Predicting Recidivism: Validity and Norming of the Static-99-R for Australian Sex Offenders

ANZSOC, PERTH, 11-13TH DECEMBER 2019

Presenter:
Caroline Spiranovic, Faculty of Law, UTAS
Caroline.spiranovic@utas.edu.au

Authors:
Caroline Spiranovic, Marie-Jeanne Buscott, Anna Ferrante, Frank Morgan, Kate Griffiths
• Any material published or made publicly available by the presenter or other researchers on this project 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.

Note: * By 'Department' we mean the departments represented by the partner organisations on this project
BACKGROUND
Actuarial risk instruments

- Commonly used to predict sexual recidivism
- Cross-cultural validity questioned in recent years
- In what circumstances are they particularly problematic?
  - Preventative detention applications
  - When assessing risk in Indigenous sex offenders in Australia.

ANZSOC, PERTH, 11-13TH DECEMBER 2019

Actuarial risk instruments use a statistical approach involving analysing offender characteristics and past history to predict future behaviour and explicit rules govern how this information is combined to produce a final estimate of risk. In the majority of cases, research has suggested that the actuarial approach to risk assessment is superior to approaches that rely upon clinician experience and judgment. However, as actuarial tools rely on statistical techniques applied with respect to a specific population of sex offenders, they require cross-validation before they can be readily applied to different populations.
### BACKGROUND

**Static-99-R (Phenix et al., 2012) - scoring**

<table>
<thead>
<tr>
<th>Qt no</th>
<th>Risk Factor</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Age at release from index sex offence (18-34.9, 35-39.9, 40-59.9, 60 +)</td>
<td>1, 0, -1, -3</td>
</tr>
<tr>
<td>2</td>
<td>Ever lived with (a lover for at least 2 years)</td>
<td>0 / 1</td>
</tr>
<tr>
<td>3</td>
<td>Index non-sexual violence – any convictions</td>
<td>0 / 1</td>
</tr>
<tr>
<td>4</td>
<td>Prior non-sexual violence – any convictions</td>
<td>0 / 1</td>
</tr>
<tr>
<td>5</td>
<td>Prior sex offences (charges/convictions)</td>
<td>0/1/2/3</td>
</tr>
<tr>
<td>6</td>
<td>Prior sentencing dates (excl. index) – 4 or more</td>
<td>0 / 1</td>
</tr>
<tr>
<td>7</td>
<td>Any convictions non-contact sex offences</td>
<td>0 / 1</td>
</tr>
<tr>
<td>8</td>
<td>Any unrelated victims</td>
<td>0 / 1</td>
</tr>
<tr>
<td>9</td>
<td>Any stranger victims</td>
<td>0 / 1</td>
</tr>
<tr>
<td>10</td>
<td>Any male victims</td>
<td>0 / 1</td>
</tr>
<tr>
<td></td>
<td><strong>TOTAL SCORE</strong></td>
<td>-3 to 12</td>
</tr>
</tbody>
</table>

ANZSOC, PERTH, 11-13<sup>th</sup> DECEMBER 2019

Static-99-R – how the items are coded and scored.
Statistically based algorithm based on norming samples comprising predominantly Caucasian sex offenders.

Nominal Risk Levels (2016 version)

Total Risk Level
-3, -2 = Very Low Risk
-1, 0 = Below Average Risk
1, 2, 3 = Average Risk
4, 5 = Above Average Risk
6 and higher = Well Above Average Risk
**BACKGROUND**

