

Long term recidivism rates among individuals at high risk to sexually reoffend

R. Karl Hanson<sup>1,2</sup>, Seung C. Lee<sup>2</sup>, and David Thornton<sup>2,3</sup>

<sup>1</sup> Department of Psychology, Carleton University

<sup>2</sup> SAARNA: Society for the Advancement of Actuarial Risk Need Assessment

<sup>3</sup> FAsTR LLC

*Sexual Abuse* (in press, September, 2022).

**This is not the version of record, which will appear on the Sage/Sexual Abuse website upon  
publication**

R. Karl Hanson  <https://orcid.org/0000-0002-0920-884x>

Seung Chan Lee  <https://orcid.org/0000-0002-6705-4535>.

David Thornton  <https://orcid.org/0000-0003--0127-1104>

R. K. Hanson and D. Thornton are co-authors and certified trainers of the Static-99R and Static-2002R risk tools. The copyright for these risk tools is held by the Government of Canada and the authors receive no royalties for their use.

The authors take responsibility for the integrity of the data, the accuracy of the data analyses, and made every effort to avoid inflating the statistical significance or nonsignificance of the results.

Correspondence concerning this article should be addressed to R. Karl Hanson, Department of Psychology, Carleton University, B552 Loeb Building, 1125 Colonel By Drive, Ottawa, Ontario, K1S 1Z3.

Email: rkarlhanson@gmail.com

### Abstract

Preventive detention provisions in the US and Canada assume we can identify, in advance, individuals at high risk for sexual recidivism. To test this assumption, 377 adult males with a history of sexual offending were followed for 20 years using Canadian national criminal history records and internet searches. Using previously collected information, a high risk/high need (HRHN) subgroup was identified based on an unusually high levels of criminogenic needs ( $n = 190$ , average age of 38 years; 83% White, 13% Indigenous, 4% other). A *well above average* subgroup of 99 individuals was then identified based on high Static-99R (6+) and Static-2002R (7+) scores. In the HRHN group, 40% reoffended sexually. STATIC HRHN norms overestimated sexual recidivism at 5 years (Static-99R,  $E/O = 1.44$ ; Static-2002R,  $E/O = 1.72$ ) but were well calibrated for longer follow-up periods (20 years: Static-99R,  $E/O = 1.00$ ; Static-2002R,  $E/O = 1.16$ ). The overall sexual recidivism rate for the *well above average* subgroup was 52.1% after 20 years, and 74.3% for any violent recidivism. The highest risk individuals (top 1%) had rates in the 60% to 70% range. We conclude that some individuals present a high risk for sexual recidivism, and can be identified using currently available methods.

**Keywords:** Sexual recidivism, risk assessment, Static-99R, Static-2002R.

### Long term recidivism rates among individuals at high risk to sexually reoffend

Individuals deemed high risk to sexually reoffend can be subject to exceptional legal measures, such as the Sexually Violent Person (SVP) provisions in many US states (Prentky et al., 2015) and the Dangerous Offender (DO) designation in Canada (Criminal Code of Canada, 1985). These provisions allow for indefinite detention, with release predicated on reassessments indicating the individual's risk has declined below some acceptable threshold. Psychologists who provide expert opinions in such cases often rely on specialized sexual recidivism risk tools (Kelley et al., 2020).

The definition of high likelihood to reoffend varies across settings, and may not be explicitly defined. SVP laws in the USA commonly use the term "likely" in their statutes, which, in some states, has been clarified to mean "more likely than not" either in statute or case law (Knighton et al., 2014). This threshold implies absolute rates of greater than 50%. Other states do not define "likely", and still other states use phrases such as "highly probable", which implies a rate greater than "more likely than not". Furthermore, some SVP laws specifically exclude quantitative interpretation. An additional complication is that the likelihoods stipulated by SVP laws are usually understood as referring to lifetime risk, which may not align with the information provided by most actuarial tables (i.e., the recidivism rate estimates associated with the scores of risk tools). Actuarial tables can be biased upwards because they include sexual crimes that do not meet the statutory definition of sexual violence (e.g., exhibitionism, possession of child abuse images). On the other hand, the rates can be biased downward because they only include *detected* offences within a discrete follow-up period (e.g., 5 years, 10 years). In practice, decision-makers routinely interpret expected rates somewhat lower than 50% as meeting the "more likely than not" threshold.

