BMJ Open

Assessing societal and offender perspectives on the value of offender health care: A research protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024899
Article Type:	Protocol
Date Submitted by the Author:	23-Jun-2018
Complete List of Authors:	Settumba, Stella; Kirby Institute, University of New South Wales, Shanahan , Marian Shanahan ; The University of New South Wales, National Drug and Alcohol Research Centre Chambers, Georgina; University of New South Wales, Center for Big Data Research in Health Schofield, PW; University of Newcastle Butler, Tony; University of New South Wales, Kirby Institute
Keywords:	Discrete choice experiment, Contingent valuation, Focus group discussions, Delphi method, Offender rehabilitation, Violence

TITLE PAGE

Title: Assessing societal and offender perspectives on the value of offender health care: A stated preference research protocol

Stella Nalukwago Settumba[1], Marian Shanahan[2], Georgina M. Chambers[3], Peter Schofield[4], Tony Butler[1]

1. Justice Health Program, Kirby Institute, University of New South Wales, Sydney, Australia
2. National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Australia
3. Centre for Big Data Research in Health, University of New South Wales, Sydney, Australia
4. School of Medicine and Public Health, University of Newcastle, Newcastle, Australia

Corresponding Author:

Stella Nalukwago Settumba

Justice Health Program, The Kirby Institute, UNSW

Wallace Wurth Building, High Street, Kensington

NSW, 2052

Sydney, Australia

Email address: snalukwago@kirby.unsw.edu.au

Telephone number: +61 (0) 29385 9007

Manuscript word count: 4338

Keywords: Discrete choice experiment, Contingent valuation, Focus group discussions, consensus methods, Delphi method, Offender rehabilitation, Impulsivity, Violence.

ABSTRACT

Introduction:

The increasing burden that offenders place on justice and health budgets necessitates better methods to determine the benefits of and value society places on offender programs to guide policy regarding resource allocation. The aim of this paper is to demonstrate how economic methods will be used to determine the strength of preferences and value of violent offender treatment programs from the perspectives of offenders, their families and the general population.

Methods and analysis:

Two stated preference economic methods, discrete choice experiment (DCE) and contingent valuation (CV), will be used to assess society's and offenders' value of treatment programs. The mixed methods process involves a literature review and qualitative methods to derive attributes and levels for the DCE and payment card values for the CV. Consensus building approaches of voting, ranking and the Delphi method will be used to further refine the findings from the qualitative phase. Attributes and their levels will be used in a D-efficient Bayesian experimental design to derive choice scenarios for the development of a questionnaire that will also include CV questions. Finally, quantitative surveys to assess societal preferences and value in terms of willingness to pay will be conducted.

Ethics and dissemination:

Ethics approval from this study was obtained from the University of New South Wales (UNSW) Human Research Ethics Committee, Corrective Services New South Wales Ethics Committee and Aboriginal Health and Medical Research Council ethics committee. The findings will be made available on the Kirby Institute UNSW website, published in peer reviewed journals and presented at national and international conferences. This study was funded by grants from the National Health and Medical Research Council, under the Centre of Research Excellence in Offender Health Australia [grant number RG124596]. It is part of the research done by the Justice and Health program, Kirby Institute.

Article Summary

Strengths and limitations of the study

- This study is the first to quantify societal and offender preferences for violent offender treatment and provides a rigorous mixed methodological approach that can be generalised for use in other DCE and CV studies of valuation of offender programs.
- The results from these studies will be used in valuing the strength of preferences of society and offenders for treatment programs to reduce reoffending.
- The study will provide an estimate of the value, in terms of willingness to pay, that society and offenders place on violent offender treatments.
- This study will also provide the basis for conducting cost-benefit analysis to indicate the relative 'value for money' for violent offender programs.
- Recruitment of violent offenders to participant in a study is often challenging and while we hope to have a large enough representative sample in the future to participate in a DCE with offenders only, in this study recruitment will be done among the general population. Questions that ask participants to self-identify as offenders and family members of offenders will be included in the survey and sub-analysis done if a large enough sample is obtained.

Funding

This study was funded by grants from the National Health and Medical Research Council, under the Centre of Research Excellence in Offender Health Australia [grant number RG124596]. It is part of the research done by the Justice and Health program, Kirby Institute.

Competing Interests

There are no competing interests for any other regarding this study.

Author contributorship statement

Stella Settumba is involved in the study design, participant recruitment for the FGDs, data collection, analysis and write-up; Georgina Chambers and Marian Shanahan are involved in the study design, data analysis and write-up, Tony Butler and Peter Schofield are involved in the study design, participant recruitment for the FGDs, data analysis, and write-up.

INTRODUCTION

Violence is a leading public health problem. It is estimated that more than 1.3 million people worldwide die each year as a result of violence accounting for 2.5% of global mortality [1]. The costs of violence to the health system in Australia, including medical costs, lost productivity, and intangible costs, are high and estimated at $AUD3.1 billion each year [2]. Imprisoning people with minimal rehabilitation has been shown to be largely ineffective as a deterrent to offending [3]. Well designed and evaluated interventions to reduce violence can save both lives and money. Research has shown that most violent crime can be classified as impulsive rather than pre-meditated and that impulsive offenders have a higher likelihood of recidivism than those offenders who commit pre-meditated crimes. Furthermore, impulsive offenders are more likely to respond positively to treatment and rehabilitation programs [4-6].

Decisions to allocate scarce resources to treat offenders, especially violent offenders, are seen by some as controversial even when the benefits of treatment extend beyond the offenders [7]. Public opinion and perception are often important determinants of the treatment and rehabilitation opportunities afforded to offenders since the justice system is financed through taxation, and politicians and other policy makers are wary of incurring the public's disapproval. However, surveys to determine the public's attitudes towards offender rehabilitation often suffer from poor methodology with poorly informed participants who lack accurate information on crime, its causes, and rehabilitation options and consequently are rarely given the opportunity to look beyond punitiveness [7 8]. Most surveys rely on snap polls with simple questions. Recommended techniques in the literature [9] for eliciting more considered and informed views from the public include: qualitative techniques such as one-to-one interviews, the Delphi technique, focus groups, citizens' juries, consensus panels and nominal group techniques; and quantitative techniques such as ranking, rating, discrete choice experiments (DCEs) and contingent valuation (CV) studies.

A variety of quantitative economic methods, including stated preferences and revealed preferences methods, have been employed to elicit patients' value for healthcare by quantifying their preferences [10]. Revealed preference methods refer to situations where people's choices are observed in actual market situations. However, in the absence of an actual market, as often found with many health programs or new interventions, stated preference techniques can be used. Stated preference methods refer to situations in which choices are made in a hypothetical market situation using a survey context. Valuation techniques using stated preference methods include the DCE [11 12] and CV method [13 14].

In one Australian state, New South Wales (NSW) after a successful pilot [15], a large randomised control trial (RCT) is underway, seeking to evaluate the effectiveness of a class of antidepressants, Selective Serotonin Reuptake Inhibitor (SSRI) (sertraline), to reduce impulsivity in men with a history of violent offending. This pharmacotherapy-based double blinded RCT is known as REINVESt ('Reducing Impulsivity in Repeat Violence Offenders Using a Selective Serotonin Reuptake Inhibitor'). Men who consent, are medically fit, have committed two or more violent offences and score highly on an impulsivity screener are randomised to receive either the SSRI or placebo for 6-12 months. If the intervention is found effective, valuation of its benefits is needed to advocate for the uptake of such treatment programs.

Using the REINVESt study as an exemplar, this paper demonstrates how economic methods will be used to assess the societal and offenders' value of treatment programs for offenders. The following are the aims of the economic study:

1. To elicit societal and offenders' preferences for treatment of impulsive violent offenders. Specifically,
    a) To assess the characteristics of treatment programs for impulsive violent offenders that could influence the uptake by offenders and support by society.
    b) To quantify the strength of preferences for and assess trade-offs between characteristics of treatment programs for impulsive violent offenders.
2. To elicit societal and offenders' value, in terms of WTP, of treatment of impulsive violent offenders. Specifically,
    a) To estimate the societal and offenders' average WTP for the treatment of impulsive violent offenders using an SSRI.
    b) To elicit the factors affecting societal and offenders WTP for offender treatment programs.

The study protocol described in this paper details the methods used in assessing offenders' and society's preference and value (stated as WTP) of violent offender treatment programs using the DCE and CV stated preference methods. To our knowledge, this will be the first study to assess both offenders' and societal preferences for offender treatment programs. Although the involvement of patients in preference measures for decision making has been advocated [16] no DCEs have been performed involving offenders.

METHODS

Aims 1(a) and 1(b) will be achieved through the qualitative (Phase 1) and quantitative (Phase II) components of the DCE respectively and objective 2 through the CV method (Phase II). The next paragraphs describe these methods.

The DCE

In a DCE, respondents' preferences are elicited based on their stated preference when faced with hypothetical choices between treatment scenarios that differ in terms of specified attributes and attribute levels. DCEs have been increasingly used in health economics to address a wide range of health policy related decisions (see [17-19] for more details on DCE methods). More recently, DCEs have been used in the justice area, for example, to explore societal preferences for alternative cannabis drug policies and to demonstrate the effect of varying cannabis policy characteristics and wider social consequences such as healthcare and criminal justice expenditures [20].

In this DCE study, participants will be asked to indicate their preference between two treatment programs for impulsive repeat-violent offenders, and a no treatment option. They will be presented with different choice scenarios comprised of differing characteristics of the treatment program (attributes) and attribute-levels. The results will be used to quantify the strengths of the preferences and assess the trade-offs between characteristics of treatment programs for impulsive violent offenders. By adding 'price' as an attribute in the DCE, the average societal WTP for a program can be estimated. A mixed methods design will be used (Figure 1). The DCE has four main steps (1)

identifying attributes and levels, (2) the experimental design, (3) the data collection survey, and (4) the analysis and interpretation of results [12]. In this study, step 1 was carried out in phase I and steps 2 to 4 will be in phase II. At the time of writing this protocol paper, phase I has been completed.

Developing attributes and levels

Phase 1 involves identifying all relevant attributes and assigning their associated levels. The generation of attributes used in DCEs is often poorly performed and reported and the need for rigorous research involving theoretical, conceptual, contextual and refined attributes has been emphasised [21]. In this study attributes were generated through a review of literature and primary qualitative research methods, Focus Group Discussions (FGDs). The consensus methods used to refine and narrow the attributes to a sizeable number included: ranking attributes, voting, and the Delphi method. The attributes characterise the factors considered by offenders and society to be most valuable for acceptance, support and uptake of treatment programs by impulsive violent offenders and the levels are the ranges over which the attributes vary.

Literature Review (1A)

A literature review developed the conceptual and theoretical attributes and levels which were further examined in FGDs. In this study, the literature searches considered existing theories that define an effective offender treatment program, positive and negative experiences of offenders with treatment programs, and views held by society regarding offender treatment programs.

Focus group discussions (IB)

Following the literature review, the next step involved developing contextual attributes and their levels using qualitative research, FGDs.

Participants recruited into the REINVESt study were invited to participate in the offender FGDs and fell into two categories: (1) current or past participants, and (2) those who were eligible for the study according to the selection criteria but declined to participate in REINVESt. Common reasons cited for non-participation in REINVESt included not wanting to take medication and current use of a psychotropic medication. All individuals invited to the offenders FGDs were: male, over 18 years of age, had a history of committing at least two violent offences, and a score over 70 on the Barratt Impulsiveness Scale [22] indicating impulsive personality traits. Purposive selection was made to ensure a balance in terms of age, and number of prior convictions. For all offenders selected to participate in the FGDs, a member of the REINVESt study team asked for their consent to be contacted for this study at one of the routine contact visits and those showing an interest were contacted by a letter requesting them to participate.

In addition, each participating offender in the REINVESt trial study was asked when they attended a routine study follow-up visit if they were happy for a family member to take part in a FGD. The REINVESt study team has good working relationships with some family members of offenders. With the offenders' consent, a member of the REINVESt study team requested the family members' consent to be contacted for this study. Those who agreed were sent an invitation letter to take part in the study including an email address and phone contact by which to contact the research team. Participants for the family members' FGD were defined as a partner or family member of a male offender participating in the REINVESt study and over 18 years of age.

Recruitment notices for the FGD with the general public were placed in libraries and community and online notice boards (e.g. Gumtree). Purposive sampling was done from those who responded, with an aim of having people with various ages, social and demographic backgrounds. Those selected were sent an invitation letter including an email address and phone number to contact the research team. Participants for the general public FGDs were required to be resident tax payers in NSW and over 18 years of age.

Recruitment and FGDs were conducted until saturation was reached, i.e. when no new data was generated with additional groups, bringing the total number of FGDs to 8 (4 offender, 2 general public and 1 family members group).

During the FGDs, after exploring participants' knowledge and views on violence, impulsivity, incarceration, recidivism and the role of treatment programs, they were provided with precise definitions of terms, examples of available interventions and contemporary statistics on violent crime, incarceration and recidivism rates. Participants then provided characteristics of treatment programs they might value if considering joining or supporting a treatment program. The levels reflected the range of situations that respondents might experience for each attribute. A semi-structured guide was used for the data collection.

After generating an exhaustive list of attributes, participants were asked to take part in a voting exercise [23] used to identify the top five characteristics generated within their FGD. Each participant was given unlimited votes and asked to vote 'yes' or 'no' if they thought a characteristic was important. 'Yes' votes were tallied for each characteristic and those with the top five most votes were noted as the top five attributes of preference for each group.

Participants then ranked the top 5 attributes in order of preference. Ranking exercises, as used in health priority setting, ask participants to give an ordinal rank to their preferences and those with the highest ranking are viewed as the most important [24-26]. The top attributes from the voting and ranking methods could now be included in the DCE. However, in this study, all attributes obtained from the FGDs were further assessed through the Delphi method and the results from the voting and ranking exercises used to provide a qualitative indication of the strength of the different attributes that will then be compared with results from the DCE.

All FGDs were recorded and Digital audio data was transcribed and then destroyed. The transcribed data and the facilitator notes were coded and analysed using thematic analysis in *NVivo* to identify all major and minor themes on characteristics of treatment programs for impulsive violent offenders. The themes were subsequently grouped to classify the similarities and differences between the different groups of offenders, their families and the general public. These themes were then summarized to create a list of attributes and levels that were discussed during the Delphi method.