Static-99-R predictive accuracy by racial group

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Comparison groups</th>
<th>AUC Caucasian (95% CI)</th>
<th>AUC non-Caucasian (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babchichin et al. (2012)</td>
<td>Canada</td>
<td>Non-Aboriginal vs. Aboriginal</td>
<td>0.74 [0.70, 0.78]</td>
<td>0.71 [0.64, 0.79]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N = 1,269</td>
<td></td>
<td>N = 319</td>
</tr>
<tr>
<td>Hanson, Lunetta et al. (2014)</td>
<td>USA</td>
<td>Caucasian vs. Hispanic</td>
<td>0.85 [0.72, 0.90]</td>
<td>0.75 [0.40, 1.00]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N = 140</td>
<td></td>
<td>N = 200</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Caucasian vs. African American</td>
<td>0.85 [0.72, 0.90]</td>
<td>0.76 [0.56, 0.97]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N = 140</td>
<td></td>
<td>N = 99</td>
</tr>
<tr>
<td>Varela et al. (2013)</td>
<td>USA</td>
<td>Caucasian vs. Hispanic</td>
<td>0.59 [0.45, 0.72]</td>
<td>0.57 [0.41, 0.73]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N = 912</td>
<td></td>
<td>N = 588</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Caucasian vs. African American</td>
<td>0.59 [0.45, 0.72]</td>
<td>0.65 [0.51, 0.78]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N = 912</td>
<td></td>
<td>N = 411</td>
</tr>
<tr>
<td>Smallbone &amp; Rallings (2013)</td>
<td>AUS</td>
<td>Non-Aboriginal vs. Aboriginal</td>
<td>0.79 [0.68, 0.91]</td>
<td>0.61 [0.45, 0.77]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N = 320</td>
<td></td>
<td>N = 67</td>
</tr>
<tr>
<td>Spiranovic (2012)</td>
<td>AUS</td>
<td>Non-Aboriginal vs. Aboriginal</td>
<td>0.71 [0.65, 0.77]</td>
<td>0.58 [0.48, 0.67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N = 660</td>
<td></td>
<td>N = 162</td>
</tr>
</tbody>
</table>

**Note.** Bolded values denote that the Static-99R scores predicted sexual recidivism at \( p < .05 \).

1Conducted analyses comparing between groups and did not find a statistically significant difference between groups in Static-99/R predictive accuracy.

2Did not conduct analyses comparing between groups but concluded there was a difference in predictive accuracy.

Despite there being a strong evidence base to support the predictive validity of risk factors and risk assessment tools for sexual reoffending, this evidence base is largely based on samples of Caucasian offenders.
Studies have found that Aboriginal (e.g., Babchishin, Blais, & Helmus, 2012), African-American (Varela, Boccaccini, Murrie, Caperton, & Gonzalez, 2013), and African-Asian offenders (Långström, 2004) score higher on Static-99R than Caucasian offenders, whereas Latino sex offenders score lower on Static-99R than Caucasian offenders (Varela et al., 2013). However, the fact that non-Caucasian offenders and Caucasian offenders exhibit differences on Static-99R risk factors does not mean that Static-99R predicts differently between these groups.

This Table presents a summary of five studies which compared the relative predictive accuracy of Static-99R scores across racial groups. Despite a general trend of Static-99R scores predicting sexual recidivism better for Caucasian offenders (estimated AUCs range from .59 to .85) than for non-Caucasian offenders (estimated AUCs range from .57 to .76), the three studies that conducted comparison analyses did not find a statistically significant difference between groups. The variability of results, however, should be a consideration in applied risk assessments with sex offenders identified with an ethnic minority.
ARC LINKAGE PROJECT
Aboriginal and non-Aboriginal sex-offenders in Australia: Assessing risk for practice & policy

Supported by ARC Linkage grant (LP140100275) & Correctives Services in each jurisdiction (partners).

Chief Investigators = Frank Morgan (UWA), Caroline Spiranovic (UTAS), Alfred Allan (ECU), Stephen Smallbone (Griffith U), Hilde Tubex (UWA) & Anna Ferrante (Curtin U).

Partner Investigators = Jayson Ware (NSW), Deborah Dawson (WA), Sarah Peart (TAS), Henry Pharo (SA), Sarah Miles (VIC), Cassandra Cowie (QLD), Mark Bartlett (ACT), Sheena Neil (NT) plus Professor Stephen Wong (Saskatchewan U, Canada).