In Canada, the Dangerous Offender (DO) designation allows for the preventive detention of individuals deemed high risk for committing a serious personal injury offence. Although the outcome of concern includes both sexual and nonsexual violence, two-thirds of individuals with DO designations

have an index sexual offence (Public Safety Canada, 2020). The threshold for high risk is that “the trial judge must be satisfied beyond a reasonable doubt that there is a “likelihood” that the offender will [commit another serious injury offence]” (R. v Currie, 1997, paragraph 42, quotation marks in the original). To our knowledge, this likelihood has never been given an explicit quantitative interpretation. In practice, however, the DO designation is restricted to a very small proportion of the highest risk individuals (Blais & Bonta, 2014). In 2017-2018, there were close to 90,000 cases involving crimes against persons, of which less than 50 resulted in a DO designation (Public Safety Canada, 2020) - a rate of less than 1 in 1,500.

Decisions about the likelihood of any future outcome should be informed by base rates (Kahneman et al., 2021). Previous reviews have observed sexual recidivism rates of between 5% and 15% after five years, increasing to 10% to 20% when the follow-up is extended to 10 years (Hanson & Bussière, 1998; Harris & Hanson, 2004; Helmus, Hanson et al., 2012). In recent routine samples, sexual recidivism rates of 2% to 8% are common (Alper & Durose, 2019; Boccaccini et al., 2017; Helmus, Hanson et al., 2021). Helmus, Hanson et al. (2021), for example, observed an average sexual recidivism rate of 4.2% among 13 studies published between 2013 and 2021 (combined sample of 48,025).

Not all individuals with a history of sexual offending are equally likely to reoffend. Rates vary based on well-established risk factors (see meta-analyses by Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2009; Mann et al., 2010). These risk factors can be divided into static, historical variables (e.g., prior sexual offences, any unrelated victims) and indicators of psychological and community adjustment that can be used to guide supervision and treatment (i.e., dynamic risk factors or criminogenic needs, such as sexual preoccupation and negative attitudes towards supervision). Although static and dynamic risk factors are based on different information, both types of variables predict recidivism because they are markers for enduring, risk relevant propensities (Mann et al., 2010).

Given the relatively low overall base rate, it is not clear what the upper end of the recidivism rates should be for individuals who have multiple, severe risk relevant propensities. Although it is easy to retrospectively identify individuals with very high recidivism rates (i.e., by post hoc analyses), recidivism rate estimates only have practical utility when they can be specified in advance. Consequently, the confidence with which evaluators can make high risk determinations depends on the *calibration* of their risk assessments (i.e., correspondence between the expected rates and the observed rates in replication studies; see Helmus & Babchishin, 2017; Hanson, 2022, Part IV).

The highest rates estimated by sexual recidivism risk tools are 60% to 80% (Helmus, Lee et al., 2021; Olver et al., 2020). The Violence Risk Scale – Sexual Offense version (VRS-SO; Olver et al., 2007), for example, estimates the 10-year sexual recidivism rate to be 80% for individuals with the maximum score (72) and no treatment gains (Olver et al., 2021). It is unclear, however, whether there are real cases with such extreme values. The highest VRS-SO score in the workbook (Olver et al., 2020) is 67, and there were only 10 cases out of 913 that had scores of 57 or higher. According to the VRS-SO, the top 1% (scores 57 or higher) would be expected to have 10-year sexual recidivism rates of 50% to 60%.

Static-99R and Static-2002R (Helmus, Thornton et al., 2012) are the most commonly used sexual recidivism risk tools, with routine use in the US (Kelley et al., 2020), Canada (Bourgon et al., 2018), Australia, New Zealand, Europe, and East Asia (Helmus et al., 2022; Neal & Grisso, 2014). As their names imply, they privilege static risk factors, such as age and criminal history. They are considered actuarial risk tools because they use an explicit method of combining the items into total scores, and these scores are associated with expected recidivism rates (Helmus, Lee et al., 2021). The user guidance for Static-99R and Static-2002R present sexual recidivism rates for three follow-up periods (5, 10, and 20 years) and for two reference groups: a) routine/complete samples (Lee & Hanson, 2021) and b) samples preselected to be high risk/high need (HRHN; Hanson et al., 2016; Helmus, Lee et al., 2021). The need for separate tables was motivated by higher-than-expected sexual recidivism rates in samples that were

