Investigating the predictive validity of Static-99/99R scores in a sample of older sexual offenders

Beatrice C. Raymond^a, Troy E. McEwan^a, Michael R. Davis^{a,b,c}, Sophie G. Reeves^d and James R. P. Ogloff^{a,b}

^aCentre for Forensic Behavioural Science, Swinburne University of Technology and Forensicare, Melbourne, Australia; ^bDepartment of Psychiatry, Monash University, Melbourne, Australia; ^cDepartment of Psychiatry, University of Melbourne, Melbourne, Australia; ^dChildren's Court Clinic, Court Services Victoria, Melbourne, Australia

There is limited information regarding the use of risk assessment tools with aging offender populations. It is known that the likelihood of offending behaviour decreases with age, a small group of men either continue or begin to offend sexually in the later decades of life. The current study investigated the predictive validity of the Static-99 and the Static-99R, in a sample of convicted Australian sex offenders aged 50 and older. A sample of 118 participants was identified, of which 17 (14.4%) re-offended within a follow-up period ranging from four months to 20 years ($M = 9.07$ years). There were seven recidivists (13.46%) above the age of 60 years ($n = 52$) and 10 aged 40 to 59.9 years ($n = 66$). Both the Static-99 and Static-99R demonstrated moderate predictive validity with both age groups. The limitations of this study are discussed as well as recommendations for future research.

Key words: age correction; aged; aged sexual offenders; Australia; predictive validity; recidivism; risk assessment; sexual offenders; Static-99; Static-99R.

Sexual offending is recognised as a significant social problem that can cause both short- and long-term harm to victims (Dworkin, Menon, Bystrynski, & Allen, 2017). As such, risk factors associated with repeated sexual offending have received considerable research attention in an attempt to guide efforts to prevent repeat sexual offending (e.g. Fazel, Sjostedt, Langstrom, & Grann, 2006; Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2005; Mann, Hanson, & Thornton, 2010). The assessment of risk for sexual recidivism has been a focus of court decisions and treatment

options for several decades. This has led to the development of a number of risk assessment tools. Some of these are mechanical in nature (i.e. actuarial instruments, e.g. Hanson, 2001), while others involve a degree of clinical discretion (i.e. structured professional judgment tools; see Davis & Ogloff, 2008). For these instruments to contribute to the legal process and treatment recommendations, their accuracy in different populations is important. One particular population that has seen little research regarding risk assessment tools is aged sex offenders.

Correspondence: Bea Raymond, Centre for Forensic Behavioural Science, Swinburne University of Technology and Forensicare, PO Box 5007, Keon Park Victoria 3073, Australia. Email: braymond@swin.edu.au

Age has been found to be a critical variable when assessing offender risk. As an offender ages, the risk of recidivism of all kinds generally decreases (Book, Starzyk, & Quinsey, 2001; Hanson, 2001; Hanson, Harris, Helmus, & Thornton, 2014; Rice & Harris, 2014). Being over the age of 50 years, in particular, is thought to be protective against non-sexual or sexual recidivism; as well as being protective against first-time sexual offending (Booth, 2016; Nicholaichuk, Olver, Gu, & Wong, 2014; Rettenberger, Haubner-Maclean, & Eher, 2013). It has been posited that this protective effect encompasses a range of issues such as reduced libido, restrictive ailments, frailty and greater likelihood of living in restrictive environments such as nursing homes (Bauer, Fetherstonhaugh, et al., 2013; Bauer, McAuliffe, Nay, & Chenco, 2013).

Despite age generally appearing to be protective against offending, the number of older offenders in correctional systems continues to increase as the populations of most industrialised countries age. In Australia, for example, prisoners over the age of 50 years accounted for 12.8% of the total prison population in 2017, compared to 11.3% in 2010 and 8.5% in 2000 (Australian Bureau of Statistics, 2018). The increasing number of older Australian prisoners reflects a worldwide trend in prison populations, with sexual offenders projected to be a particularly common group among older prisoners (Chua, Cheung, Friedman, & Taylor, 2018). In Australia, the majority of prisoners over the age of 50 years are imprisoned due to sexual offending (Australian Bureau of Statistics, 2018; Chua et al., 2018). Given the increasing number of aged sexual offenders in prison and probation systems, it is essential that common risk assessment instruments that are used to determine security and rehabilitation programme placements, and even influence release decision making, can be shown to be valid with this group. The most commonly used actuarial instruments, both in Australia and worldwide, are the Static-99 family of instruments developed by Hanson and Thornton (2000).

The Static-99 family of instruments

The first edition of the Static-99 was developed in the late 1990s by Hanson and Thornton (2000). The measure was developed to estimate the likelihood of recidivism among adult male sex offenders and was an amalgam of two previous instruments, the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997) and the Structured Anchored Clinical Judgement (SACJ; Grubin, 1998). The Static-99 was developed using large datasets from Canada and the United Kingdom to evaluate the optimal combination of risk factors that could most effectively predict sexual recidivism. It ultimately included 10 historical/static factors that were positively associated with increased risk of sexual recidivism. These were: having a stranger victim, having a male victim, prior sexual offending, prior non-sexual offending, index non-sexual offending, youth (under the age of 25 years), having ever lived with a partner for two years, prior sentencing dates and whether they had victimised someone unrelated to them. Scores on the instrument range from a theoretical zero to a maximum of 12, although no one in the development samples scored more than nine. The Static-99 total score corresponds to a descriptive risk category (low, moderate-low, moderate-high, high) each of which is characterised by a group probability of sexual reoffending over five and 10 years. This group probability was the proportion of recidivists in the development sample.

