University of Tasmania Open Access Repository

Cover sheet

The Promises and Perils of Developing a National Sex Offender Recidivism Database in Australia

Caroline Spiranovic[a]*, Anna Ferrante[b], Marie-Jeanne Buscot[c], Kate Griffiths[d], Alfred Allan[e], Stephen Wong[f], Hilde Tubex[g] and Frank Morgan[h]

[a]*Faculty of Law, College of Arts, Law & Education, University of Tasmania, Hobart, Australia, ORCiD http://orcid.org/0000-0002-5270-8719; [b] Centre for Data Linkage, School of Public Health, Curtin University, Perth, Australia, ORCiD https://orcid.org/0000-0002-9479-7908; [c]Faculty of Law, College of Arts, Law & Education, University of Tasmania, Hobart, Australia; [d] School of Population and Global Health, University of Western Australia, ORCiD https://orcid.org/0000-0003-1347-0119; [e] School of Arts and Humanities, Edith Cowan University, Joondalup, Australia, orcid.org/ 0000-0001-7039-797X; [f]Department of Psychology, University of Saskatchewan, Canada, Swinburne University of Technology, Melbourne, Australia; [g]Law School, University of Western Australia, Perth, Australia, [h]School of Population and Global Health, University of Western Australia and Law School, University of Tasmania.*

* Corresponding author - Caroline Spiranovic, Senior Lecturer in Law, Faculty of Law, College of Arts, Law and Education, University of Tasmania, Private Bag 89, Hobart, Tasmania, Australia 7001; caroline.spiranovic@utas.edu.au.

The Promises and Perils of Developing a National Sex Offender Recidivism Database in Australia

Much of what we know about sexual offenders and risk management is derived from empirical studies on sex offender populations in North America. In comparison to Canada and the U.S., the evidence base in Australia on sexual offender risk management is under-developed. In this paper, we describe a current research project tasked with developing a national sex offender recidivism database to advance the evidence base in Australia. It is argued that a national database would advance knowledge and practice in the field of sex offender risk management in Australia in a multitude of ways. Yet there are many obstacles and difficulties in developing such a database. After putting forward a case for the need for such a database, we outline the issues we have encountered and the approaches we have adopted to develop this database. It is intended that this contemporary comment may not only alert readers to this emerging data resource in Australia but also function as a road map to guide future empirical research on offender population databases in Australia.

Keywords: sex offenders; recidivism; risk management; evidence base; national database

Sex Offenders and Risk Management: The Evidence base

Australia is a federation and each of the nine state and territory jurisdictions therefore has its own legislation and criminal justice system. In the absence of any risk assessment tools specifically for sexual offenders in Australia (see Allan et al. 2019), practitioners, expert witnesses, parole boards and correctional service agencies in all these jurisdictions use a range of internationally and empirically supported tools when making decisions about Australian sex offenders' risk of recidivism. Table 1 provides a brief overview of some of the most used risk assessment tools in Australia specifically designed for sexual offenders.

Table 1. Risk Assessment Tools Specifically Designed for Sexual Offenders and Commonly Applied in Australia

Name of tool	Brief description of tool	Country of validation samples	Authors of tool
Static-99Revised	10 item actuarial tool, derived from the Static-99, to characterise adult male sexual offenders in terms of relative risk of sexual recidivism based on demographic and criminal history variables.	Predominantly United States and Canada but also England, Austria, Sweden.	Phenix, Helmus, and Hanson 2012
RSVP[1]	Structured Professional Judgment guidelines to assess adult sex offenders for risk of recidivism and to formulate risk management strategies. Comprises 22 items for assessing sexual violence history, clinical and other risk factors.	Limited validation studies thus far – guidelines developed based on systematic review of scientific literature	Hart et al. 2003
VRS-SO[2]	24-item rating scale, comprising 7 static and 17 dynamic items, to assess risk of sexual recidivism (including change in risk) as well as identify treatment targets for convicted sexual offenders.	Canada, New Zealand, United States, Germany	Wong et al. 2003
STABLE-2007	13-items measuring stable but dynamic/changeable factors (including interpersonal and contextual/situational factors) so that change in risk for sexual recidivism in adult male sex offenders may be assessed over time.	Canada, United States	Hanson et al. 2007
ACUTE-2007	7-items designed to measure highly transient states and situations associated with imminent risk in adult male sex offenders and to be assessed regularly.	Canada, United States	Hanson et al. 2007

Notes:
[1] RSVP = The Risk for Sexual Violence Protocol (Hart et al. 2003 but see also Hart and Boer 2010)
[2] VRS-SO = Violence Risk Scale – Sex Offender Version (Wong et al. 2003 but see also Olver et al. 2007)

These tools are typically used in combination with other tools, in order to form a well-rounded and comprehensive assessment of the risk of recidivism and the risk management needs pertaining to an offender.