preselected as high risk. Examples of the high risk/high need samples include individuals referred to high intensity treatment programs, or for assessments to determine whether they qualify as Dangerous Offenders (based on Canadian criminal law). According to user guidance, HRHN samples are conceptually distinguished from routine samples based on an unusually high density of dynamic risk factors/criminogenic needs (Hanson et al., 2016). The major differences between the sexual recidivism rates for the routine/complete and HRHN samples are for the lower risk and average risk levels; in these ranges, the recidivism rates are materially higher for the HRHN groups. In contrast, the expected recidivism rates associated with the highest scores are similar for both reference groups. The highest named risk level for Static-99R and Static-2002R is Level IVb – *well above average*, representing the top 10%, for which the expected sexual recidivism rates are 25% to 45% after 5 years, and 45% to 65% after 20 years (Helmus, Lee et al., 2021).

There are several significant qualifications to STATIC (Static-99R and Static-2002R) expected recidivism rate tables. First, the HRHN norms were based on a limited number of studies ( $k = 5$  for Static-99R;  $k = 2$  for Static-2002R). Second, the 20-year rates are not observed values; instead, they are projected from the observed rates at 5-year and 10-year follow-up periods using the procedure described in Thornton et al. (2021). A further limitation is that there is considerable variability in the recidivism base rates across studies for reasons not fully understood. Recent calibration studies of the 2016 Static-99R recidivism norms (Boccaccini et al., 2017; Baudin et al., 2021; Helmus, Hanson et al., 2021; Lee et al., 2016, 2018), have found adequate to poor calibration; the general pattern is that new samples display lower recidivism rates than suggested by the 2016 norms (e.g., average *E/O* index of 1.88 in five field validity studies; Helmus, Hanson et al., 2021). In response to such findings, the routine/complete norms for Static-99R have been revised (Lee & Hanson, 2021). There is insufficient research, however, to determine whether the 2016 HRHN norms similarly need updating.

The STATIC 20-year estimates in particular could benefit from calibration studies. In some contexts, the relevant question is the lifetime risk of sexual offending. Given that long term studies are difficult (and always dated), a reliable method of extrapolating from shorter follow-up periods would be of considerable utility. Thornton and colleagues (2021) have proposed a generic mathematical model that can be used to translate observed sexual recidivism rates at any specific follow-up period (e.g., 5-years) to rates at any other follow-up period (e.g., 1-year, 15-years). Thornton's model is based on a predictable and consistent decline in sexual recidivism rates the longer that individuals remain sexual offence free in the community (Hanson et al., 2018). In the model, the 20-year rates are treated as equivalent to the lifetime rates. Residual risk after 20 years is not zero; however, it is so small as to be indistinguishable from the rate of spontaneous out-of-the-blue sexual offences in the general male population. Although general time free effects have been replicated (e.g., Moore, 2018), the specific 20-year projections for Static-99R and Static-2002R have not.

### **Overview of Current Study**

A primary goal was to estimate the upper range of sexual recidivism rates that could be specified in advance. Given that the expected rates were based on norms for Static-99R and Static-2002R risk tools, this research question could also be stated as testing the calibration of these measures. Twenty-year estimates for the high risk/high need (HRHN) reference group (Helmus, Lee et al., 2021) for the highest STATIC scores (95<sup>th</sup> percentile) are in the 50% to 65% range, which is substantially higher than rates typically observed in contemporary follow-up studies (< 10%). Finding observed rates over 50%, however, would require exceptionally long follow-up of individuals who were preselected to be high risk for sexual recidivism.

The opportunity to address these research questions was provided by an extended follow-up (median of 20.5 years) of a research sample of approximately 400 men who had committed sexual offences in the 1990s. The original sample was preselected to be above average risk - half had already

sexually reoffended while on community supervision (Hanson & Harris, 2000a, 2000b). Using variables collected during the initial research, a subgroup was identified that matched the definition of a high risk/high need sample asserted by the STATIC norms (i.e., unusually high levels of criminogenic needs; Hanson et al., 2016). Static-99R and Static-2002R scores were coded retrospectively based on case file and criminal history information.