Personnel = Dr Marie-Jeanne Buscot (statistician/research fellow), Kate Griffiths (project manager)
THE PROJECT
Aboriginal and non-Aboriginal sex-offenders in Australia: Assessing risk for practice & policy

Three major research questions:
1. Are established actuarial risk tools (e.g. Static-99) valid for Indigenous and non-Indigenous sex offenders and are there inter-jurisdictional differences in validity?
2. What culturally valid methods and tools can be used to examine risk for reoffending in Indigenous sex offenders?
3. Do rates and patterns of recidivism vary systematically according to other factors?

ANZSOC, PERTH, 11-13TH DECEMBER 2019

Focusing in this presentation on RQ1 and PRELIMINARY FINDINGS ON PREDICTIVE VALIDITY OF THE STATIC-99-R FOR A NATIONAL AUSTRALIAN SAMPLE OF SEX OFFENDERS
PRELIMINARY FINDINGS
Data cleaning and analysis is still in progress

ANZSOC, PERTH, 11-13TH DECEMBER 2019

Focusing in this presentation on RQ1 and PRELIMINARY FINDINGS on predictive validity of the Static-99-R for a national Australian sample of sex offenders
Data units from correctional service (CS) agencies in each jurisdiction have securely sent the requested data for the national sex offender recidivism database to the Population Health Research Network (PHRN) Centre for Data Linkage (CDL) at Curtin University. The CDL is an Authorised Data Linkage Unit which implements a range of protocols that mitigate the risks associated data linkage/integration projects.

The CDL have merged data from all corrective services agencies and data cleaning and validation is still underway. This work has resulted in a national sex offender recidivism database. The PHRN Centre for CDL at Curtin University have established a secure remote access server which contains a de-identified version of the national sex offender database and associated data files from all agencies. The researchers (namely statistician Marie-Jeanne Buscot) works off this deidentified version of the database through the CDL’s secure environment.
THE NATIONAL SEX OFFENDER RECIDIVISM DATABASE

161 data variables (from approx. 6,400 male sex offenders) including:

- Demographics – e.g. age, Aboriginal and/or Torres Strait Islander status;
- Prior criminal history - e.g. first and subsequent court appearances incl. offence type and dates
- Details ‘index’ sex offence – e.g. offence type, CJS outcomes - sentence type/length, dates;
- Static-99/R total & item scores index offence, 
- Post-index offence treatment, supervision and release variables
- Post-index recidivism records (sex, violent, other) incl. dates, offence type.

ANZSOC, PERTH, 11-13TH DECEMBER 2019

The database comprises 161 data items and approximately 7,000 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 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 & Thornton, 1999; Harris, Phenix, Hanson & Thornton, 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.

ANZSOC = Australian and New Zealand Standard Offence Classification, Third Edition (ABS Catalogue No. 1234.0)
ATSI = Aboriginal and Torres Strait Islander

1229 offenders in the national sample identified as indigenous (i.e. indigenous and/or ATSI) (19%) and 5115 as non-indigenous (indigenous status was missing for 55 offenders (<1%)).

For the purpose of norming, offenders for whom estimated “street time” was missing (N=625) or negative (N= 938) could not be included in the five year norming analyses (total of n=1563 or 24.4% of offenders excluded).

In addition, among those offenders who committed a sex reoffense (N=560) those for whom the time elapsed between start of follow-up and time of sex reoffense (i.e. “time_to_1st sex reoffense”) could not be determined (N=67) were also excluded from the norming analyses. (For these 67 offenders, no date was available to anchor the reoffense to the follow-up time (i.e. either actual reoffense date and/or court date for sex reoffense was missing)

A total of 1630 offenders (i.e 25.4%) were excluded for the 5 year norming analyses.

2913 of the 4769 included offenders were followed-up for 5 years or more.
2913 of the 4769 included offenders were followed-up for 5 years or more. Among those 2913, 296 offenders committed a sex reoffense. Off the 296 offenders who did reoffend, **166 reoffended within 5 years** (89 offenders reoffended after more than 5 years; time-to reoffense was missing for 33 reoffenders, and time to reoffense was negative for 8 reoffenders).

For the 5 year norming: those 89 offenders who reoffended after 5 years were recoded as “non-reoffenders” since they did not reoffended within the 5 years window.

The 41 reoffenders with either negative or missing time to reoffense were **treated as missing** in the 5 years norming analyses.