Since its development, the Static-99 has been well researched over nearly three decades. A revised coding manual was published in 2003 (Harris, Phenix, Hanson, & Thornton, 2003), and several variations of the instrument have also been published. These include the Static-2002 (Phenix, Doren, Helmus, Hanson, & Thornton, 2008) and the Static-2002R (Helmus, Thornton, Hanson, & Babchishin, 2012). The Static-99 itself received new normative data in 2009 (Helmus, Hanson, & Thornton, 2009) and was quickly followed by a new variant known as the Static-99R

(Phenix, Helmus, & Hanson, 2009). This new version had just one minor change to the original tool – namely, an expanded age item that considered the protective value of advancing age in more detail. This change occurred as a result of limitations of the earlier version suggesting that it was largely insensitive to this protective effect and, furthermore, that the total score and risk categories over-estimated the probability of reoffending among aged offenders (Barbaree, Blanchard, & Langton, 2006; Helmus, Hanson, et al., 2012; Lennings, Seidler, Heard, Collins, & Nasr, 2011; Rettenberger et al., 2013; Wollert, Cramer, Waggoner, Skelton, & Vess, 2010). A further revised manual for both instruments was published in 2016 (Phenix et al., 2016), and normative data for the Static-99R have been regularly and incrementally updated over the past decade.

In the Static-99, the only age-related item is whether the person is under 25 years of age (which is scored 1) or over 25 years (scored 0) at exposure to risk. In the Static-99R, those aged 18 to 34.9 years at time of release receive a score of 1; those 35 to 39.9 years old score 0; those 40 to 59.9 years old score –1; and those aged 60 and over receive a reduction of 3. The use of negative scoring means that older offenders' total scores are reduced due to their age, therefore altering the final risk score, and potentially the risk category to which they are assigned. These variable age weightings were based on empirical research (Helmus, Hanson, et al., 2012).

The validity of the Static-99 and its risk factors with older offenders

While the Static-99R with its age correction has been found to produce more accurate group recidivism rates for older offenders (Helmus, Hanson, et al., 2012), there is no evidence that it is more accurate at differentiating recidivists from non-recidivists at any age than the original Static-99. Indeed, Hanson, Helmus, and Thornton (2010) found no

differences in predictive validity for those aged less than 50 years, or more than 50, 60 or 70 years of age. This is not particularly surprising given that Hanson (2006) found that the original Static-99 had a non-statistically significant trend towards better predictive validity for those aged 60 and older than other age groups. However, the recidivism rates for the 60 and older group were lower. This underscores the fact that older sexual offenders in general do have lower recidivism rates, but that within the group of older offenders the Static-99 risk factors still differentiate recidivists from non-recidivists in a comparable fashion to how they do with younger offenders.

It should be noted that some research has specifically considered the protective factor of age and concluded that a universal age correction may lead to under-estimation of risk in some cases. While not specifically using the Static-99, Nicholaichuk, Olver, Gu, and Wong (2014) created a brief actuarial scale (BARS) that was composed of six variables that were the same or similar to the 10 items on the Static-99. They investigated reoffending outcomes among 2,158 Canadian sexual offenders scored on this tool, of whom 517 (23.9%) were aged over 50 years. Interestingly, while the older offenders did reoffend at a significantly lower rate than the younger offenders, this effect was only present for those who were not in the highest risk category. The authors concluded that while age did appear to have a general effect of lowering sexual recidivism risk, older offenders assessed as high risk continued to present an increased risk similar to that of high-risk younger adult offenders. Based on this finding, Nicholaichuk et al. (2014) suggested that both age and risk matter, and not all older offenders can be assumed to be a lower risk. Those with offending that persisted into older age were generally characterised by sexual deviance rather than violence (see also Davis, 2017).

Recent research by Hanson (2018) noted that the longer high risk sexual offenders live

in the community without reoffending sexually, the more likely it is that their risk of reoffending decreases or lowers. For an aged sample this is yet to be tested for first time offenders; however, it appears to be a pertinent point for those offenders convicted of historical offences who have not reoffended since that time.

In the Australian context, there have been three published studies investigating the predictive validity of the Static-99 and the Static-99R inventories (Allan, Dawson, & Allan, 2006; Reeves, Ogloff, & Simmons, 2018; Smallbone & Rallings, 2013), the findings of which have been largely consistent with studies from North America. Reeves et al. (2018) investigated predictive validity with a sample of 621 Australian sex offenders assessed as part of pre- or post-sentence proceedings, using police data to identify recidivism. Over the five years following initial assessment, 13% were reconvicted for another sexual offence, growing to 17% at 10 years and 21% at 15 years. The resulting area under the curve (AUC) values for Static-99/99R scores ranged between .67 and .69 across all three follow-up time periods. Interestingly, Smallbone and Rallings (2013) found that the Static-99 and Static-99R both showed predictive validity with non-indigenous sexual offenders, but only the original version was related to recidivism amongst indigenous Australian sexual offenders. Given that the only difference between the two instruments is the age item, it can thus be posited that the protective effect of age did not apply to the indigenous offenders in this admittedly small sample (Davis, 2017).

Aims

There are currently no Australian published studies that consider the predictive validity of the Static-99 and Static-99R for sex offenders who are aged 50 years and above. Accordingly, this paper aimed to address this under-studied aspect of sex offender risk assessment and specifically to address the

efficacy of the age correction on the Static-99R amongst those aged 50–59.9 years and those over the age of 60 years. It was hypothesised that there would be a protective effect of age resulting in reduced recidivism amongst the older of the two groups. It was also hypothesised that while both tools would discriminate between recidivists and non-recidivists at a similar rate, the Static-99R would reflect a total score and risk category that better encapsulates the likelihood of recidivism of the aged offender.