Delphi Method (IC)

The attributes and their levels obtained from the literature review and FGDs were further deliberated on by a team of experts using the Delphi method to generate a final list of attributes that will be used for the experimental design of the DCE. The Delphi research method is widely used in healthcare research to achieve consensus from a panel on issues of selected subjects [27 28]. It has also been recommended for use in deliberating on issues raised through FGDs and literature reviews

[29] and for further refining of attributes and levels to be used in a DCE [30]. It is popular because, in addition to providing an opportunity for everyone's views to be taken into consideration by the group, it allows anonymous voting and avoids the domination of the consensus process by experts [31]. Delphi, in contrast to other data gathering and analysis techniques, involves heterogeneous expertise, motivated and involved participants and employs multiple iterations/rounds in the form of feedback giving participants an opportunity to make informed decisions with good reasons for judgments or preferences [32]. Using iterative qualitative methods to refine attributes for a DCE also enables the rewriting of attributes to incorporate all relevant concepts [21]. Although there are no strict guidelines on the number of rounds needed to achieve consensus, the basic principle of the Delphi technique is to have as many rounds as are required or until the 'law of diminishing returns' occurs but generally at least two rounds are required [33]. Figure 2 describes the Delphi method process that was used in this study.

The aim of the Delphi process was:

1. To further refine the attributes that had been gathered from the FGDs.
2. To reach consensus on the levels for each attribute.
3. To arrive at a consensus of 5 – 8 attributes that would be evaluated in a DCE survey.

All participants, identified through research, academic and program implementation networks of people in the justice space, received an invitation e-mail, together with an information sheet explaining the study, the Delphi method, and an online informed consent form. Non-responders were approached by phone after one week. Before enrolment, it was confirmed that participants had the intention to complete all rounds of the study and had access to the internet. Participants for the Delphi method included criminologists, nurses from the justice health sector, psychologists working in criminal justice, health economists, forensic psychiatrists, members of the Australian Indigenous community, Corrective Services NSW staff, and police officers.

Experimental design and pilot

Scenarios will be constructed using the final attributes and levels ascertained from the Delphi method. A full factorial design takes on all possible combinations of attributes and their levels. Given the large number of attributes and multiple levels obtained from phase I, it is not feasible for respondents to assess all possible choices. An experimental design, which involves selecting through the use of statistical software (*NGENE* [34]) a subset of scenarios for respondents to complete, will be used to construct a fractional factorial design [35] . This helps to minimise the number of choice sets presented to respondents while still obtaining the maximum amount of information.

Attributes in this study will be described by a continuous, discrete or categorical scale. Effects coding will be used for all categorical attributes and parameters estimated for each level. The design will be unlabelled, which means that the treatments in the scenarios will be generic and labelled as treatment 1 and treat 2.

A D-efficient experimental design that maximises model statistical efficiency by minimising the parameter standard errors will be employed [36]. To optimise D-efficiency, prior assumptions on model parameter estimates will be used. A pilot study will be carried out to obtain priors and to guide development and testing of the questionnaire. This will include testing of the appropriateness of the questions such as determining the number of respondents willing to answer personal

questions on exposure to violence, respondents' understanding and the correct balance between attributes and levels, task complexity, and timing of the length of response rates. Priors and their signs for the pilot will be based on data from the literature, or knowledge of known parameters [37]. Coefficients from the pilot testing will then be used as Bayesian priors for the Bayesian efficient experimental design [38], and the refined questionnaire will then be created. The design will be optimised for a multinomial logit model and this will then be evaluated in *NGENE* using a panel mixed logit model, which accounts for the parameter distribution, and a latent class model which accounts for non-uniformity of respondents.

Scenario presentation

Scenarios constructed from the experimental design will be presented to respondents in a survey delivered via a web-based questionnaire to elicit preferences. Respondents will be directed to read a description of all attributes prior to answering the questionnaire. Respondents will be then asked to choose between two treatment choice sets with different levels of attributes and a no treatment option. Those who chose the no treatment option will also be presented with a forced choice. The total number of choice sets per participant will be determined during the pilot and care will be taken to reduce cognitive burden. Generally 6-8 choice sets are recommended. Figure 3 is an example of a choice set.

Future work using the above methodology will involve conducting three separate DCEs, one with offenders, one with their families and one with members of the general public. Currently, the DCE will sample only from the general public. However, a question will be included in the survey to identify participants who are themselves offenders (having been accused of violence and having been in contact with the justice system for a violent offence) and family members of offenders. If an adequate number of participants self-identify as offenders and family members of offenders, sub-analyses for each group will be undertaken.

There is no agreed rule on the correct sample size required for a DCE [39]. However, research has shown that in all DCE studies with efficient designs, model estimate precision increases rapidly at sample sizes greater than 150 and then flattens out at around 300 observations [35]. It is also estimated that a minimum sample size of 200 respondents per sub-group be used for studies involving an analysis of differences between samples [40]. Furthermore, the s-efficiency measure in the experimental design in *NGENE* will estimate the required sample size for the study [41]. Recruitment, for the first DCE, will be from an outsourced online panel provider where respondent duplication and fraudulent completion of surveys is monitored. Participants are recruited via verified, certified sources and methods to create a large pool of potential research respondents for our clients. These participant panels have agreed and provided consent to participate in research conducted by the commercial survey company.

Data analysis and result interpretation

The data derived from the DCE surveys will be analysed to estimate attribute preference weights, also known as parameters, denoting the relative strength of each attribute in the choice of treatment programs for the offenders. This is done using the random utility maximisation framework [42].

The econometrics software *Nlogit* [43] will be used to perform the analysis. A multinomial logit model, a mixed logit model, or a latent class model will be estimated [44]. The final model will depend on

which model best fits the data. An assessment of how each model predicts the data will be made using the likelihood ratio index. Sub group analysis will be performed to analyse the differences in parameter strengths between the three groups: offenders, family members of offenders and members of the general population.

WTP for an attribute will be defined as the ratio of the change in marginal utility of an attribute (attribute k in the equation) to marginal utility for the price attribute (p in the equation), as follows:

$$\text{WTP} = \frac{Change\ in\ X_k}{Change\ in\ X_p} = \frac{\frac{d}{dx_k}\beta_k x_k}{\frac{d}{dx_p}\beta_c x_p} = -\frac{\beta_k}{\beta_p}$$

An estimation of WTP for a treatment program that is described by the attributes in the DCE model will be calculated as the sum of marginal WTP for each attribute.

The Contingent Valuation method

Design

The contingent valuation method will also be used to solicit respondents' willingness to pay for a defined treatment program for impulsive violent offenders. Obtaining accurate WTP estimates using CV method requires detailed descriptions of the treatment being valued. This is evident from the name of the method, which produces values, contingent upon, the description of treatment. A description of the REINVESt study treatment program will be provided as an exemplar of a treatment program for impulsive violent offenders.

The payment card will be used as the WTP elicitation question. Respondents will be presented with a range of bid amounts and asked to choose the maximum amount in the form of an additional tax levy that they are willing to pay to have a described treatment available to impulsive repeat violent offenders. This reflects real life by allowing individuals to 'shop around' for the value closest to their maximum WTP [45]. The dollar values used on the payment cards were also explored in the FGD qualitative interviews and in the pilot study.

The CVM has been widely criticised for bias in terms of the validity of its results. Therefore, care will be taken in the design and analysis to reduce any bias that may arise. This will include randomisation of positioning of the dollar values of the payment cards to reduce anchoring or starting point bias [46] [47]. Furthermore, to reduce the point bias or range bias [48] one of the options in the WTP payment card values will be 'none of these amounts' and respondents will then be asked to state how much they would be willing to pay.

Data Collection

The results from the qualitative methods in phase I and the pilot survey described in the DCE method will be used to describe the treatment to be valued in the CVM. Through an additional question to the DCE survey, participants will be asked to state their WTP for a described intervention similar to REINVESt. The respondents and sample will therefore be the same as explained in the DCE study.

Data Analysis

Mean and/or median WTP values will be calculated. Logistic regression models will be used To identify the factors affecting both zero and positive WTP and to estimate the independent effects of demographic characteristics on the WTP for offender treatments. The outcome of the model will be specified as the probability of agreeing to pay for offender treatment. The model fit will be estimated using the maximum likelihood function.

Patient and Public involvement

This protocol is about a study that seeks to assess offender and public preferences and therefore greatly involves the two groups. Phase 1 of this study involves the eliciting of offender and general public preferences through focus group discussions and the Delphi method. The offenders were voluntarily recruited through REINVESt, a study by the Justice Health Program at Kirby Institute UNSW. Phase II is a quantitative general population survey that will quantify the strength of preferences and assess the value of the treatment program. Participants for the survey will be representative of the NSW population and will be voluntarily recruited through a marketing survey company.

ETHICS AND DISSEMINATION

Ethics approval for this study has been provided for the two phases. Phase 1 ethics approval has been provided by UNSW – Higher Risk Ethics Committee, NSW Corrective Services Ethics Committee, and Aboriginal Health and Medical Research Council (AH&MRC) ethics committee. Phase II ethics approval has been provided by UNSW – higher risk ethics committee for the DCE general population sample. If in future the DCE is to be conducted with a sample from offenders, further ethics applications will be made to NSW corrective services and AH&MRC ethics committees.

The findings of this study will be made available on the Kirby Institute UNSW website, published in peer reviewed journals and presented at national and international conferences.

IMPORTANCE OF THIS PAPER

This research will provide a significant contribution to the assessment and evaluation of offender programs. In the DCE, an understanding of the trade-offs made and the strengths of preferences of society in the provision of healthcare for violent offenders will help provide valuable information for policy makers, treatment providers and other practitioners in designing treatment options.

Eliciting societal willingness to pay for offender treatment programs will be used to assess the value/benefit of the programs to both offenders and the public. When deciding whether to fund an intervention, policy makers need to consider how much the public values the benefits - hence how much they would be willing to pay. If the costs of interventions similar to REINVESt are known, the results (benefit values) of this study can be used in cost-benefit analyses.

The average WTP obtained using the DCE method can be compared with the average WTP obtained using the payment card CVM [49]. This can allow for testing of convergent validity of the two WTP methods i.e. the degree to which the results of the two methods are related.

This paper outlines a rigorous methodological approach that can be used to assess societal preferences and generalised for use in other DCE and CVM valuation for offender treatment

programs as opposed to the traditional methods of opinion polls, which often only emphasise punitiveness of the public towards offenders, especially those who commit violent offences.

We outline a mixed methods process that involves qualitative methods, consensus approaches and economic methods of preference setting. We also provide a study context where the methods are applied: the REINVESt study. The rich qualitative component of this study will contribute to the literature concerned with the development of attributes for DCEs.

REFERENCES

1. Allen R. Global prison trends 2015. *Penal Reform International* 2015
2. Rollings K. Counting the costs of crime in Australia: a 2005 update: Australian Institute of Criminology Canberra, ACT, 2008.
3. Weatherburn D, Froyland G, Moffatt S, et al. Prison populations and correctional outlays: The effect of reducing re-imprisonment. *BOCSAR NSW Crime and Justice Bulletins* 2009:12.
4. Kockler TR, Stanford MS, Nelson CE, et al. Characterizing aggressive behavior in a forensic population. *American Journal of Orthopsychiatry* 2006;76(1):80.
5. Stanford MS, Houston RJ, Mathias CW, et al. Characterizing aggressive behavior. *Assessment* 2003;10(2):183-90.
6. Heilbrun AB. Psychopathy and violent crime. *Journal of Consulting and Clinical Psychology* 1979;47(3):509.
7. Cullen FT, Fisher BS, Applegate BK. Public opinion about punishment and corrections. *Crime and justice* 2000:1-79.
8. Hutton N. Beyond populist punitiveness? *Punishment & Society* 2005;7(3):243-58. doi: 10.1177/1462474505053828
9. Ryan M, Scott D, Reeves C, et al. Eliciting public preferences for healthcare: a systematic review of techniques, 2001.
10. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ : British Medical Journal* 2000;320(7248):1530-33.
11. Ryan M, Gerard K, Amaya-Amaya M. Using discrete choice experiments to value health and health care: Springer Science & Business Media 2007.
12. Lancsar E, Louviere J. Conducting Discrete Choice Experiments to Inform Healthcare Decision Making. *PharmacoEconomics* 2008;26(8):661-77. doi: 10.2165/00019053-200826080-00004
13. Klose T. The contingent valuation method in health care. *Health policy* 1999;47(2):97-123.
14. Diener A, O'brien B, Gafni A. Health care contingent valuation studies: a review and classification of the literature. *Health economics* 1998;7(4):313-26.
15. Butler T, Schofield PW, Greenberg D, et al. Reducing impulsivity in repeat violent offenders: an open label trial of a selective serotonin reuptake inhibitor. *Australian and New Zealand Journal of Psychiatry* 2010;44(12):1137-43. doi: 10.3109/00048674.2010.525216
16. Cleary PD. The increasing importance of patient surveys : Now that sound methods exist, patient surveys can facilitate improvement. *BMJ : British Medical Journal* 1999;319(7212):720-21.
17. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ* 2012;21(2):145-72. doi: 10.1002/hec.1697
18. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2003;2(1):55-64.
19. Clark M, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. *Pharmacoeconomics* 2014;32(9):883-902. doi: 10.1007/s40273-014-0170-x
20. Shanahan M, Gerard K, Ritter A. Preferences for policy options for cannabis in an Australian general population: A discrete choice experiment. *International Journal of Drug Policy* 2014;25(4):682-90. doi: http://dx.doi.org/10.1016/j.drugpo.2014.03.005
21. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Economics* 2012;21(6):730-41. doi: 10.1002/hec.1739
22. Barratt E, Stanford M. Impulsiveness, in the personality characteristics of the Personality Disordered client. Edited by Costello, CG New York: Wiley, 1995.
23. Mullen PM. Public involvement in health care priority setting: an overview of methods for eliciting values. *Health Expectations* 1999;2(4):222-34. doi: 10.1046/j.1369-6513.1999.00062.x

