

THE UNIVERSITY OF SYDNEY  
FACULTY OF LAW

**CURRENT ISSUES IN**  

---

**CRIMINAL JUSTICE**

**Journal of the Institute of Criminology**

Issue Editor: Professor Duncan Chappell  
University of Sydney Law School

---

Volume 20 Number 2

November 2008

---

## **CURRENT ISSUES IN CRIMINAL JUSTICE**

### **Volume 20 Number 2 November 2008**

*Issue Editor:* Professor Duncan Chappell

*Executive Editor:* Professor Mark Findlay

*Production Editor:* Dawn Koester

*Book Review Editor:* Dr Arlie Loughnan

*Current Issues in Criminal Justice* is published three times per year, in July, November and March. This issue of *Current Issues in Criminal Justice* forms the second of three issues published in the twentieth volume. Volume 20 Number 3 will be published in March 2009. *Current Issues in Criminal Justice* Number 1 (October 1989) and Number 2 (March 1990) should be regarded as forming Volume 1, Numbers 1 and 2, respectively.

The views expressed in this issue are those of the individual authors and do not necessarily reflect those of the Institute of Criminology or the institution with which they are affiliated. Whilst every reasonable care has been taken in the preparation of this publication, no liability is assumed for any errors or omissions. *Current Issues in Criminal Justice* is a peer reviewed journal. The Editorial Board would like to thank the reviewers of submissions to this issue of the Journal.

ISSN 1034-5329

© 2008 Institute of Criminology University of Sydney. This work is copyright. Apart from any fair dealing for the purposes of private study, research, criticism or review as permitted under *the Copyright Act*, no part may be reproduced by any process without written permission. Inquiries should be addressed to the Director, Institute of Criminology, University of Sydney, Faculty of Law, 173–175 Phillip Street, Sydney NSW 2000, Australia.

Typeset by the Institute of Criminology.

Printed and bound by Ligare Printery, Sydney, Australia.

### **Subscriptions**

Subscription rates within Australia (gst incl.): \$110 (1 year), \$209 (2 years), \$308 (3 years). Overseas subscriptions: \$115 (1 year), \$220 (2 years), \$325 (3 years). Single copies of the current issue are \$44. Back issues are available for \$27.50. For subscriptions in North America, contact Gaunt, Inc, who are sole distributors of the journal in that country (Gaunt Building, 3011 Gulf Drive, Holmes Beach, Florida 34217-2199 USA).

### **Membership**

Membership to the Institute of Criminology includes a Journal subscription, entry to seminars and 10% discount on Institute publications. Contact the Institute for details.

### **Contact details**

Institute of Criminology, Sydney University Law School. 173–175 Phillip Street, Sydney, NSW 2000, Australia Ph: 61 2 9351 0239 Fax: 61 2 9351 0200  
e-mail: [d.koester@usyd.edu.au](mailto:d.koester@usyd.edu.au) <[www.criminology.law.usyd.edu.au](http://www.criminology.law.usyd.edu.au)>

*Current Issues in Criminal Justice* is indexed in APAIS: Australian Public Affairs Information Service, produced by the National Library of Australia, and in CINCH, the Australian Criminology Database, produced by the Australian Institute of Criminology. Both databases are available on-line or in CD-ROM format. Contact: INFORMIT, phone: (03) 9341 3265; email: <[info@rmitpublishing.com.au](mailto:info@rmitpublishing.com.au)>.

*Current Issues in Criminal Justice* is published online at Informit: <<http://search.informit.com.au>>, and at HeinOnline: <<http://heinonline.org>>.



# *Enhancing Evidence-Based Policy: Principles and Practice from a Case Study of Australian Firearms Legislation*

Samara McPhedran\* and Jeanine Baker\*\*

---

## *Abstract*

As the value of evidence-based policy is increasingly recognised in Australia and internationally, it is incumbent upon researchers to ensure that their work is robust, appropriate in its selection of statistical methods, and based upon testable hypotheses. This caution is particularly apt for contentious areas within justice policy. This review examined strengths and weaknesses in the application of research methodology, using the often controversial example of Australian firearms legislation. Particular emphasis was given to comparing and contrasting different analysis methods. The key conclusions of each paper are examined, and contextualised against the statistical methods used. The application of basic principles such as assessing congruence between results and conclusions, as well as treating evidence as a cumulative rather than definitive process, has the ability to enhance the quality of research and policy.