Method

Sample

The selected dataset was derived from secondary data that were collected at Forensicare (Victorian Institute of Forensic Mental Health) between 1987 and 2011 (see Reeves et al., 2018). Forensicare is the state-wide specialist provider of forensic mental health services in Victoria, Australia. When the original data collection occurred, participants were identified through databases and file searches at Forensicare. The participant data identified were then reviewed to confirm that arrest and charge information regarding sexual offending behaviour was detailed. Only adult male participants were included. These data were then reviewed for the purposes of the current study, and only those who were 50 years and over at the time of contact with Forensicare were included. Of the 653 sexual offenders in the original dataset, 124 (19%) were 50 years and over. This group was divided into two subsamples. One consisted of 66 offenders (55.9%) aged 50 to 59.9 years, and the 52 (44%) who were aged 60 years and above were the second subsample.

The cut-off of 50 years and older as the defining criterion for ‘aged offenders’ was chosen for consistency with existing research into aged sexual offenders (e.g. Maschi et al., 2011; Nicholaichuk et al., 2014). While 50 years and over might not be considered particularly ‘aged’ in a normal population within

the community, the sex offender populations in prison have been shown to age at a faster rate than their community counterparts (Dawes, 2009; Fazel, Hope, O'Donnell, Piper, & Jacoby, 2001; Williams, Stern, Mellow, Safer, & Greifinger, 2012). It is likely that life experiences that contribute to sexual offending (e.g. young age at abuse, substance use, mental illness) lead to faster aging, and for those who have been incarcerated, the aging process occurs at a different rate from that within the community. Therefore, the age of 50 years is considered equivalent to 65 years in the non-offending community (Maschi et al., 2011). From a more pragmatic perspective, the Static-99R age weights provide reductions from the age of 50 years, which also suggested that 50 years should be the criterion for the current study.

The information required to code the Static-99/R risk factors was derived from detailed clinical files that included demographic and offence-specific details. There were seven (5.6%) participant files that lacked sufficient information for scoring purposes, reducing the sample to 118. The mean age at time of contact with Forensicare was 59.6 years ($SD = 7.07$, range = 50–78 years). It was noted that while the Static-99/R developers ascertained absolute risk estimates as well as relative risk estimates, due to the small size of the current sample, relative risk was the only type of risk that could be established.

At the time of coding the Static-99/R, fewer than one third of the sample had committed an indecent act (36; 30.3%); 27 committed indecent assault (22.7%); 19 committed sexually penetrative acts (16%); and the remaining 37 committed other sexual offences such as Incest (11.8%). Fifty-nine (50%) victims were aged between 0 and 11 years, another 39 (31.1%) were aged between 12 and 17 years, and the remaining 20 (16.9%) were aged over 18 years. A total of 86 (72.3%) victims were related or known to the offender for longer than 24 hours at the time of offence. Thirty (25.2%) victims were male, and 73

(61.3%) were female. The remaining 15 (13.5%) index offences involved victims of both genders.

Follow-up recidivism data were collated from the dataset developed by Reeves et al. (2018). Reoffending data were collected in 2011 from Corrections Victoria, identifying subsequent convictions for sexual offences and also accounting for time at risk data regarding periods of incarceration. Follow-up time ranged from four months to 20 years ($M = 9.07$ years; $SD = 5.5$), which reflects deaths ($n = 17$), reoffending ($n = 17$) or the end of the study ($n = 86$). Seventeen (14.4%) men in the sample were reconvicted of a sexual offence during the follow-up period.

Measures

The Static-99 and Static-99R were used for this study. The Static-99 and Static-99R are actuarial risk assessment tools that both contain 10 historical/static risk factors. The items on the Static-99 are scored dichotomously except for one (Prior Sentencing Dates), which is coded on an ordinal scale ranging from zero to three. Possible scores range from zero to 12. The total score is used to mechanically place an individual into one of four nominal risk categories (0–1 = low; 2–3 = moderate–low risk; 4–5 = moderate–high risk; 6+ = high).

The Static-99R is scored identically to the Static-99 except for the age item. Rather than the dichotomous over or under 25 years coding on the Static-99, the Static-99R includes four possible age scores based on time at release from prison ($-3 = 60+$; $-1 = 40-59.9$; $0 = 35-39.5$; $1 = 18-34.9$). The highest score that can be obtained is again 12 but the lowest score that can be obtained is now -3 . The total score is used to mechanically place an individual into one of five nominal risk categories (-3 to $-2 =$ very low risk; $-1-0 =$ below average risk; $1-3 =$ average risk; $4-5 =$ above average risk; and $6+ =$ well above average risk).

Procedure

Human research ethics approval was granted by the relevant ethic committees including Swinburne University of Technology and Justice Human Research Ethics Committee. Clinical files of participants were accessed by the first author and the Static instruments scored, with the Static scores from the original data collection in 2011 (published in Reeves et al., 2018) used for inter-rater reliability analyses.

Statistical analysis

Statistical analyses were primarily conducted using the Statistical Package for the Social Sciences (SPSS) Version 23. Medcalc was also used to compare the results of receiver operating characteristics (ROC) analysis between the two variants of the Static-99. All analyses were conducted three times, once for the total sample, and once each for the two age ranges.

Chi-square tests of homogeneity were utilised to investigate the differences in recidivism at each category on both the Static-99 and Static-99R. Pearson's Φ and Cramer's V were used as effect sizes for the chi-square tests. Predictive validity was assessed by means of the area under the curve (AUC) statistic in ROC analysis and the more traditional point-biserial correlation coefficient (r_{pb}).