24. Bowling A, Jacobson B, Southgate L. Explorations in consultation of the public and health professionals on priority setting in an inner London health district. *Social science & medicine* 1993;37(7):851-57.
25. Furnham A, Meader N, McClelland A. Factors affecting nonmedical participants' allocation of scarce medical resources. *Journal of Social Behavior and Personality* 1998;13(4):735.
26. Rosko MD, McKenna W. Modeling consumer choices of health plans: a comparison of two techniques. *Social Science & Medicine* 1983;17(7):421-29.
27. Alexander J, Kroposki M. Outcomes for community health nursing practice. *J Nurs Adm* 1999;29(5):49-56. doi: 10.1097/00005110-199905000-00007
28. Green B, Jones M, Hughes D, et al. Applying the Delphi technique in a study of GPs' information requirements. *Health Soc Care Community* 1999;7(3):198-205. doi: 10.1046/j.1365-2524.1999.00176.x
29. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32(4):1008-15. [published Online First: 2000/11/30]
30. Dachary-Bernard J, Rambonilaza T. Choice experiment, multiple programmes contingent valuation and landscape preferences: How can we support the land use decision making process? *Land Use Policy* 2012;29(4):846-54.
31. Jairath N, Weinstein J. The Delphi methodology (Part one): A useful administrative approach. *Canadian journal of nursing administration* 1994;7(3):29-42.
32. Bolger F, Wright G. Improving the Delphi process: Lessons from social psychological research. *Technological Forecasting and Social Change* 2011;78(9):1500-13. doi: http://dx.doi.org/10.1016/j.techfore.2011.07.007
33. Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing* 2006;53(2):205-12. doi: 10.1111/j.1365-2648.2006.03716.x
34. NGENE software, version: 1.1.2 [program]: ChoiceMetrics, 2014.
35. Johnson FR, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health* 2013;16(1):3-13.
36. Bliemer M, Rose JM, Chorus CG. Detecting dominancy and accounting for scale differences when using stated choice data to estimate logit models. 2015
37. Kanninen BJ. Optimal design for multinomial choice experiments. *Journal of Marketing Research* 2002;39(2):214-27.
38. Bliemer MCJ, Rose JM, Hess S. Approximation of bayesian efficiency in experimental choice designs. *Journal of Choice Modelling* 2008;1(1):98-126. doi: https://doi.org/10.1016/S1755-5345(13)70024-1
39. de Bekker-Grob EW, Donkers B, Jonker MF, et al. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. *The Patient-Patient-Centered Outcomes Research* 2015;8(5):373-84.
40. Johnson R, Orme B. Sample size issues for conjoint analysis. *Getting started with conjoint analysis: strategies for product design and pricing research Madison: Research Publishers LLC* 2010:57-66.
41. Rose JM, Bliemer MCJ. Sample size requirements for stated choice experiments. *Transportation* 2013;40(5):1021-41. doi: 10.1007/s11116-013-9451-z
42. McFadden D. Conditional logit analysis of qualitative choice behavior. 1973
43. Greene W. NLOGIT version 6.0: EconometricsSoftware, Inc., 2015.
44. Hensher DA, Greene WH. The mixed logit model: the state of practice. *Transportation* 2003;30(2):133-76.
45. Donaldson C, Thomas R, Torgerson DJ. Validity of open-ended and payment scale approaches to eliciting willingness to pay. *Applied Economics* 1997;29(1):79-84.

46. Cummings RG, Harrison GW, Osborne LL. Can the bias of contingent valuation be reduced? Evidence from the laboratory. *Economics Working Paper B-95* 1995;3
47. Boyle KJ. Contingent Valuation in Practice. A Primer on Nonmarket Valuation: Springer 2017:83-131.
48. Heinzen RR, Bridges JF. Comparison of four contingent valuation methods to estimate the economic value of a pneumococcal vaccine in Bangladesh. *International journal of technology assessment in health care* 2008;24(4):481-7. doi: 10.1017/s026646230808063x [published Online First: 2008/10/03]
49. Ryan M, Watson V. Comparing welfare estimates from payment card contingent valuation and discrete choice experiments. *Health Economics* 2009;18(4):389-401. doi: 10.1002/hec.1364

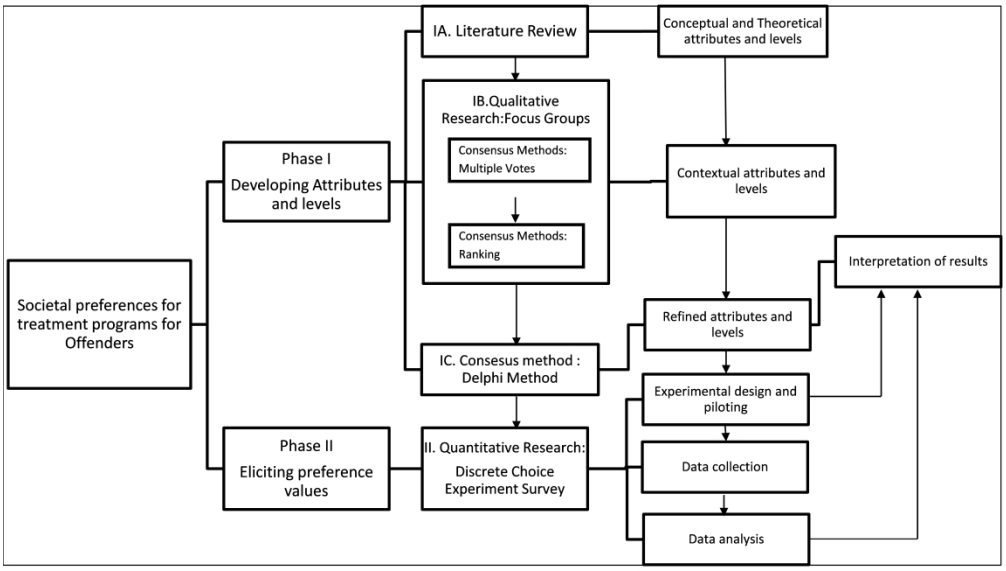

Figure 1: The mixed methods design of the Discrete Choice Experiment

494x296mm (300 x 300 DPI)

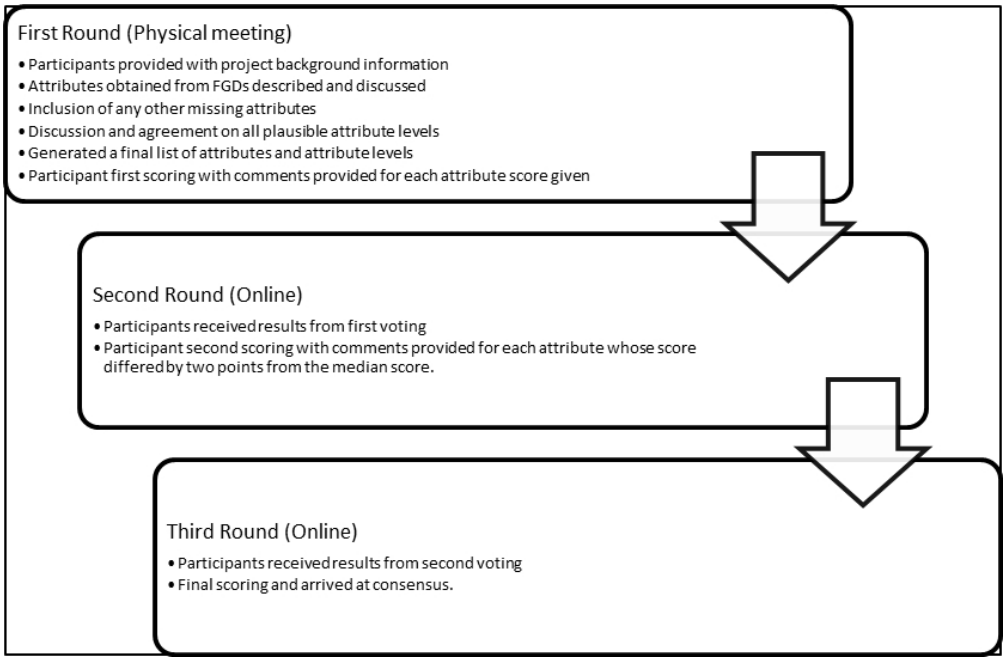

Figure 2: The Delphi method used to refine the attributes for the impulsive violent offender DCE

70x46mm (300 x 300 DPI)

Characteristic of treatment	Treatment 1	Treatment 2	No treatment
Effectiveness of the treatment	30% reduction in crime	50% reduction in crime	_
Treatment of co-occurring health conditions/addictions	Minimal treatment provided in program	Full treatment of all co-occurring morbidities both within program and at referral facilities	_
Type of treatment	Offender group counselling sessions only	Individual and Family counselling with Medication	_
Treatment provider	Prison/Probation & parole officers with Counsellors/Psychologists	Prison/Probation & Parole officer with Counsellors/Psychologists with Health Professional	_
Flexibility of appointments	Not flexible	Flexible	_
Compulsory/Voluntary participation	Voluntary	Compulsory	_
Cost per tax payer per year	$50	$75	_
Which treatment would you prefer to be given to Impulsive violent offenders?	○	○	○
If you had to choose between treatment 1 and treatment 2, which one would you prefer?	○	○	

Figure 3: An example of a choice set for the DCE

321x178mm (300 x 300 DPI)

BMJ Open

Assessing societal and offender perspectives on the value of offender health care: A stated preference research protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024899.R1
Article Type:	Protocol
Date Submitted by the Author:	29-Nov-2018
Complete List of Authors:	Settumba, Stella; Kirby Institute, University of New South Wales, Shanahan , Marian Shanahan ; The University of New South Wales, National Drug and Alcohol Research Centre Chambers, Georgina; University of New South Wales, Center for Big Data Research in Health Schofield, PW; University of Newcastle Butler, Tony; University of New South Wales, Kirby Institute
Primary Subject Heading:	Health economics
Secondary Subject Heading:	Health economics
Keywords:	Discrete choice experiment, Contingent valuation, Delphi method, Offender rehabilitation, Violence, Impulsivity

TITLE PAGE

Title: Assessing societal and offender perspectives on the value of offender health care: A stated preference research protocol

Stella Nalukwago Settumba[1], Marian Shanahan[2], Georgina M. Chambers[3], Peter Schofield[4], Tony Butler[1]

1. Justice Health Program, Kirby Institute, University of New South Wales, Sydney, Australia
2. National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Australia
3. Centre for Big Data Research in Health, University of New South Wales, Sydney, Australia
4. School of Medicine and Public Health, University of Newcastle, Newcastle, Australia

Corresponding Author:

Stella Nalukwago Settumba

Justice Health Program, The Kirby Institute, UNSW

Wallace Wurth Building, High Street, Kensington

NSW, 2052

Sydney, Australia

Email address: snalukwago@kirby.unsw.edu.au

Telephone number: +61 (0) 29385 9007

Manuscript word count: 4338

Keywords: Discrete choice experiment, Contingent valuation, Delphi method, Offender rehabilitation, Impulsivity, Violence.

ABSTRACT

Introduction:
The increasing burden that offenders place on justice and health budgets necessitates better methods to determine the benefits of and value society places on offender programs to guide policy regarding resource allocation. The aim of this paper is to demonstrate how economic methods will be used to determine the strength of preferences and value of violent offender treatment programs from the perspectives of offenders, their families and the general population.

Methods and analysis:
Two stated preference economic methods, discrete choice experiment (DCE) and contingent valuation (CV), will be used to assess society's and offenders' value of treatment programs. The mixed methods process involves a literature review and qualitative methods to derive attributes and levels for the DCE and payment card values for the CV. Consensus building approaches of voting, ranking and the Delphi method will be used to further refine the findings from the qualitative phase. Attributes and their levels will be used in a D-efficient Bayesian experimental design to derive choice scenarios for the development of a questionnaire that will also include CV questions. Finally, quantitative surveys to assess societal preferences and value in terms of willingness to pay will be conducted.

Ethics and dissemination:
Ethics approval from this study was obtained from the University of New South Wales (UNSW) Human Research Ethics Committee, Corrective Services New South Wales Ethics Committee and Aboriginal Health and Medical Research Council ethics committee. The findings will be made available on the Kirby Institute UNSW website, published in peer reviewed journals and presented at national and international conferences. This study was funded by grants from the National Health and Medical Research Council, under the Centre of Research Excellence in Offender Health Australia [grant number RG124596]. It is part of the research done by the Justice and Health program, Kirby Institute.

Article Summary

Strengths and limitations of the study

- This study is the first to quantify societal and offender preferences for violent offender treatment and provides a rigorous mixed methodological approach that can be generalised for use in other discrete choice experiment (DCE) and contingent valuation studies of valuation of offender programs.
- The results from these studies will be used in valuing the strength of preferences of society and offenders for treatment programs to reduce reoffending.
- The study will provide an estimate of the value, in terms of willingness to pay, that society and offenders place on violent offender treatments.
- This study will also provide the basis for conducting cost-benefit analysis to indicate the relative 'value for money' for violent offender programs.
- Recruitment of violent offenders to participant in a study is often challenging and while we hope to have a large enough representative sample in the future to participate in a DCE with offenders only, in this study recruitment will be done among the general population. Questions that ask participants to self-identify as offenders and family members of offenders will be included in the survey and sub-analysis done if a large enough sample is obtained.

Funding

This study was funded by grants from the National Health and Medical Research Council, under the Centre of Research Excellence in Offender Health Australia [grant number RG124596]. It is part of the research done by the Justice and Health program, Kirby Institute.

Competing Interests

There are no competing interests for any other regarding this study.

Author contributorship statement

Stella Settumba is involved in the study design, participant recruitment for the FGDs, data collection, analysis and write-up; Georgina Chambers and Marian Shanahan are involved in the study design, data analysis and write-up, Tony Butler and Peter Schofield are involved in the study design, participant recruitment for the FGDs, data analysis, and write-up. All authors have provided approval for the publication of this work and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

INTRODUCTION

Violence is a leading public health problem. It is estimated that more than 1.3 million people worldwide die each year as a result of violence accounting for 2.5% of global mortality [1]. The costs of violence to the health system in Australia, including medical costs, lost productivity, and intangible costs, are high and estimated at $AUD3.1 billion each year [2]. Imprisoning people with minimal rehabilitation has been shown to be largely ineffective as a deterrent to offending [3]. Well designed and evaluated interventions to reduce violence can save both lives and money. Research has shown that most violent crime can be classified as impulsive rather than pre-meditated and that impulsive offenders have a higher likelihood of recidivism than those offenders who commit pre-meditated crimes. Furthermore, impulsive offenders are more likely to respond positively to treatment and rehabilitation programs [4-6].

Decisions to allocate scarce resources to treat offenders, especially violent offenders, are seen by some as controversial even when the benefits of treatment extend beyond the offenders [7]. Public opinion and perception are often important determinants of the treatment and rehabilitation opportunities afforded to offenders since the justice system is financed through taxation, and politicians and other policy makers are wary of incurring the public's disapproval. However, surveys to determine the public's attitudes towards offender rehabilitation often suffer from poor methodology with poorly informed participants who lack accurate information on crime, its causes, and rehabilitation options and consequently are rarely given the opportunity to look beyond punitiveness [7 8]. Most surveys rely on snap polls with simple questions. Recommended techniques in the literature [9] for eliciting more considered and informed views from the public include: qualitative techniques such as one-to-one interviews, the Delphi technique, focus groups, citizens' juries, consensus panels and nominal group techniques; and quantitative techniques such as ranking, rating, discrete choice experiments (DCEs) and contingent valuation (CV) studies.

