Research Bulletin

Automated assessment of sexual recidivism risk for custody-based sex offenders

Alexander Bell & Mark Howard

Aim

To explore the feasibility of an automated risk assessment tool for sexual recidivism using static variables that are routinely available through Corrective Services NSW (CSNSW) operational databases.

Methods

Automated tools were modelled on a sample of 3,824 custody-based sex offenders, using variable definitions and scoring rules from the Static-99R (Helmus, Thornton, Hanson, & Babchishin, 2012) as our benchmark. The Automated Static Tool (AST) was developed by adapting coding for six items of the Static-99R and summing item scores. A second model, named the Weighted AST, regressed each of the six items from the AST onto sexual recidivism outcomes to derive an estimate of offenders' predicted probability of reoffending.

Results

Inter-rater reliability testing indicated that automated coding tended to have good agreement with examples of manual scoring on the Static-99R. For the total sample, area under the curve (AUC) statistics for the AST were .72 and .68 for sexual recidivism within two years and five years respectively, indicating fair to acceptable discrimination. The Weighted AST showed better discrimination performance than the AST (AUC = .72 - .74). When assessed among offenders who also had a manual risk assessment, the original 10-item manual Static-99R showed the strongest predictive validity; however variability across measures was relatively modest.

Conclusion

The results indicate that automated coding of Static-99R items from the CSNSW operational database has promise for estimating sexual recidivism risk, and may confer additional benefits by allowing for more complex modelling of relationships between predictors and outcomes. While the AST and Weighted AST may not currently be considered replacements for the Static-99R, such tools could support treatment delivery processes by triaging higher risk offenders into more comprehensive assessment by skilled clinicians.

CORRECTIONS RESEARCH EVALUATION AND STATISTICS | CORRECTIVE SERVICES NSW
RESEARCH BULLETIN No. 47 | OCTOBER 2020 | ISSN 2207-0850

INTRODUCTION

Effective delivery of interventions to reduce sex offenders' likelihood of sexual recidivism is a key challenge for corrections systems across jurisdictions. Sexual offending has severe impacts on victims and the broader community, and there is substantial public interest in efforts to prevent repeat offending. Demand for interventions is also relatively high, with many offenders coming into contact with the criminal justice system as a result of sexual offending. For example, a total of 13,635 adult offenders were housed in correctional centres across the state of New South Wales (NSW) as at December 2019 (NSW Bureau of Crime Statistics and Research, 2019). Within this population, 17.2% of offenders' most serious offence was categorised as relating to sexual offending, which was the second largest category behind acts intended to cause injury (19.6%).

Reducing sexual reoffending requires careful consideration of how to best provide interventions to sex offenders. The Risk Need Responsivity (RNR: Bonta & Andrews, 2007) model of correctional intervention posits that individuals at a higher chance of reoffending should be prioritised (risk); that treatment should target dynamic factors related to reoffending (need); and that interventions should be tailored to individuals in a way that optimises their engagement and progress (responsivity).

In the case of offence-specific interventions such as sex offender treatment programs, it is important that the risk principle is met by delivering higher intensity interventions to offenders who are at higher risk of sexual recidivism in particular. As such, accurate assessment of sex offenders' likelihood of sexual recidivism is critical to effective management of these offenders.

Risk assessment for sexual reoffending

The risk assessment tools used to estimate sexual recidivism differ to the tools used to estimate other types of recidivism (e.g. general recidivism; violent recidivism). There is an established body of literature indicating that risk factors can differ across or be specific to categories of offending (Craig, Browne, Beech & Stringer, 2006). As a result, tools developed to assess one type of recidivism may not have comparable predictive validity for other types of recidivism. For example, studies have found that general risk assessments, such as the Level of Service Inventory - Revised (LSI-R: Andrews & Bonta, 1995), have utility in predicting general and violent recidivism but often have accuracy for predicting sexual recidivism that is not significantly better than chance (e.g. Regusa-Salerno, Ostermann, & Thomas, 2013).

A common feature of tools that assess sexual and other recidivism risk is that they estimate probability of recidivism using static or dynamic risk factors, or a combination of both. Static risk factors are unchangeable historical or demographic variables such as the individual's age, gender, or prior criminal behaviours. Dynamic risk factors (also known as criminogenic needs) are factors that have a causal relationship with reoffending and are amenable to change, such as antisocial attitudes or substance use (Andrews & Bonta, 2010).

Static factors (particularly age and criminal history) have been found to be the strongest predictors of future offending behaviours, and may perform better than dynamic risk factors because they are more stable over time and tend to be more easily measured using standard data sources and definitions (Beech, Wakeling, Szumski, & Freemantle, 2016; Raudino, Corben, Galouzis, Mahajan, & Howard, 2019). Conversely, assessment of dynamic risk factors is a critical part of offender case management because they reflect changeable contributors to risk that can be targeted in treatment.

Corrective Services NSW (CSNSW) predominantly uses the Static-99R (Helmus, Thornton, Hanson, & Babchishin, 2012) to assess risk of sexual recidivism. As the name suggests, the Static-99R uses a number of static demographic and criminal history variables to assess risk. Static variables across 10 items are scored to indicate the offender's risk of sexual recidivism relative to other sex offenders. In accordance with RNR principles, Static-99R scores can then be used to determine offenders' eligibility for interventions and the intensity of those interventions.

The case for automated risk assessment

The Static-99R and its predecessors have established validity in predicting risk of sexual recidivism and are widely used to inform case management of sex offenders (e.g. Boccaccini et al., 2017; Hanson & Morton-Bourgon, 2009). However, like many risk assessments there are challenges in using the Static-99R as part of system-wide strategies for delivering interventions to sex offenders. The Static-99R is resource intensive, in that it needs to be completed by a trained psychologist through extensive review of historical documentation, and has complex coding rules for a number of items. Increasing resource and time costs of risk assessment can have negative flow on effects on the likelihood that treatment can successfully be delivered to sex offenders, particularly in custody settings where offenders have discrete windows of opportunity for intervention before they are released (Howard, 2016).

One solution to the limitations of manual risk assessments, that has received increasing attention over recent years, is to automate the assessment process. Given that many risk assessments use static variables that are scored from formal offence records and other archival information, there is the potential to generate tools that directly access and calculate such variables from relevant databases.

A number of tools have previously been developed to predict general recidivism among offender populations using finalised court convictions and other databases (e.g., Francis, Soothill, & Humphreys, 2007; Stavrou & Poynton, 2016; Xie, Neto, Corben, Galouzis, Kevin, & Eyland, 2018). CSNSW has recently developed two automated risk assessment tools for operational use, named the Community Triage Risk Assessment Scale (Community TRAS: Raudino, Corben, van Doorn, & Galouzis, 2018) and the Custody TRAS (Raudino et al., 2019). These automated tools were developed in order to predict general recidivism for offenders serving sentences in the community or in custody, respectively, using variables routinely recorded in the CSNSW operational database.

