Research Article

Check for updates

Predictive Validity of Stable-2007 in Incarcerated Samples

Jan Looman [1], Joshua Goldstein [2], Brian R. Abbott [3], Jeff Abracen [3]

[1] *Forensic Behaviour Services, Kingston, Ontario, Canada.* [2] *Toronto South Detention Centre, Toronto, Canada.*
[3] *Independent Practice, San Jose, CA, USA.*

Abstract

Some are unclear whether risk assessment instruments, specifically dynamic risk instruments, have demonstrated utility in the risk estimation, treatment recommendations, and monitoring change over time in men at risk for or under sentence of Indeterminate Detention (ID) for sexual offenses. We compare two datasets, the first consisting of individuals representing a routine sample of persons convicted of a sexual offense and the second of men representative of a high risk/needs sample. These two distinct samples (n = 442, mean Static-99R score = 2.4; n = 168, mean Static-99R score 4.5) were then also scored on the Stable-2007. For both groups this scoring occurred in an institutional setting. The Stable-2007 predicted sexual recidivism in Sample 1 independently and in conjunction with the Static-99R. In the high-risk sample the results were the same. In both samples a compound outcome variable (Sexual + Violent reoffense) was also calculated with the Stable-2007 predicting the compound outcome variable in Sample 1 but not Sample 2. This is interesting in that it suggests that the Stable-2007 assesses constructs specific to sexual re-offense in higher risk offenders and not general traits of violence or common anti-social behaviour. Limitations and directions for further research are discussed.

Keywords

sexual offenders, dynamic risk assessment, incarcerated samples

Non-Technical Summary

Background

The Static-99R and Stable-2007 are widely used actuarial instruments in the domain of risk assessment for men who have committed sexual offences. They have been found to be valid indicators of risk in countries around the world, in a variety of settings.

Why was this study done?

While the instruments are widely used in incarcerated populations, the Stable-2007 was normed on a community sample, and there has been concern raised regarding the validity of this instrument for men who have served lengthy prison sentences.

What did we find?

The Static-99R and the Stable-2007 were scored on two independent samples of men incarcerated for having committed sexual offences. One sample (n = 442, mean Static-99R score = 2.4) was a Routine sample, while the second (n = 168, mean Static-99R score 4.5) was comprised of High Risk/Needs offenders. Results showed that in both samples the Static-99R added incrementally to the predictive validity of the Static-99R for the prediction of sexual recidivism. For serious (i.e., sexual + violent) recidivism the Stable-2007 was only useful in the Routine sample.

What do these findings mean?

The results indicate that the Static-99R and the Stable-2007 can be used in the prediction of sexual, but not serious reoffence for men who have been incarcerated for lengthy periods of time.

Highlights

- The Stable-2007 can be validly scored on incarcerated samples.
- The predictive validity of the Stable-2007 scored on incarcerated samples is the same as found in the developmental sample.
- The Stable-2007 adds incremental value to the prediction obtained with the Static-99R alone.
- Length of sentence does not appear to be related to the predictive validity of the Stable-2007.

The assessment of risk for sexual recidivism serves an important function when considering public safety. Efforts to increase the accuracy of prediction through the use of static factors appears to have reached a ceiling, as most extant measures are approximately equal in effectiveness and efforts to improve on these predictive values have not been successful (Hanson & Morton-Bourgon, 2009).

PsychOpen GOLD

Over the past decade efforts have increasingly turned to dynamic measures (Mann et al., 2010) to improve the predictive validity of risk assessments. Dynamic instruments have the added advantage of assessing putative treatment targets (Mann et al., 2010) and thus have greater potential to measure change over time including those changes subsequent to treatment or other intervention. This hypothesis is supported in the meta-analysis of Hanson and Morton-Bourgon (2009) who reported that measures that contained dynamic factors (treatment targets) were "more accurate for the prediction of sexual recidivism than were measures based on static, historical factors" (p. 7) alone.

One commonly used measure of dynamic risk is the Stable-2007 (Fernandez et al., 2014; Hanson & Harris, 2001). Research has demonstrated that this instrument can be scored reliably and predicts sexual recidivism (Hanson et al., 2015). However, this instrument was originally developed using only a community sample, thus its validity with an incarcerated population must be determined.

History of the Stable-2007

Hanson and Harris (2001) developed the Sex Offender Need Assessment rating (SONAR) using the dynamic factors identified in the Dynamic Predictors Project (Hanson & Harris, 2000), to evaluate risk change in sex offenders. The SONAR consisted of five stable and four acute items. Although the measure distinguished between recidivists and non-recidivists (r = .43; ROC area under the curve, AUC = .74) and found general self-regulation deficits to have the strongest effect on sexual recidivism (r = .41), the researchers suggested caution when interpreting the results since the same data set was used to develop and test the items on the measure.