## **Introduction**

As the role of evidence-based policy is increasingly recognised in Australia and internationally, it is incumbent upon social scientists to ensure that the research they produce aims to be robust, appropriate in its selection of statistical methods, and based upon testable hypotheses. This caution, recognised elsewhere through initiatives such as the Campbell Collaboration,<sup>1</sup> is particularly apt for contentious issues found within broader areas such as justice policy. It is therefore beneficial to seek ways to enhance the quality of evidence-based policy. Within the scope of the current study, basic principles and key concepts were identified by examining strengths and weaknesses in the application and interpretation of methodology using the often controversial example of Australian firearms legislation.

---

\* PhD, Position at the time of writing: Postdoctoral Research Fellow, School of Psychology, University of Sydney, Camperdown, NSW, 2006. Email: [chair@ic-wish.org](mailto:chair@ic-wish.org). Affiliations: School of Psychology, University of Sydney; International Coalition for Women in Shooting and Hunting (WiSH).

\*\* PhD. Affiliations: International Coalition for Women in Shooting and Hunting (WiSH).

<sup>1</sup> The Campbell Collaboration promotes the view that interventions should be systematically reviewed, in order to inform and improve policy and services.

In 1996, Australia enacted sweeping reforms to its gun laws after the Port Arthur mass shooting. The National Firearms Agreement (NFA) was ratified by Federal Parliament in 1996 and implemented across all states and territories by the end of 1997. A key component of the NFA was the prohibition of certain types of firearms; primarily self-loading rifles and self-loading and pump-action shotguns. A taxpayer funded 'buy-back' scheme was introduced to compensate owners for the surrender of their firearms. Over 640 000 firearms were handed in to Police for destruction, at an estimated cost of AU\$500 million.

Additionally, the NFA introduced stringent requirements governing the possession of firearms. These included the need to have a 'genuine reason' for firearm ownership (self defence was explicitly excluded), compulsory written safety tests, and the stipulation that all privately owned firearms must be registered through a state-controlled authority. Elements such as safe storage of firearms when not in use, and 28-day waiting periods for acquisitions of firearms were also included in the legislative changes.

The 1996 gun laws represent an 'intervention point' in a time series, creating a natural experimental design. Therefore, the issues raised herein are applicable to other policy contexts where 'before' and 'after' data are available. Since 1996, a number of studies have set out to evaluate the impacts of the NFA, either directly (using 1996 as an intervention point and undertaking 'before' and 'after' comparisons) or indirectly (by examining general trends over time, for example). The literature uses a wide variety of statistical methods, from simple percentages through to modelling techniques, and the conclusions that have been drawn frequently appear to contradict one another.

To date there has not been an attempt to bring together the available research and critique the variety of analyses used. Similarly, identification of analytical shortcomings has not been undertaken, and the potential for methodological differences to influence research outcomes has not been widely discussed. The current review does not seek to establish whether the NFA did or did not have an impact on firearm related deaths. Rather, it utilises the case study of post-NFA research as a focus point for broader questions of methodological rigour and statistical analysis in social science – particularly in areas with significant policy implications. Particular emphasis is given to comparing and contrasting the analysis methods used. The key conclusions of each paper are examined, and their robustness is contextualised against the statistical methods used.

The criterion for inclusion of a study for evaluation was the attempt of some degree of quantitative analysis (whereas general review or discussion papers were excluded). Using this criterion, the following papers were considered:

- Firearm-related violence: The impact of the Nationwide Agreement on Firearms (Mouzos 1999).
- Gun control in Tasmania (Auditor-General Special Report No. 55 2005).
- Firearm related deaths: the impact of regulatory reform (Ozanne-Smith et al 2004).
- Gun laws and sudden death: Did the Australian firearms legislation of 1996 make a difference? (Baker & McPhedran 2007).
- Australia's 1996 gun law reforms: faster fall in firearm deaths, firearm suicides and a decade without mass shootings (Chapman et al 2006).

## **Study 1: Firearm-related violence: The impact of the Nationwide Agreement on Firearms (Mouzos 1999)**

Mouzos (1999) provides a preliminary analysis of post-1996 data. The intent of the study was simply to look at firearm-related deaths in 1997, and establish whether they were less frequent than in previous years. Despite the title, the study itself does not set out to directly evaluate whether the NFA had an effect. Mouzos (1999) concluded that, nationally, firearm-related deaths declined in 1997 relative to previous years, mostly due to a decline in the suicides and accidents. It was suggested that in the case of suicides, method substitution may have occurred.