A variety of quantitative economic methods, including stated preferences and revealed preferences methods, have been employed to elicit patients' value for healthcare by quantifying their preferences [10]. Revealed preference methods refer to situations where people's choices are observed in actual market situations. However, in the absence of an actual market, as often found with many health programs or new interventions, stated preference techniques can be used. Stated preference methods refer to situations in which choices are made in a hypothetical market situation using a survey context. Valuation techniques using stated preference methods include the DCE [11 12] and CV method [13 14].

In one Australian state, New South Wales (NSW) after a successful pilot [15], a large randomised control trial (RCT) is underway, seeking to evaluate the effectiveness of a class of antidepressants, Selective Serotonin Reuptake Inhibitor (SSRI) (sertraline), to reduce impulsivity in men with a history of violent offending. This pharmacotherapy-based double blinded RCT is known as REINVESt ('Reducing Impulsivity in Repeat Violence Offenders Using a Selective Serotonin Reuptake Inhibitor'). Men who consent, are medically fit, have committed two or more violent offences and score highly on an impulsivity screener are randomised to receive either the SSRI or placebo for 6-12 months. If the intervention is found effective, valuation of its benefits is needed to advocate for the uptake of such treatment programs.

Using the REINVESt study as an exemplar, this paper demonstrates how economic methods will be used to assess the societal and offenders' value of treatment programs for offenders. The following are the aims of the economic study:

1. To elicit societal and offenders' preferences for treatment of impulsive violent offenders. Specifically,
    a) To assess the characteristics of treatment programs for impulsive violent offenders that could influence the uptake by offenders and support by society.
    b) To quantify the strength of preferences for and assess trade-offs between characteristics of treatment programs for impulsive violent offenders.
2. To elicit societal and offenders' value, in terms of willingness to pay (WTP), of treatment of impulsive violent offenders. Specifically,
    a) To estimate the societal and offenders' average WTP for the treatment of impulsive violent offenders using an SSRI.
    b) To elicit the factors affecting societal and offenders WTP for offender treatment programs.

The study protocol described in this paper details the methods used in assessing offenders' and society's preference and value (stated as WTP) of violent offender treatment programs using the DCE and CV stated preference methods. To our knowledge, this will be the first study to assess both offenders' and societal preferences for offender treatment programs. Although the involvement of patients in preference measures for decision making has been advocated [16] no DCEs have been performed involving offenders. This study will be conducted between July 2017 and April 2019.

METHODS

Aims 1(a) and 1(b) will be achieved through the qualitative (Phase 1) and quantitative (Phase II) components of the DCE respectively and objective 2 through the CV method (Phase II). The next paragraphs describe these methods.

The DCE

In a DCE, respondents' preferences are elicited based on their stated preference when faced with hypothetical choices between treatment scenarios that differ in terms of specified attributes and attribute levels. DCEs have been increasingly used in health economics to address a wide range of health policy related decisions (see [17-19] for more details on DCE methods). More recently, DCEs have been used in the justice area, for example, to explore societal preferences for alternative cannabis drug policies and to demonstrate the effect of varying cannabis policy characteristics and wider social consequences such as healthcare and criminal justice expenditures [20].

In this DCE study, participants will be asked to indicate their preference between two treatment programs for impulsive repeat-violent offenders, and a no treatment option. They will be presented with different choice scenarios comprised of differing characteristics of the treatment program (attributes) and attribute-levels. The results will be used to quantify the strengths of the preferences and assess the trade-offs between characteristics of treatment programs for impulsive violent offenders. By adding 'price' as an attribute in the DCE, the average societal WTP for a program can be estimated. A mixed methods design will be used (Figure 1). The DCE has four main steps (1)

identifying attributes and levels, (2) the experimental design, (3) the data collection survey, and (4) the analysis and interpretation of results [12]. In this study, step 1 was carried out in phase I and steps 2 to 4 will be in phase II. At the time of writing this protocol paper, phase I has been completed.

Developing attributes and levels

Phase 1 involves identifying all relevant attributes and assigning their associated levels. The generation of attributes used in DCEs is often poorly performed and reported and the need for rigorous research involving theoretical, conceptual, contextual and refined attributes has been emphasised [21]. In this study attributes were generated through a review of literature and primary qualitative research methods, Focus Group Discussions (FGDs). The consensus methods used to refine and narrow the attributes to a sizeable number included: ranking attributes, voting, and the Delphi method. The attributes characterise the factors considered by offenders and society to be most valuable for acceptance, support and uptake of treatment programs by impulsive violent offenders and the levels are the ranges over which the attributes vary. An example of a common attribute used in DCE studies is 'cost of the treatment program'. The attribute levels would be the various dollar amounts the treatment might cost.

Literature Review (1A)

A literature review developed the conceptual and theoretical attributes and levels which were further examined in FGDs. In this study, the literature searches considered existing theories that define an effective offender treatment program, positive and negative experiences of offenders with treatment programs, and views held by society regarding offender treatment programs.

Focus group discussions (IB)

Following the literature review, the next step involved developing contextual attributes and their levels using qualitative research, FGDs. For the FGDs we had three different participant group types: offenders, family members of offenders and people from the general public. We hypothesise that the attributes or attribute levels for treatment programs preferred will be different for the various groups. The FGDs were facilitated by an experienced qualitative researcher with vast experience working with offender populations.

Participants recruited into the REINVESt study were invited to participate in the offender FGDs and fell into two categories: (1) current or past participants, and (2) those who were eligible for the study according to the selection criteria but declined to participate in REINVESt. Common reasons cited for non-participation in REINVESt included not wanting to take medication and current use of a psychotropic medication. All individuals invited to the offenders FGDs were: male, over 18 years of age, had a history of committing at least two violent offences, and a score over 70 on the Barratt Impulsiveness Scale [22] indicating impulsive personality traits. Purposive selection was made to ensure a balance in terms of age, and number of prior convictions. For all offenders selected to participate in the FGDs, a member of the REINVESt study team asked for their consent to be contacted for this study at one of the routine contact visits and those showing an interest were contacted by a letter requesting them to participate.

In addition, each participating offender in the REINVESt trial study was asked when they attended a routine study follow-up visit if they were happy for a family member to take part in a FGD. The REINVESt study team has good working relationships with some family members of offenders. With the offenders' consent, a member of the REINVESt study team requested the family members'

consent to be contacted for this study. Those who agreed were sent an invitation letter to take part in the study including an email address and phone contact by which to contact the research team. Participants for the family members' FGD were defined as a partner or family member of a male offender participating in the REINVESt study and over 18 years of age.

Recruitment notices for the FGD with the general public were placed in libraries and community and online notice boards (e.g. Gumtree). Purposive sampling was done from those who responded, with an aim of having people with various ages, social and demographic backgrounds. Those selected were sent an invitation letter including an email address and phone number to contact the research team. Participants for the general public FGDs were required to be resident tax payers in NSW and over 18 years of age.

Recruitment and FGDs were conducted until saturation was reached, i.e. when no new data was generated with additional groups, bringing the total number of FGDs to 8 (4 offender, 2 general public and 1 family members group).

During the FGDs, after exploring participants' knowledge and views on violence, impulsivity, incarceration, recidivism and the role of treatment programs, they were provided with precise definitions of terms, examples of available interventions and contemporary statistics on violent crime, incarceration and recidivism rates. Participants then provided characteristics of treatment programs they might value if considering joining or supporting a treatment program. The levels reflected the range of situations that respondents might experience for each attribute. A semi-structured guide was used for the data collection.

After generating an exhaustive list of attributes, participants were asked to take part in a voting exercise [23] used to identify the top five characteristics generated within their FGD. Each participant was given unlimited votes and asked to vote 'yes' or 'no' if they thought a characteristic was important. 'Yes' votes were tallied for each characteristic and those with the top five most votes were noted as the top five attributes of preference for each group.

Participants then ranked the top 5 attributes in order of preference. Ranking exercises, as used in health priority setting, ask participants to give an ordinal rank to their preferences and those with the highest ranking are viewed as the most important [24-26]. The top attributes from the voting and ranking methods could now be included in the DCE. However, in this study, all attributes obtained from the FGDs were further assessed through the Delphi method and the results from the voting and ranking exercises used to provide a qualitative indication of the strength of the different attributes that will then be compared with results from the DCE.

All FGDs were recorded and Digital audio data was transcribed and then destroyed. The transcribed data and the facilitator notes were coded and analysed using thematic analysis in *NVivo* to identify all major and minor themes on characteristics of treatment programs for impulsive violent offenders. The themes were subsequently grouped to classify the similarities and differences between the different groups of offenders, their families and the general public. These themes were then summarized to create a list of attributes and levels that were discussed during the Delphi method.

Delphi Method (IC)

The attributes and their levels obtained from the literature review and FGDs were further deliberated on by a team of experts using the Delphi method to generate a final list of attributes that will be used for the experimental design of the DCE. The Delphi research method is widely used in healthcare research to achieve consensus from a panel on issues of selected subjects [27 28]. It has also been recommended for use in deliberating on issues raised through FGDs and literature reviews [29] and for further refining of attributes and levels to be used in a DCE [30]. It is popular because, in addition to providing an opportunity for everyone's views to be taken into consideration by the group, it allows anonymous voting and avoids the domination of the consensus process by a few individuals[31]. Delphi, in contrast to other data gathering and analysis techniques, involves heterogeneous expertise, motivated and involved participants and employs multiple iterations/rounds in the form of feedback giving participants an opportunity to make informed decisions with good reasons for judgments or preferences [32]. Using iterative qualitative methods to refine attributes for a DCE also enables the rewriting of attributes to incorporate all relevant concepts [21]. Although there are no strict guidelines on the number of rounds needed to achieve consensus, the basic principle of the Delphi technique is to have as many rounds as are required or until the 'law of diminishing returns' occurs but generally at least two rounds are required [33]. Figure 2 describes the Delphi method process that was used in this study.

The aim of the Delphi process was:

1. To further refine the attributes that had been gathered from the FGDs.
2. To reach consensus on the levels for each attribute.
3. To arrive at a consensus of 5 – 8 attributes that would be evaluated in a DCE survey.

All participants, identified through research, academic and program implementation networks of people in the justice space, received an invitation e-mail, together with an information sheet explaining the study, the Delphi method, and an online informed consent form. Non-responders were approached by phone after one week. Before enrolment, it was confirmed that participants had the intention to complete all rounds of the study and had access to the internet. Participants for the Delphi method included criminologists, nurses from the justice health sector, psychologists working in criminal justice, health economists, forensic psychiatrists, members of the Australian Indigenous community, Corrective Services NSW staff, and police officers.

Experimental design and pilot

Scenarios will be constructed using the final attributes and levels ascertained from the Delphi method. A full factorial design takes on all possible combinations of attributes and their levels. Given the large number of attributes and multiple levels obtained from phase I, it is not feasible for respondents to assess all possible choices. An experimental design, which involves selecting through the use of statistical software (*NGENE* [34]) a subset of scenarios for respondents to complete, will be used to construct a fractional factorial design [35] . This helps to minimise the number of choice sets presented to respondents while still obtaining the maximum amount of information.

Attributes in this study will be described by a continuous, discrete or categorical scale. Effects coding will be used for all categorical attributes and parameters estimated for each level. The design will be

unlabelled, which means that the treatments in the scenarios will be generic and labelled as treatment 1 and treatment 2.

A D-efficient experimental design that maximises model statistical efficiency by minimising the parameter standard errors will be employed [36]. To optimise D-efficiency, prior assumptions on model parameter estimates will be used. A pilot study will be carried out to obtain priors and to guide development and testing of the questionnaire. This will include testing of the appropriateness of the questions such as determining the number of respondents willing to answer personal questions on exposure to violence, respondents' understanding and the correct balance between attributes and levels, task complexity, and timing of the length of response rates. Priors and their signs for the pilot will be based on data from the literature, or knowledge of known parameters [37]. Coefficients from the pilot testing will then be used as Bayesian priors for the Bayesian efficient experimental design [38], and the refined questionnaire will then be created. The design will be optimised for a multinomial logit model and this will then be evaluated in *NGENE* using a panel mixed logit model, which accounts for the parameter distribution, and a latent class model which accounts for non-uniformity of respondents.

Scenario presentation

Scenarios constructed from the experimental design will be presented to respondents in a survey delivered via a web-based questionnaire to elicit preferences. Respondents will be directed to read a description of all attributes prior to answering the questionnaire. Respondents will then be asked to choose between two treatment choice sets with different levels of attributes and a no treatment option. Those who chose the no treatment option will also be presented with a forced choice. The total number of choice sets per participant will be determined during the pilot and care will be taken to reduce cognitive burden. Generally 6-8 choice sets are recommended. Figure 3 is an example of a choice set.

Future work using the above methodology will involve conducting three separate DCEs, one with offenders, one with their families and one with members of the general public. Currently, the DCE will sample only from the general public. However, a question will be included in the survey to identify participants who are themselves offenders (having been accused of violence and having been in contact with the justice system for a violent offence) and family members of offenders. If an adequate number of participants self-identify as offenders and family members of offenders, sub-analyses for each group will be undertaken.

There is no agreed rule on the correct sample size required for a DCE [39]. However, research has shown that in all DCE studies with efficient designs, model estimate precision increases rapidly at sample sizes greater than 150 and then flattens out at around 300 observations [35]. It is also estimated that a minimum sample size of 200 respondents per sub-group be used for studies involving an analysis of differences between samples [40]. Furthermore, the s-efficiency measure in the experimental design in *NGENE* will estimate the required sample size for the study [41]. Recruitment, for the first DCE, will be from an outsourced online panel provider where respondent duplication and fraudulent completion of surveys is monitored. Participants are recruited via verified, certified sources and methods to create a large pool of potential research respondents for our clients. These participant panels have agreed and provided consent to participate in research conducted by the commercial survey company. The panel

Data analysis and result interpretation

The data derived from the DCE surveys will be analysed to estimate attribute preference weights, also known as parameters, denoting the relative strength of each attribute in the choice of treatment programs for the offenders. This is done using the random utility maximisation framework [42].

The econometrics software *Nlogit* [43] will be used to perform the analysis. A multinomial logit model, a mixed logit model, or a latent class model will be estimated [44]. The final model will depend on which model best fits the data. An assessment of how each model predicts the data will be made using the likelihood ratio index. Sub group analysis will be performed to analyse the differences in parameter strengths between the three groups: offenders, family members of offenders and members of the general population.