Automated risk assessments have clear administrative advantages by reducing assessment time and personnel costs. There is also the possibility that they could deliver improved predictive validity over their manual counterparts. For example, standardised coding of variables could result in less measurement error compared to coding by clinicians. Significantly for the purposes of this study, automated scoring also allows for more advanced calculations of risk-relevant variables and modelling of the relationships between predictors and outcomes when compared to the more intuitive, simplified values (e.g. scoring a variable as 0 or 1) that many assessments adopt to facilitate manual coding. Consistent with this, model validation studies indicated that the Custody TRAS and Community TRAS tools had better predictive validity compared to the LSI-R, which is routinely administered by Corrective Services NSW as a primary measure of general recidivism risk (Raudino et al., 2018; 2019).

While there have been multiple automated tools created to assess general recidivism risk, there are less examples examining sexual recidivism. Barbaree, Seto, Langton, and Peacock (2001) reported on an initial attempt to automatically code

Hanson's (1997) four item actuarial tool, the Rapid Risk Assessment for Sexual Offence Recidivism (RRASOR). They used variables from the Canadian Police Information Centre database (which records criminal charges and convictions incurred in Canada) to code each item of the RRASOR. Results indicated that whereas the manually scored RRASOR significantly predicted sexual recidivism, the automated version did not. Barbaree and colleagues (2001) related these outcomes with inter-rater reliability issues associated with differences in scoring between the manual and automated versions of the RRASOR. This highlights potential challenges in automation relating to poor capture of critical information when adapting existing coding frameworks to local data streams, or having insufficient data available to sufficiently replicate important variables.

More recently, Skelton, Riley, Wales and Vess (2006) created the Automated Sexual Recidivism Scale (ASRS) as part of efforts to automate assessments of sexual reoffending risk for applications within the New Zealand Department of Corrections. The ASRS was derived from the Static-99 (Hanson & Thornton, 1998) and used the New Zealand national correctional database to automatically code seven of the ten items from the measure. A later revision (the ASRS-R: Grace, 2014) updated the ASRS to reflect coding rules of the Static-99R, specifically around re-weighting of the offender age variable (see Helmus, Thornton et al., 2012). Both the ASRS and the ASRS-R showed promising results with predictive validity that was comparable to their manually scored counterparts among custody-based sex offenders in New Zealand (Grace, 2014; Skelton et al., 2006).

The present study

This study aimed to explore the feasibility of automated risk assessment for sex offenders, using static variables derived from the central CSNSW operational database. We based development of the assessment tool on the Static-99R, following previous evidence of success in automating items from this measure (Grace, 2014; Skelton et al., 2006). The Static-99R was also deemed a strong foundation for the tool given its common use by CSNSW and other jurisdictions, and established evidence base for the predictive validity of items.

We addressed the aim of this study through two areas of inquiry. The first examined whether it was possible to reliably replicate items from the Static-99R based on data that is commonly available through the CSNSW database. This was intended to assess the quality and comprehensiveness of the automated coding approach for variables that have an established relationship with sexual recidivism risk, and provide context to interpretation of subsequent analyses.

The second tested whether it was possible to generate models based on available CSNSW data that had adequate predictive validity, relative to benchmarks set by manual assessments. This firstly involved comparing the discrimination performance of automated and manual versions of the Static-99R, using the same coding and scoring rules. Following the example of other tools (e.g. Raudino et al., 2018; 2019) we also hypothesised that performance of the automated model may be further improved by incorporating more statistically advanced weightings of the relationship between predictor variables and sexual recidivism. In this regard we considered automated models to be viable if they showed good discrimination accuracy, irrespective of their fidelity to the items and coding rules for the manual Static-99R.

METHODS

Sample

The sample for this study was derived from a cohort of all adult male offenders who were convicted of one of more sex offences and commenced a custodial episode with CSNSW from January 1999.

Offenders were required to have been released from custody before September 2017 to allow for a minimum follow-up period of two years in the community.

In accordance with Static-99R assessment eligibility criteria (Phenix et al., 2017), offenders were also required to have at least one conviction for ‘category A’ sex offences in their criminal history. All offenders with a history of only ‘category B’ sex offences were excluded from analysis. Categories of sex offences were calculated by applying Static-99R coding rules to existing ANZSOC divisions used by CSNSW to define offence types (see Appendix 1).

To maintain independence of observations in the study, we only considered the first custodial episode for each offender and any subsequent episodes were excluded. After applying these criteria the final sample comprised n = 3,824 offenders.

Materials

Data

The data used in the creation of our new tool were retrieved from the CSNSW Offender Information Management System (OIMS). OIMS is an operational database used to maintain information about all offenders under the supervision of CSNSW in custody and in the community. For the purposes of this study, we aimed to derive all predictor variables from OIMS so that automated tools could be calculated without the need for any additional external data.

Key variables extracted from OIMS included date of birth, offence and sentencing information attached to the index and historical corrections episodes, and counts of prior custodial episodes. We also extracted previous assessment data for offenders in the sample to support primary analyses in the study.

Recidivism data used in assessing reoffending outcomes for sex offenders in the sample was retrieved from the Bureau of Crime Statistics and Research (BOCSAR) Reoffending Database (ROD). Key variables obtained from the reoffending data included date of offence and offence type. The most recent data available from ROD included information on finalised convictions in NSW criminal courts up to the end of September 2019.

Measures

Static-99 / Static-99R

The Static-99 (Hanson & Thornton, 1999) and its subsequent revision, the Static-99R (Helmus, Thornton et al., 2012), are actuarial risk assessment tools used to estimate sex offenders’ risk of sexual recidivism. Both tools comprise 10 items that assess static variables related to demographic and criminal history characteristics of the offender. These include the offender’s age at release, presence of prior and index non-sexual violence convictions, number of prior sex offences and sentencing dates, victim details, and if they had ever lived with a partner for over two years.

The Static-99R is identical to the Static-99 with the exception of the age variable, which was revised following additional research into patterns of association with sexual recidivism (Helmus, Thornton et al., 2012). Whereas the Static-99 scores age as a dichotomous variable so that older offenders receive a score of 0 and younger offenders receive a score of 1, the Static-99R scores age on a more incremental four-point scale ranging between -3 and 1. For both tools, eight items are also scored dichotomously with offenders either receiving a score of 0 or 1. The remaining item, which pertains to the number of prior sex offences, has four scoring options ranging from 0 to 3. Individual items are added together to provide a total score ranging between 0 and 12 for the Static-99, and between -3 and 12 for the Static-99R. Total scores can then be compared to normative tables as a means of estimating absolute and relative sexual recidivism risk.

