QT**)	1	Y- expected finish June 2006	-underake surveys of tertiary students re drink-walking	-observational surveys of drink-walkers at tertiary institutes	-consider alcohol-related pedestrian crashes at tertiary institutions		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
					and involving tertiary students		
30. Work with the Liquor Licensing Commission, Queensland Hotels Association & Queensland Health to research and trial innovative counter measures for drink walking. (**QT* / QPS**)	1	Y - expected completion December 2006	Vulnerable road users (including drink walking) campaign	-observational surveys of pedestrians during HAH times and at specific (drinking) locations	has there been a reduction in number & severity of alcohol-related pedestrian crashes ?		
Fatigue related crashes							
Improve community awareness of fatigue as a road safety problem							
31. Launch a new public education campaign to promote the dangers of driving tired, or without due care and attention. (**QT**)	1	Y- The MICROSLEEPS campaign ran over the 2004 Christmas holiday period and other key risk periods. -87% prompted awareness (Feb 2005)	Surveys will measure changes in drivers attitudes and behaviour towards driving tired, and awareness of campaign material.	-on-road survey of drivers asking questions about driving whilst tired (with a focus on heavy vehicle drivers) -combine with other on-road surveys	Has there been a reduction in crashes where fatigue was a factor? - need to define how to measure fatigue		
Reduce the impact of fatigue on road crashes							
33. Continue to take a leadership and innovation role in updating the national heavy vehicle driving hours provisions. (**QT**)	2	Y-draft model. Legislation endorsed by Aug 2005			-has there been a reduction in fatigue-related		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
					heavy vehicle crash costs ?		
Speed related crashes							
Continue to implement Queensland's Speed Management Strategy							
38. Undertake key enforcement activities such as the delivery of at least 43,800 hours of speed camera activity and 140,000 hours of non-camera speed enforcement activity during each year of the Action Plan. (**QPS**)	**2**	Y- 2004 complete (In 2004 QPS conducted 214,072 officer hours of non camera speed enforcement activity."	-no. of hours of Speed Camera Operation by month/district -no. of officer hours of non-camera speed enforcement by month/district – specific & general deterrent effect of speed camera program in Qld -no. of speeding offences	-undertake on road speed surveys	- has there been a reduction in number and severity of crashes due to QLD speed camera program?		Qld Speed camera program has resulted in ave. annual savings of 110 fatal crashes; 1100 hospitalisation crashes; 2200 medically treated crashes; 500 other injury crashes; 1600 non-injury crashes - $2.8 billion savings in social costs to community during implementation period
39. Deliver a targeted public education campaign with complementary enforcement activity that will position speeding as a socially unacceptable behaviour. (**QT* / QPS**)	**1**	Y- ongoing (campaign to run through 2005/2006) New campaign	-community attitude surveys. Decreases in the % of people	-speed surveys to determine if speed behaviour has changed	-has there been a reduction in number of speed-related crashes?		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
		Pram 1&2 implemented	reporting that they speed. Increase in the percentage of people reporting that they find speeding socially unacceptable.	before and after campaign and enforcement changes			
40. Increase awareness and understanding in the community about the philosophy of the speed camera program, the manner in which the program operates and, as evaluations occur, the outcomes the program achieves. (**QT* / QPS**)	**2**	Y- ongoing	-community attitude surveys -increased positive attitudes towards the speed camera program -enforcement measures (speed camera hours, infringements,	- conduct speed surveys	- has there been a reduction in road toll/serious injuries due to speed camera program?		Qld Speed camera program has resulted in ave. annual savings of 110 fatal crashes; 1100 hospitalisation crashes; 2200 medically treated crashes; 500 other injury crashes; 1600 non-injury crashes - $2.8 billion savings in social costs to community during implementation period
42. Ensure network wide adoption of the guidelines to ensure speed limits	2	Y- ongoing –	-increased	Before/after			

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
are consistent and credible across Queensland. (**QT* / MR / LG**)		adoption of guidelines expected by mid 2005 -speed limit reviews are completed on an ongoing basis	compliance with speed limits - community attitude surveys (increased level of agreement that speed limits in Qld are applied consistently)	study on speed limit compliance			
Develop innovative ways to ensure that motorists travel at safe and legal speeds							