However, these tools have not been developed for, and are rarely cross-validated on, Australian sex offenders. The application of these tools to make risk management decisions about Australian sex offenders raises a number of concerns, including cross-cultural validity and bias and procedural justice, particularly where Aboriginal and Torres Strait Islander offenders are subjected to assessment (e.g. see Allan et al. 2019; Day et al. 2018; Harris and McPhredran 2018; Shepherd and Lewis-Fernandez 2016). These concerns are shared by correctional agencies, policy makers, the judiciary and researchers which led to the research project described herein.

Part of the solution to ensuring that decisions made about sex offenders in Australia are valid and fair is to develop a robust evidence base that makes use of Australian research data. In Australia there is in fact already potentially rich and comprehensive data available comprising predominantly retrospective administrative data collected by correctional agencies on a broad range of static, dynamic, and situational offence variables predictive of recidivism. However, this administrative data was not necessarily collected for research purposes and there are various challenges in drawing this data together to build an Australian database sufficiently robust to support advancement of the Australian evidence base. In the following sections we outline a current research project that is developing a national sex offender recidivism database based on retrospectively collected agency data to examine the effectiveness of risk assessment tools for sex offenders. First, we examine the benefits envisaged from developing such a database, particularly if it were to be made available to a broader range of researchers. However, this would need to be accompanied by policies and

procedures agreed by State Corrections Agencies to ensure confidentiality and privacy safeguards regarding access. Next, we describe the nature and scope of the database we are developing. We then outline the issues we have encountered in constructing this database and the approaches we have employed to address these issues. Finally, we discuss lessons learned and a way forward.

Current Australian research project: Developing a National Sex Offender Recidivism Database

Funded by an Australian Research Council Linkage grant (LP140100275), a team of researchers from Australian universities and partnering correctional service agencies from all Australian states and territories have embarked on an ambitious research project exploring the validity of risk assessment tools for use with Australian sex offenders. A key project outcome is the development of a retrospective national sex offender database to be used in current and future recidivism research.

The national sex offender recidivism database supports the following research objectives:

- To assess the validity of the Static-99R (Phenix et al. 2012) and the Static-99 (Hanson and Thornton 1999; Harris et al. 2003), where data are provided[1], to

---

[1] The Static-99/R is the only risk assessment tool used consistently across all correctional service agencies in Australia and therefore the only tool from which we could collect data across all jurisdictions. In some jurisdictions, retrospectively collected data includes the Static-99, which is the predecessor to the Static-99R. The only item which differs between the two versions is the way in which the age variable is scored; Static-99 scores can be converted using available data, if desired, to Static-99R compliant scores.

predict recidivism in Australian sex offenders, particularly Aboriginal and Torres Strait Islander offenders.

- To derive Australian norms for the Static-99/R including absolute and relative risk estimates for sexual recidivism based on scores on the Static-99/R for the entire Australian sexual offender sample and for cohorts, where possible (e.g. Aboriginal and Torres Strait Islander, jurisdiction and offender type based on victim demographics).

- To establish the unique additive value of dynamic variables (e.g. post-release treatment and planning variables), over and above the contribution of static items from the Static-99/R items, in predicting risk of recidivism in Australian sex offenders.

In addition, the database will provide a rich source of Australian empirical data to explore additional lines of enquiry including the early offence trajectories of convicted sexual offenders to examine the link and cross-over between non-contact and contact crimes and the connection between non-sexual and sexual offending.