In ROC analysis, the hit rate (true positives) is plotted against the false-alarm rate (false positives) for all observed predictor values. The ROC curve graphs the trade-offs in specificity that occur as sensitivity increases. ROC analyses generate AUCs that describe the probability that a random recidivist will score higher on the risk assessment tool than a random non-recidivist. As such, an AUC value of .5 indicates chance prediction, and scores of 0 and 1 represent perfect positive and perfect negative prediction, respectively. The AUC statistic thus provides a useful summary of the overall discriminating power of the instrument (Grann & Langstrom, 2007; Mossman, 1994). The predictive validity of Static-99 and

Static-99R scores was then compared to ascertain the overall effects of the age-weight correction that was added to Static-99R.

While ROC analyses are generally preferred in risk assessment research because they are less affected by base rates than correlations, it is useful to report both in primary studies as point-biserial correlation coefficients can be used in subsequent meta-analyses (Lopez, Boer, Kirby, & Davis, 2017; Strub, Douglas, & Nicholls, 2014). Due to the limited size of the sample, it was decided that the use of a survival analysis would not accurately display the results over the time span. The nature of the data, as well as the size, limited the usability of this statistical analysis.

A random selection of 30 cases was used to test interrater reliability by using data already scored during Reeves et al. (2018). Interrater reliability (IRR) was assessed using intraclass correlation coefficients ($ICC_{A,1}$; two-way, random effects model, single ratings, absolute agreement method). Reliability for the Static-99 total scores was good ($ICC_{A,1} = .91$, 95% confidence interval, CI [.81, .96]) and similarly for Static-99R total scores ($ICC_{A,1} = .93$, 95% CI [.84, .96]).

Results

Descriptive analysis

The mean age of participants was 59.66 years ($SD = 7.07$; range = 50–78 years) at the time they were assessed. Of the 118 participants, two thirds had prior sexual offences ($n = 76$, 64.4%), and fewer than one third ($n = 17$, 14%) went on to perpetrate further sexual offences. In the 50 to 59.9 year age group ($n = 66$, 55.9%), a total of 11 (16.7%) reoffended sexually and/or non-sexually. Of these, one (1.5%) reoffended non-sexually only; eight participants (12.1%) sexually reoffended only; and two participants (3%) committed some form of non-sexual and sexual reoffence. Among those 60 years and above ($n = 52$, 44%), 10 (19.2%) participants reoffended sexually and/or non-sexually. Of these, three

Table 1. Total scores and categories of samples using Static-99 and Static-99R.

Static-99 outcome			Static-99R outcome		
Category	<i>N</i> (%) <i>in category</i>	Further sexual offences (%) within category)	Category	<i>N</i> (%) <i>in category</i>	Further sexual offences (%) within category)
Low	33 (28)	2 (6)	Very low	13 (11)	2 (15)
Mod-low	40 (34)	3 (7)	Below average	39 (33)	0
Mod-high	25 (21)	5 (20)	Average	46 (39)	7 (15)
High	20 (17)	7 (35)	Above average	15 (13)	5 (33)
			Well above average	5 (4)	3 (60)

Note: Percentages in parentheses. Mod = moderate.

(5.8%) reoffended with non-sexual offences; four participants (7.6%) recidivated sexually; and three (5.8%) committed some form of non-sexual and sexual reoffence. In the current study there were five participants (4.2%) who were convicted of offences that were perpetrated 10 or more years prior. This might reflect ‘desistance’ as described by Hanson (2018), and this time in the community offence free would reflect a lowered overall risk.

The number of offenders in each of the nominal risk categories on the Static-99 and Static-99R are shown in Table 1, along with the number that obtained each score. Static-99 scores for the total sample ranged between 0 and 9 ($M = 2.00$, $SD = 1.05$), and Static-99R scores ranged from -3 to 8 ($M = 3.00$, $SD = 0.98$). Overall there were few offenders with scores falling in the high category of the Static-99 ($n = 20$, 17%) or its Static-99R equivalent the well above average category ($n = 5$, 4%).

Twelve out of the 17 (70.6%) subjects who recidivated fell into either the moderate–high or high risk categories of the Static-99. A chi-square test of homogeneity indicated that there were significantly more subjects with sexual recidivism that fell in the moderate–high or high categories on the Static-99 than in

the lower categories, $\chi^2(3, N = 118) = 10.92$, $p = .012$, with a small effect size, $V = .30$. Seven recidivists (35%) fell into the Static-99 high risk category alone. At the individual score level there were increasingly more recidivists as the Static-99 score increased, $\chi^2(9, N = 118) = 32.31$, $p < .001$, with a concomitantly medium effect size, $V = .52$. Specifically, those with a score of 7+ were highly likely to recidivate (5.9%).

In regard to the Static-99R, eight of the 17 participants (47%) with further sexual offences fell into either the above average or well above average risk categories, showing that there were more subjects with sexual recidivism that fell in the above average and well above average on the Static-99R than in the lower categories, $\chi^2(4, N = 118) = 19.39$, $p = .001$, and reflected in a small effect size, $V = .41$. There were increasingly more recidivists as the scores increased, which was statistically significant, $\chi^2(11, N = 118) = 32.76$, $p = .001$, and had a medium effect size, $V = .53$. In the well above average risk category alone, there were a total three (60%) recidivists.

Age and recidivism

The number of offenders in each of the nominal risk categories on the Static-99 and

Table 2. Scores for each category in subsample 50 years to 59.9 years.

Static-99 outcome			Static-99R outcome		
Category	<i>N</i> (%) <i>in category</i>	Further sexual offences(% within category)	Category	<i>N</i> (%) <i>in category</i>	Further sexual offences(% within category)
			Very low	0	0
Low	17 (27)	0	Below average	17 (27)	0
Mod-low	21 (34)	3 (14)	Average	27 (44)	3 (11)
Mod-high	13 (21)	3 (23)	Above average	13 (21)	3 (23)
High	11 (18)	4 (36)	Well above average	5 (8)	3 (60)

Note: Percentages in parentheses. Mod = moderate.

Table 3. Scores for each category in subsample 60 years and above.