WTP for an attribute will be defined as the ratio of the change in marginal utility of an attribute (attribute k in the equation) to marginal utility for the price attribute (p in the equation), as follows:

$$\text{WTP} = \frac{Change\ in\ X_k}{Change\ in\ X_p} = \frac{\frac{d}{dx_k}\beta_k x_k}{\frac{d}{dx_p}\beta_c x_p} = -\frac{\beta_k}{\beta_p}$$

An estimation of WTP for a treatment program that is described by the attributes in the DCE model will be calculated as the sum of marginal WTP for each attribute.

The Contingent Valuation method

Design

The CV method will also be used to solicit respondents' willingness to pay for a defined treatment program for impulsive violent offenders. Obtaining accurate WTP estimates using CV method requires detailed descriptions of the treatment being valued. This is evident from the name of the method, which produces values, contingent upon, the description of treatment. A description of the REINVESt study treatment program will be provided as an exemplar of a treatment program for impulsive violent offenders.

The payment card will be used as the WTP elicitation question. Respondents will be presented with a range of bid amounts and asked to choose the maximum amount in the form of an additional tax levy that they are willing to pay to have a described treatment available to impulsive repeat violent offenders. This reflects real life by allowing individuals to 'shop around' for the value closest to their maximum WTP [45]. The dollar values used on the payment cards were also explored in the FGD qualitative interviews and in the pilot study.

The CV method has been widely criticised for bias in terms of the validity of its results. Therefore, care will be taken in the design and analysis to reduce any bias that may arise. This will include randomisation of positioning of the dollar values of the payment cards to reduce anchoring or starting point bias [46 47]. Furthermore, to reduce the point bias or range bias [48] one of the options in the WTP payment card values will be 'none of these amounts' and respondents will then be asked to state how much they would be willing to pay.

Data Collection

The results from the qualitative methods in phase I and the pilot survey described in the DCE method will be used to describe the treatment to be valued in the CV method. Through an additional question to the DCE survey, participants will be asked to state their WTP for a described intervention similar to REINVESt. The respondents and sample will therefore be the same as explained in the DCE study.

Data Analysis

Mean and/or median WTP values will be calculated. Logistic regression models will be used to identify the factors affecting both zero and positive WTP and to estimate the independent effects of demographic characteristics on the WTP for offender treatments. The outcome of the model will be specified as the probability of agreeing to pay for offender treatment. The model fit will be estimated using the maximum likelihood function.

Patient and Public involvement

This protocol is about a study that seeks to assess offender and public preferences and therefore greatly involves the two groups. Phase 1 of this study involves the eliciting of offender and general public preferences through focus group discussions and the Delphi method. The offenders were voluntarily recruited through REINVESt, a study by the Justice Health Program at Kirby Institute UNSW. Phase II is a quantitative general population survey that will quantify the strength of preferences and assess the value of the treatment program. Participants for the survey will be representative of the NSW population and will be voluntarily recruited through a marketing survey company.

ETHICS AND DISSEMINATION

Ethics approval for this study has been provided for the two phases. Phase 1 ethics approval has been provided by UNSW – Higher Risk Ethics Committee, NSW Corrective Services Ethics Committee, and Aboriginal Health and Medical Research Council (AH&MRC) ethics committee. Phase II ethics approval has been provided by UNSW – higher risk ethics committee for the DCE general population sample. If in future the DCE is to be conducted with a sample from offenders, further ethics applications will be made to NSW corrective services and AH&MRC ethics committees.

The findings of this study will be made available on the Kirby Institute UNSW website, published in peer reviewed journals and presented at national and international conferences.

IMPORTANCE OF THIS PAPER

This research will provide a significant contribution to the assessment and evaluation of offender programs. In the DCE, an understanding of the trade-offs made and the strengths of preferences of society in the provision of healthcare for violent offenders will help provide valuable information for policy makers, treatment providers and other practitioners in designing treatment options.

Eliciting societal willingness to pay for offender treatment programs will be used to assess the value/benefit of the programs to both offenders and the public. When deciding whether to fund an intervention, policy makers need to consider how much the public values the benefits - hence how much they would be willing to pay. If the costs of interventions similar to REINVESt are known, the results (benefit values) of this study can be used in cost-benefit analyses.

The average WTP obtained using the DCE method can be compared with the average WTP obtained using the payment card CV method [49]. This can allow for testing of convergent validity of the two WTP methods i.e. the degree to which the results of the two methods are related.

This paper outlines a rigorous methodological approach that can be used to assess societal preferences and generalised for use in other DCE and CV studies of societal value of offender treatment programs as opposed to the traditional methods of opinion polls, which often only emphasise punitiveness of the public towards offenders, especially those who commit violent offences. To test external validity, we will use convergent validity to compare the results from the DCE to those of the CV method.

We outline a mixed methods process that involves qualitative methods, consensus approaches and economic methods of preference setting. We also provide a study context where the methods are applied: the REINVESt study. The rich qualitative component of this study will contribute to the literature concerned with the development of attributes for DCEs.

REFERENCES

1. Allen R. Global prison trends 2015. *Penal Reform International* 2015
2. Rollings K. Counting the costs of crime in Australia: a 2005 update: Australian Institute of Criminology Canberra, ACT, 2008.
3. Weatherburn D, Froyland G, Moffatt S, et al. Prison populations and correctional outlays: The effect of reducing re-imprisonment. *BOCSAR NSW Crime and Justice Bulletins* 2009:12.
4. Kockler TR, Stanford MS, Nelson CE, et al. Characterizing aggressive behavior in a forensic population. *American Journal of Orthopsychiatry* 2006;76(1):80.
5. Stanford MS, Houston RJ, Mathias CW, et al. Characterizing aggressive behavior. *Assessment* 2003;10(2):183-90.
6. Heilbrun AB. Psychopathy and violent crime. *Journal of Consulting and Clinical Psychology* 1979;47(3):509.
7. Cullen FT, Fisher BS, Applegate BK. Public opinion about punishment and corrections. *Crime and justice* 2000:1-79.
8. Hutton N. Beyond populist punitiveness? *Punishment & Society* 2005;7(3):243-58. doi: 10.1177/1462474505053828
9. Ryan M, Scott D, Reeves C, et al. Eliciting public preferences for healthcare: a systematic review of techniques, 2001.
10. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ : British Medical Journal* 2000;320(7248):1530-33.
11. Ryan M, Gerard K, Amaya-Amaya M. Using discrete choice experiments to value health and health care: Springer Science & Business Media 2007.
12. Lancsar E, Louviere J. Conducting Discrete Choice Experiments to Inform Healthcare Decision Making. *PharmacoEconomics* 2008;26(8):661-77. doi: 10.2165/00019053-200826080-00004
13. Klose T. The contingent valuation method in health care. *Health policy* 1999;47(2):97-123.
14. Diener A, O'brien B, Gafni A. Health care contingent valuation studies: a review and classification of the literature. *Health economics* 1998;7(4):313-26.
15. Butler T, Schofield PW, Greenberg D, et al. Reducing impulsivity in repeat violent offenders: an open label trial of a selective serotonin reuptake inhibitor. *Australian and New Zealand Journal of Psychiatry* 2010;44(12):1137-43. doi: 10.3109/00048674.2010.525216
16. Cleary PD. The increasing importance of patient surveys : Now that sound methods exist, patient surveys can facilitate improvement. *BMJ : British Medical Journal* 1999;319(7212):720-21.
17. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ* 2012;21(2):145-72. doi: 10.1002/hec.1697
18. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2003;2(1):55-64.
19. Clark M, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. *Pharmacoeconomics* 2014;32(9):883-902. doi: 10.1007/s40273-014-0170-x
20. Shanahan M, Gerard K, Ritter A. Preferences for policy options for cannabis in an Australian general population: A discrete choice experiment. *International Journal of Drug Policy* 2014;25(4):682-90. doi: http://dx.doi.org/10.1016/j.drugpo.2014.03.005
21. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Economics* 2012;21(6):730-41. doi: 10.1002/hec.1739
22. Barratt E, Stanford M. Impulsiveness, in the personality characteristics of the Personality Disordered client. Edited by Costello, CG New York: Wiley, 1995.
23. Mullen PM. Public involvement in health care priority setting: an overview of methods for eliciting values. *Health Expectations* 1999;2(4):222-34. doi: 10.1046/j.1369-6513.1999.00062.x

24. Bowling A, Jacobson B, Southgate L. Explorations in consultation of the public and health professionals on priority setting in an inner London health district. *Social science & medicine* 1993;37(7):851-57.
25. Furnham A, Meader N, McClelland A. Factors affecting nonmedical participants' allocation of scarce medical resources. *Journal of Social Behavior and Personality* 1998;13(4):735.
26. Rosko MD, McKenna W. Modeling consumer choices of health plans: a comparison of two techniques. *Social Science & Medicine* 1983;17(7):421-29.
27. Alexander J, Kroposki M. Outcomes for community health nursing practice. *J Nurs Adm* 1999;29(5):49-56. doi: 10.1097/00005110-199905000-00007
28. Green B, Jones M, Hughes D, et al. Applying the Delphi technique in a study of GPs' information requirements. *Health Soc Care Community* 1999;7(3):198-205. doi: 10.1046/j.1365-2524.1999.00176.x
29. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32(4):1008-15. [published Online First: 2000/11/30]
30. Dachary-Bernard J, Rambonilaza T. Choice experiment, multiple programmes contingent valuation and landscape preferences: How can we support the land use decision making process? *Land Use Policy* 2012;29(4):846-54.
31. Jairath N, Weinstein J. The Delphi methodology (Part one): A useful administrative approach. *Canadian journal of nursing administration* 1994;7(3):29-42.
32. Bolger F, Wright G. Improving the Delphi process: Lessons from social psychological research. *Technological Forecasting and Social Change* 2011;78(9):1500-13. doi: http://dx.doi.org/10.1016/j.techfore.2011.07.007
33. Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing* 2006;53(2):205-12. doi: 10.1111/j.1365-2648.2006.03716.x
34. NGENE software, version: 1.1.2 [program]: ChoiceMetrics, 2014.
35. Johnson FR, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health* 2013;16(1):3-13.
36. Bliemer M, Rose JM, Chorus CG. Detecting dominancy and accounting for scale differences when using stated choice data to estimate logit models. 2015
37. Kanninen BJ. Optimal design for multinomial choice experiments. *Journal of Marketing Research* 2002;39(2):214-27.
38. Bliemer MCJ, Rose JM, Hess S. Approximation of bayesian efficiency in experimental choice designs. *Journal of Choice Modelling* 2008;1(1):98-126. doi: https://doi.org/10.1016/S1755-5345(13)70024-1
39. de Bekker-Grob EW, Donkers B, Jonker MF, et al. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. *The Patient-Patient-Centered Outcomes Research* 2015;8(5):373-84.
40. Johnson R, Orme B. Sample size issues for conjoint analysis. *Getting started with conjoint analysis: strategies for product design and pricing research Madison: Research Publishers LLC* 2010:57-66.
41. Rose JM, Bliemer MCJ. Sample size requirements for stated choice experiments. *Transportation* 2013;40(5):1021-41. doi: 10.1007/s11116-013-9451-z
42. McFadden D. Conditional logit analysis of qualitative choice behavior. 1973
43. Greene W. NLOGIT version 6.0: EconometricsSoftware, Inc., 2015.
44. Hensher DA, Greene WH. The mixed logit model: the state of practice. *Transportation* 2003;30(2):133-76.
45. Donaldson C, Thomas R, Torgerson DJ. Validity of open-ended and payment scale approaches to eliciting willingness to pay. *Applied Economics* 1997;29(1):79-84.

46. Cummings RG, Harrison GW, Osborne LL. Can the bias of contingent valuation be reduced? Evidence from the laboratory. *Economics Working Paper B-95* 1995;3
47. Boyle KJ. Contingent Valuation in Practice. A Primer on Nonmarket Valuation: Springer 2017:83-131.
48. Heinzen RR, Bridges JF. Comparison of four contingent valuation methods to estimate the economic value of a pneumococcal vaccine in Bangladesh. *International journal of technology assessment in health care* 2008;24(4):481-7. doi: 10.1017/s026646230808063x [published Online First: 2008/10/03]
49. Ryan M, Watson V. Comparing welfare estimates from payment card contingent valuation and discrete choice experiments. *Health Economics* 2009;18(4):389-401. doi: 10.1002/hec.1364

Figure legends

Figure 1: The mixed methods design of the Discrete Choice Experiment
The mixed methods design is in two phases: Phase I and II. At the time of writing this protocol, only phase I has been completed.

Figure 2: The Delphi method used to refine the attributes for the impulsive violent offender DCE
Three iterative rounds of the Delphi method process were used to refine the attributes and attribute levels obtained in the qualitative research.

Figure 3: An example of a choice set for the DCE
This is an example of a choice set for the DCE that will be the result of scenarios generated using experimental design. Attribute examples are the characteristics of treatment and attribute levels are the ranges for each characteristic shown under treatment 1 and treatment 2.