The study highlighted that ‘as a result of the many issues associated with evaluation research, it is still too soon to determine definitively whether Australia’s uniform firearms laws have achieved their aim in reducing firearm-related violence and misuse’ (Mouzos 1999:1). However, it was also noted that 1997 saw a decrease in firearm-related violence and misuse, especially firearm suicides.

Basic rates and raw numbers of firearm-related incidents are presented, particularly for 1996 and 1997. No formal statistical analysis was undertaken, nor was it claimed that the lower incidence of firearm-related deaths in 1997 was evidence of an ‘impact’ of the NFA. Therefore, this study’s primary contribution is in establishing background information about what was occurring in the years immediately prior to 1996. Specifically, it shows that firearm-related deaths in Australia were already trending downwards before the introduction of the legislative reforms in 1996, with a steady and ongoing decrease in the incidence of firearm-related deaths over time.

The report enables studies seeking to directly investigate whether the NFA had an impact to make an informed choice of statistical method, by demonstrating the trended nature of the data. It also underscores the need to avoid drawing conclusions based on limited post-intervention data, reinforcing the need for longitudinal study. However, this paper is best viewed – as the author acknowledges – as a preliminary investigation, rather than an effort to quantify the impacts of the NFA. Importantly, the paper illustrates the conceptual difference between looking at trends over time versus scientifically examining the effect of a specific intervention upon those trends.

## **Study 2: Gun control in Tasmania (Auditor-General Special Report No. 55 2005)**

The objective of the Special Report was to determine whether the ‘Tasmania Together’ goal of ‘Safer communities’ was being furthered by the 1996 firearms legislation. The report concluded that the 1996 reforms have had an impact on reducing the incidence of firearms use in violent crime and suicide in one Australian state. This Report purports to offer statistical evidence for an impact of the NFA, in contrast to the work of Mouzos (1999) which did not seek to statistically analyse impacts.

The statistical methods used in this report represent one of the most basic levels of analysis in the available literature. For firearm suicides, three different legislative epochs were examined: 1985-1992 (‘no legislation’), 1993-1996 (*Guns Act* 1991), and 1997-2004 (*Firearms Act* 1996). The percentage of firearm suicides, as a proportion of all suicides, was

calculated within each epoch. This generated a simple average (i.e., the sum of the percentage of firearm suicides in each year, divided by the number of years in the epoch).

A simple averaging method was also used when examining the proportion of 'crimes against the person' in which a firearm was used. Regarding crimes against the person, only two epochs were considered: 1994-1996, and 1997-2004. In both instances, the average incidence of firearms use was lower in the 1997-2004 period than in the earlier period/s. On the basis of this observation, it was concluded that the legislative reform reduced the incidence of firearm misuse. It is unclear what the precise statistical methods used to test this assumption were (if any were used), given the scant detail provided in relation to analysis.

A major shortcoming in the report is that while averages can provide a useful tool for quantifying changes, they are generally unsuitable for assessing examples where long term, ongoing trends are apparent. If a consistent trend is present in the data, then 'grouping' data and comparing averages of those groups will produce predictable but severely constrained results.

The apparently dramatic decreases 'caused' by each successive legislative period are simply a byproduct of ongoing declines. Whenever the occurrence of an event shows a decline over time (for whatever reason/s), it follows that the 'later' end of the trendline will be lower than the 'earlier' end of the trendline. Groups combining 'later' years (1997-2003) will always produce lower averages than groups formed by combining 'middle' years (1993-1996), which in turn will always produce lower averages than groups composed of 'earlier' years (1985-1992).

Therefore, the method used in the Special Report would be more suitable for application to data without any clear trend, and would require additional analysis to test for the statistical significance of the observed differences between groups. The methods used in the Special Report cannot provide the level of analysis necessary to ascertain whether the ongoing downward trends changed as a result of the legislative intervention. The conclusion that the NFA has an impact in Tasmania is, therefore, not robust in the case of the Special Report due to methodological misspecification.

### **Study 3: Firearm related deaths: the impact of regulatory reform (Ozanne-Smith et al 2004)**

The objective of Ozanne-Smith and colleagues' (2004) study was to examine trends in rates of firearm-related deaths in Victoria, Australia, over 22 years in the context of legislative reform. It was concluded that there were 'dramatic declines in rates of firearm-related deaths in the state of Victoria, and for the rest of Australia ... in the context of strong legislative reform' (Ozanne-Smith et al 2004:285).