Static-99 outcome			Static-99R outcome		
Category	<i>N</i> (%) <i>in category</i>	Further sexual offences (%) within category)	Category	<i>N</i> (%) <i>in category</i>	Further sexual offences (%) within category)
Low	13 (25)	2 (15)	Very low	13 (25)	2 (15)
Mod-low	19 (37)	0	Below average	19 (37)	0
Mod-high	11 (21)	2 (18)	Average	18 (36)	3 (17)
High	9 (17)	3 (33)	Above average	2 (100)	2 (100)
			Well above average		

Note: Percentages in parentheses. Mod = moderate.

Static-99R for the two different age groups are shown in Tables 2 and 3. In the 50 to 59.9 years age group there were a total of 10 (15.2%) recidivists, while in the 60 years and above age group there were seven (13.5%) recidivists. This was not a statistically significant difference, $\chi^2(1, N=118) = 0.07, p = .795$, with a concomitantly small effect size, $\Phi = -.02$. Accordingly, those aged 60 years and above did not recidivate at a greater rate than those aged 50 to 59.9 years.

The 50 to 59.9 years age group

Using the Static-99, the number of recidivists differed across the categories in the 50 to 59.9 years group, $\chi^2(3, N=66) = 7.86, p = .049$, with a small effect size, $V = .35$. The proportion of recidivists significantly increased as the Static-99 score increased, $\chi^2(9, N=66) = 22.22, p = .008$, with a medium effect size, $V = .58$. Using the Static-99R, the number of recidivists also differed significantly across the categories, $\chi^2(3, N=66) = 12.05, p = .007$,

and with a small effect size, $V = .43$. As was the case with the Static-99, as scores on the Static-99R increased so did the proportion of recidivists. This was statistically significant, $\chi^2(9, N=66) = 22.22, p = .008$, with a medium effect size, $V = .58$.

The 60 years and over age group

Differences in the number of recidivists across Static-99 categories in this sub-sample were not statistically significant, $\chi^2(3, N=52) = 6.26, p = 0.1$, and the effect size was small, $V = .35$. This would appear to be a consequence of two recidivists scoring in the low risk category. Despite these two cases, the proportion of recidivists did increase to some degree as scores on the Static-99 increased. This was statistically significant, $\chi^2(8, N=52) = 19.84, p = .011$, and had a concomitantly medium effect size, $V = .62$.

For the Static-99R, the number of recidivists significantly differed across the categories in the 60 years and above group, $\chi^2(3, N=52) = 16.01, p = .001$, with a medium effect size, $V = .56$. The age-correction for the Static-99R resulted in none of the offenders in this age group falling into the well above average risk category. Nonetheless, as scores on the Static-99R increased, so did the proportion of recidivists. This increase was statistically significant, $\chi^2(8, N=52) = 19.84, p = .011$, with a medium effect size, $V = .62$. The ROC curves for sexual recidivism for the two instruments across the entire sample are shown in Figure 1. The overall AUC and correlational results for sexual and non-sexual recidivism at the different age ranges are presented in Table 4.

Some authors argue a preference to not label AUC effects (Rice & Harris, 2005; Mossman, 2013) as it can be misleading due to differing definitions of what is considered small, moderate or large effects. However, the use of labels for AUCs is useful to contextualise the utility of the statistic. With this in mind, the AUCs (Table 4) indicate that total scores on both tools have a moderate to large

(Mossman, 2013) ability to discriminate between those who recidivate and those who do not recidivate in the total sample (AUC values = .74 and .75).

When examining the age-related sub-samples, there were no discernible differences between the discriminatory performance of the Static-99 and Static-99R. The subsample of 50 to 59.9 years produced the same sexual reoffending AUC values (AUCs = .80) for both the Static-99 and the Static-99R, as did the 60 years and over subsample (AUCs = .68). Among those aged 60 and over, neither total score was able to differentiate between recidivists and non-recidivists in a statistically significant manner, despite the moderate AUC values of .68. Statistical power was, of course, one contributor to this lack of statistical significance, as the number of offenders in the over 60 years group was unfortunately smaller. However, the very wide asymptotic confidence intervals for this group also reflects the fact that in this age category, several low scorers on both instruments were sexual recidivists. In regard to non-sexual recidivism, both instruments had comparable predictive validity for each age group. However, both failed to have a statistically significant relationship with non-sexual recidivism amongst the older offenders.

The statistical significance of any differences in the predictive validity of the two instruments was assessed using the method of DeLong, DeLong, and Clarke-Pearson (1988). This involved a pairwise comparison of each ROC curve to determine whether the difference between the two areas itself was significantly different from zero. As the error is never consistent across each point of a ROC curve, this is a more powerful method than simply comparing the overlap between asymptotic confidence intervals. Perhaps unsurprisingly, the differences in predictive validity between the Static-99 and Static-99R for sexual recidivism were not statistically significant for the total sample ($z = 0.197, p = .844$), the 50 to 59.9 years age group ($z = 4.016, p =$

Figure 1. Receive operating characteristics (ROC) for the Static-99 and Static-99R in relation to sexual recidivism.

Table 4. Predictive validity of static measures for sexual reoffending.

	Sexual reoffending			
	AUC	SE	95% CI	r_{pb}
1. Total sample				
Static-99	.74**	.08	[.60, .89]	.36**
Static-99R	.75*	.08	[.60, .89]	.34**
2. Age 50 to 59.9 years				
Static-99	.80**	.07	[.66, .93]	.43**
Static-99R	.80**	.07	[.66, .93]	.43**
3. Age 60+ years				
Static-99	.68	.15	[.39, .97]	.26
Static-99R	.68	.15	[.39, .97]	.26

Note: AUC = area under the curve (confidence intervals are asymptotic); r_{pb} = point-biserial correlation coefficient.