The Dynamic Supervision Project (DSP) was launched as a more comprehensive, prospective study, led by Hanson and colleagues (2007). The DSP assessed the risk of 997 men on community supervision for sexual offending from all Canadian provinces and territories as well as Alaska and Iowa, USA. The project used the Static-99 (Hanson & Thornton, 2000) and the Stable-2000/Acute-2000 (modified and renamed version of SONAR; Hanson & Harris, 2001) to assess and predict risk by trained professionals. Although the Stable-2000 was positively correlated with sexual recidivism (AUC = .64), it did not increase the accuracy of predicting sexual recidivism beyond the accuracy of the Static-99 (β = .085, SE = .058, p = .141).

The results of the above analyses led to changes made to improve incremental validity and scoring in the development of the Stable-2007. First, the removal of all attitude items since they yielded the lowest relationship with recidivism (AUC = .47 to .54). Second, the scoring criteria for the following three items were refined: (a) deviant sexual interests (b) lovers/intimate partners (c) emotional identification with children. Lastly, the total score criteria were simplified. As a result, the Stable-2007 provided incremental validity to the prediction of sexual recidivism (β = .059, SE = .030, p = .049) and all other types of recidivism.

PsychOpen GOLD

Saum (2007), in a study conducted within the Department of Corrections in North Dakota, examined the predictive validity of the Static-R and the Stable-2000 in a sample of 175 persons convicted of a sexual offense. Overall, 35.7% of the sample sexually re-offended over an average of 42 months of follow-up. Both the Static-99 and the Stable-2000 significantly predicted sexual recidivism on their own (AUC = .72 and .68 respectively). In addition, analyses indicated that although the Stable-2000 failed to add to the prediction through a Cox Regression analysis; however, the combination of the Static-99 and the Stable-2000 did predict recidivism significantly.

Eher et al. (2012) examined the predictive and incremental validity of the Stable-2000/2007 in a sample of 263 sex offenders released into the community in Austria. Eher et al. used the Static-99, the Sex Offender Risk Appraisal Guide (SORAG: Quinsey, Harris, Rice, & Cormier, 1998), and both the Stable-2000 and Stable-2007. The average follow-up was 6.4 years after release resulting in 10.3% (n = 27) of the sample being reconvicted of a sexual crime, 24.3% (n = 64) reconvicted for a violent crime, and 35.9% (n = 104) reconvicted for a general crime. Analysis indicated that the Stable-2007 was more accurate than the Stable-2000 in predicting any type of recidivism but less accurate in predicting violent recidivism than the SORAG. After controlling for Static-99, the Stable-2007 provided incremental validity for the prediction of violent and general recidivism but not for sexual recidivism. After controlling for SORAG, the Stable-2007 only provided incremental validity for the prediction of sexual recidivism.

In a second study Eher, Olver, Heurix, Schilling, and Rettenberger (2015) followed a group of 189 pedophilic men convicted of child sex offenses. Some of these men were included in the sample above, however the current study added to that sample and extended the follow-up period. They reported predictive accuracy for the Static-99R, Stable-2007, the combination of the Static-99R and Stable-2007 as well as the Violence Risk Scale-Sexual Offense version (VRS-SO; Wong et al., 2003, 2017). They found that the Static-99R predicted sexual recidivism; however, neither the Stable-2007 nor the combination of the Static-99R and Stable-2007 were predictive. The VRS-SO was predictive at a non-significantly higher rate than the Static-99R. The Dynamic factor scale from the VRS-SO on its own was also predictive of sexual recidivism.

Sowden and Olver (2017) also reported on a study in which the Stable-2007 was scored on a group of 180 incarcerated males who committed sexual offenses, who were involved in a high-intensity sexual offender treatment program. The VRS-SO and Static-99R were also scored on each of the participants. The authors reported that the Stable-2007 did not significantly predict increased sexual violence, but its pre- and posttreatment scores consistently significantly predicted increased general, nonsexual violent, and any violent recidivism. They also found that the Stable-2007 pretreatment and posttreatment scores significantly uniquely predicted nonsexual violent and any violent recidivism after controlling for the Static-99R; however, the Stable-2007 did not

PsychOpen GOLD

significantly add to the prediction of sexual violence and only posttreatment scores uniquely predicted general recidivism.

A recent meta-analysis (Brankley et al., 2021) found that, using 12 unique samples (*N* = 6955) the Stable-2007 was significantly and incrementally related to sexual recidivism, violent nonsexual recidivism and any crime, Exp(β) = 1.12; 1.09 and 1.10 respectively. They also found that the Stable-2007 added incrementally to the prediction of recidivism; with the fixed effect weight hazard ratio of 1.07, 95% CI [1.04, 1.09] for the Stable-2007 after controlling for the effect of the Static-99R. While the authors examined the influence of a number of moderator variables, they did not examine the effect of incarceration in the validity of the results.