Ozanne-Smith et al (2004) examined data from 1979-2000, and divided that series into three epochs: 1979-1988 (period 1, no legislation), 1989-1996 (period 2, legislative change in one state, Victoria), and 1997-2000 (period 3, legislative change across Australia). Analysis was undertaken using a Poisson regression model that incorporated a fixed offset. Death counts were treated as the dependent variable in the model, while the offset (population counts by region and year) enabled the analysis to focus on rates rather than absolute death counts.

The Poisson model allowed comparison of relative rates of firearm-related deaths between Victoria and the rest of Australia over the three time periods. Data from the rest of Australia acted as the 'control' for periods 1 and 2, with Victorian legislative reform introduced in period 2. In period 3, the roles of the 'treatment' and 'control' in the analysis were reversed. The decrease in the rate of firearm-related deaths in Victoria in period 2 relative to Australian trends, as well as the decrease in rates of firearm-related deaths in Australia in period 3 relative to Victorian trends, was assessed. The study found significant differences in trends for firearm suicide, but not for firearms homicides, across different periods of firearm legislative reform.

A difficulty with the application of Poisson modelling is that although it is widely used to model count data, it assumes equidispersion (i.e., the variance of the dependent variable equals its mean). In many situations, the variance of the dependent variable is greater than its mean – referred to as 'overdispersion'. Consequently, alternative versions of the Poisson model, such as negative binomial regression, have been used to take into account overdispersion. Although Ozanne-Smith et al (2004) provide goodness of fit data for their model, there is no discussion of the suitability of the Poisson model for the data. The treatment of the data as equidispersed conflicts with Chapman et al's (2006) statement that the data were viewed as overdispersed (see Study 5).

A further methodological consideration is the comparison of one state with the rest of Australia, without any apparent 'weighting' process applied. The other Australian states display substantial variation in trends for firearm-related deaths, and it would be desirable to adjust the analyses to take this variation into account. Although the use of combined other states as a control group for Victoria enabled different periods of reform to be examined, it is prudent to consider whether grouping data from other states forms a reliable control, given that the resulting variability of the control group exceeds the variability of the Victorian data.

A drawback acknowledged by Ozanne-Smith et al (2004) was that non-firearm deaths were not considered, and that the possibility of method substitution was not taken into account. As such, the study provides a limited tool in terms of understanding overall impacts of legislative change, because it does not discuss the connotations of declines in non-firearm suicides (noted by both Baker & McPhedran 2007 and Chapman et al 2006). These declines raise the possibility that changes above and beyond firearm legislative reform (e.g., improved suicide prevention programs introduced in Victoria in the late 1980s and nationally in the late 1990s) may have impacted on suicides across the board, irrespective of method. This casts doubt on the robustness of Ozanne-Smith and colleagues' (2004) conclusion that firearm legislation was responsible for the declines in firearm suicide.

Additionally, the study failed to realise the implications of the substantial over-representation of firearm suicides among all firearm-related deaths (approximately 80%). Basing speculation about the efficacy of the reforms on observations of overall firearm-related deaths, without acknowledging the fundamental contribution of suicide trends in dictating those outcomes, generates conclusions that are inconsistent with the results for firearm homicides or accidental deaths. To conclude that legislative reform was associated with significant declines in firearm-related deaths is therefore relevant only to firearm suicides.

#### **Study 4: Gun laws and sudden death: Did the Australian firearms legislation of 1996 make a difference? (Baker & McPhedran 2007)**

The objective of Baker and McPhedran's (2007) research was to evaluate the benefits of buying back legally held firearms and increasing restrictions on firearm owners. This was operationalised by testing whether post-1996 trends in firearm homicide, suicide, and accidental death differed from forecasts based on pre-1996 trends. Baker and McPhedran (2007) also looked at trends in non-firearm homicide and suicide.

Baker and McPhedran (2007) standardised the number of deaths against population, generating a rate expressed per 100,000 population. They applied ARIMA (AutoRegressive Integrated Moving Average) modelling to the time series. ARIMA belongs to a flexible family of time series models that allow for the inclusion of an intervention (e.g., legislative change). Using the period 1979-1996, Baker and McPhedran (2007) forecast death rates out to 2004. The forecast and observed rates were then statistically compared using t-tests.