Both the Static-99 and Static-99R have been well validated across different settings and samples, and have consistently demonstrated moderate predictive validity (typically AUC = .69- .70; Helmus, Hanson, Thornton, Babchishin, & Harris, 2012) and excellent interrater reliability (typically α > 0.9; Phenix, Helmus, & Hanson, 2012).

Recidivism

As our study was specifically interested in predicting sexual recidivism, recidivism was defined as when an offender was convicted of a sexual offence following release from their index custodial episode. The commission of a sexual offence was identified using the Australian Bureau of Statistics Australian and New Zealand Standard Offence Classification (ANZSOC) codes (see Appendix 1). To exclude instances of pseudorecidivism, we only counted convictions where the date of reoffending was recorded as occurring after release from the index custodial episode.

Following previous recommendations for assessment of sexual recidivism outcomes (e.g., Quinsey, Rice, & Harris, 1995), our benchmark measure for model development and validation purposes was sexual reoffending within 5 years' free time following release from the index custodial episode. To give additional insights about the performance of risk estimates over time, we also examined sexual recidivism within 2 years' free time when conducting model validation analyses.

Analytical plan

Model development

The first step in developing our tool, which we named the Automated Static Tool (AST), was to review the OIMS database to determine which items from the Static-99R could be meaningfully scored using routinely available data. We concluded that it was possible to score six of the 10 items, including age at release; index conviction for non-sexual violence; prior conviction for non-sexual violence; number of prior sentencing dates; number of prior sexual offences; and any conviction for a non-contact sex offence[1].

We then derived coding systems that best adapted available OIMS data to the item scoring rules and variable definitions for the Static-99R (Phenix et al., 2017). The age item was calculated by subtracting the offender's date of birth from their custodial episode end date. The number of prior sentencing dates was derived from all sentences attached to the offender's previous corrections episodes with CSNSW. For the remaining four items, ANZSOC subdivision codes attached to the offender's corrections episodes were used to categorise what constituted a violent offence, a category A or category B sexual offence, and a non-contact sex offence as defined in the Static-99R manual.

As the ANZSOC codes group offences together it was not always possible to perfectly align definitions of individual offence types in the Static-99R to ANZSOC subdivisions. For example, a subdivision may contain multiple types of Category A offence in addition to a single example of Category B offence. In these cases, we reviewed the prevalence of different offence types within a subdivision and assigned relevant Static-99R definitions on a probabilistic basis. Details of ANZSOC codes used in offence definitions are given in Appendix 1[2].

---

[1] This subset of variables that can be automated by OIMS is the same as that used for the ASRS-R (Grace, 2014), with the exception of the male victim item.

[2] We adopted this probabilistic approach to offences represented in ANZSOC subdivisions in preference to coding of Lawpart codes for each individual offence type. This is because Lawpart codes were not consistently available and had different definitions over the study timeframe, making them unreliable for the purposes of scoring criminal history information about offenders in the sample.

After developing coding rules, we generated two iterations of the predictive model from OIMS. The first (the AST) scored each of the 6 items of the Static-99R, and summed the values of each of the items to derive a total numerical score, in a similar procedure to the manual version.

The second (which we hereafter refer to as the Weighted AST) also scored each of the 6 items using Static-99R scoring rules. Scores for each item were then individually entered into a binary logistic regression model as predictors for sexual reoffending within 5 years. The model generated weights for the relationship between scores on each item and reoffending outcomes (beta coefficients), which were then used to generate a regression equation predicting each offender's probability of sexual reoffending. This predicted probability score[3] comprised the overall index of sexual recidivism risk for the Weighted AST.

Model reliability

To compare the consistency of coding between the automated and manual scoring systems we examined the inter-rater reliability between individual items in the AST and the Static-99R, for those offenders who had been administered a valid Static-99R during their index custodial episode (n = 1,174). Total scores from the AST were also compared to total scores derived from a reduced version of the Static-99R, comprising the manually scored versions of the 6 items that were also represented in the AST.

Inter-rater reliability statistics were calculated using Intraclass Correlation Coefficients (ICCs) in a two-way mixed-effects model specified for absolute agreement of scores.

[3] For the purposes of analysis and reporting in this study, predicted probability scores were multiplied by 100 and rounded to the nearest integer to give a numerical score for the Weighted AST.

Model validity

Predictive validity of the AST and Weighted AST was assessed through Receiver Operating Characteristics (ROC) area under the curve (AUC) statistics. AUC statistics are a widely used measure of discrimination accuracy which assess the probability that a score randomly selected from a group with the signal of interest (in this instance, offenders who recidivate) is higher than a score selected from a group without the signal of interest (in this instance, offenders who do not recidivate). AUC values can be interpreted so that .50 indicates discrimination that is no better than chance and 1.0 indicates perfect discrimination. Hosmer and Lemeshow (2000) suggest rules of thumb to classify discrimination performance so that AUC values of .50 to .60 are 'poor', .60 to .70 are 'fair', .70 to .80 are 'acceptable', .80 to .90 are 'excellent' and .90 or above are 'outstanding'.

Predictive validity of the AST and Weighted AST was also compared to that of manually scored risk assessments. To do this we derived AUC statistics for the AST, Weighted AST, and Static-99R assessments, for those offenders in the sample who had been administered a Static-99 or Static-99R during their index custodial episode[4]. To facilitate interpretation of differences in predictive validity, we compared automated models to both the full 10-item Static-99R as well as the reduced 6-item Static-99R comprising only those items that were also represented in the AST and Weighted AST.

[4] Offenders in the sample had often been administered a Static-99 or a Static-99R, but not both. To allow for consistency in interpreting results and maintain optimal sample sizes, we converted all Static-99 scores into Static-99R scores by applying the revised age scoring rules to item 1. Predictive validity analyses involving Static-99R assessments used offenders with original or recoded versions of the assessment. Conversely, to minimise any biases associated with recoding the original manual scoring of the age item, reliability analyses only included offenders who were administered a Static-99R.

RESULTS

Sexual recidivism

Among all offenders in the sample who met criteria for minimum follow-up periods, 4.5% (172/3,824) were convicted for sexual recidivism within 2 years, and 7.4% (212/2,864) were convicted for sexual recidivism within 5 years.

It was noted that recidivism rates varied for offenders in the sample who did or did not have a valid manual risk assessment attached to their index custodial episode. Convictions for sexual recidivism were detected for 3.0% (47/1,585) within 2 years and 5.9% (61/1,042) within 5 years for offenders with a Static-99 or Static-99R. By comparison, offenders who did not have a risk assessment for their index episode returned higher recidivism rates of 5.6% (125/2,238) within 2 years and 8.3% (151/1,822) within 5 years.