One issue which has not been sufficiently addressed in the extant research is the extent to which the Stable-2007 is valid with offenders who have served lengthy sentences, given that the instrument was developed in a population of men under supervision is the community. This issue is an important one, given that in some jurisdictions (e.g., the United States), offenders may serve lo ng periods of incarceration prior to release, thus it is important to determine whether the items included in risk assessment tools are relevant for such offenders.

Thus, the purpose of this research was to examine the predictive validity in two samples of incarcerated persons convicted of a sexual offense. One sample consisted of 442 men assessed and/or treated in the Ontario Region of the Correctional Service of Canada while the second sample consisted of 168 offenders assessed and/or treated at the Regional Treatment Centre (Ontario) prior to 1992.

Method

Measures

Stable-2000

The Stable-2000 (Hanson et al., 2007) is a 16-item mechanical risk tool (i.e., items summed to allow the user to rank the offender in terms of risk, but there are no associated risk estimates) assessing dynamic risk factors among adult male sex offenders. The are organized into six subsections: significant social influences, intimacy deficits, sexual self-regulation, attitudes, general self-regulation, and cooperation with supervision; and are assessed using a 3-point rating system, where 0 refers to no problem, 1 refers to some concern/slight problem, and 2 refers to present/definite concern. Total scores on the Stable-2000 are calculated by summing the highest score on each subsection, resulting in total scores ranging from 0 to 12, where scores of 0 to 4 indicate low risk, 5 to 8 indicate moderate risk, and scores of 9 or higher indicate high risk.

PsychOpen GOLD

Stable-2007

The Stable-2007 (Fernandez et al., 2014; Hanson et al., 2007) is an empirical actuarial risk tool assessing dynamic risk factors among men with a history of sexual offending. It was developed by revising the Stable-2000 scale based on results from the Dynamic Supervision Project research (Hanson et al., 2007). The attitudes domain was removed and the scoring instructions for lovers/relationship stability and deviant sexual interest were revised. The emotional identification with children item was restricted to apply only to offenders with at least one victim less than 14 years old. Thus, the Stable-2007 has 13 items organized into five subsections (significant social influences, intimacy deficits, sexual self-regulation, general self-regulation, and cooperation with supervision) and the total score is simply the sum of the item scores. Using the total score, offenders can be assigned to one low risk (0-3), moderate risk (4-11), and high risk (12+) categories.

For those individuals for whom the Stable-2000 was scored the necessary information to convert the score to the Stable-2007 was gathered from file information. While inter-rater reliability was not calculated specifically for this study, since it is a retrospective analysis of data gathered for evaluation purposes, previous research conducted at the Millhaven Assessment Unit (Fernandez, 2008) indicated high levels of interrater reliability when scoring Stable-2007. Fernandez had 55 persons convicted of sexual offenses assessed on the Stable-2007 by two independent raters and obtained ICCs = .92 for total score on the Stable-2007. The ICCs for the individual Stable-2007 items ranged from .56 to .91 with a median value of .83.

Static-99R

The Static-99R (Hanson & Thornton, 2000; Helmus et al., 2012) is an empirical actuarial risk assessment tool designed to assess risk for sexual recidivism in adult males with a history of sexual offending. The Static-The scale is composed of 10 items assessing criminal history, victim characteristics, age, and relationship history. It has been found to have moderate predictive accuracy for sexual recidivism (AUC = .69; Helmus et al., 2012). Previous research by our group (Looman & Abracen, 2010) has indicated acceptable interrater reliability in scoring the Static-99R ($r = .84$).

Participants

Sample 1 consisted of 442 persons convicted of sexual offenses assessed and/or treated in the Ontario Region of the Correctional Service of Canada. The sample consisted of two groups: For 376 of the men, the Static-99/99R and the Stable-2000/2007 were scored as part of specialized sexual offense assessment completed within three to five months of their entry to the Correctional Service of Canada at the Millhaven Assessment Unit. Of these men, 247 went on to complete a sexual offense treatment program during their sentence while 43 refused treatment, 22 were discharged from treatment prior to completion (typically for failure to comply with program rules), and for 24 there

PsychOpen GOLD

was no evidence that they were offered treatment prior to release. Data concerning the treatment status for the 40 remaining individuals was not available. Information used to score the instruments included police reports and court documents related to their trial/sentencing and when available pre-sentence reports, psychological/psychiatric assessments completed prior to sentencing and any documents available for those who had previous sentences.