47. Implement a multi-deployment schedule for speed camera vehicles to enable deployment of speed camera vehicles to a number of sites within one shift*.* ***(*QPS**)	**2**	Y- Training has taken place across the state to assist with implementation of this action (expected to be finished July 2005)	-specific & general deterrence effects -greater coverage of road network for speed camera enforcement	- undertake speed surveys	- Evaluation results of the speed camera program indicate that crash reductions are strongly related to the size of the overall program as well as to the density of enforcement. By increasing the number of sites visited, the area targeted will		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
					also increase, thereby increasing crash reductions		
Unrestrained occupants							
Encourage better use of restraints							
51. Introduce new penalties and sanctions for non-use of restraints. (**QT**)	**1**	Y- 100% complete	-increased deterrent effect of new penalties/sanctions -community awareness surveys of restraint use since new penalties introduced	- measure seatbelt wearing rates -has there been an increase in compliance since new penalties introduced - (Has there been a rreduction in seatbelt offence data)	- has there been a reduction in no. of fatal crashes due to reduction in non-restraint use - seat belt info. in crash data		
52. Ensure that enforcement and education campaigns are targeted and coordinated. (**QT* / QPS**)	**2**	Y- 100% complete The new seat belts public education campaign commenced in November and includes a new television commercial and publicity	Surveys to measure changes in drivers attitudes to restraint use;speeding	-intended behaviour change through observational surveys	Has there been a reduction in (fatal) crashes where traveling unrestrained was a contributing factor		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	Action Type	Action taken (Y=Yes/N= No)?	Know-ledge &/or attitude change	Behav-iour change (specific)	Crash rate reduction	Injury outcome per crash	Long-term conse-quences per injury
		statewide.62 % of motorists are aware that tougher seatbelt penalties have been introduced in Queensland. To maintain the perception among 8 in 10 motorists that they will be caught by police if they don't wear a seat belt (80% Dec '04 vs 80% Oct '04). 84% of motorists are awareness of the campaign (Dec '04). Self-reported wearing rates have reduced from 7% (Feb '05) to 4% (July '05). 87% agree that "if I don't wear a seatbelt, I increase my					

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
		risk of harming myself in a crash."					
53. Develop culturally targeted campaigns to improve the installation and use of suitable child restraints. (**DES**)	**1**	Y- completed July 2005	-Promoted the QAS Baby Capsule Hire Scheme amongst Indigenous communities within SE Queensland. Established a community based hire scheme in collaboration with the Inala Justice Group. Expected to result in an increase in Indigenous children correctly restrained while travelling in a motor vehicle and thus a decline in injury sustained by Indigeous	-observational survey of child restraint use in carparks (specifically targeting areas with high indigenous populations)	- has there been a reduction in crashes where children were not restrained		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
			children who have been involved in a car crash				
54. Promote the inclusion of seat belt information in Occupational Health and Safety and fleet management guidelines. (**QT**)	**1**	Y- expected completion August 2005	To include restraint fact sheets in QT websites and in "Safe Driving" policy in fleet safety kits for the driving industry.		- has there been a reduction in fatal crashes		
55. Promote the fitting of integrated seat belts in truck seats. (**QT**)	**1**	Y- expected to be completed end of 2005	-survey of truck drivers re restraint use	-conduct a wearing rate survey of truck drivers	- has there been a reduction in road fatalities/serious injuries in unrestrained truck drivers		
56. Implement the recommendations approved by government from the School Transport Safety Interdepartmental Working Group on seat belts in school buses. (**QT**)	**1**	Y ? – 20% completed - funding secured for fiitment of seatbelts in school buses in steep and hazardous areas. Rollout commenced July 2005.	Community awareness survey of child restraint use in school buses	- survey of children on buses in conjunction with schools?	- has there been a reduction in fatalities/serious injuries involving children on school buses. (before/after introduction of seatbelts)		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
Rural and remote road users							
Reduce the involvement of rural and remote road users in road trauma							
58. Increase licensing rates among rural and remote communities by improving access to licensing. (**QT**)	**2**	Y- 50% complete – development of interventions to identify problems by Dec2005	-no. of licence holders in rural communities		-crashes by licence type (rural areas, indigenous community) – reduction in fatalities involving indigenous people in rural areas		