The AST and Weighted AST

Figure 1 shows the distribution of the AST for the total study sample. Scores on the 6 item AST ranged from -3 and 8 and showed a normal distribution, with an average (median) of 1.

As previously described, the Weighted AST was developed by modelling each of the 6 items from the AST as multivariate predictors of sexual recidivism within 5 years, and using the resulting regression coefficients to generate each offender's predicted probability of recidivism. Results from the logistic regression model are given in Table 1.

The overall model showed goodness of fit in predicting sexual recidivism outcomes ($\chi 2$ (10) = 132.06; $p < .0005$). It can be seen that individual items tended to have the expected relationships with likelihood of sexual recidivism, so that increasing scores were associated with increasing odds of recidivism within 5 years. The single exception to this was index non-sexual violence, which was found to have a (non-significant) negative association with outcomes. This indicates that after adjusting for other factors, offenders with index convictions for non-sexual violence tended to be less likely to reoffend than those without index convictions for non-sexual violence.

Table 1. Regression coefficients for items of the AST as predictors of sexual recidivism within 5 years

Item (Score)	B (SE)	Wald χ^2	p-value	Odds Ratio [95% CI]
Age (-3)		10.64	.014	
Age (-1)	.27 (.29)	.83	.363	1.30 [.74-2.31]
Age (0)	.74 (.32)	5.26	.022	2.09 [1.11-3.94]
Age (1)	.70 (.30)	5.44	.020	2.01 [1.12-3.60]
Index non-sexual violence (1)	-.26 (1.86)	1.85	.165	.77 [.54 - 1.11]
Prior non-sexual violence (1)	.11 (.21)	.36	.594	1.12 [.74 - 1.68]
Prior sex offences (0)		9.01	.018	
Prior sex offences (1)	.43 (.23)	3.30	.068	1.53 [.97- 2.42]
Prior sex offences (2)	.41 (.23)	3.31	.069	1.51 [.97-2.35]
Prior sex offences (3)	.60 (.22)	7.70	.006	1.82 [1.19-2.77]
Prior sentencing dates (1)	.76 (.21)	13.64	<.0005	2.13 [1.43 - 3.18]
Non-contact sex offences (1)	1.04 (.16)	43.58	<.0005	2.82 [2.07 - 3.83]

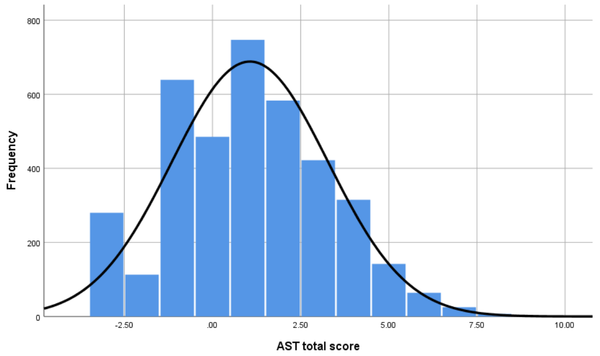

Figure 1. Distribution of AST scores across the total sample (n = 3,824)

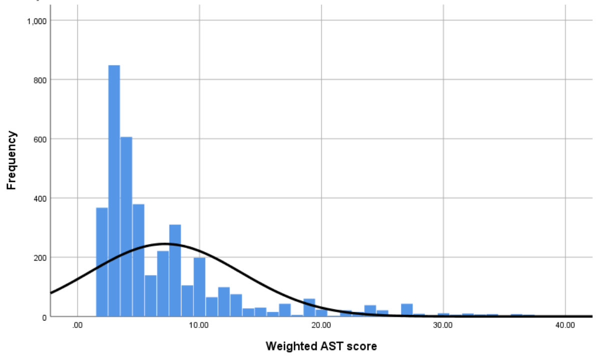

Figure 2. Distribution of Weighted AST scores across the total sample (n = 3,824)

Items with the strongest associations with recidivism included prior sentencing dates and a history of non-contact sexual offences. Items assessing prior and index non-sexual violence had the weakest predictive validity for sexual recidivism within 5 years and returned non-significant multivariate associations with the outcome.

Figure 2 shows the distribution of Weighted AST predicted probability scores for the total sample. Scores ranged from 2 to 37 (corresponding to estimated probability of sexual recidivism between 2% and 37% within 5 years) with an average (median) of 5. The Weighted AST scores also showed substantial skew, whereby three in five offenders (57.5%) received a score of 5 or less[5].

[5] Differences in distribution between the AST and Weighted AST are expected, because AST scores reflect Static-99R scores in representing risk relative to other offenders (with scores roughly in the middle of the range representing average risk), whereas Weighted AST scores represent absolute risk of sexual recidivism, which tends to be low among sex offenders on average.

Model reliability

Inter-rater reliability testing was conducted to compare consistency in scores across manual and automated coding methods for items in the AST and Weighted AST. Only offenders in the sample who had a manually scored Static-99R were included in analyses (n = 1,174).

ICC statistics are given in Table 2. It can be seen that inter-rater reliability ranged from moderate to excellent (.67 – .99)[6] between automated and manual scoring approaches to individual items. The age item (item 1) showed the highest correlation between scoring methods. Conversely, the only other item involving a multiple category as opposed to a dichotomous scoring system, relating to counts of prior sex offences (item 5), showed the lowest absolute agreement between scores.

Table 2. Inter-rater reliability statistics for selected items of the Static-99R when scored manually, and when scored automatically in the AST (n = 1,174)

Item	Description	Scoring (Range)	ICC [95% CI]
1	Age at release	-3, -1, 0, 1	.99 [.99 – .99]
3	Index non-sexual violence	0, 1	.81 [.79 – .84]
4	Prior non-sexual violence	0, 1	.76 [.68 – .81]
5	Prior sex offences	0, 1, 2, 3	.67 [.63 – .70]
6	Prior sentencing dates	0, 1	.72 [.61 – .79]
7	Non-contact sexual offences	0, 1	.75 [.72 – .78]
Total	Summed total of the 6 items	(-3 – 8)	.90 [.87 – .92]

Notes. ICC = intraclass coefficient; CI = confidence interval

[6] As a rule of thumb, ICC statistics can be interpreted so that values less than .5, between .5 and .75, between .75 and .90, and greater than .90 are indicative of poor, moderate, good, and excellent reliability, respectively (e.g. Koo & Li, 2016).

The 6-item total scores achieved an ICC of .90, which suggests a good to excellent level of absolute agreement between manual and automated scored total scores.