The remaining 66 individuals in Sample 1 were assessed as part of the pre-treatment assessment for a sexual offense treatment program (RTCSOTP; Abracen & Looman, 2015). This group consisted of men who entered the correctional system prior to the use of the Stable on intake; however, the Stable 2000/2007 was scored as part of the pre-treatment assessment for the sex offender program. Of these men, three were assessed only; 45 completed treatment; 14 were discharged from treatment; and 4 withdrew from treatment. Information used for scoring the instruments for these men included the sources noted above, as well as any information which became available while serving their sentence before entering treatment. This may have included reports from other programs, reports regarding behavior during institutional employment, and so on. These men were assessed between 2000 and 2011.

While approximately 15% of the sample was derived from a High Intensity Sexual Offender Treatment Program, most the men included in this analysis were consecutive admissions to the Federal Prison system in Ontario, thus the sample can be considered to represent a routine sample of men convicted of sexual offenses (Hanson et al., 2015).

Sample 2 was derived from a group of 506 men who committed sexual offences treated or assessed at the RTC(O) prior to 1992. Of these 334 were released and available for follow-up. For this sample, the Static-99R and Stable-2007 were scored from assessment/pretreatment file information. There was sufficient information to score the Static-99R on 326 cases, and for the Stable- 2007 for 168. Therefore, there were 168 cases with complete information.

Recidivism

Recidivism data for both samples were collected from official criminal records maintained by the Royal Canadian Mounted Police (RCMP). The Fingerprint Service (FPS) sheets (official Canadian criminal history) for each offender were obtained electronically and new convictions were coded according to the Cormier–Lang system (Harris, Rice, Quinsey, & Cormier, 2015). Violent convictions were those convictions listed as Group 1 offenses according to the Cormier–Lang system (e.g., assault, robbery with violence). New sexual offenses were those offenses clearly of a sexual nature according to the recorded conviction (e.g., sexual assault, gross indecency, invitation to sexual touching). Harris et al. (2015) make the case that due to plea bargaining and under-reporting of sexual recidivism, the composite outcome of violent (including sexual) recidivism may be a more valid outcome for recidivism research with persons convicted of sexual offenses.

PsychOpen GOLD

Where possible, both outcomes were examined in the current study. Outcome data was collected during the summer of 2014. The average follow-up time for Sample 1 was 6.1 (*SD* = 2.9) years (6 days to 12.9) years, while the average follow-up time for Sample 2 was 22.4 years (*SD* = 5.1).

Analytic Plan

For each sample we conducted a series of analyses, starting with calculating ROC AUCs for the Static-99R and Stable-2007 to examine the ability of the instruments to discriminate recidivists from non-recidivists in the current sample. This was followed by Cox Regression analysis to evaluate the extent to which the Stable-2007 discriminates recidivists from non-recidivists over time, first on its own and then together this the Static-99R, to determine incremental validity. Finally, we examined calibration with normative data via E/O indices.

Results

Sample 1

Descriptive Statistics

The average age at release for Sample 1 was 43.3 (*SD* = 13.1) years. The average Static-99R score was 2.4 (*SD* = 3.2), while the average Stable-2007 score was 8.8 (*SD* = 5.1); placing this group of individuals, on average in the Level III (average) range (Brankley et al., 2017). The Pearson correlation coefficient between the total Static-99R score and the total Stable-2007 score was .38, *p* = .0001. In terms of offense type, 164 (39.0%) of the individuals had adult (i.e., 16 years of age or higher) victims, while 37 (8.8%) had extra-familial victims age 13 to 15 inclusive. Ninety-two (31.9%) had extra-familial victims 12 years of age or younger while 40 (9.5%) had victims in more than one age group. Finally, 86 (20.4%) had victims within their biological family. Offense type information was missing for 21 individuals. The average time served prior to release for this sample was 35.5 months (*SD* = 27.9, range 10 to 342 months).

In terms of recidivism, overall, 7.3% of the sample was detected to have sexually recidivated over an average 5.8 (*SD* = 2.9) year follow-up. For the outcome of violent+sexual recidivism the corresponding proportion was 16.6% over an average of 5.5 (*SD* = 3.1) years of follow-up.

Discrimination

Examining the relationship of the Static-99R and Stable-2007 to recidivism using the AUC statistic indicated that both measures were significantly related to sexual and sexual + violent recidivism. For this analysis a fixed follow-up of 5 years was used. The AUC for the Static-99R for sexual recidivism was .69, 95% CI [.54, .83] and for the combined

PsychOpen GOLD

sexual + violent recidivism outcome AUC = .72, 95% CI [.64, .79]. For the Stable-2007, the corresponding values for recidivism type were AUCs = .77, 95% CI [.65, .90] and .67, 95% CI [.58, .77].