The study found a significant difference between observed and forecast rates for firearm suicide, but no difference for firearm homicide. The authors noted that non-firearm suicides began declining in the late 1990s, and concluded that 'the only category of sudden death that may have been influenced by the introduction of the NFA was firearm suicide. However, this effect must be considered in light of the findings for suicide (non-firearm)' (Baker & McPhedran 2007:461-463).

Although widely used, a caveat upon ARIMA is that the longer the forecast, the less reliable that forecast becomes (particularly for variable data). Hence, forecasting over periods of, for example, many decades is not likely to be a sound application of ARIMA. Another limitation of the ARIMA model is that it is stationary, therefore requiring differencing passes to transform the data to meet this assumption. Also, the parameters selected for the different components of the model may not produce a model of best fit for all the data to which that model is applied. A shortcoming acknowledged by Baker and McPhedran (2007) is that their model was not a good fit for non-firearm homicide. This may explain why Baker and McPhedran (2007) did not find a significant difference between observed and predicted rates for non-firearm homicide, in contrast with Chapman et al (2006) who found changes in non-firearm homicide pre- and post-1996.

As a result of the variability in the data and relatively low sample sizes for accidental deaths and homicides, Baker and McPhedran's (2007) lower confidence limit for accidental deaths falls below zero. It would therefore not be possible for the observed rate to fall below the lower limit, given that this would represent a negative death value. While important in its emphasis on allowing rejection of hypotheses, the negative value observation does not apply in practical terms to Baker and McPhedran's (2007) univariate time series.

With regard to accidental firearm-related deaths, the incidence increased post-1996. The comparison was not between an observed and predicted decline, but between a predicted decline and an observed increase. This allows for the possibility of observed values to fall outside the upper limit of the confidence interval, enabling rejection of the null hypothesis. For firearm homicides, the lower interval remains positive throughout the 1997-2004 period. The observed rate could, therefore, have fallen below this level at any time; the fact that it *did* not fall below that limit does not entail that it *could* not have fallen below the limit. Nonetheless, the importance of maintaining the possibility of rejection of the null hypothesis should be emphasised. Generating testable hypotheses that are open to falsification is a central premise for ensuring scientific rigour.

Concerning method substitution, Baker and McPhedran (2007) suggest that there is little evidence of displacement from firearms to other methods. However, although they used non-firearm deaths as a proxy indicator of method substitution, this commonly used technique does not equate to a direct test of substitution. Indeed, direct tests for substitution require a level of experimental control often unattainable in social research. An experimental difficulty is that directly testing for the presence of method substitution is problematic in the context of firearms legislation, therefore statements about the occurrence or otherwise of displacement should be regarded with caution.

### **Study 5: Australia's 1996 gun law reforms: faster fall in firearm deaths, firearm suicides and a decade without mass shootings (Chapman et al 2006)**

The objective of Chapman and colleagues' (2006) research was to determine whether Australia's 1996 major gun law reforms were associated with changes in rates of mass firearm homicides, total firearm deaths, firearm homicides and firearm suicides. They also examined whether there were any apparent method substitution effects for homicides and suicides. It was concluded that the 'Australian example provides evidence that removing large numbers of firearms from a community can be associated with a sudden and ongoing decline in mass shootings and accelerating declines in total firearm-related deaths, firearm homicides and firearm suicides' (Chapman et al 2006:371).

Chapman et al (2006) applied a negative binomial model to national data, to compare rate ratios for 1979-1996 with 1997-2003. Although the model used by Chapman and colleagues (2006) treats deaths as count data, in practical terms it expresses deaths as rates, by using population as an offset. Like Ozanne-Smith et al (2004) for Victorian trends, and Baker and McPhedran (2007) for national trends, Chapman et al (2006) found a significant acceleration in the rate of decline for firearm suicides post-1996, but no significant difference in the rate ratios for firearm homicide pre- and post-1996. There was a significant difference pre- and post-1996 for non-firearm suicides, with faster declines post-1996. The authors also found a faster decline in non-firearm homicides post-1996.

Chapman et al (2006) used non-firearm deaths as a proxy assessment of method substitution, expecting that substitution would manifest as an 'increasing downward trend in firearm deaths after the introduction of gun control laws but a compensatory lesser downward or even upward trend in non-firearm-related deaths over the same period' (Chapman et al 2006:366). No signs of this were observed, with declines in non-firearm homicides and suicides apparent. Chapman et al (2006) suggest there is no evidence for substitution from one method to another. The decline in non-firearm suicides was also noted by Baker and McPhedran (2007), who viewed it as a confound when attempting to draw conclusions about the accelerated decline in firearm suicides post-1996. Those authors suggest a need for differentiation between any effects of legislation, versus method-independent impacts of wider social changes (for example, improved funding for mental healthcare).