Model validity

Associations between AST and Weighted AST scores and recidivism

Figures 3 and 4 depict the five year observed sexual reoffending rates for offenders (n = 2,864) who received each of the total risk assessment scores derived by the AST and the Weighted AST respectively. In both figures it can be seen that increasing scores on each of the measures was associated with increasing rates of recidivism in a broadly linear trend.

For the AST, greater variability in the relationship between increasing scores and increasing rates of recidivism was observed for offenders who received higher scores than those who received scores in the lower and middle ranges. This is likely attributable to the relatively low numbers of offenders who were assessed as having scores in the upper risk bands, meaning that the outcomes of each individual offender had an increased impact on overall reoffending rates.

For the Weighted AST, it can be seen that there was a less ‘smooth’ linear association between scores and reoffending rates, or greater variation around the slope, compared to the AST on average. This is related to the Weighted AST having a greater range of possible scores than the AST, resulting in lower numbers of offenders on any single given score. For example, the prominent spike in the Weighted AST distribution is a result of only a single offender (who was also eligible for five year follow-up analyses) receiving a score of 21, who ultimately reoffended.

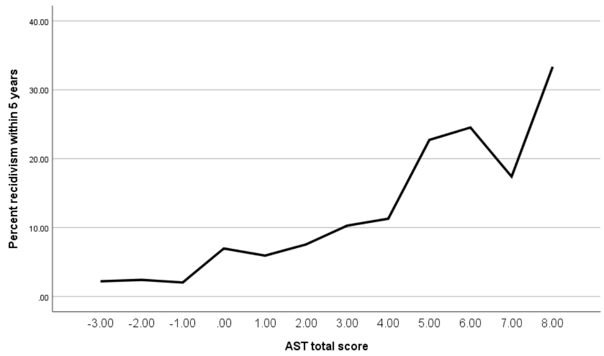

Figure 3. Rates of sexual recidivism within 5 years corresponding to scores on the AST (n = 2,864)

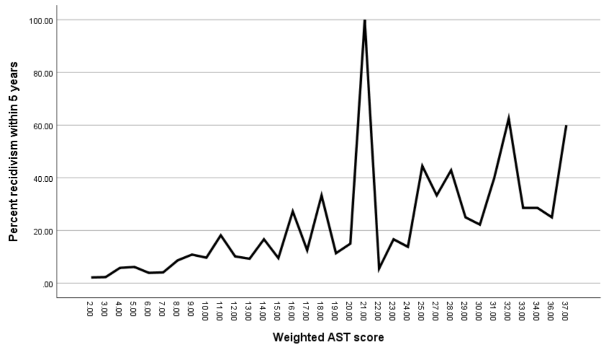

Figure 4. Rates of sexual recidivism within 5 years corresponding to scores on the Weighted AST (n = 2,864).

Predictive validity of the AST and Weighted AST

Discrimination performance of the AST and Weighted AST for the total study sample was assessed against sexual recidivism within 2 years’ and 5 years’ free time following release from the index custodial episode. The variable minimum survival periods resulted in different sample sizes available for each analysis, so that the total sample for analyses of reoffending within 2 years was n = 3,824 and for analyses of reoffending within 5 years was n = 2,864.

The AST showed fair to acceptable discrimination performance with AUC = .72 (95% CI = .68 – .76) for sexual recidivism within 2 years and AUC = .68 (95% CI = .65 – .72) for sexual recidivism within 5 years. The Weighted AST was found to perform better than the AST and showed more consistently acceptable discrimination for both outcome measures, with AUC = .74 (95% CI = .71 – .78) for sexual recidivism within 2 years and AUC = .72

(95% CI = .68 - .75) for sexual recidivism within 5 years.

To give an indication of the effect sizes associated with the AST and Weighted AST, we also conducted a series of binary logistic regression analyses whereby sexual recidivism was predicted by total scores derived by each of the measures. For the AST, the exponentiated beta coefficient, or odds ratio, was 1.44 for recidivism within 2 years and 1.34 for recidivism within 5 years. This indicates that each unit increase in AST score was associated with a 44% increase and 34% increase in the odds of sexual recidivism within 2 years and 5 years, respectively. For the Weighted AST, the odds ratios derived from these models were 1.11 for sexual recidivism within 2 years and 1.10 for sexual recidivism within 5 years. This corresponds to around a 10% increase in risk of recidivism for each unit increase in Weighted AST score[7].

Comparisons with manual assessments

Discrimination performance of the AST and Weighted AST was compared to that of the manually scored Static-99R. To facilitate interpretation, we compared the automated tools to total scores derived from both the original 10-item version of the Static-99R, as well as the reduced 6-item version of the Static-99R which summed manual coding for the 6 items that were used in the AST and Weighted AST. We also restricted analyses to samples of offenders who had manual assessments attached to their index episode and were eligible for detection of sexual recidivism within 2 years (n = 1,585) and within 5 years (n = 1,401) of release.

AUC statistics for each of the measures are shown in Table 3. The pattern of results indicates that the 10-item version of the manual Static-99R had the strongest predictive validity. Incremental losses to discrimination performance were observed when reducing the manual Static-99R to the abbreviated 6-item version, and again when automating scoring of those 6 items in the form of the AST. While some predictive validity was lost during automation, this was offset by gains when reweighting the contribution of individual items to overall risk estimates in the Weighted AST. As a result, the predictive power of the Weighted AST approached that of the 6-item manual Static-99R. Differences in performance across models became relatively more pronounced, and performance for all models declined, when assessing recidivism within 5 years compared to recidivism within 2 years.

[7] We note that the odds ratios for the AST and Weighted AST are not directly comparable because scores have different units of measurement.

Table 3. AUC statistics for sexual recidivism within 2 and 5 years for each of the measures

Measure	Recidivism within 2 years		Recidivism within 5 years	
	AUC	[95% CI]	AUC	[95% CI]
Static-99R (10 item)	.81	[.75-.86]	.76	[.70-.82]
Static-99R (6 item)	.79	[.73-.84]	.71	[.65-.78]
AST	.75	[.68-.81]	.67	[.60-.74]
Weighted AST	.77	[.70-.83]	.72	[.65-.78]

DISCUSSION

This study aimed to explore the feasibility of developing an automated risk assessment tool for sexual recidivism using data that is routinely stored in the CSNSW operational database. To achieve this we developed automated models based on variable coding and scoring rules from an existing gold standard for sex offender risk assessment in the form of the Static-99R. Primary analyses examined whether the available data within OIMS and our coding framework for that data allowed for reliable replication of items from the Static-99R when scored manually, and whether resulting models showed adequate predictive validity for sexual recidivism.