Table 1 displays the results of Cox regression for the Stable-2007 on its own, and with the Static-99R as the predictor for sexual and sexual + violent recidivism. For sexual recidivism, the Stable-2007 was a significant predictor, $\chi^2(1) = 20.78$, $p < .001$. The *Exp(B)* indicates that the probability of re-offense increased by 17% with each one-point increment in the Stable-2007 score. When the Static-99R was entered on the first block and the Stable-2007 was entered on the second block, as displayed in the Table the Static-99R was a significant predictor of recidivism on its own. When the Stable-2007 was added to the equation on the second step it added significantly to the prediction, $\Delta\chi^2(1) = 7.23$, $p = .007$.

Table 1

Sample 1: Cox Regression Analysis for Stable-2007 and Recidivism

Regression model	*B*	*SE*	Wald	*df*	*p*	Exp(*B*)	95% CI Exp(*B*)
Sexual Recidivism							
Step 1							
Stable-2007	.16	.04	19.47	1	< .001	1.17	1.09 − 1.25
Step 2							
Static-99R	.18	.06	8.16	1	.004	1.20	1.06 − 1.36
Stable-2007	.10	.04	6.85	1	.009	1.11	1.03 − 1.20
Sexual + Violent Recidivism							
Step 1							
Stable-2007	.12	.02	27.58	1	< .001	1.13	1.08 − 1.18
Step 2							
Static-99R	.17	.04	18.39	1	.001	1.19	1.10 − 1.29
Stable-2007	.07	.03	7.41	1	.006	1.07	1.02 − 1.13

For sexual + violent recidivism once again the Stable-2007 is a significant predictor of recidivism $\chi^2(1) = 28.89$, $p = .00001$. The *Exp(B)* indicates that with each one-point increment in the Stable-2007 score, the probability of serious re-offense increases by 13%. When the Static-99R was entered on its own it was a significant predictor. When the Stable-2007 was added in the second step the effect was significant, $\Delta\chi^2(1) = 7.58$, $p = .006$; for the model $\chi^2(2) = 51.73$, $p < .001$.

Calibration

Table 2 displays the recidivism rates for individuals grouped according to the guidelines provided by Brankley et al. (2017). For example, the Low Priority group is made up of men who score in the low-risk group on the Static-99R and in Low or Moderate risk groups on the Stable-2007. The Very High Priority group is made up of men who score in the High range on both instruments. To allow for direct comparison to the data provided by Brankley et al., recidivism rates for a fixed 5-year follow-up are provided. For this analysis, data for 263 individuals were available. In addition, E/O indices with 95% confidence intervals were computed using procedures described by Hanson (2017) for sexual recidivism. When the 95% CI for the E/O index does not include 1.0, the index is considered statistically significant at $p < .05$ (Hanson, 2017), which reflects poor calibration between the expected and observed recidivism rates. Examining the 95% confidence intervals for the sexual recidivism E/O analysis all confidence intervals included 1.0, signifying no statistically significant differences between the expected and observed recidivism rates for the individual risk levels (Hanson, 2017).

Table 2

Sample 1: E/O Index and 95% CIs for Static-99R/Stable-2007 Priority Categories by Recidivism Type

Priority Category	n	Observed recidivists n (%)	Expected recidivists n (%)	E/O Index	95% CI
Below Average	18	1 (5.6)	0.95 (5.3)	0.95	0.13 − 6.77
Average	140	5 (3.6)	10.5 (7.5)	2.10	0.87 − 5.05
Above Average	54	4 (7.4)	7.3 (13.6)	1.84	0.69 − 4.89
Well-Above Average	51	14 (27.5)	13.7 (26.8)	0.98	0.58 − 1.65
Total	263	24 (9.1)	28.4 (10.8)	1.18	0.79 − 1.76

Sample 2

Descriptive Statistics

The average age at release for Sample 2 was 35.5 (SD = 8.7) years. The average Static-99R score was 4.5 (SD = 2.2), while the average Stable-2007 score was14.5 (SD = 4.3); placing this group of individuals, on average in the Level IVb (well above average) range (Brankley et al., 2017). The Pearson correlation coefficient between the total Static-99R score and the total Stable-2007 score was .57, $p < .001$. In terms of offense type, 57 (37.7%) of the individuals had adult (i.e., 16 years of age or higher) victims, while 6 (4.0%) had extra-familial victims age 13 to 15 inclusive. Twenty-one (13.9%) had extra-familial child victims 12 years of age or younger while 67 (44.4%) had victims in more than one age group. Offense type information was missing for 14 individuals. While information for

PsychOpen GOLD

mental health diagnoses was not available for this sample whether or not they were being followed by a psychiatrist was available for 141 of these men (i.e., 84.9%). Of these 141 men, 78 (55.3%) were receiving regular psychiatric care.