**The final sample size for the 5 years norming was N=2872 offenders**

The N=2871 five year norming sample comprised 2278 non-indigenous offenders (79.4%) and 559 indigenous offenders (19.4%). (Indigenous status missing for n=34 offenders included in the five year norming analyses).
PRELIMINARY 5-YEAR SAMPLE
For 5 year Static-99R predictive validity/norms

• 107 non-ATSI (out of 2278) committed sex-reoffense (base rate = 4.7%)
• 58 ATSI offenders (out of 559) committed sex reoffense (base rate = 10.3%)
• Recoded Static 99-R scores in the five year norming sample
  • Non-ATSI: average static 99-R = 2.26 (sd=2.6) (median: 2, range: -3 to 11, missing: n=102)
  • ATSI: average static 99-R = 4.12 (sd=2.3) (median: 4, range: 3 to 11, missing: n=5)

ANZSOC, PERTH, 11-13TH DECEMBER 2019
Predictive validity for 5 years follow-up for sexual recidivism was calculated using non-parametric estimates of the Area Under the Curve (AUC) statistic of the Receiver Operating Characteristic (otherwise referred to as ROC analysis). The ROC curve for assessing the predictive accuracy of a measure plots the hit rate (sensitivity or true positives) against the false alarm rate (specificity or false positives) for predicting recidivist/high risk and non-recidivist/low risk offenders (Hanson & Thornton, 1999). The AUC is a summary statistic derived from the ROC curve that may be interpreted as the probability that a randomly selected recidivist would have a higher risk score than a randomly selected non-recidivist (Craig, Browne, Beech, & Stringer, 2006 AUCs can range from 0 to 1.0 where .50 indicates chance level of accuracy in prediction and 1.0 indicates perfect accuracy, or 100 percent hit rate, in prediction (Craig et al., 2003).

The predictive accuracy of the non-indigenous model was acceptable (i.e. based on our data, there is a close to 70% chance that a predictive model based on the static 99-R only was able to predict those non-indigenous offenders who did reoffend vs. those who did not).

However, our data reveals that the static 99-R has close to no diagnostic ability for predicting 5 years sexual recidivism in indigenous Australian offenders (AUC close to 0.5).
WHAT ELSE HAVE WE TRIED

• Currently working with data custodians to address missing dates and negative street time issues to boost sample
• Survival analysis
• 2 year norms
### PRELIMINARY 2-YEAR SAMPLE

For 2 year Static-99R predictive validity/norms

<table>
<thead>
<tr>
<th>National sample N = 6399</th>
</tr>
</thead>
<tbody>
<tr>
<td>19% (n = 1,229) ATSI</td>
</tr>
</tbody>
</table>

N = 1,677 excluded from 2 yr norms as missing key dates

N = 361 had follow-up period shorter than 2 yrs

2 year follow-up data (n= 4,361)

| 19% (n = 809) ATSI   | 81% (n = 3,516) non-ATSI |

ATSI status missing < 1%

**ANZSOC, PERTH, 11-13TH DECEMBER 2019**

**ATSI = Aboriginal and Torres Strait Islander**

In the 2 year norming sample, we still lose 1,677 in total due to missing key dates... But we gain approx. 1,500 offenders that were excluded from the 5 year follow-up due to insufficient follow-up time.

4408 of the 4769 included offenders were followed-up for 2 years or more (68.9% of original nationwide sample).

Among those 4408 offenders follow-up for >2 years, 395 committed a sex reoffense.

**Off those 395 offenders who did reoffend, 137 reoffended within 2 years** (i.e. <2 years elapsed since end of prison episode for index offense (custodial episodes) or final court appearance date for index offense (non-custodial episodes)). 211 offenders reoffended after >2 years, ; time-to-reoffense was missing for 37 reoffenders, and time-to-reoffense was negative or null for 10 reoffenders.

For the 2 year norming: those 211 offenders who reoffended after 2 years were recoded as "non-reoffenders" (i.e. treated as 'successes') since they did not reoffend within the 2 years window.
For the two years norming analyses, the N=211 offenses that happened after 2 years were recoded as “non-reoffense”. The negative time to reoffense (10 cases) and missing time to reoffense (n=37 cases) were treated as missing.