* $p < .05$. ** $p < .01$.

.0001), or the 60 years and above age group ($z = 1.11$, $p = .267$). In regard to non-sexual recidivism, there were also no statistically

significant differences between the predictive validity of the Static-99 and that of the Static-99R for the total sample ($z = 0.825$, $p = .41$),

the 50 to 59.9 years age group ($z = 0.00$, $p = 1.00$), or those aged 60 and above ($z = 0.00$, $p = 1.00$).

Discussion

The current study assessed the utility of the Static-99 and Static-99R with aged sexual offenders. Specifically, the aim was to ascertain the predictive validity of scores obtained using each measure, and to investigate protective effects of age in reducing recidivism. It was hypothesised that each measure would be able to differentiate between recidivists and non-recidivists similarly, and that scores would show moderate predictive validity with aged offenders, consistent with prior research. To explore the impact of the aging process on recidivism, the sample for this study was divided into two groups as it was hypothesised that sexual recidivism would be lower among sex offenders aged 60+ years than among those aged 50 to 59 years.

The results indicated no statistically significant differences between the predictive validity of Static-99 and Static-99R scores for sexual recidivism in the entire sample or within either of the two age sub-groups. Both tools had a moderate ability to distinguish recidivists from non-recidivists in the whole sample and a large ability to distinguish between recidivists and non-recidivists in the 50–59 sample. While the size of the AUC in the 60+ sample indicated a moderate classification ability, neither tool's score had a statistically significant relationship with recidivism. While this was undoubtedly a partial artefact of the necessarily smaller sample size, closer inspection of the results revealed that it also reflected the fact that some lower scoring individuals reoffended. These results provide support for the validity of the Static-99 and Static-99R scores as predictors of recidivism in an aged population, which is generally consistent with results of other Australian studies of sexual offenders of all ages (Reeves et al., 2018; Smallbone & Rallings, 2013). They

indicate the need for further validation of both tools in a larger sample of offenders aged 60+ as it is possible that the non-significant finding in this study was due to limited sample size.

While the AUC values were similar between the Static-99 and Static-99R, the spread of scores across the risk categories of each tool differed. The age weighting of the Static 99R had a clear impact on the spread of scores, with 83% of offenders placed in the three lowest risk categories on this tool (compared to 62% in these categories on the Static-99). Nonetheless, the distribution of recidivists across risk categories remained largely as expected for both tools. The high risk category of the Static-99 contained the greatest proportion of recidivists (41% of all recidivists), with a further 29% of recidivists classified as moderate–high. On the Static-99R, 47% of recidivists were classified as above or well above average risk and 41% as average risk.

The distribution of Static-99 scores in the 50–59 years age group reflected the spread of scores across the categories for the overall sample and resembles the distribution of recidivism across categories as detailed in the wider literature. However, this was not the case for the spread of scores across the categories with those in the 60+ years age group. In this age group, the distribution of scores across the categories appears to be relatively even on the Static-99. While the age correction applied on the Static-99R had a positive effect on the distribution of recidivism across the categories, the rates of recidivism in the very low and above average categories were similar. This effect could be the result of using a small sample and highlights the need for further research to occur in relation to offenders aged 60+ years, to clarify whether the age correction distributes recidivists across the risk categories as intended.

This study adds to the small pool of Australian research on the Static-99 and Static-99R. The three Australian published research articles using samples with wide age ranges obtained similar AUC values to those

identified during the current research (Allan et al., 2006; Reeves et al., 2018; Smallbone and Rallings, 2013), suggesting that both the Static-99 and Static-99R perform similarly in older samples as in younger samples. The AUC values obtained in this study were also consistent with international studies using samples of varied age, showing a moderate ability to discriminate between recidivists and non-recidivists using the Static-99 and Static-99R (see Boccaccini et al., 2017; de Vogel et al., 2004; Goncalves et al., 2020; Helmus, Hanson, et al., 2012).

The current study investigated the predictive validity of the Static-99 and Static-99R specifically in an aged sample. The results confirm that these tools demonstrate a moderate to large ability to discriminate between recidivists and non-recidivists, consistent with the findings of previous research (Boccaccini et al., 2017; Helmus, Hanson, et al., 2012; Reeves et al., 2018). Interestingly, in this study both tools performed better for the 50–59 years age group than the 60+ years age group, contrary to the findings of Hanson et al. (2010) and Hanson (2006); however, the limited sample size in the current study prevents any firm conclusions being drawn about this difference. The current results are also consistent with those of Nicholaichuk et al. (2014), who concluded that sexual recidivism rates among higher risk aged offenders are comparable to those of higher risk younger offenders (also see Hanson et al., 2014). While advanced age does appear to be a protective factor for some men who sexually offend, there is clearly a subset of individuals who continue to present an increased risk of sexual offending into old age. Further research focusing specifically on this subset and what dynamic risk factors are relevant to this group would be beneficial.

Limitations

The small sample size, particularly in regard to the group of sexual offenders aged 60 years or higher, limits the ability to make confident conclusions that could be generalised to a

larger population. The sample would also need to be much larger to obtain sufficient data to estimate absolute risk; therefore this study could not replicate the findings of the tools' authors (Phenix et al., 2016).

Conclusion

The current study adds to the literature regarding the utility of the Static-99 and Static-99R and to the body of research on Australian men who sexually offend. This study indicated that both the Static-99 and Static-99R have a moderate to large ability to discriminate between aged recidivists and non-recidivists overall (Helmus, Hanson, et al., 2012; Rice & Harris, 2014). For those aged between 50 and 59 years this effect was reduced, and was non-significant when specifically examining offenders aged 60 years and older. While age is a protective factor against sexual offending and reoffending, there remains a small group of men who continue to sexually offend into their late life. The scholarly literature is in its development stages in its ability to explain the recidivism of such individuals, beyond simply noting that sexual deviance is an important consideration.