Nature and scope of the National Database

Correctional service agencies in each jurisdiction of Australia have contributed to this nationally constructed and retrospective sex offender recidivism database. The database comprises 161 data items and 6,399 male sex offenders. The 161 data items can be broken down into the following variable categories:

- Demographics (including age, education level, Aboriginal and/or Torres Strait Islander status);

- Prior criminal history (including first and subsequent court appearances, offence type based on ANZSOC code[2] and number of charges (or counts)). Dates for these priors were also included;

- Details of 'index' offence, that is, particulars for the sex offence on which a Static-99/R assessment was based, including offence type and date of commission, victim demographics, criminal justice system outcomes in terms of sentence type, sentence length and key dates for episodes of contact (where contact could comprise custodial sentence or community based order);

- Static-99/R (Hanson and Thornton 1999; Harris et al. 2003; Phenix et al. 2012) total and item scores for the index offence;

- Post-index offence treatment, supervision and release variables (i.e. type and length of supervision/treatment and release planning including postcodes/suburbs where released, employment and accommodation arrangements); and

- Post-index recidivism records were further disaggregated into sex, violent and other types of reoffending including dates of re-offending and subsequent court appearances and outcomes, ANZSOC codes for associated offences, and the location where sexual re-offences were committed.

---

[2] ANZSOC = Australian and New Zealand Standard Offence Classification, Third Edition (Pink 2011)

National Database: Data Issues and Approaches

Notwithstanding the nature of the data we have used, including the fact that this retrospective data was collected primarily for administrative purposes, the categories identified above have presented various roadblocks in our mission to establish a robust national database. In general, demographics, prior criminal history, and Static-99/R data have presented no issues and were provided in a comprehensive manner. On the other hand, the index offence and post-index recidivism records have proven more challenging to work with. The data issues we have experienced can be categorised into one or more of the following types:

(1) Inconsistencies in data entry

(2) Missing items

(3) Duplicate records

Here we discuss the implications of the above issues for some key questions in this research and share solutions we have so far employed. Above all else though, these issues impact on our ability to establish appropriate "time at risk" periods for recidivism follow-up. For 3,235 of 6,399 sexual offender records, we were required to develop and implement 'business rules' to obtain estimates of follow-up for recidivism for each individual offender. However, solutions to the data issues described are still being actively pursued at the time of writing this paper.

1. Inconsistencies in data entry

Inconsistencies in data entry within -and between- jurisdictions were identified across some key variables. Inconsistencies could be categorised primarily as:

(a) Unexpected data codes or out of range values. For instance, instead of coding 'episode type' as a binary variable – imprisonment versus 'other'

- some jurisdictions provided unspecified codes to delineate between 'other' outcomes,

(b) Use of different conventions to report variables. For instance, inconsistencies were observed in the entry of "time periods" (notably length of head sentence or actual time served) reported in days, months, years or an undefined mix of these, and

(c) implausible values when comparing different date variables. For example, in some instances, the reported date for end of data collection pre-dated either the court finalisation date and/or the prison episode end date for the index offence (on which the Static-99/R assessment and recidivism data were based).

2. Missing items

Each state and territory agreed to provide complete data where possible on all 161 variables for a smaller subset of their offender sample, as time and resources did not permit for data to be collected on all 161 variables for the full sample (N = 6,399). In some cases, data collection required manual data coding using paper-based files. Due to these considerations, larger jurisdictions sought to provide complete data for 400 sex offenders (evenly split between Aboriginal and Torres Strait Islander and non-Aboriginal and Torres Strait Islander), prioritising those with longer follow-up times. For smaller jurisdictions, the undertaking was to provide as much data as possible on the sample supplied. As such, missing items were anticipated from the outset.

For the purposes though of sharing lessons learnt from working with this data, we outline as follows the reasons, as some were not anticipated, that items appeared as missing in the nationwide database:

(a) data from some jurisdictions were truly missing for the prescribed variables,

(b) the data were provided but appeared as "missing" when collating the nationwide dataset due to errors at the data collection stage (i.e. data entry or coding errors such as misspelling of the variable labels, or misalignment of variables across columns of the dataset),

(c) unspecified and inconsistent codes were used to report missingness, and/or

(d) there were errors in the database compiling stage (i.e. automated scripts for data cleaning/formatting were applied to uncomplying and/or incorrect variables for a given jurisdiction) when merging data from different jurisdictions.

To distinguish between truly missing data and missing data caused by miscoding and other errors, the researchers conducted a data review comprising screening and manually checking of some of the raw data provided by each jurisdiction. Due to the breadth and depth of the data (i.e. 6,399 individual offender records and 161 unique variables) this was an unanticipated challenge to our progress.