The commonly used negative binomial method can overcome problems associated with overdispersed data. Poisson estimates are inefficient in such instances, due to standard errors being biased downwards and yielding spuriously high z-values (thereby increasing the likelihood of finding a significant result). In contrast, negative binomial models allow relaxation of the assumptions about the mean-variance relationship and introduce an

additional parameter ( $\gamma$ ) into their structure. However, by introducing an extra parameter into the model, interpretability of the results becomes more problematic, and it is difficult to assign clear meaning to the variance. Nakashima (1997) highlighted that negative binomial models can be ineffective if the underlying distribution is not negative binomial, even if the variance satisfies the negative binomial form. Also, negative binomial models are generally not appropriate in situations where the data are autocorrelated,<sup>2</sup> which often occurs within time series. In such instances, methods such as ARIMA offer a preferable alternative, because they can accommodate autocorrelated data (see Barron 1992 for a helpful discussion of overdispersion and autocorrelation in count data).

Similar to Ozanne-Smith et al (2004), Chapman and associates' (2006) paper did not recognise that a significant acceleration in the downwards trend in total firearm deaths reflects a statistical byproduct of the large number of suicides and the acceleration in the decline in firearm suicides post-1996. Also, Chapman et al's (2006) study contains incongruence between the results and conclusions. For example, the assertion that there was an accelerated decline in firearm homicides conflicts with the statistical findings, which showed that there was no significant acceleration in the decline in firearm homicides post-1996. Therefore, the evidence in Chapman and others' (2006) does not support their conclusions about firearm homicides.

Despite its stated objectives, the study did not test for whether the incidence of mass shootings differed significantly pre- and post-1996. The conclusion that there was a sudden and ongoing decline in mass shootings does not follow from the analyses presented. Indeed, the analyses undertaken would not be appropriate for analysing extremely rare events such as mass shootings (for further discussion of methodological issues associated with rare events, see e.g. Bhati 2004; Frei & Schar 2001; Luft & Brown 1993).

## Conclusions

The case study of research into firearm-related deaths in Australia leads to the question of which method is most suitable for the study of a legislative intervention such as the 1996 NFA. However, it would be mistaken to assert that any given method is invariably superior to another, or should be used across the board. Rather, the above critique demonstrates that the selection of a particular analysis method should be driven by an integrated quality framework that gives careful consideration to the nature and shape of the data, the hypotheses being tested and potential confounds associated with those hypotheses, and an awareness of the relative strengths and limitations of each analysis method that could be applied to a specific dataset. Consideration should also be given to whether a certain model or statistical test can be validly applied to a particular dataset, or whether key assumptions of that model or test are violated.

Given the downward trends in firearm related homicides and suicides in Australia, the analysis methods chosen must, if they are to produce meaningful outcomes, be able to be applied to time series data that contains a trend. The three studies that have attempted this – using time series modelling or variants on the Poisson method – all contain shortcomings of some form. However, this observation highlights that evidence-based policy should incorporate the principle that knowledge is a cumulative process. Given that no single study should wisely be treated as definitive, it is recommended that the way forward for

---

<sup>2</sup> Briefly, autocorrelation refers to the correlation of a variable (such as observed deaths) with itself over successive time periods. Autocorrelation invalidates assumptions about the independence of observations.

policymaking overall is to examine (where possible) the accumulated body of knowledge for consistency.

Regarding the study of firearms legislation and its impact in Australia, there is a strong level of consistency in the results of the three studies that examine trends (as opposed to average rates) before and after periods of legislative reform. Although the methods used in each study can be critiqued at the individual level, the consistent statistical outcomes suggest that collectively the research points in the same direction; no impact of the reforms on firearm homicide, with any conclusions about possible impacts on firearm suicides confounded by the accompanying decline in non-firearm suicides. However, in cases where this level of consistency between studies is not observed, it may be helpful to refer to other indices such as the number of years of data included in the study, whether the chosen analysis method is suitable for the data at hand, and whether the statistical models used are a good fit for the data.