Ethical standards

Declaration of conflicts of interest

Beatrice C. Raymond has declared no conflicts of interest

Troy E. McEwan has declared no conflicts of interest

Michael R. Davis has declared no conflicts of interest

Sophie Reeves has declared no conflicts of interest

James R. P. Ogloff has declared no conflicts of interest

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Allan, A., Dawson, D., & Allan, M.M. (2006). Prediction of the risk of male sexual reoffending in Australia. *Australian Psychologist*, 41(1), 60–68. doi:10.1080/00050060500391886
- Australian Bureau of Statistics. (2018). Retrieved from <http://abs.gov.au>
- Barbaree, H.E., Blanchard, R., & Langton, C.M. (2006). The development of sexual aggression through the life span: The effect of age on sexual arousal and recidivism among sex offenders. *Annals of the New York Academy of Sciences*, 989(1), 59–71. doi:10.1111/j.1749-6632.2003.tb07293.x
- Bauer, M., Fetherstonhaugh, D., Tarzia, L., Nay, R., Wellman, D., & Beattie, E. (2013). I always look under the bed for a man.' Needs and barriers to the expression of sexuality in residential aged care: The views of residents with and without dementia. *Psychology and Sexuality*, 4(3), 296–309. doi:10.1080/19419899.2012.713869
- Bauer, M., McAuliffe, L., Nay, R., & Chenco, C. (2013). Sexuality in older adults: Effect of an education intervention on attitudes and beliefs of residential aged care staff. *Educational Gerontology*, 39(2), 82–91. doi:10.1080/03601277.2012.682953
- 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. doi:10.1037/pas0000377
- Book, A.S., Starzyk, K.B., & Quinsey, V.L. (2001). The relationship between testosterone and aggression: A meta-analysis. *Aggression and Violent Behavior*, 6(6), 579–599. doi:10.1016/S1359-1789(00)00032-X
- Booth, B.D. (2016). Elderly sexual offenders. *Current Psychiatry Reports*, 18(4), 34. doi:10.1007/s11920-016-0678-1
- Chua, J., Cheung, G., Friedman, S.H., & Taylor, S. (2018). What are the characteristics of first time older sexual offenders? *International Psychogeriatrics*, 30(5), 647–662. doi:10.1017/S104161021700186
- Davis, M.R. (2017). *A clinician's guide to the Static-99 family of instruments*. Report prepared for the Specialised Offender Assessment and Treatment Service, Corrections Victoria, Department of Justice and Regulation, Melbourne, Australia.
- Davis, M.R., & Ogloff, J.R.P. (2008). Risk assessment and the dangerous offender. In K. Fritzon & P. Wilson (Eds). *Forensic and criminal psychology: An Australasian perspective*. North Ryde, NSW: McGraw Hill Australia.
- Dawes, J. (2009). Ageing prisoners: Issues for social work. *Australian Social Work*, 62(2), 258–271. doi:10.1080/03124070902803475
- De Vogel, V., De Ruiter, C., van Bee, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior*, 28(3), 235–251. doi:10.1023/B:LAHU.0000029137.41974.eb
- DeLong, E.R., DeLong, D.M., & Clarke-Pearson, D. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3), 837–845.
- Dworkin, E.R., Menon, S.V., Bystrynski, J., & Allen, N.E. (2017). Sexual assault victimization and psychopathology: A review and meta-analysis. *Clinical Psychology Review*, 56, 65–81. doi:10.1016/j.cpr.2017.06.002
- Fazel, S., Hope, T., O'Donnell, I., Piper, M., & Jacoby, R. (2001). Health of elderly male prisoners: Worse than the general population, worse than younger prisoners. *Age and Ageing*, 30(5), 403–407. doi:10.1093/ageing/30.5.403
- Fazel, S., Sjostedt, G., Langstrom, N., & Grann, M. (2006). Risk factors for criminal recidivism in older sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 18(2), 159–167. doi:10.1177/107906320601800204
- Goncalves, L.C., Gerth, J., Rossegger, A., Noll, T., & Endrass, J. (2020). Predictive validity of the Static-99 and Static-99R in Switzerland. *Sexual Abuse*, 32(2), 203–217. doi:10.1177/10790618821117
- Grann, M., & Langstrom, N. (2007). Actuarial assessment of violence risk: To weigh or not weight. *Criminal Justice and Behavior*, 34(1), 22–36. doi:10.1177/0093854806290250
- Grubin, D. (1998). *Sex offending against children: Understanding the risk* (Home Office Research Development and Statistics Directorate Research Findings, Police Research Series, Paper 99). London: Home Office, Policing and Reducing Crime Unit Research, Development and Statistics Directorate.
- Hanson, R.K. (1997). *The development of a brief actuarial scale for sex offender recidivism*. Ottawa, Ontario: Department of the Solicitor General of Canada. Retrieved from <http://www.publicsafety.gc.ca/res/cor/rep/cprmin-dex-eng.aspx#a97>