Can automated models replicate items from the Static-99R?

Inter-rater reliability analyses indicated a good to excellent level of agreement between our coding rules for the six items calculated for the automated models and examples of manual scoring for those items in the Static-99R. The ICC for the summed total of the items was .90, which indicates that both models returned the same absolute score for 90% of the sample. This high level of absolute agreement can be partly attributed to very high reliability for the age item (item 1), which has a substantial contribution to variability in overall scores following revisions to the Static-99R (Helmus, Thornton et al., 2012).

A more modest degree of agreement between coding methods was found for items relating to criminal history, and prior sex offences in particular. This latter finding may be related to the use of a multiple category scoring system for the prior sex offences item, which is more sensitive to variation from absolute agreement compared to the dichotomous scores derived for most other items. More broadly, we acknowledge that the automated coding system used in this study may be less sensitive to nuances in scoring prior sexual and other offence items compared to skilled clinicians. For example, because our automation process was based on CSNSW data only it could not detect offending that occurred in other jurisdictions, or where a sentence was imposed that did not involve oversight from CSNSW (e.g. a suspended sentence without supervision). In addition, coding was linked to explicit definitions of sex offences and would not be able to identify non-sexual offences with sexual elements, such as sexually motivated murder. The probabilistic method used to allocate offence definitions to ANZSOC codes may have also been less accurate compared to manual reviews.

Challenges in accurately replicating variables or constructs from manual assessments at the time of automation have been identified as a risk to validity (Barbaree et al., 2001) and may have a bearing on subsequent analyses of the models' discrimination performance, which will be discussed in further detail in the next section.

Can automated models have predictive validity?

When assessed for the total sample of sex offenders in our study, both the AST and the Weighted AST showed fair to acceptable predictive validity. AUC statistics for the AST and Weighted AST were .72 and .74 after two years, and .68 and .72 after five years respectively, which is in similar ranges of discrimination performance to established manual measures such as the Static-99R (e.g. Helmus, Hanson et al., 2012).

The Weighted AST tended to perform slightly better than the AST, which supports proposals that automation could confer advantages by allowing for more complex calculation of the relationships between predictor variables and outcomes. Such advantages are also implied by the results of regression modelling (see Table 1) which showed that after automation, some items in the AST had poor multivariate associations with sexual recidivism. In particular, the index non-sexual violence item had a negative association with recidivism outcomes, suggesting that the presence of index violent offences would contribute to a higher total raw score on the AST or the Static-99R while potentially being indicative of (non-significantly) lower risk. By revising item scores by their regression coefficients, the Weighted AST was better able to account for the influence of such items on global estimates of risk. Further research is needed to better understand findings for the association between non-sexual violence and sexual recidivism, with potential applications for both automated and manual revisions of risk assessments (see also Helmus & Thornton, 2015; Sjostedt & Langstrom, 2001).

Comparisons of predictive validity across models for offenders who also had a manual assessment showed that the manual, 10-item Static-99R had the strongest discrimination performance. Predictive validity declined for manual assessments using the reduced set of 6 items that were the basis of the AST and Weighted AST; however this decline was relatively minor. These results support previous indications that briefer versions of the Static-99R derived primarily from demographic and criminal history variables can have adequate utility in predicting sexual recidivism risk (e.g. Skelton et al., 2006). Conversely, there is also the implication that predictive validity could be optimised by more comprehensive and systematic recording of offence details in corrections and other databases, such as victim gender and relationships with the offender.

Additional losses to predictive validity were observed when comparing the reduced version of the Static-99R to the AST. When considered in conjunction with the results of inter-rater reliability testing, the pattern of results suggests that some degree of risk-relevant information was lost when converting manual Static-99R coding rules to the AST. The relatively poor absolute agreement in scoring for prior sex offences may be particularly influential to validity outcomes, because this variable is among the strongest predictors of sexual recidivism (Helmus & Thornton, 2015)[8]. Following from previous discussions of reliability, we note that difficulties scoring prior sex offences may be a function of local factors such as the coding rules used or availability of data on OIMS as opposed to a limitation of automation more broadly.

[8] It was beyond the scope of this study to report extensively on normative and other statistics for the manual version of the Static-99R. However, preliminary analyses were conducted to examine multivariate associations between the reduced set of 6 items from the manual Static-99R and sexual recidivism within 5 years. The prior sex offences item was the most significant predictor of recidivism; $\chi^2(3) = 34.79$; $p < .0005$.

Similar to analyses using the full sample, the Weighted AST showed stronger discrimination performance than the AST for offenders who also had a manually scored assessment, and predictive validity for the Weighted AST approached that of the reduced 6-item Static-99R. This suggests that by applying more sophisticated modelling of the relationships between predictors and outcomes, the Weighted AST may have served to offset limitations of the AST resulting from imperfect coding or capture of risk-relevant variables. Given evidence for the utility of re-weighting the contribution of items to overall estimates of risk, there is the implication that automated models could show further improvements in predictive validity by more fully modelling the relationship between raw data pertaining to risk-relevant constructs and recidivism outcomes, as opposed to using the specific scoring rules for items adopted by the Static-99R.

Notwithstanding these considerations, we note that differences in predictive validity across each of the models in this study were relatively minor. Confidence intervals for the AUC statistics indicated that all models showed discrimination performance for sexual recidivism that was significantly better than chance, and no model showed superior performance that reached thresholds for statistical significance. As previously mentioned, AUC statistics observed for most analyses of models in this study are broadly commensurate to those of established sex offender risk assessments, and may be argued to represent relatively large effect sizes for risk assessment involving recidivism outcomes with low base rates more generally (Raynor, Kynch, Roberts, & Merrington, 2000; Rice & Harris, 2005).

Applications and future directions

While the results of this study show promise for automated prediction of sexual recidivism risk, we note that the AST and Weighted AST would not be currently considered viable replacements for established assessments such as the manual Static-

99R. Importantly, this study indicates that the Static-99R has superior predictive validity over other tested models and therefore continues to be an optimal standard of assessment. The Static-99R is also supported by an established body of validation research which cannot be readily generalised to adaptations such as the AST and Weighted AST. These are critical considerations given the substantial implications of risk assessment on legal as well as intervention outcomes. We also acknowledge that further work to establish categories of relative risk is needed for the AST and Weighted AST to assist in identifying offenders who are priorities for case management relative to the sex offender population; this may be particularly challenging for the Weighted AST because it generates an index of absolute risk that is highly skewed towards lower bounds for most sex offenders.