Another consideration is statistical power. In the case of firearm homicides, the low number of incidents may not be sufficient to enable detection of a statistically significant effect. However, it is also possible that there is not an effect to detect. Interestingly, the number of accidental firearm deaths is lower than firearm homicides, yet both Baker and McPhedran (2007) and Chapman et al (2006) found statistically significant differences pre- and post-1996 for that type of death. Nonetheless, common scientific practice entails the assumption that there is no effect, unless persuasive evidence to the contrary is found.

This in turn relates to Type I versus Type II error. It is possible that a researcher may conclude that there is no effect, when an effect does exist (Type II error). More seriously, a researcher may reject the null hypothesis (the hypothesis of 'no effect') and conclude there is an effect, when there is not an effect (Type I error). Given that Type I error is widely considered the more serious error to make,<sup>3</sup> various methods are commonly employed to reduce the likelihood of this occurring – for example, corrections can be made to the level of statistical significance required to achieve rejection of the null hypothesis.

There is value in adopting a cautious approach to research which recognises the need to minimise the likelihood of Type I error. It is beneficial to take a conservative perspective that considers the characteristics of the data and the potential for committing a Type I error. In the context of firearms research, this translates to a caution against statements that the laws had an impact, given the lack of evidence for this proposition. Applying this principle in the broader context necessitates that in the absence of evidence that an effect has occurred, it is prudent to continue to work on the assumption that there is no effect, until and unless a body of convincing evidence to the contrary emerges.

Additionally, it is important to recognise that the conclusions of the latter three studies presented in this critique differ not as a result of different statistical outcomes, but as a consequence of varied interpretations of those results. Therefore, evidence-based policy would be enhanced through recognition that the conclusions presented do not necessarily reflect or relate logically to the outcomes of the actual analyses undertaken. It is suggested that this possibility be taken into account during any evaluation processes.

---

<sup>3</sup> An example is useful to illustrate this principle. If a researcher conducted a clinical trial of a new cancer drug and wrongly concluded that drug was more effective than other treatments, the practical implications of that error are more severe than if the researcher wrongly concluded that the drug was not more effective than other treatments.

In summary, the case study of firearms legislation provides a number of cautions regarding method choice and rigour in policy relevant social science. It demonstrates the need for appropriate method selection that takes into account the type of data being used and its specific characteristics (such as the presence of a trend over time), and the need to recognise how various different analysis methods may affect the outcomes obtained. The importance of evaluating not just the conclusions of research, but the results themselves, is highlighted, as is the potential for incongruence between results and conclusions. Similarly, if multiple studies exist then there is a need to evaluate their consistency, and the value of conservatism as a counter for error should be considered during any assessment of policy relevant research. The application of these basic principles has the ability to enhance the quality of both research and evidence-based policy.

## References

- Auditor-General of Tasmania 2005 Gun control in Tasmania, Auditor-General Special Report No. 55 Tasmanian Audit Office
- Baker J & McPhedran S 2007 'Gun laws and sudden death: Did the Australian firearms legislation of 1996 make a difference?' *British Journal of Criminology* vol 47 pp 455-469
- Barron D 1992 'The analysis of count data: Overdispersion and autocorrelation' *Sociological Methodology* vol 22 pp 179-220
- Bhati AS 2004 Robust spatial analysis of rare crimes, Technical Report, Mapping and Analysis for Public Safety (MAPS) Program, National Institute of Justice (NIJ), United States
- Chapman S, Alpers P, Agho K & Jones M 2006 'Australia's 1996 gun law reforms: faster fall in firearm deaths, firearm suicides and a decade without mass shootings' *Injury Prevention* vol 12 pp 365-372
- Frei C & Schar C 2001 'Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region' *Journal of Climate* vol 14 pp 1568-1584
- Luft HS & Brown B 1993 'Calculating rare events: Why settle for an approximation?' *Health Services Research* vol 28 no 4 pp 419-439
- Mouzos J 1999 'Firearm-related violence: The impact of the Nationwide Agreement on Firearms' *Trends and Issues in Crime and Criminal Justice No. 116* Australian Institute of Criminology
- Nakashima E 1997 'Some methods for estimation in a negative-binomial model' *Annals of the Institute of Statistical Mathematics* vol 49 no 1 pp 101-115
- Ozanne-Smith J, Ashby K, Newstead S, Stathakis VZ & Clapperton A 2004 'Firearm related deaths: the impact of regulatory reform' *Injury Prevention* vol 10 pp 280-286