3. Duplicate records

Some of the jurisdictions appeared to provide more than one record (or in other words more than one row of data) for a given offender. In some cases, because each offender was allocated a truly unique identifier, it was relatively easy to identify and consolidate their 'duplicate' records into a single offender record for the nationwide dataset. In some other instances, however, duplicates arose because data had been sourced from different agencies (within the same jurisdiction) and, without the

existence of a shared unique identifier, it was difficult to ascertain whether the records truly referred to the same offender. This issue is not specific to this research but rather reflects more broadly on the lack of uniformity in identifiers used across different agencies in Australia (see for instance Stewart et al. 2015 discussion of 'technical challenges' of data linkage using administrative data). In our case, 150 records were excluded from the national database because they could not be matched with a high degree of certainty.

Implications for deriving key variables in the project

While most jurisdictions provided very complete and high quality data for static scores and recidivism-related variables, the abovementioned issues that we encountered affected a number of other variables that were necessary for the purposes of Static-99/R norming and assessment of predictive validity for recidivism. For instance, we encountered difficulties in establishing "street time" (i.e. the time spent by offenders in the community after the index offence and excluding any prison episodes and/or time on remand), which is a crucial requirement for norming. We explored various solutions for resolving the difficulties we had in some cases of establishing street time. Two options proved viable. The first was to check with States whose data on 'street time' seemed inconsistent with their generally high standard of data provision. This allowed us to clarify misunderstandings about how the relevant variables should originally have been provided and to resolve many apparent errors. A second option to accurately estimate "street time" in some cases was to define and use proxy variables for the commencement and end of censoring periods (i.e. start date and end date for recidivism follow-up periods). In the case of commencement, we typically used either the date of release from prison for those serving a prison sentence for the index offence or the date

of the final court appearance for the index offence and therefore community release date for those receiving non-custodial orders. The end date was the date when the court history report was compiled for data coding purposes and, as such, the criminal history data for a given offender included all recidivism events up until this date.

Lessons learned and a Way Forward

In this paper, we have outlined the need to develop a national sex offender recidivism database in Australia to advance the knowledge base on sex offender recidivism, and the issues we have encountered in doing so. Development of this database has proved to be a challenging process and several obstacles have hindered our progress. There are several lessons to be learned:

a) To ensure consistency in data entry, there needs to be a high level of specificity and detailed guidance around variable definitions and data coding. Hence, we provided a 'data dictionary' to guide data collection for the 161 variables for which we requested data. This data dictionary was possibly insufficient though to address what were some unforeseen issues. It is likely that moving forwards, based on lessons learnt, we could produce a refined data dictionary addressing these previously unexpected issues and this may reduce data issues. However, we would highly recommend that similar work would ideally include funding for a small team of highly trained coders to collect data from all jurisdictions to ensure consistency in data coding.

b) Cleaning and standardisation of data is of critical importance, particularly when considering the compilation of data at a national level. In our experience, it was the comprehensive checking of patterns in the data across variables and between jurisdictions that highlighted data issues that might

have been over-looked had data checking been conducted in a piecemeal fashion only. There is no doubt that being a federation and having nine separate jurisdictions each with its own legislation, correctional agencies, data collection systems, counting and reporting rules contributes to the challenge of compiling a nationally comparable resource for Australia. The Australian Bureau of Statistics (ABS) for instance is a major and resourced agency yet the ABS has experienced ongoing issues in compiling a nationally comparable recorded crime collection. The Differences in Recorded Crime Statistics (DiRCS) project (National Crime Statistics Unit 2005) illustrated the problems State/Territory specific legislation and practices, including recording of crime, present in providing uniform national crime statistics. As a consequence of these ongoing issues, the ABS have been unable to provide, for instance, truly national statistics on assault since 2002[3].

c) The data collection burden was significant for all parties concerned. It has been a huge and ambitious undertaking in many respects, and within a relatively short period of time, to build a national harmonised and cleaned national recidivism database comprising over 6,399 sexual offenders and 161 variables. But the burden was not just on the team who built the national database from state and territory data but also on our jurisdictional partners who had to coordinate and collect this comprehensive and challenging data. Due to the absence of electronic records in some smaller jurisdictions, this

---

[3] The most recent ABS (2019) recorded crime statistics still provide data on assault for a limited range of jurisdictions.

was particularly difficult as manual coding from paper-based records was required.

Conclusion

Some important questions about risk assessment of Australian sex offenders can now be answered using data from this national sex offender database and key analyses are already well underway. To maximise the value and impact of this data on research in Australia, we anticipate engaging in discussions and planning in the near future to explore the feasibility of making the data available through a data repository upon completion of the project.

on this project.