59. Make licensing processes and road safety materials more accessible and culturally appropriate to Indigenous peoples. (**QT* / QPS***)	**2**	Y- 25% complete? Development of sustainable Interventions to identify problems by June 2007	-no.of licence holders in rural indigenouscommunity by licence type; - education campaigns targeted towards indigenous people in rural areas focusing on road safety and licensing processes		-analysis of rural crashes by licence type		
60. Develop an integrated state-wide emergency services communication network across the emergency service agencies. (**DES* / QPS***)	**1**	N- scheduled to start 2007					

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
Develop intelligence on crashes in rural and remote areas							
62. In conjunction with other jurisdictions, develop a website and CD ROM resource for Indigenous road safety resources. (**QT**)	**2**	Y- 80% complete- scheduled completion date Jan2008	-has website/ CDRom been accessed by targeted audience -before/ after website development survey on road safety issues by indigenous persons				
Young adults							
Reduce young adults' involvement in serious crashes							
64. Trial the re-introduction of L-plates on all vehicles driven or ridden by holders of learner licenses. (**QT**)	**1**	Y – 100% implementatio n –ongoing pilot project	Surveyof learner and other drivers re:attitudes/w areness of towards learner drivers after L-plate re-intro	See Mike K?	-reduction in crashes involving learner drivers before/after trial		
65. Promote resources to assist schools and community groups in delivering mentor driving programs. (**QT**)	**1**	Y- expected completion Dec2005 -aim is to reach 100% of young drivers issued		Survey of young learner & p-plate drivers re safe driving behaviours	Reduction in crashes involving young drivers (p-plate drivers) before & after		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
		with provisional licence over a two year period and influence the adoption and			mentor driving programs		
66. Update the Road Accident Awareness Program to align with contemporary road safety issues and approaches. (**DES**)	**2**	Y? – proposed to start late 2004		-use of new methods to track behaviour change?			
67. Review school-based education materials to include a greater focus on risk taking and harm minimisation. (**QT* / EQ**)	**2**	Y- 100% complete					
Unlicensed drivers and riders							
Reduce the involvement of unlicensed drivers and riders in crashes							
72. Trial the use of Automatic Number Plate Recognition technology. (**QPS* / QT**)	**1**	Y- expected to start Dec2004 – aim is to identify unregistered/stolen vehicles			-has there been a reduction in crashes involving unregistered vehicles		
73. Use crash data and information on unregistered vehicles and unlicensed drivers and riders from the mobile data initiative to focus enforcement. (**QPS* / QT**)	**1**	Y- ongoing	-no. of unregistered vehicles -no. of unlicensed drivers -reduction in these numbers	--	-reduction in crashes involving unregistered vehicles; unlicensed drivers etc..		
74. Implement a public education campaign highlighting the increased	**1**	Y- ongoing	-survey public	-change the	-has there		

TABLE A: PROGRAMS, ACTIONS AND EVALUATION CRITERIA BASED ON THE 2004-2005 QLD ROAD SAFETY ACTION PLAN							
ACTION TYPES: 1.New key initiatives 2.Modifications to existing key initiatives							
		Feasible Evaluation Criteria					
Evaluation framework based on the Qld Road Safety ACTION PLAN for 2004-2005	**Action Type**	**Action taken (Y=Yes/N= No)?**	**Know-ledge &/or attitude change**	**Behav-iour change (specific)**	**Crash rate reduction**	**Injury outcome per crash**	**Long-term conse-quences per injury**
capability to detect unlicensed drivers and riders using the Mobile Integrated Network Data Access System (MINDA) and Automatic Number Plate Recognition technology. (**QT* / QPS**)			on awareness of improved technolog to detect unlicensed drivers	behaviour of drivers to become more law-abiding?	been a reduction in crashes involving unregistered vehicles and unlicensed drivers		
Seniors							
Improve the safety of seniors in crashes							
80. Promote and expand initiatives, such as RACQ's 'Years Ahead' program, which educate older drivers how to drive safely for longer. (**RACQ**)	**2**	N?	-survey of older drivers re: driving safely/health/		- has there been a reduction in number of crashes involving older drivers		
Fleet safety							
Reduce fleet related crashes							
83. Use available information from WorkCover and police to guide high-risk industries and companies to implement Occupational Health and Safety policies for workrelated driving. (**QT**)	1	Y- scheduled completion Aug2005	-survey of fleet drivers re OHS policies		-has there been a reduction in no. of crashes involving fleet vehicles		
Develop intelligence about fleet related crashes							
84. Trial a comprehensive range of tailored fleet safety initiatives in varying conditions throughout Australia, aimed at improving road safety specific circumstances for: - volunteer and charity service providers; and - large commercial fleets. (**CARRS-Q**)	**1**	N?					