- Hanson, R.K. (2001). *Age and sexual recidivism: A comparison of rapists and child molesters* (Report No. 2001-01). Ottawa: Solicitor General Canada.
- Hanson, R.K. (2006). Does Static-99 predict recidivism among older sexual offenders? *Sexual Abuse: A Journal of Research and Treatment*, 18(4), 343–355. doi:10.1007/s11194-006-9027-y
- Hanson, R.K. (2018). Long-term recidivism studies show that desistance is the norm. *Criminal Justice and Behavior*, 45(9), 1340–1346. doi:10.1177/0093854818793382
- Hanson, R.K., & Bussiere, M.T. (1998). Predicting relapse: A meta-analysis of sexual offending recidivism studies. *Journal of Consulting and Clinical Psychology*, 66(2), 348–362. doi:10.1037/0022-006X.66.2.348
- Hanson, R.K., Harris, A.J.R., Helmus, L., & Thornton, D. (2014). High-risk sex offenders may not be high risk forever. *Journal of Interpersonal Violence*, 29(15), 2792–2813. doi:10.1177/0886260514526062
- Hanson, R.K., Helmus, L., & Thornton, D. (2010). Predicting recidivism amongst sexual offenders: A multi-site study of Static-2002. *Law and Human Behavior*, 34(3), 198–211. doi:10.1007/s10979-009-9180-1
- Hanson, R.K., & Morton-Bourgon, K.E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, 73(6), 1154–1163. doi:10.1037/0022-0006X.73.6.1154
- Hanson, R.K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, 24(1), 119–136. doi:10.1023/A:1005482921333
- Harris, A., Phenix, A., Hanson, R.K., & Thornton, D. (2003). Static-99 coding rules revised-2003. Retrieved from http://amyphenix.com/wp-content/uploads/2016/02/Static-99_Coding_Rules.pdf
- Helmus, L., Hanson, R.K., & Thornton, D. (2009). Reporting Static-99 in light of new research on recidivism norms. *The Forum*, 21(1), 38–45.
- Helmus, L., Hanson, R.K., Thornton, D., Babchishin, K.M., & Harris, A.J.R. (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. doi:10.1177/0093854812443648
- 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: A Journal of Research and Treatment*, 24(1), 64–101. doi:10.1177/1079063211409951
- Lennings, C., Seidler, K., Heard, R., Collins, E., & Nasr, R. (2011). Age and the Static-99R. *Sexual Abuse in Australia and New Zealand*, 3, 34–41.
- Lopez, K.L., Boer, D.P., Kirby, E.M., & Davis, M.R. (2017). Assessing risk in a law enforcement context: The SHARP sexual risk protocol. *The Association for the Treatment of Sexual Abusers Forum*, 30, 1–9.
- Mann, R.E., Hanson, R.K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, 22(2), 191–217. doi:10.1177/1079063210366039
- Maschi, T., Dennis, K.S., Gibson, S., MacMillan, T., Sternberg, S., & Hom, M. (2011). Trauma and stress among older adults in the criminal justice system: A review of the literature with implications for social work. *Journal of Gerontological Social Work*, 54(4), 390–424. doi:10.1080/01634372.2011.552099
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62(4), 783–792. doi:10.1037/0022-006X.62.4.783
- Mossman, D. (2013). Evaluating risk assessments using receiver operating characteristic analysis: Rationale, Advantages, insights, and limitations. *Behavioral Sciences & the Law*, 31(1), 23–39. doi:10.1002/bsl.2050
- Nicholaichuk, T.P., Olver, M., Gu, D., & Wong, S.C. (2014). Age, actuarial risk, and long-term recidivism in a national sample of sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 26(5), 406–423. doi:10.1177/1079063213492340
- Phenix, A., Doren, D., Helmus, L., Hanson, R.K., & Thornton, D. (2008). Coding rules for Static-2002. Retrieved from <http://www.static99.org/pdffdocs/static2002codingrules>
- Phenix, A., Fernandez, Y., Harris, A.J.R., Helmus, M., Hanson, R.K., & Thornton, D. (2016). Static-99R coding rules revised – 2016. Retrieved from http://static99.org/pdffdocs/Coding_manual_2016_InPRESS

- Phenix, A., Helmus, L., & Hanson, R.K. (2009). Static-99R and static-2002R evaluator's workbook. Retrieved from http://www.static99.org/pdfdocs/Evaluators_Workbook_2016-10-19
- Reeves, S.G., Ogloff, J.R.P., & Simmons, M. (2018). The predictive validity of the Static-99, Static-99R, and Static-2002/R: Which one to use? *Sexual Abuse: A Journal of Research and Treatment*, 30(8), 887–907. doi:10.1177/1079063217712216
- Rettenberger, M., Haubner-Maclean, T., & Eher, R. (2013). The contribution of age to the Static-99 risk assessment in a population-based prison sample of sexual offenders. *Criminal Justice and Behavior*, 40(12), 1413–1433. doi:10.1177/0093854813492518
- 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(5), 615–620. doi:10.1007/s10979-005-6832-7
- Rice, M.E., & Harris, G.T. (2014). What does it mean when age is related to recidivism among sex offenders? *Law and Human Behavior*, 38(2), 151–161. doi:10.1037/lhb000052
- Smallbone, S., & Rallings, M. (2013). Short-term predictive validity of the Static-99 and Static-99R for indigenous and nonindigenous Australian sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 25(3), 302–316. doi:10.1177/109063212472937
- Strub, D.S., Douglas, K.S., & Nicholls, T.L. (2014). The validity of version 3 of the HCR-20 Violence Risk Assessment scheme amongst offenders and civil psychiatric patients. *International Journal of Forensic Mental Health Services*, 13(2), 148–159. doi:10.1080/14999013.2014.911785
- Williams, B.A., Stern, M.F., Mellow, J., Safer, M., & Greifinger, R.B. (2012). Aging in correctional custody: Setting a policy agenda for older prisoner health care. *American Journal of Public Health*, 102(8), 1475–1481. doi:10.2105/AJPH.2012.300704
- Wollert, R., Cramer, E., Waggoner, J., Skelton, A., & Vess, J. (2010). Recent research ($N=9,305$) underscores the importance of using age-stratified actuarial tables in sex offender risk assessments. *Sexual Abuse: A Journal of Research and Treatment*, 22(4), 471–490. doi:10.1177/1079063210384633