However, there may be scope for automated models such as the AST and Weighted AST to support intervention delivery processes by acting as an initial triaging tool. By working to screen out very low risk offenders and prioritising higher risk offenders for more comprehensive assessment, such applications can substantially mitigate the resource costs of manual assessments (see Raudino et al., 2018; 2019). For example, among offenders in the study sample with a Static-99R, being placed in the 'average category' of risk was associated with a 3.3% rate of sexual recidivism within 5 years[9]. Adopting thresholds for manual follow-up assessment based on scores on the AST and Weighted AST that correspond with a lower than 3.3% recidivism risk (AST < 0; Weighted AST $< .033$) would hypothetically circumvent manual assessment of 32.7% (based on the AST) to 33.5% (based on the Weighted AST) of offenders in the sample with a Static-99R. Sex offender risk assessment may be particularly amenable to this triaging approach because sexual recidivism is relatively infrequent and large proportions of offenders tend to have a low predicted probability of reoffending; therefore efficiency gains can be achieved even when setting a highly sensitive threshold for further assessment.

To facilitate the use of automated models in this capacity, it would be beneficial for future studies to further establish appropriate recidivism norms. A relevant consideration is that most offenders in the study sample did not have a Static-99R assessment; the reasons for this are not clearly understood although could relate to a number of factors ranging from missing data to ineligibility for custody-based intervention pathways (e.g. lack of time for programs prior to release). Interestingly, these offenders also had higher rates of sexual recidivism than those with a Static-99R, which suggests that further work may be required to identify the appropriate target populations for assessment and adapting screening thresholds to the recidivism patterns of that target group.

Additional research would also be beneficial to test the stability of automated assessments over time and across different offender groups. This is particularly relevant in the case of Indigenous offenders, given their overrepresentation among sex offenders in NSW as well as previous indications that the Static-99R may have poorer predictive validity for Indigenous compared to non-Indigenous Australian offenders (Spiranovic, 2012). As previously mentioned, the results of this study also suggest that future refinements of automated tools may improve predictive validity by better accounting for the role of non-sexual violence items and more fully modelling the relationships between risk-relevant variables and recidivism.

[9] This 'average' rate of sexual recidivism is lower than that observed in Static-99R normative data and elsewhere; it has previously been noted that the relationship between Static-99R scores and absolute recidivism rates are not stable across samples (Helmus, Hanson et al., 2012; Hanson et al., 2016).

Conclusions

Decades of research have established the utility of measures such as the Static-99R in assessing sexual recidivism risk. However, manual assessments often have substantial time and other resource costs, which can impact system-wide strategies for delivering interventions to sex offenders. Given that many assessments estimate risk based on static variables, efficiency gains may be achieved by replicating or adapting assessments to automation within local databases.

The results of this study indicate that automation of Static-99R items from the CSNSW operational database has promise for estimating sexual recidivism risk with acceptable predictive validity. This approach may also confer benefits by allowing for more complex modelling of relationships between predictors and outcomes. While the results suggest that the AST and Weighted AST may not be considered replacements for the Static-99R as a standard for optimal risk assessment, such tools could feasibly support treatment delivery by helping to identify and prioritise higher risk sex offenders for more comprehensive assessment by skilled clinicians.

REFERENCES

Andrews, D.A., & Bonta, J. (1995). Level of Supervision-Revised (LSI-R): An offender *assessment system. User's guide*. Toronto: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct (5th ed.).* New Providence: LexisNexis Group.

Bonta, J., & Andrews, D. A. (2007). Risk-need-responsivity model for offender assessment and rehabilitation. *Rehabilitation, 6*(1), 1-22.

Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior, 28*(4), 490-521.

Beech, A. R., Wakeling, H. C., Szumski, F., & Freemantle, N. (2016). Problems in the measurement of dynamic risk factors in sexual offenders. *Psychology, Crime & Law, 22*(1-2), 68-83.

Boccaccini, M. T., Rice, A. K., Helmus, L. M., Murrie, D. C., & Harris, P. B. (2017). Field validity of Static-99/R scores in a statewide sample of 34,687 convicted sexual offenders. *Psychological Assessment, 29*(6), 611-623.

Craig, L. A., Browne, K. D., Beech, A., & Stringer, I. A. N. (2006). Differences in personality and risk characteristics in sex, violent and general offenders. *Criminal Behaviour and Mental Health, 16*(3), 183-194.

Francis, B., Soothill, K., & Humphreys, L. (2007). *Development of a reoffending measure using the Police National Computer database*. Lancashire: Lancaster University.

Grace, R. C. (2014). *The Automated Sexual Recidivism Scale (ASRS): Revised age weights*. Unpublished manuscript commissioned by NZ Department of Corrections.

Hanson, R. K. (1997). *The development of a brief actuarial risk scale for sexual offense recidivism (No. 1997).* Ottawa: Solicitor General Canada.

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*(1), 1-21.

Hanson, R. K., & Thornton, D. (1999). *Static 99: Improving actuarial risk assessments for sex offenders (2).* Ottawa: Solicitor General Canada.

Hanson, R. K., Thornton, D., Helmus, L. M., & Babchishin, K. M. (2016). What sexual recidivism rates are associated with Static-99R and Static-2002R scores? *Sexual Abuse, 28*(3), 218-252.

Helmus, L., Hanson, R. K., Thornton, D., Babchishin, K. M., & Harris, A. J. (2012). Absolute recidivism rates predicted by Static-99R and Static-2002R sex offender risk assessment tools vary across samples: A meta-analysis. *Criminal Justice and Behavior, 39*(9), 1148-1171.

Helmus, L. M., & Thornton, D. (2015). Stability and predictive and incremental accuracy of the individual items of Static-99R and Static-2002R in predicting sexual recidivism. *Criminal Justice and Behavior, 42*(9), 917-937.

Helmus, L., Thornton, D., Hanson, R. K., & Babchishin, K. M. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse, 24*(1), 64-101.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression (2nd ed.)*. New York: John Wiley & Sons.

Howard, M. (2016). *Process Evaluation of the Custody Based Intensive Treatment (CUBIT) Program for Sex Offenders*. Sydney: Corrective Services NSW.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163.

NSW Bureau of Crime Statistics and Research (2019). *New South Wales custody statistics: Quarterly update December 2019*. Sydney: NSW Bureau of Crime Statistics and Research.

Phenix, A., Fernandez, Y., Harris, A. J., Helmus, M., Hanson, R. K., & Thornton, D. (2017). *Static-99R Coding Rules, Revised-2016*. Ottawa: Public Safety Canada.

Phenix, A., Helmus, L., & Hanson, R. K. (2012). *Static-99R & Static-2002R evaluators' workbook*. Ottawa: Public Safety Canada.

Quinsey, V.L., Rice, M.E., & Harris, G.T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence, 10*, 85-105.