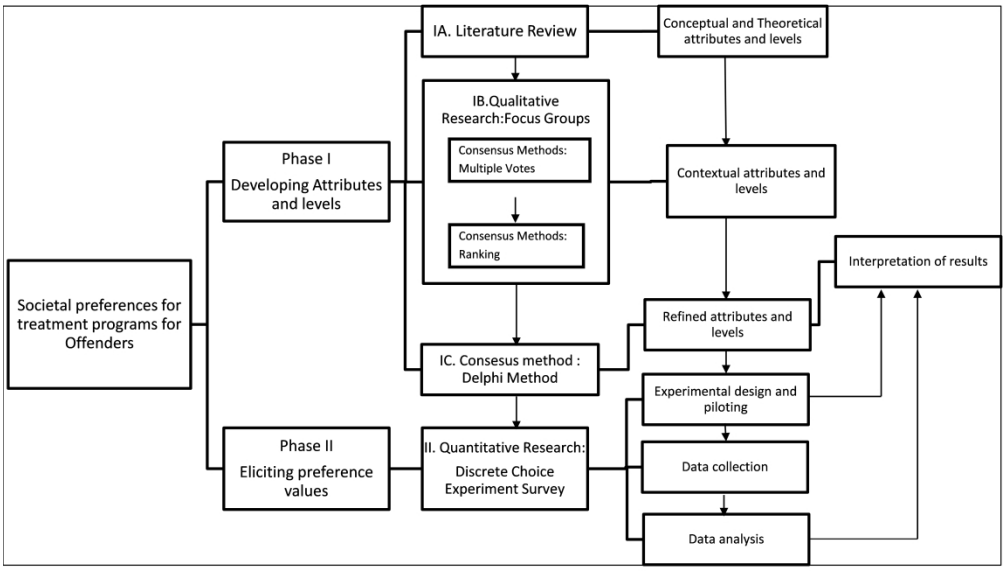

Figure 1: The mixed methods design of the Discrete Choice Experiment

494x296mm (300 x 300 DPI)

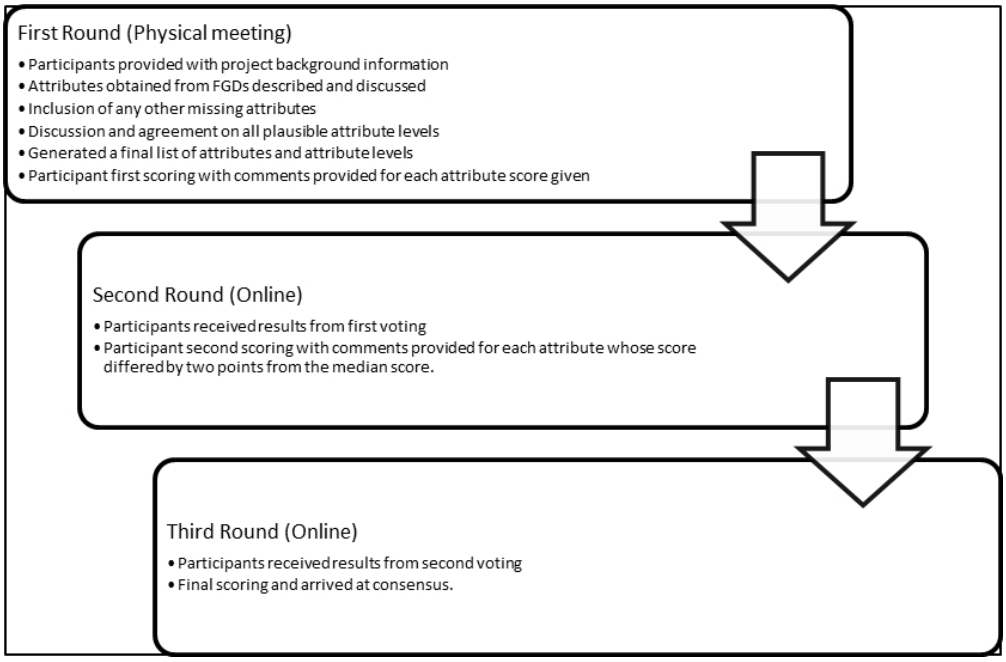

Figure 2: The Delphi method used to refine the attributes for the impulsive violent offender DCE

70x46mm (300 x 300 DPI)

Characteristic of treatment	Treatment 1	Treatment 2	No treatment
Effectiveness of the treatment	30% reduction in crime	50% reduction in crime	_
Treatment of co-occurring health conditions/addictions	Minimal treatment provided in program	Full treatment of all co-occurring morbidities both within program and at referral facilities	_
Type of treatment	Offender group counselling sessions only	Individual and Family counselling with Medication	_
Treatment provider	Prison/Probation & parole officers with Counsellors/Psychologists	Prison/Probation & Parole officer with Counsellors/Psychologists with Health Professional	_
Flexibility of appointments	Not flexible	Flexible	_
Compulsory/Voluntary participation	Voluntary	Compulsory	_
Cost per tax payer per year	$50	$75	_
Which treatment would you prefer to be given to Impulsive violent offenders?	○	○	○
If you had to choose between treatment 1 and treatment 2, which one would you prefer?	○	○	

Figure 3: An example of a choice set for the DCE

321x178mm (300 x 300 DPI)

BMJ Open

Assessing societal and offender perspectives on the value of offender health care: A stated preference research protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024899.R2
Article Type:	Protocol
Date Submitted by the Author:	13-Feb-2019
Complete List of Authors:	Settumba, Stella; Kirby Institute, University of New South Wales, Shanahan , Marian Shanahan ; The University of New South Wales, National Drug and Alcohol Research Centre Chambers, Georgina; University of New South Wales, Center for Big Data Research in Health Schofield, PW; University of Newcastle Butler, Tony; University of New South Wales, Kirby Institute
Primary Subject Heading:	Health economics
Secondary Subject Heading:	Health economics
Keywords:	Discrete choice experiment, Contingent valuation, Delphi method, Offender rehabilitation, Violence, Impulsivity

TITLE PAGE

Title: Assessing societal and offender perspectives on the value of offender health care: A stated preference research protocol

Stella Nalukwago Settumba[1], Marian Shanahan[2], Georgina M. Chambers[3], Peter Schofield[4], Tony Butler[1]

1. Justice Health Program, Kirby Institute, University of New South Wales, Sydney, Australia
2. National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Australia
3. Centre for Big Data Research in Health, University of New South Wales, Sydney, Australia
4. School of Medicine and Public Health, University of Newcastle, Newcastle, Australia

Corresponding Author:

Stella Nalukwago Settumba

Justice Health Program, The Kirby Institute, UNSW

Wallace Wurth Building, High Street, Kensington

NSW, 2052

Sydney, Australia

Email address: snalukwago@kirby.unsw.edu.au

Telephone number: +61 (0) 29385 9007

Manuscript word count: 4338

Keywords: Discrete choice experiment, Contingent valuation, Delphi method, Offender rehabilitation, Impulsivity, Violence.

ABSTRACT

Introduction:
The increasing burden that offenders place on justice and health budgets necessitates better methods to determine the benefits of and value society places on offender programs to guide policy regarding resource allocation. The aim of this paper is to demonstrate how economic methods will be used to determine the strength of preferences and value of violent offender treatment programs from the perspectives of offenders, their families and the general population.

Methods and analysis:
Two stated preference economic methods, discrete choice experiment (DCE) and contingent valuation (CV), will be used to assess society's and offenders' value of treatment programs. The mixed methods process involves a literature review and qualitative methods to derive attributes and levels for the DCE and payment card values for the CV. Consensus building approaches of voting, ranking and the Delphi method will be used to further refine the findings from the qualitative phase. Attributes and their levels will be used in a D-efficient Bayesian experimental design to derive choice scenarios for the development of a questionnaire that will also include CV questions. Finally, quantitative surveys to assess societal preferences and value in terms of willingness to pay will be conducted.

Ethics and dissemination:
Ethics approval from this study was obtained from the University of New South Wales (UNSW) Human Research Ethics Committee, Corrective Services New South Wales Ethics Committee and Aboriginal Health and Medical Research Council ethics committee. The findings will be made available on the Kirby Institute UNSW website, published in peer reviewed journals and presented at national and international conferences. This study was funded by grants from the National Health and Medical Research Council, under the Centre of Research Excellence in Offender Health Australia [grant number RG124596]. It is part of the research done by the Justice and Health program, Kirby Institute.

Article Summary

Strengths and limitations of the study

- This study is the first to quantify societal and offender preferences for violent offender treatment and provides a rigorous mixed methodological approach that can be generalised for use in other discrete choice experiment (DCE) and contingent valuation studies of valuation of offender programs.
- The results from these studies will be used in valuing the strength of preferences of society and offenders for treatment programs to reduce reoffending.
- The study will provide an estimate of the value, in terms of willingness to pay, that society and offenders place on violent offender treatments.
- This study will also provide the basis for conducting cost-benefit analysis to indicate the relative 'value for money' for violent offender programs.
- Recruitment of violent offenders to participant in a study is often challenging and while we hope to have a large enough representative sample in the future to participate in a DCE with offenders only, in this study recruitment will be done among the general population. Questions that ask participants to self-identify as offenders and family members of offenders will be included in the survey and sub-analysis done if a large enough sample is obtained.

Funding

This study was funded by grants from the National Health and Medical Research Council, under the Centre of Research Excellence in Offender Health Australia [grant number RG124596]. It is part of the research done by the Justice and Health program, Kirby Institute.

Competing Interests

There are no competing interests for any other regarding this study.

Author contributorship statement

Stella Settumba is involved in the study design, participant recruitment for the FGDs, data collection, analysis and write-up; Georgina Chambers and Marian Shanahan are involved in the study design, data analysis and write-up, Tony Butler and Peter Schofield are involved in the study design, participant recruitment for the FGDs, data analysis, and write-up. All authors have provided approval for the publication of this work and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

INTRODUCTION

Violence is a leading public health problem. It is estimated that more than 1.3 million people worldwide die each year as a result of violence accounting for 2.5% of global mortality [1]. The costs of violence to the health system in Australia, including medical costs, lost productivity, and intangible costs, are high and estimated at $AUD3.1 billion each year [2]. Imprisoning people with minimal rehabilitation has been shown to be largely ineffective as a deterrent to offending [3]. Well designed and evaluated interventions to reduce violence can save both lives and money. Research has shown that most violent crime can be classified as impulsive rather than pre-meditated and that impulsive offenders have a higher likelihood of recidivism than those offenders who commit pre-meditated crimes. Furthermore, impulsive offenders are more likely to respond positively to treatment and rehabilitation programs [4-6].

Decisions to allocate scarce resources to treat offenders, especially violent offenders, are seen by some as controversial even when the benefits of treatment extend beyond the offenders [7]. Public opinion and perception are often important determinants of the treatment and rehabilitation opportunities afforded to offenders since the justice system is financed through taxation, and politicians and other policy makers are wary of incurring the public's disapproval. However, surveys to determine the public's attitudes towards offender rehabilitation often suffer from poor methodology with poorly informed participants who lack accurate information on crime, its causes, and rehabilitation options and consequently are rarely given the opportunity to look beyond punitiveness [7 8]. Most surveys rely on snap polls with simple questions. Recommended techniques in the literature [9] for eliciting more considered and informed views from the public include: qualitative techniques such as one-to-one interviews, the Delphi technique, focus groups, citizens' juries, consensus panels and nominal group techniques; and quantitative techniques such as ranking, rating, discrete choice experiments (DCEs) and contingent valuation (CV) studies.

A variety of quantitative economic methods, including stated preferences and revealed preferences methods, have been employed to elicit patients' value for healthcare by quantifying their preferences [10]. Revealed preference methods refer to situations where people's choices are observed in actual market situations. However, in the absence of an actual market, as often found with many health programs or new interventions, stated preference techniques can be used. Stated preference methods refer to situations in which choices are made in a hypothetical market situation using a survey context. Valuation techniques using stated preference methods include the DCE [11 12] and CV method [13 14].

In one Australian state, New South Wales (NSW) after a successful pilot [15], a large randomised control trial (RCT) is underway, seeking to evaluate the effectiveness of a class of antidepressants, Selective Serotonin Reuptake Inhibitor (SSRI) (sertraline), to reduce impulsivity in men with a history of violent offending. This pharmacotherapy-based double blinded RCT is known as REINVESt ('Reducing Impulsivity in Repeat Violence Offenders Using a Selective Serotonin Reuptake Inhibitor'). Men who consent, are medically fit, have committed two or more violent offences and score highly on an impulsivity screener are randomised to receive either the SSRI or placebo for 6-12 months. If the intervention is found effective, valuation of its benefits is needed to advocate for the uptake of such treatment programs.

Using the REINVESt study as an exemplar, this paper demonstrates how economic methods will be used to assess the societal and offenders' value of treatment programs for offenders. The following are the aims of the economic study:

1. To elicit societal and offenders' preferences for treatment of impulsive violent offenders. Specifically,
    a) To assess the characteristics of treatment programs for impulsive violent offenders that could influence the uptake by offenders and support by society.
    b) To quantify the strength of preferences for and assess trade-offs between characteristics of treatment programs for impulsive violent offenders.
2. To elicit societal and offenders' value, in terms of willingness to pay (WTP), of treatment of impulsive violent offenders. Specifically,
    a) To estimate the societal and offenders' average WTP for the treatment of impulsive violent offenders using an SSRI.
    b) To elicit the factors affecting societal and offenders WTP for offender treatment programs.

The study protocol described in this paper details the methods used in assessing offenders' and society's preference and value (stated as WTP) of violent offender treatment programs using the DCE and CV stated preference methods. To our knowledge, this will be the first study to assess both offenders' and societal preferences for offender treatment programs. Although the involvement of patients in preference measures for decision making has been advocated [16] no DCEs have been performed involving offenders. This study will be conducted between July 2017 and April 2019.

METHODS

Aims 1(a) and 1(b) will be achieved through the qualitative (Phase 1) and quantitative (Phase II) components of the DCE respectively and objective 2 through the CV method (Phase II). The next paragraphs describe these methods.

The DCE

In a DCE, respondents' preferences are elicited based on their stated preference when faced with hypothetical choices between treatment scenarios that differ in terms of specified attributes and attribute levels. DCEs have been increasingly used in health economics to address a wide range of health policy related decisions (see [17-19] for more details on DCE methods). More recently, DCEs have been used in the justice area, for example, to explore societal preferences for alternative cannabis drug policies and to demonstrate the effect of varying cannabis policy characteristics and wider social consequences such as healthcare and criminal justice expenditures [20].

In this DCE study, participants will be asked to indicate their preference between two treatment programs for impulsive repeat-violent offenders, and a no treatment option. They will be presented with different choice scenarios comprised of differing characteristics of the treatment program (attributes) and attribute-levels. The results will be used to quantify the strengths of the preferences and assess the trade-offs between characteristics of treatment programs for impulsive violent offenders. By adding 'price' as an attribute in the DCE, the average societal WTP for a program can be estimated. A mixed methods design will be used (Figure 1). The DCE has four main steps (1)

identifying attributes and levels, (2) the experimental design, (3) the data collection survey, and (4) the analysis and interpretation of results [12]. In this study, step 1 was carried out in phase I and steps 2 to 4 will be in phase II. At the time of writing this protocol paper, phase I has been completed.

Developing attributes and levels

Phase 1 involved identifying all relevant attributes and assigning their associated levels. The generation of attributes used in DCEs is often poorly performed and reported and the need for rigorous research involving theoretical, conceptual, contextual and refined attributes has been emphasised [21]. In this study attributes were generated through a review of literature and primary qualitative research methods, Focus Group Discussions (FGDs). The consensus methods used to refine and narrow the attributes to a sizeable number included: ranking attributes, voting, and the Delphi method. The attributes characterise the factors considered by offenders and society to be most valuable for acceptance, support and uptake of treatment programs by impulsive violent offenders and the levels are the ranges over which the attributes vary. An example of a common attribute used in DCE studies is 'cost of the treatment program'. The attribute levels would be the various dollar amounts the treatment might cost.