The 47 reoffenders with either negative or missing time to reoffense where treated as missing in the 2 years norming analyses.

The final sample size for the 2 years norming was N=4361 offenders. Nsex\text{recid} in this sample was 137 (2 years base-rate: 3.14%)

The N=4361 two year norming sample comprised 3516 non-indigenous offenders (80.6%) and 809 indigenous offenders (18.5%). (indigenous status missing for n=36 offenders included in the two year norming analyses)
PRELIMINARY 2-YEAR SAMPLE
For 2 year Static-99R predictive validity/norms

- 98 non-ATSI (out of 3,516) committed sex-reoffense (base rate = 2.8%)
- 38 ATSI offenders (out of 809) committed sex reoffense (base rate = 4.7%)
- Recoded Static 99-R in the two year norming sample
  - Non-ATSI: average static 99-R = 2.16 (sd=2.4) (median: 2, range: -3 to 11, missing: n=140)
  - ATSI: average static 99-R = 4.23 (sd=2.5) (median: 4, range: 3 to11, missing: n=4)
### PRELIMINARY 2-YEAR FINDINGS

#### Area Under the Curve (AUC) for sexual recidivism

<table>
<thead>
<tr>
<th></th>
<th>AUS national Non-ATSI</th>
<th>AUS national ATSI</th>
<th>Helmus et al 2011 N = 8,106</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 yrs</td>
<td>N = 3,516</td>
<td>2 yrs</td>
<td>5 yrs</td>
</tr>
<tr>
<td>N = 809</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Static-99-R         | 0.68                  | 0.61             | 0.72                         |

Predictive accuracy of the static 99-R instrument was assessed by Area Under the curve statistics.

- The predictive accuracy of the non-indigenous model was acceptable and comparable to the one observed for the 5 years prediction (i.e. based on our data, there is a ~ 68% chance that a predictive model based on the static 99-R only was able to predict those non-indigenous offenders who did reoffend vs. those who did not).

- Although the AUC is <0.7 still, compared to the 5 years norming model, the diagnostic value of the static 99-R has significantly increased for the 2 year prediction of sex recidivism in ATSI ((Delong correlated ROC curves test, p-value = 0.03).
TENTATIVE CONCLUSIONS

From data available so far

• For Australian non-ATSI sample, levels of predictive accuracy for Static-99-R similar to that observed in the Helmus at al (2011) norming sample.
• For Australian ATSI sample, Static-99-R predicted risk sexual recidivism no greater than chance at 5 years but improved predictive validity (AUC = 0.61) at 2 years.
• Limitations of preliminary data = missing data in particular due to missing dates or negative street time – currently addressing these issues.
• Tentatively suggests IF using Static-99-R with ATSI, should base predictions on 2 years not 5 years (base rates sexual recidivism lower at 2 years but score distributions on Static-99-R similar at 2 & 5 years).
REFERENCES


• Spiranovic, C. (2012). The Static-99 and Static-99-R Norms Project: Developing Norms Based on Western Australian Sex Offenders. Crawley, Western Australia: University of Western Australia.


ANZSOC, PERTH, 11-13TH DECEMBER 2019
ACKNOWLEDGMENTS

We wish to extend our deeply felt gratitude and thanks to our partner investigators/organisations and to all the staff from each correctional agency who have contributed towards the development of the national database. We also wish to acknowledge in particular the expertise and time contributions of our project manager, Kate Griffiths (University of Western Australia), our postdoctoral statistician Marie-Jeanne Buscott (University of Tasmania) and our data support officer Davie Botes (Curtin University).

This work was supported by the Australian Research Council through a funded Linkage Project [LP140100275; Aboriginal and non-Aboriginal sex offenders in Australia: Assessing risk for practice and policy]. We received also in-kind and cash contributions from each correctional agency in Australia who were our official partner organisations on this project.

ANZSOC, PERTH, 11-13th DECEMBER 2019