In terms of recidivism, overall, 45.5% of the sample was detected to have sexually recidivated over an average 14.2-year follow-up (*SD* = 9.5; range 2 days to 28.6 years). For the outcome of violent + sexual recidivism the corresponding proportion was 63.0% over an average of 16.9 years of follow-up (*SD* = 10.0, range 2 days to 36.2 years).

Discrimination

Examining the relationship of the Static-99R and Stable-2007 to recidivism using the AUC statistic for a fixed 20-year follow-up indicated that both measures were significantly related to sexual and sexual + violent recidivism. The AUC for the Static-99R for sexual recidivism was .67, 95 CI% [.59, .75] and for the combined sexual + violent recidivism outcome it was .68, 95% CI [.59, .77]. For the Stable-2007, the corresponding AUCs = .68, 95% CI [.60, .77] and .65, 95% CI [.55, .74].

Table 3 displays the results of Cox regression survival analysis with the Static-99R and Stable-2007 as predictors for sexual and sexual + violent recidivism. For sexual recidivism, the Stable-2007 on its own was a significant predictor, $\chi^2(1)$ = 19.97, p < .001. The *Exp(B)* indicates that the probability of re-offense increases by 16% with each one-point increment in the Stable-2007 score. When the Static-99R was entered with the Stable-2007 the Stable-2007 added significantly to the prediction of outcome, $\Delta\chi^2(1)$ = 7.23, p = .007.

Table 3

Sample 2: Cox Regression Analysis for Stable-2007 and Recidivism

Regression model	B	SE	Wald	df	p	Exp(B)	95% CI Exp(B)
Sexual Recidivism							
Step 1							
Stable-2007	.15	.03	19.88	1	< .001	1.16	1.09 − 1.24
Step 2							
Static-99R	.20	.08	6.30	1	.012	1.20	1.04 − 1.42
Stable-2007	.10	.04	6.88	1	.009	1.11	1.02 − 1.19
Sexual + Violent Recidivism							
Step 1							
Stable-2007	.05	.03	3.45	1	.063	1.05	0.99 − 1.10
Step 2							
Static-99R	.11	.06	3.31	1	.069	1.11	0.99 − 1.26
Stable-2007	.02	.03	0.31	1	.57	1.02	0.96 − 1.08

For sexual + violent recidivism once again the Stable-2007 by itself approaches significance as a predictor of recidivism $\chi^2(1) = 3.46$, $p = .063$. The *Exp(B)* indicates that with each one-point increment in the Stable-2007 score, the probability of serious re-offense increases by 5%. When the Static-99R and Stable-2007 are entered the effect was not significant, $\Delta\chi^2(1) = 0.32$, $p = .57$; for the model $\chi^2(2) = 6.52$, $p = .038$.

Calibration

Table 4 displays the recidivism rates for individuals grouped according to the guidelines provided by Brankley et al. (2017) The groups represent combinations of the Static-99R and Stable-2007 risk groups. To allow for direct comparison to the data provided by Hanson and Helmus (2013), recidivism rates for a fixed 5-year follow-up are provided. For this analysis data on 163 individuals were available. Note that due to the high-risk nature of the sample none of the individuals fell into categories below Average risk.

Table 4

Sample 2: E/O Index and 95% CIs for Static-99R/Stable-2007 Priority Categories by Recidivism Type

Priority Category	N	Observed Recidivists n (%)	Expected Recidivists n (%)	E/O Index	95% CI
Average	23	3 (13.0)	1.7 (7.5)	0.58	0.19, 1.78
Above Average	33	4 (12.1)	4.5 (13.6)	1.13	0.42, 2.99
Well-Above Average	107	34 (31.8)	28.7 (26.8)	0.84	0.60, 1.18
Total	163	41 (25.2)	17.6 (10.8)	0.43	0.32, 0.58

Examining the 95% confidence intervals for the sexual recidivism E/O analysis in Table 4 indicates that all of the intervals, with the exception of that for the Total recidivism rate, include the value of 1.0. This signifies no statistically significant differences between the expected and observed recidivism rates for the individual risk levels (Hanson, 2017); however, for the total recidivism rate the observed recidivism rate is over twice the expected. This appears to be due to the lack of low-risk men in this sample.