Literature Review (1A)

A literature review developed the conceptual and theoretical attributes and levels which were further examined in FGDs. In this study, the literature searches considered existing theories that define an effective offender treatment program, positive and negative experiences of offenders with treatment programs, and views held by society regarding offender treatment programs.

Focus group discussions (IB)

Following the literature review, the next step involved developing contextual attributes and their levels using qualitative research, FGDs. For the FGDs we had three different participant group types: offenders, family members of offenders and people from the general public. We hypothesise that the attributes or attribute levels for treatment programs preferred will be different for the various groups. The FGDs were facilitated by an experienced qualitative researcher with vast experience working with offender populations.

Participants recruited into the REINVESt study were invited to participate in the offender FGDs and fell into two categories: (1) current or past participants, and (2) those who were eligible for the study according to the selection criteria but declined to participate in REINVESt. Common reasons cited for non-participation in REINVESt included not wanting to take medication and current use of a psychotropic medication. All individuals invited to the offenders FGDs were: male, over 18 years of age, had a history of committing at least two violent offences, and a score over 70 on the Barratt Impulsiveness Scale [22] indicating impulsive personality traits. Purposive selection was made to ensure a balance in terms of age, and number of prior convictions. For all offenders selected to participate in the FGDs, a member of the REINVESt study team asked for their consent to be contacted for this study at one of the routine contact visits and those showing an interest were contacted by a letter requesting them to participate.

In addition, each participating offender in the REINVESt trial study was asked when they attended a routine study follow-up visit if they were happy for a family member to take part in a FGD. The REINVESt study team has good working relationships with some family members of offenders. With the offenders' consent, a member of the REINVESt study team requested the family members'

consent to be contacted for this study. Those who agreed were sent an invitation letter to take part in the study including an email address and phone contact by which to contact the research team. Participants for the family members' FGD were defined as a partner or family member of a male offender participating in the REINVESt study and over 18 years of age.

Recruitment notices for the FGD with the general public were placed in libraries and community and online notice boards (e.g. Gumtree). Purposive sampling was done from those who responded, with an aim of having people with various ages, social and demographic backgrounds. Those selected were sent an invitation letter including an email address and phone number to contact the research team. Participants for the general public FGDs were required to be resident tax payers in NSW and over 18 years of age.

Recruitment and FGDs were conducted until saturation was reached, i.e. when no new data was generated with additional groups, bringing the total number of FGDs to 8 (4 offender, 3 general public and 1 family members group).

During the FGDs, after exploring participants' knowledge and views on violence, impulsivity, incarceration, recidivism and the role of treatment programs, they were provided with precise definitions of terms, examples of available interventions and contemporary statistics on violent crime, incarceration and recidivism rates. Participants then provided characteristics of treatment programs they might value if considering joining or supporting a treatment program. The levels reflected the range of situations that respondents might experience for each attribute. A semi-structured guide was used for the data collection.

After generating an exhaustive list of attributes, participants were asked to take part in a voting exercise [23] used to identify the top five characteristics generated within their FGD. Each participant was given unlimited votes and asked to vote 'yes' or 'no' if they thought a characteristic was important. 'Yes' votes were tallied for each characteristic and those with the top five most votes were noted as the top five attributes of preference for each group.

Participants then ranked the top 5 attributes in order of preference. Ranking exercises, as used in health priority setting, ask participants to give an ordinal rank to their preferences and those with the highest ranking are viewed as the most important [24-26]. The top attributes from the voting and ranking methods could now be included in the DCE. However, in this study, all attributes obtained from the FGDs were further assessed through the Delphi method and the results from the voting and ranking exercises used to provide an - indication from the FGDs of the strength of the different attributes that will then be compared with results from the DCE.

All FGDs were recorded and Digital audio data was transcribed and then destroyed. The transcribed data and the facilitator notes were coded and analysed using thematic analysis in *NVivo* to identify all major and minor themes on characteristics of treatment programs for impulsive violent offenders. The themes were subsequently grouped to classify the similarities and differences between the different groups of offenders, their families and the general public. These themes were then summarized to create a list of attributes and levels that were discussed during the Delphi method.

Delphi Method (IC)

The attributes and their levels obtained from the literature review and FGDs were further deliberated on by a team of experts using the Delphi method to generate a final list of attributes that will be used for the experimental design of the DCE. The Delphi research method is widely used in healthcare research to achieve consensus from a panel on issues of selected subjects [27 28]. It has also been recommended for use in deliberating on issues raised through FGDs and literature reviews [29] and for further refining of attributes and levels to be used in a DCE [30]. It is popular because, in addition to providing an opportunity for everyone's views to be taken into consideration by the group, it allows anonymous voting and avoids the domination of the consensus process by a few individuals[31]. Delphi, in contrast to other data gathering and analysis techniques, involves heterogeneous expertise, motivated and involved participants and employs multiple iterations/rounds in the form of feedback giving participants an opportunity to make informed decisions with good reasons for judgments or preferences [32]. Using iterative qualitative methods to refine attributes for a DCE also enables the rewriting of attributes to incorporate all relevant concepts [21]. Although there are no strict guidelines on the number of rounds needed to achieve consensus, the basic principle of the Delphi technique is to have as many rounds as are required or until the 'law of diminishing returns' occurs but generally at least two rounds are required [33]. Figure 2 describes the Delphi method process that was used in this study.

The aim of the Delphi process was:

1. To further refine the attributes that had been gathered from the FGDs.
2. To reach consensus on the levels for each attribute.
3. To arrive at a consensus of 5 – 8 attributes that would be evaluated in a DCE survey.

All participants, identified through research, academic and program implementation networks of people in the justice space, received an invitation e-mail, together with an information sheet explaining the study, the Delphi method, and an online informed consent form. Non-responders were approached by phone after one week. Before enrolment, it was confirmed that participants had the intention to complete all rounds of the study and had access to the internet. Participants for the Delphi method included criminologists, nurses from the justice health sector, psychologists working in criminal justice, health economists, forensic psychiatrists, members of the Australian Indigenous community, Corrective Services NSW staff, and police officers.

Experimental design and pilot

Scenarios will be constructed using the final attributes and levels ascertained from the Delphi method. A full factorial design takes on all possible combinations of attributes and their levels. Given the large number of attributes and multiple levels obtained from phase I, it is not feasible for respondents to assess all possible choices. An experimental design, which involves selecting through the use of statistical software (*NGENE* [34]) a subset of scenarios for respondents to complete, will be used to construct a fractional factorial design [35] . This helps to minimise the number of choice sets presented to respondents while still obtaining the maximum amount of information.

Attributes in this study will be described by a continuous, discrete or categorical scale. Effects coding will be used for all categorical attributes and parameters estimated for each level. The design will be

unlabelled, which means that the treatments in the scenarios will be generic and labelled as treatment 1 and treatment 2.

A D-efficient experimental design that maximises model statistical efficiency by minimising the parameter standard errors will be employed [36]. To optimise D-efficiency, prior assumptions on model parameter estimates will be used. A pilot study will be carried out to obtain priors and to guide development and testing of the questionnaire. This will include testing of the appropriateness of the questions such as determining the number of respondents willing to answer personal questions on exposure to violence, respondents' understanding and the correct balance between attributes and levels, task complexity, and timing of the length of response rates. Priors and their signs for the pilot will be based on data from the literature, or knowledge of known parameters [37]. Coefficients from the pilot testing will then be used as Bayesian priors for the Bayesian efficient experimental design [38], and the refined questionnaire will then be created. The design will be optimised for a multinomial logit model and this will then be evaluated in *NGENE* using a panel mixed logit model, which accounts for the parameter distribution, and a latent class model which accounts for non-uniformity of respondents.

Scenario presentation

Scenarios constructed from the experimental design will be presented to respondents in a survey delivered via a web-based questionnaire to elicit preferences. Respondents will be directed to read a description of all attributes prior to answering the questionnaire. Respondents will then be asked to choose between two treatment choice sets with different levels of attributes and a no treatment option. Those who chose the no treatment option will also be presented with a forced choice. The total number of choice sets per participant will be determined during the pilot and care will be taken to reduce cognitive burden. Generally 6-8 choice sets are recommended. Figure 3 is an example of a choice set.

Future work using the above methodology will involve conducting three separate DCEs, one with offenders, one with their families and one with members of the general public. Currently, the DCE will sample only from the general public. However, a question will be included in the survey to identify participants who are themselves offenders (having been accused of violence and having been in contact with the justice system for a violent offence) and family members of offenders. If an adequate number of participants self-identify as offenders and family members of offenders, sub-analyses for each group will be undertaken.

There is no agreed rule on the correct sample size required for a DCE [39]. However, research has shown that in all DCE studies with efficient designs, model estimate precision increases rapidly at sample sizes greater than 150 and then flattens out at around 300 observations [35]. It is also estimated that a minimum sample size of 200 respondents per sub-group be used for studies involving an analysis of differences between samples [40]. Furthermore, the s-efficiency measure in the experimental design in *NGENE* will estimate the required sample size for the study [41]. Recruitment, for the first DCE, will be from an outsourced online panel provider where respondent duplication and fraudulent completion of surveys is monitored. Participants are recruited via verified, certified sources and methods to create a large pool of potential research respondents. These participant panels have agreed and provided consent to participate in research conducted by the commercial survey company.

Data analysis and result interpretation

The data derived from the DCE surveys will be analysed to estimate attribute preference weights, also known as parameters, denoting the relative strength of each attribute in the choice of treatment programs for the offenders. The theoretic underpinning of the DCE analysis is based on Lancaster's theory of choice [42] and the random utility maximisation framework [43]. As shown in equation 1, the utility (*U*) that an individual n derives from the treatment alternative j in the choice set c is explained by an observed component V_{ncj} and an unobserved component ε_{ncj}.

$$U_{ncj} = V_{ncj} + \varepsilon_{ncj} \quad \text{(Equation 1)}$$

The observed component of the utility associated with alternative j, V_{ncj}, is a function of a vector of k attributes that describe treatment alternative t, x_{ncjk}, with associated preference weights, β, to be estimated. Such that:

$$V_{ncj} = \sum_{k=1}^{k} \beta_k x_{ncjk} \quad \text{(Equation 2)}$$

When faced with a choice task with treatment alternatives an individual will choose i over j if the utility obtained from i is greater than that from$j$. Such that:

$$(V_{nci} + \varepsilon_{nci}) > (V_{ncj} + \varepsilon_{ncj}) \quad \text{(Equation 3)}$$

The econometrics software *Nlogit* [44] will be used to perform the analysis. A multinomial logit model, a mixed logit model, or a latent class model will be estimated [45]. The final model will depend on which model best fits the data. An assessment of how each model predicts the data will be made using the likelihood ratio index. Sub group analysis will be performed to analyse the differences in parameter strengths between the three groups: offenders, family members of offenders and members of the general population.

WTP for an attribute will be defined as the ratio of the change in marginal utility of an attribute (attribute k in the equation) to marginal utility for the price attribute (p in the equation), as shown in equation 4:

$$\text{WTP} = \frac{Change\ in\ X_k}{Change\ in\ X_p} = \frac{\frac{d}{dx_k}\beta_k x_k}{\frac{d}{dx_p}\beta_c x_p} = -\frac{\beta_k}{\beta_p} \quad \text{(Equation 4)}$$

An estimation of WTP for a treatment program that is described by the attributes in the DCE model will be calculated as the sum of marginal WTP for each attribute.

The Contingent Valuation method

Design

The CV method will also be used to solicit respondents' willingness to pay for a defined treatment program for impulsive violent offenders. Obtaining accurate WTP estimates using CV method requires detailed descriptions of the treatment being valued. This is evident from the name of the method, which produces values, contingent upon, the description of treatment. A description of the REINVESt study treatment program will be provided as an exemplar of a treatment program for impulsive violent offenders.

The payment card will be used as the WTP elicitation question. Respondents will be presented with a range of bid amounts and asked to choose the maximum amount in the form of an additional tax levy that they are willing to pay to have a described treatment available to impulsive repeat violent offenders. This reflects real life by allowing individuals to 'shop around' for the value closest to their maximum WTP [46]. The dollar values used on the payment cards were also explored in the FGD qualitative interviews and in the pilot study.

The CV method has been widely criticised for bias in terms of the validity of its results. Therefore, care will be taken in the design and analysis to reduce any bias that may arise. This will include randomisation of positioning of the dollar values of the payment cards to reduce anchoring or starting point bias [47 48]. Furthermore, to reduce the point bias or range bias [49] one of the options in the WTP payment card values will be 'none of these amounts' and respondents will then be asked to state how much they would be willing to pay.

Data Collection

The results from the qualitative methods in phase I and the pilot survey described in the DCE method will be used to describe the treatment to be valued in the CV method. Through an additional question to the DCE survey, participants will be asked to state their WTP for a described intervention similar to REINVESt. The respondents and sample will therefore be the same as explained in the DCE study.

Data Analysis

Mean and/or median WTP values will be calculated. Logistic regression models will be used to identify the factors affecting both zero and positive WTP and to estimate the independent effects of demographic characteristics on the WTP for offender treatments. The outcome of the model will be specified as the probability of agreeing to pay for offender treatment. The model fit will be estimated using the maximum likelihood function.

Patient and Public involvement

This protocol is about a study that seeks to assess offender and public preferences and therefore greatly involves the two groups. Phase 1 of this study involves the eliciting of offender and general public preferences through focus group discussions and the Delphi method. The offenders were voluntarily recruited through REINVESt, a study by the Justice Health Program at Kirby Institute UNSW. Phase II is a quantitative general population survey that will quantify the strength of preferences and assess the value of the treatment program. Participants for the survey will be representative of the NSW population and will be voluntarily recruited through a marketing survey company.

ETHICS AND DISSEMINATION

Ethics approval for this study has been provided for the two phases. Phase 1 ethics approval has been provided by UNSW – Higher Risk Ethics Committee, NSW Corrective Services Ethics Committee, and Aboriginal Health and Medical Research Council (AH&MRC) ethics committee. Phase II ethics approval has been provided by UNSW – higher risk ethics committee for the DCE general population sample. If in future the DCE is to be conducted with a sample from offenders, further ethics applications will be made to NSW corrective services and AH&MRC ethics committees.

The findings of this study will be made available on the Kirby Institute UNSW website, published in peer reviewed journals and presented at national and international conferences.

IMPORTANCE OF THIS PAPER

This research will provide a significant contribution to the assessment and evaluation of offender programs. In the DCE, an understanding of the trade-offs made and the strengths of preferences of society in the provision of healthcare for violent offenders will help provide valuable information for policy makers, treatment providers and other practitioners in designing treatment options.