Please note, any material published or made publicly available by the researchers cannot be considered as either endorsed by the Department or an expression of the policies or view of the Department. Any errors of omission or commission are the responsibility of the researchers. By 'Department' we mean the departments represented by the partner organisations on this project.

References

ABS (Australian Bureau of Statistics). 2019. *Recorded Crime - Victims, Australia, 2018.* Cat. No. 4510.0. Canberra, ACT: Australian Bureau of Statistics.

Allan, Alfred, Catherine L. Parry, Anna Ferrante, Christine Gillies, Catherine S. Griffiths, Francis Morgan, Caroline Spiranovic, Stephen Smallbone, Hilde Tubex, and Stephen CP Wong. 2019. "Assessing the risk of Australian indigenous sexual offenders reoffending: a review of the research literature and court decisions." *Psychiatry, Psychology and Law 26* (*2*): 274-294. doi: 10.1080/13218719.2018.1504242.

Day, Andrew, Armon J. Tamatea, Sharon Casey, and Lynore Geia. 2018. "Assessing violence risk with Aboriginal and Torres Strait Islander offenders: considerations for forensic practice." *Psychiatry, psychology and law 25* (*3*): 452-464. doi: 10.1080/13218719.2018.1467804.

Hanson, R. Karl, Andrew JR Harris, Terri-Lynne Scott, and Leslie Helmus. 2007. *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project*. Cat. No. PS3-1/2007-5. Ottawa, Ontario: Public Safety Canada.

Hanson, Robert Karl, and David Thornton. 1999. *Static 99: Improving actuarial risk assessments for sex offenders*. Ottawa, Ontario: Solicitor General Canada.

Harris, Danielle Arlanda, and Samara McPhedran. 2018. "The sentencing and supervision of individuals convicted of sexual offences in Australia." *Sexual Offender Treatment 13* (1/2). URL: http://www.sexual-offender-treatment.org/167.html

Harris, Andrew, Amy Phenix, Robert Karl Hanson, and David Thornton. 2003. *Static 99: Coding rules revised 2003*. Ottawa, Ontario: Solicitor General Canada.

Hart, Stephen D. and Douglas P. Boer. 2010. "Structured professional judgment guidelines for sexual violence risk assessment: The Sexual Violence Risk-20 (SVR-20) and Risk for Sexual Violence Protocol (RSVP)." In *Handbook of violence risk assessment*, edited by Randy K. Otto and Kevin S. Douglas, 269-294. New York: Routlegde.

Hart, Stephen D., Randall Kropp, Richard D. Laws, Jessica Klaver, Caroline Logan, and Kelly A. Watt. 2003. *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, BC, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

National Crime Statistics Unit. 2005. *Differences in Recorded Crime Statistics*. Belconnen ACT, Australia: National Crime Statistics Unit.

Olver, Mark E., Stephen CP Wong, Terry Nicholaichuk, and Audrey Gordon. 2007. "The validity and reliability of the Violence Risk Scale-Sexual Offender version: Assessing sex offender risk and evaluating therapeutic change." *Psychological assessment 19* (*3*): 318-29. doi: 10.1037/1040-3590.19.3.318.

Phenix, Amy, Leslie Helmus, and Karl R. Hanson. 2012. *Static-99R and Static-2002R evaluators' workbook*. URL: http://www.static99.org/pdfdocs/Static-99RandStatic-2002R_EvaluatorsWorkbook2012-07-26.pdf.

Pink, Brian. 2011. *Australian and New Zealand standard offence classification (ANZSOC)*. 3rd ed. Cat. No. 1234.0. Canberra, ACT: Australian Bureau of Statistics.

Shepherd, Stephane M., and Roberto Lewis-Fernandez. 2016. "Forensic risk assessment and cultural diversity: Contemporary challenges and future directions." *Psychology, Public Policy, and Law 22* (*4*): 427–438. doi: 10.1037/law0000102

Stewart, Anna, Susan Dennison, Troy Allard, Carleen Thompson, Lisa Broidy, and April Chrzanowski. 2015. "Administrative data linkage as a tool for developmental and life-course criminology: The Queensland Linkage Project." *Australian & New Zealand Journal of Criminology 48 (3)*: 409-428. doi: 10.1177/0004865815589830

Wong, Stephen, Mark E. Olver, Terry P. Nicholaichuk, and Audrey Gordon. 2003. *The violence risk scale: Sexual offender version (VRS-SO)*. Saskatoon, Saskatchewan, Canada: Regional Psychiatric Centre and University of Saskatchewan.