Ragusa-Salerno, L. M., Ostermann, M., & Thomas, S. S. (2013). Does the LSI-R have utility for sex offenders? *Criminal Justice and Behavior, 40*(9), 952-969.

Raudino, A., Corben S., van Doorn, G., & Galouzis, J. (2018). *The Community Triage Risk Assessment Scale (Community TRAS): A statistical model for predicting risk of reimprisonment*. Sydney: Corrective Services NSW.

Raudino, A., Corben S., Galouzis, J., Mahajan, Y., & Howard, M. (2019). *The Custody Triage Risk Assessment Scale (Custody TRAS): An updated statistical model for predicting risk of return to custody*. Sydney: Corrective Services NSW.

Raynor, P., Kynch, J., Roberts, C., and Merrington, M. (2000). *Risk and need assessment in probation services: An evaluation*. London: Home Office.

Rice, M.E., & Harris, G.T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r. *Law and Human Behavior, 29*, 615-620.

Sjostedt, G. & Langstrom, N. (2001). Actuarial assessment of sex offender recidivism risk: A cross-validation of the RRASOR and the Static-99 in Sweden. *Law and Human Behaviour , 25*, 629-645.

Skelton, A., Riley, D., Wales, D., & Vess, J. (2006). Assessing risk for sexual offenders in New Zealand: Development and validation of a computer-scored risk measure. *Journal of Sexual Aggression, 12*(3), 277-286.

Spiranovic, C. (2012). *The Static-99 and Static-99-R norms project: Developing norms based on Western Australian sex offenders*. Perth: Crime Research Centre.

Stavrou, E., & Poynton, S. (2016). *The revised Group Risk Assessment Model (GRAM 2): Assessing risk of reoffending among adults given non-custodial sanctions*. Sydney: NSW Bureau of Crime Statistics and Research.

Xie, Z., Neto, A., Corben, S., Galouzis, J., Kevin, M., and Eyland, S. (2018). *The Criminal Reimprisonment Estimate Scale (CRES): A statistical model for predicting risk of reimprisonment*. Sydney: Corrective Services NSW.

APPENDIX 1

Static-99R variable definitions adapted to OIMS data

Variable	OIMS coding definition
Sex offence category*	
Category A	ANZSOC subdivisions 0311; 0312; 0329; 1325
Category B	ANZSOC subdivisions 0322; 0323; 1324
Static-99R item	
1. Age at release	Index episode end – date of birth
3. Index non-sexual violence – any convictions	Convictions for any of the following ANZSOC subdivisions attached to index episode: 0111; 0121; 0131; 0211; 0212; 0213; 0291; 0299; 0511; 0521; 0532; 0611; 0612; 0621; 1211; 1334; 1531
4. Prior non-sexual violence – any convictions	Convictions for any of the following ANZSOC subdivisions attached to prior episodes: 0111; 0121; 0131; 0211; 0212; 0213; 0291; 0299; 0511; 0521; 0532; 0611; 0612; 0621; 1211; 1334; 1531
5. Prior sex offences	Count of convictions for the following ANZSOC subdivisions attached to prior episodes: 0311; 0312; 0321; 0322; 0329; 0323; 1324; 1325
6. Prior sentencing dates	Count of sentences attached to prior episodes
7. Non-contact sex offences – any convictions	Convictions for any of the following ANZSOC subdivisions attached to index or prior episodes: 0322; 0329; 1325

*Refer to the Static-99R manual page 21 (Phenix et al., 2017) for conceptual definitions

Other CRES Research Titles

Oct 2020	The predictive validity of general risk assessment tools for offence-specific recidivism among domestic violence offenders –
Oct 2020	Access to programs and services among Culturally and Linguistically Diverse (CALD) offenders: The case of EQUIPS
Oct 2020	Process evaluation of Custodial Case Management: Case plans
Aug 2020	Understanding the spectrum of domestic violence: Risk factors, treatment pathways and recidivism among offenders who commit intimate partner or non-intimate partner violence
Aug 2020	Process evaluation of the High Intensity Program Units (HIPUs): Within-treatment change
Aug 2020	Evaluation of the Practice Guide for Intervention (PGI): Relationship between offender needs and PGI use in case planning and supervision practice
Dec 2019	Effectiveness of the Initial Transitional Support (ITS) Service 2014-2017
Sept 2019	Evaluation of EQUIPS treatment pathways for domestic violence offenders in New South Wales
Sept 2019	Process evaluation of the Practice Guide for Intervention (PGI): Staff experiences of implementation and continuing service delivery
Sept 2019	Desistance in an ageing inmate population: An examination of trends in age, assessed risk of recidivism and criminogenic needs
Aug 2019	The Custody Triage Risk Assessment Scale (Custody TRAS): An updated statistical model for predicting risk of return to custody
May 2019	Effects of the Practice Guide for Intervention (PGI) on behaviour change intervention dosage among community-based offenders
May 2019	Blending care and control in delivery of the Practice Guide for Intervention (PGI): An assessment of the quality of dual role relationships between offenders and supervising officers in the community
Feb 2019	Process evaluation of the Practice Guide for Intervention (PGI): Staff perceptions of community supervision in the context of change
Oct 2018	The Community Triage Risk Assessment Scale: A Statistical model for predicting recidivism among community-based offenders
Aug 2018	Assessing offender change over treatment: The influence of treatment context on self-reported antisocial attitudes
Aug 2018	Forty is the new thirty (for recidivism): Trends in offender age, reimprisonment, and time to desistance among NSW custodial population
Aug 2018	The Criminal Reimprisonment Estimate Scale (CRES): A Statistical model for predicting risk of reimprisonment
Aug 2017	Evaluation of vocational training in custody: Offenders' experiences of training and pathways to post-release employment
Aug 2017	Evaluation of vocational training in custody: Relationships between Training, Post-Release Employment and Recidivism
Dec 2017	The Case Quantify and Search Tool (C-QST)
Aug 2017	Increase in the community corrections population
Oct 2016	Process Evaluation of the Custody Based Intensive Treatment (CUBIT) Programs for Sex Offenders
Mar 2015	A Process Evaluation of the Intensive Drug & Alcohol Treatment Program (IDATP) – Study One
June 2013	Evaluation of the Getting SMART Program
June 2013	Drug Use in the Inmate Population – prevalence, nature and context
Apr 2012	Maintaining the Link: A Survey of Visitors to New South Wales Correctional Centres

Research Bulletin No.47
ISSN 2207 0850

Corrections Research, Evaluation & Statistics
Governance & Continuous Improvement
Corrective Services W
GPO Box 31
Sydney NSW Australia

Telephone: (02) 8346 1556
Email: research.enquiries@justice.nsw.gov.au