Eliciting societal willingness to pay for offender treatment programs will be used to assess the value/benefit of the programs to both offenders and the public. When deciding whether to fund an intervention, policy makers need to consider how much the public values the benefits - hence how much they would be willing to pay. If the costs of interventions similar to REINVESt are known, the results (benefit values) of this study can be used in cost-benefit analyses.

The average WTP obtained using the DCE method can be compared with the average WTP obtained using the payment card CV method [50]. This can allow for testing of convergent validity of the two WTP methods i.e. the degree to which the results of the two methods are related.

This paper outlines a rigorous methodological approach that can be used to assess societal preferences and generalised for use in other DCE and CV studies of societal value of offender treatment programs as opposed to the traditional methods of opinion polls, which often only emphasise punitiveness of the public towards offenders, especially those who commit violent offences. To test external validity, we will use convergent validity to compare the results from the DCE to those of the CV method.

We outline a mixed methods process that involves qualitative methods, consensus approaches and economic methods of preference setting. We also provide a study context where the methods are applied: the REINVESt study. The rich qualitative component of this study will contribute to the literature concerned with the development of attributes for DCEs.

REFERENCES

1. Allen R. Global prison trends 2015. *Penal Reform International* 2015
2. Rollings K. Counting the costs of crime in Australia: a 2005 update: Australian Institute of Criminology Canberra, ACT, 2008.
3. Weatherburn D, Froyland G, Moffatt S, et al. Prison populations and correctional outlays: The effect of reducing re-imprisonment. *BOCSAR NSW Crime and Justice Bulletins* 2009:12.
4. Kockler TR, Stanford MS, Nelson CE, et al. Characterizing aggressive behavior in a forensic population. *American Journal of Orthopsychiatry* 2006;76(1):80.
5. Stanford MS, Houston RJ, Mathias CW, et al. Characterizing aggressive behavior. *Assessment* 2003;10(2):183-90.
6. Heilbrun AB. Psychopathy and violent crime. *Journal of Consulting and Clinical Psychology* 1979;47(3):509.
7. Cullen FT, Fisher BS, Applegate BK. Public opinion about punishment and corrections. *Crime and justice* 2000:1-79.
8. Hutton N. Beyond populist punitiveness? *Punishment & Society* 2005;7(3):243-58. doi: 10.1177/1462474505053828
9. Ryan M, Scott D, Reeves C, et al. Eliciting public preferences for healthcare: a systematic review of techniques, 2001.
10. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ : British Medical Journal* 2000;320(7248):1530-33.
11. Ryan M, Gerard K, Amaya-Amaya M. Using discrete choice experiments to value health and health care: Springer Science & Business Media 2007.
12. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making. *Pharmacoeconomics* 2008;26(8):661-77.
13. Klose T. The contingent valuation method in health care. *Health policy* 1999;47(2):97-123.
14. Diener A, O'brien B, Gafni A. Health care contingent valuation studies: a review and classification of the literature. *Health economics* 1998;7(4):313-26.
15. Butler T, Schofield PW, Greenberg D, et al. Reducing impulsivity in repeat violent offenders: an open label trial of a selective serotonin reuptake inhibitor. *Australian and New Zealand Journal of Psychiatry* 2010;44(12):1137-43. doi: 10.3109/00048674.2010.525216
16. Cleary PD. The increasing importance of patient surveys : Now that sound methods exist, patient surveys can facilitate improvement. *BMJ : British Medical Journal* 1999;319(7212):720-21.
17. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. *Health Econ* 2012;21(2):145-72. doi: 10.1002/hec.1697
18. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2003;2(1):55-64.
19. Clark M, Determann D, Petrou S, et al. Discrete Choice Experiments in Health Economics: A Review of the Literature. *Pharmacoeconomics* 2014;32(9):883-902. doi: 10.1007/s40273-014-0170-x
20. Shanahan M, Gerard K, Ritter A. Preferences for policy options for cannabis in an Australian general population: A discrete choice experiment. *International Journal of Drug Policy* 2014;25(4):682-90. doi: http://dx.doi.org/10.1016/j.drugpo.2014.03.005
21. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Economics* 2012;21(6):730-41. doi: 10.1002/hec.1739
22. Barratt E, Stanford M. Impulsiveness, in the personality characteristics of the Personality Disordered client. Edited by Costello, CG New York: Wiley, 1995.
23. Mullen PM. Public involvement in health care priority setting: an overview of methods for eliciting values. *Health Expectations* 1999;2(4):222-34. doi: 10.1046/j.1369-6513.1999.00062.x

24. Bowling A, Jacobson B, Southgate L. Explorations in consultation of the public and health professionals on priority setting in an inner London health district. *Social science & medicine* 1993;37(7):851-57.
25. Furnham A, Meader N, McClelland A. Factors affecting nonmedical participants' allocation of scarce medical resources. *Journal of Social Behavior and Personality* 1998;13(4):735.
26. Rosko MD, McKenna W. Modeling consumer choices of health plans: a comparison of two techniques. *Social Science & Medicine* 1983;17(7):421-29.
27. Alexander J, Kroposki M. Outcomes for community health nursing practice. *J Nurs Adm* 1999;29(5):49-56. doi: 10.1097/00005110-199905000-00007
28. Green B, Jones M, Hughes D, et al. Applying the Delphi technique in a study of GPs' information requirements. *Health Soc Care Community* 1999;7(3):198-205. doi: 10.1046/j.1365-2524.1999.00176.x
29. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32(4):1008-15. [published Online First: 2000/11/30]
30. Dachary-Bernard J, Rambonilaza T. Choice experiment, multiple programmes contingent valuation and landscape preferences: How can we support the land use decision making process? *Land Use Policy* 2012;29(4):846-54.
31. Jairath N, Weinstein J. The Delphi methodology (Part one): A useful administrative approach. *Canadian journal of nursing administration* 1994;7(3):29-42.
32. Bolger F, Wright G. Improving the Delphi process: Lessons from social psychological research. *Technological Forecasting and Social Change* 2011;78(9):1500-13. doi: http://dx.doi.org/10.1016/j.techfore.2011.07.007
33. Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing* 2006;53(2):205-12. doi: 10.1111/j.1365-2648.2006.03716.x
34. NGENE software, version: 1.1.2 [program]: ChoiceMetrics, 2014.
35. Johnson FR, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health* 2013;16(1):3-13.
36. Bliemer M, Rose JM, Chorus CG. Detecting dominancy and accounting for scale differences when using stated choice data to estimate logit models. 2015
37. Kanninen BJ. Optimal design for multinomial choice experiments. *Journal of Marketing Research* 2002;39(2):214-27.
38. Bliemer MCJ, Rose JM, Hess S. Approximation of bayesian efficiency in experimental choice designs. *Journal of Choice Modelling* 2008;1(1):98-126. doi: https://doi.org/10.1016/S1755-5345(13)70024-1
39. de Bekker-Grob EW, Donkers B, Jonker MF, et al. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. *The Patient-Patient-Centered Outcomes Research* 2015;8(5):373-84.
40. Johnson R, Orme B. Sample size issues for conjoint analysis. *Getting started with conjoint analysis: strategies for product design and pricing research Madison: Research Publishers LLC* 2010:57-66.
41. Rose JM, Bliemer MCJ. Sample size requirements for stated choice experiments. *Transportation* 2013;40(5):1021-41. doi: 10.1007/s11116-013-9451-z
42. Lancaster KJ. A New Approach to Consumer Theory. *Journal of Political Economy* 1966;74(2):132-57.
43. McFadden D. Conditional logit analysis of qualitative choice behavior. 1973
44. Greene W. NLOGIT version 6.0: EconometricsSoftware, Inc., 2015.
45. Hensher DA, Greene WH. The mixed logit model: the state of practice. *Transportation* 2003;30(2):133-76.

46. Donaldson C, Thomas R, Torgerson DJ. Validity of open-ended and payment scale approaches to eliciting willingness to pay. *Applied Economics* 1997;29(1):79-84.
47. Cummings RG, Harrison GW, Osborne LL. Can the bias of contingent valuation be reduced? Evidence from the laboratory. *Economics Working Paper B-95* 1995;3
48. Boyle KJ. Contingent Valuation in Practice. A Primer on Nonmarket Valuation: Springer 2017:83-131.
49. Heinzen RR, Bridges JF. Comparison of four contingent valuation methods to estimate the economic value of a pneumococcal vaccine in Bangladesh. *International journal of technology assessment in health care* 2008;24(4):481-7. doi: 10.1017/s026646230808063x [published Online First: 2008/10/03]
50. Ryan M, Watson V. Comparing welfare estimates from payment card contingent valuation and discrete choice experiments. *Health Economics* 2009;18(4):389-401. doi: 10.1002/hec.1364

Figure legends

Figure 1: The mixed methods design of the Discrete Choice Experiment
The mixed methods design is in two phases: Phase I and II. At the time of writing this protocol, only phase I has been completed.

Figure 2: The Delphi method used to refine the attributes for the impulsive violent offender DCE
Three iterative rounds of the Delphi method process were used to refine the attributes and attribute levels obtained in the qualitative research.

Figure 3: An example of a choice set for the DCE
This is an example of a choice set for the DCE that will be the result of scenarios generated using experimental design. Attribute examples are the characteristics of treatment and attribute levels are the ranges for each characteristic shown under treatment 1 and treatment 2.

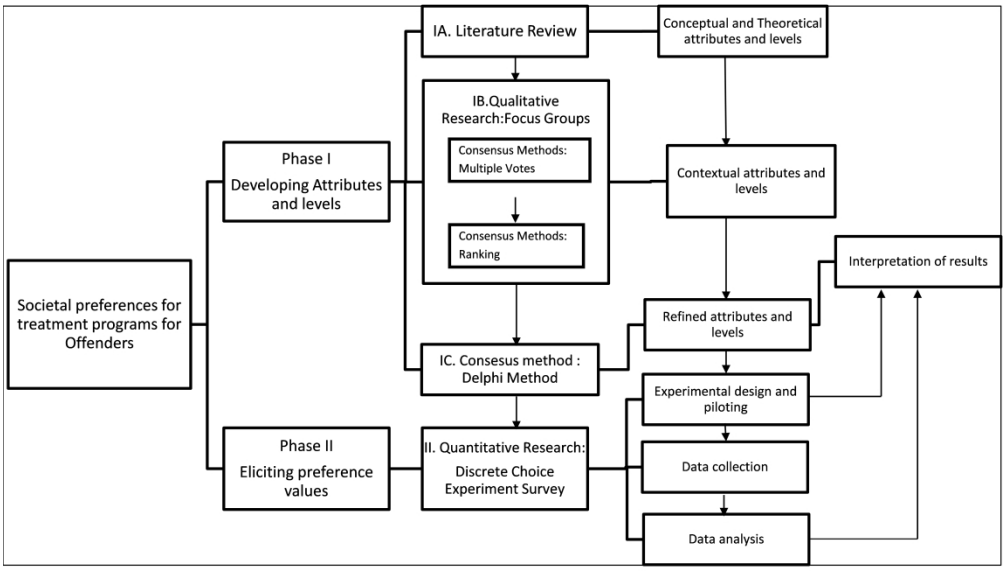

Figure 1: The mixed methods design of the Discrete Choice Experiment

494x296mm (300 x 300 DPI)

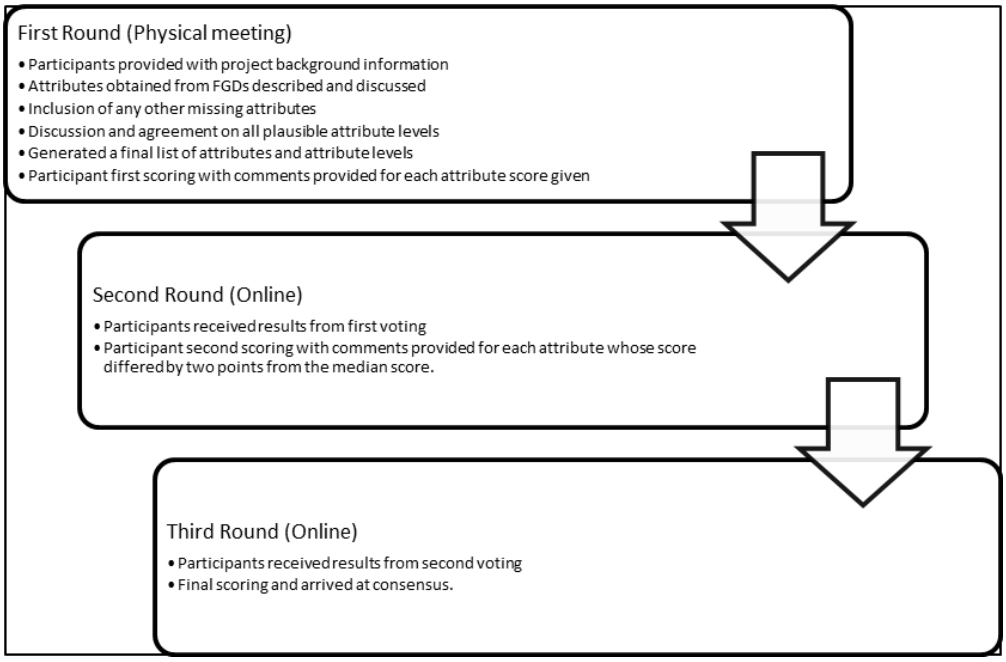

Figure 2: The Delphi method used to refine the attributes for the impulsive violent offender DCE

70x46mm (300 x 300 DPI)

Characteristic of treatment	Treatment 1	Treatment 2	No treatment
Effectiveness of the treatment	30% reduction in crime	50% reduction in crime	_
Treatment of co-occurring health conditions/addictions	Minimal treatment provided in program	Full treatment of all co-occurring morbidities both within program and at referral facilities	_
Type of treatment	Offender group counselling sessions only	Individual and Family counselling with Medication	_
Treatment provider	Prison/Probation & parole officers with Counsellors/Psychologists	Prison/Probation & Parole officer with Counsellors/Psychologists with Health Professional	_
Flexibility of appointments	Not flexible	Flexible	_
Compulsory/Voluntary participation	Voluntary	Compulsory	_
Cost per tax payer per year	$50	$75	_
Which treatment would you prefer to be given to Impulsive violent offenders?	○	○	○
If you had to choose between treatment 1 and treatment 2, which one would you prefer?	○	○	

Figure 3: An example of a choice set for the DCE

321x178mm (300 x 300 DPI)