Discussion

An interesting comparison is presented here – one sample being an average group of persons convicted of a sexual offense versus a high-risk comparison group, a group sent for assessment or treatment at a Regional Treatment Centre (RTC). In Canada, a federal RTC is generally a multi-level security institution, (low or medium security through high security and special units) operating within an institution as a legally acknowl-

edged hospital generally offering psychiatric treatment and assessment that assists the Correctional Service of Canada by providing special services to people transferred from their home institution for assessment and treatment of a mental disorder or a physical condition that limits their ability to benefit from the general correctional regime. This is quite a high risk sample as noted by the initial number of men who populated this cohort (N = 506), yet after well more than 20 years later only 334, about two-thirds, had been released and were available for follow-up and by their high Static-99R score (4.5, SD = 2.2) when compared to the standard intake sample (Sample 1) who had an average Static-99R score of 2.4 (SD = 3.2). Looking at Sample 2, we see not only an elevated Static-99R score but also a high frequency of mental health concerns; with 55% of the sample receiving regular psychiatric care. By way of comparison, we can compare the Static-99R scores of these two samples to the expected values based upon those seen in the coding instructions for the Static-99R (Phenix et al., 2016). A score of 2 on the Static-99R would yield a 5-year prediction of sexual recidivism of 5.6%, 95% CI [4.8, 6.5] and we report observed sexual recidivism for this group at 7.3% over an average 5.8-year follow-up. It is important to note that the Static-99R estimate is censored at 5 years while many in sample one would have been followed for much longer. For Sample 2, the high-risk group, the average Static-99R score was 4.5 and Static-99R instructions would expect a 5-year recidivism rate of 11.0%, 95% CI [10.0, 12.1] while observed for Sample 2 was 45.5% over 14.2 years. Yet, as demonstrated in the results, the Stable-2007 provided additional predictive power.

Secondly, as shown in Harris (2021, this section) Table 3 – in 946 high risk men, the majority of whom were involved in or being considered for either a Dangerous Offender (DO) or Sexually Violent Predator/Person (SVP) the average Static-99R score was 4.22. Therefore at least in terms of risk this sample is comparable to DO and SVP samples, given consideration of the factors noted under Study Limitations.

The results of this study show that Stable-2007 scores reliably predict sexual recidivism for both a sample of incarcerated average risk men with sexual offenses and a high-risk sample. Further useful information is provided in the analysis of Sample 2 when the Stable-2007 predicts for person convicted of a sexual offense but this predictive ability fades when the compound variable of sexual + violent recidivism is examined as seen in Table 4. This is important information as it speaks to the specificity of Stable-2007, that it taps sexual offense specific constructs and not more generic constructs of inter-personal violence and criminality.

The current research indicates that although the Stable-2007 was developed and normed on a community-based sample, it has the ability to predict recidivism in an incarcerated sample. The findings also reinforce the importance of utilizing both static and dynamic risk measures to estimate future risk. Although static measures of risk provide generally robust predictions of future risk, the addition of a measure of dynamic risk, such as the Stable-2007, will enhance the accuracy of such estimations. These results

indicated that those who conduct SVP assessment who have been resultant to adopt the Stable-2007 in their practice have no reason for concern.

Study Limitations

The principle limitation to this data is the age of Sample 2, all entering the system prior to 1992. While some believe it is important to have the longest possible follow-up, these efforts are confounded by change in correctional climate and practice, including loss of data partially due to the transition from paper-based to digital records. This is compounded by historical record keeping practices that would "cull" inactive paper records after a given period of inactivity due to space and personnel limitations. It must be kept in mind that of the 506 men who were assessed or treated, complete data was only available for 168, approximately one-third of the identified sample. As seen in other work within mental health institutions (Saum, 2007) there is strong reason to believe that those not included in the present analysis would have had a higher percentage of men who would not have reoffended. Recidivism outcomes from Sample 2 should be interpreted with this consideration in mind.

In addition, institutional treatment in the 1990's would have focused almost totally on developing a personal crime cycle as part of a relapse prevention approach. Both are based upon avoidance techniques that have not been seen to be as promising as they were once presumed to be. While it has to be quickly added that current, more approach goal oriented, techniques may prove in time to be more effective – this only goes to underscore the point that research of this nature is presenting a picture from a particular timeframe that has variable relevance to current practices.

Future Research

Research efforts in this area would greatly benefit from multi-jurisdictional research agreements allowing follow-up of DO and SVP such that the small number of men released in each jurisdiction could be compiled into composite datasets, followed for fixed time intervals, and then reliably re-assessed as to recidivism status. The large problem here is that this process would take not only researcher sophistication and energy (neither of these factors are lacking) but it would take some political will to allow for inter-jurisdictional cooperation. Given the current climate it is unfortunate that the conditions necessary for advanced research cooperation on this topic are highly unlikely. This being said, larger jurisdictions could easily implement one of the standard dynamic risk assessment instruments throughout their jurisdiction and develop and publish in-house norms and guidelines that would test, replicate, and ultimately prove the utility of dynamic assessment in high-risk high-needs samples.

Data Availability: The data used in this paper will be made available to others. Please contact the first author.

References

Abracen, J., & Looman, J. (2015). *Treatment of high-risk sexual offenders: An integrated approach.* Wiley-Blackwell.

Brankley, A. E., Babchishin, K. M., & Hanson, R. K. (2021). Stable-2007 demonstrates predictive and incremental validity in assessing risk-relevant propensities for sexual offending: A meta-analysis. *Sexual Abuse, 33*(1), 34-62. https://doi.org/10.1177/1079063219871572

Brankley, A. E., Helmus, L. M., & Hanson, R. K. (2017). *Stable-2007 Evaluator Workbook: Revised 2017.* Public Safety Canada, May 8, 2017.

Eher, R., Matthes, A., Schilling, F., Haubner-MacLean, T., & Rettenberger, M. (2012). Dynamic risk assessment in sexual offenders using STABLE-2000 and the STABLE-2007: An investigation of predictive and incremental validity. *Sexual Abuse, 24*(1), 5-28. https://doi.org/10.1177/1079063211403164

Eher, R., Olver, M. E., Heurix, I., Schilling, F., & Rettenberger, M. (2015). Predicting reoffense in pedophilic child molesters by clinical diagnoses and risk assessment. *Law and Human Behavior, 39*(6), 571-580. https://doi.org/10.1037/lhb0000144

Fernandez, Y. (2008). *An examination of the inter-rater reliability of the STATIC-99 and Stable-2007.* Poster presentation at the 27th Annual Research and Treatment conference of the Association for the Treatment of Sexual Abusers. Atlanta, GA, USA.

Fernandez, Y., Harris, A. R. J., Hanson, R. K., & Sparks, J. (2014). *Stable-2007 coding manual revised 2012.* Public Safety Canada.

Hanson, R. K. (2017). Assessing the calibration of actuarial risk scales: A primer on the E/O index. *Criminal Justice and Behavior, 44*(1), 26-39. https://doi.org/10.1177/0093854816683956

Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sexual offense recidivism. *Criminal Justice and Behavior, 27*(1), 6-35. https://doi.org/10.1177/0093854800027001002

Hanson, R. K., & Harris, A. J. R. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse, 13*(2), 105-122. https://doi.org/10.1177/107906320101300204

Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project.* Public Safety Canada.

Hanson, R. K., & Helmus, L. (2013). *Stable-2007: Updated Recidivism Rates.* Public Safety Canada.

Hanson, R. K., Helmus, L., & Harris, A. J. R. (2015). Assessing the risk and needs of supervised sexual offenders: A prospective study using STABLE-2007, Static-99R, and Static-2002R. *Criminal Justice and Behavior, 42*(12), 1205-1224. https://doi.org/10.1177/0093854815602094

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. https://doi.org/10.1037/a0014421

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. https://doi.org/10.1023/A:1005482921333

Harris, A. J. R. (2021). STABLE-2007 and indeterminate detention. *Sexual Offending: Theory, Research, and Prevention, 16*, Article e4587. https://doi.org/10.5964/sotrap.4587

Harris, G. T., Rice, M. E., Quinsey, V. L., & Cormier, C. A. (2015). *Violent offenders: Appraising and managing risk* (3rd ed.). Washington, DC, USA: American Psychological Association.

Helmus, L., Thornton, D., Hanson, R., & Babchishin, K. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse, 24*(1), 64-101. https://doi.org/10.1177/1079063211409951

Looman, J., & Abracen, J. (2010). Comparison of measures of risk for recidivism in person convicted of a sexual offenses. *Journal of Interpersonal Violence, 25*(5), 791-807. https://doi.org/10.1177/0886260509336961

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, 22*(2), 191-217. https://doi.org/10.1177/1079063210366039

Phenix, A., Helmus, L. M., & Hanson, R. K. (2016). *Static-99R and Static-2002R Evaluator's Workbook.* Retrieved from www.static99.org

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk.* Washington, DC, USA: American Psychological Association.

Saum, S. (2007). *A comparison of an actuarial risk prediction measure (Static-99) and a Stable dynamic risk prediction measure (Stable-2000) in making risk predictions for a group of person convicted of a sexual offenses* (Unpublished doctoral dissertation). Fielding Graduate University, Ann Arbor, MI, USA. (3255539)

Sowden, J. N., & Olver, M. E. (2017). Use of the violence risk scale-sexual offender version and the stable 2007 to assess dynamic sexual violence risk in a sample of treated sexual offenders. *Psychological Assessment, 29*(3), 293-303. https://doi.org/10.1037/pas0000345

Wong, S., Olver, M. E., Nicholaichuk, T. P., & Gordon, A. (2003, 2017). *The Violence Risk Scale— Sexual Offense version (VRS–SO).* Saskatoon, Saskatchewan, Canada: Regional Psychiatric Centre and University of Saskatchewan.

PsychOpen GOLD