### **Hypotheses**

We expected the predictive accuracy of Static-99R and Static-2002R prediction tools to be acceptable both in terms of discrimination (*AUC* values around .70) and calibration (nonsignificant *E/O* values close to 1). The coders had been trained by certified trainers and used research data that has proven informative in previous studies (Hanson & Harris, 2000b). Furthermore, Static-99R and Static-2002R have performed well in this jurisdiction (Canada; Helmus, Thornton, et al., 2012).

Given the preselection, we expected observed sexual recidivism rates to be closer to the HRHN norms than to the routine/complete norms. Nevertheless, our confidence was tempered by knowledge that the 2016 Static-99R routine/complete norms had overestimated the observed rates in recent samples (Helmus, Hanson et al., 2021). We had more confidence in the calibration for the 5-year and 10-year rates than we did in the 20-year rates because the very long-term estimates require additional assumptions concerning the shape of the recidivism survival function. Although the general shape of the hazard function has been replicated, the accuracy (calibration) of the specific method used by the STATIC measures for estimating the 20-year rates (Thornton et al., 2021) has yet to be empirically tested in a new sample.

### **Method**

#### **Subjects**

The current sample was selected from men previously identified for Hanson and Harris' (2000a, 2000b) study of dynamic (changeable) risk factors for sexual recidivism. The original sample included

208 men known to have sexually reoffended while on community supervision, roughly matched on criminal history and location/jurisdiction to 201 men who were on community supervision for a sexual crime without reoffending. Cases were drawn from Canada's federal prison service (Correctional Service of Canada [CSC]) and all provincial correctional systems (except for Prince Edward Island). From the original sample of 409, 7.8% ( $n = 32$ ) were excluded due to lack of recidivism information; the sample was further reduced to 190 individuals preselected to be high risk using the criteria described below.

On average, the presumptively high-risk individuals ( $n = 190$ ) were 38 years old ( $SD = 10.8$ , range from 18 to 78 years) when released into the community between 1986 and 2016 ( $Mdn = 1996$ ). Eighty-three percent (157/190) were White, 13% ( $n = 24$ ) were of Indigenous heritage, 4% were from other ethnocultural groups (e.g., Black, Asian). The average Static-99R and Static-2002R total scores were 5.1 ( $SD = 2.29$ ; range of -3 to 10;  $Mdn = 5$ ) and 6.9 ( $SD = 2.31$ ; range of -1 to 11;  $Mdn = 7$ ), respectively.

The recidivism dataset used in this study has been previously used by Aelick et al. (2020) in a study of mental health variables, and by Lee et al.'s (2020) in a meta-analysis on the predictive validity of sexual recidivism risk tools for individuals of Indigenous heritage in Canada.

## **Measures**

### ***Structured Risk Assessment (SRA) Need Framework***

The SRA framework was used to identify a presumptively high risk/high need (HRHN) subgroup from the original sample. Some measure of dynamic risk factors was necessary because the STATIC HRHN norms are intended to apply to individuals who have been preselected to have unusually high levels of criminogenic needs. Step two of the SRA system provides a flexible method of quantifying the intensity and diversity of long-term vulnerabilities for sexual recidivism risk (see Thornton, 2016). The variables available did not allow for reliable scoring of other validated measures of criminogenic needs, such as STABLE-2007 (Brankley et al., 2021) or the VRS-SO (Olver et al., 2007).

The long term vulnerabilities from the SRA system were scored from indicators for each of the four Need Domains postulated by SRA (Sexual Interests, Distorted Attitudes, Relational Style, and Self-Management). For each domain a summary score between 0 and 2 was constructed. These domain summaries were then summed, yielding a Need Score from 0 to 8. See Supplemental Appendix for further details.

For the present purposes, it was also necessary to define a threshold Need score to classify individuals as “high need”. A score of 4.0 or above was used for this purpose. Such a score can be obtained by having serious, multi-need problems within two domains; having a serious single-need problem in all four domains; or some mixture of serious, single and multi-need problems in three domains. It also corresponds to someone halving at least half of the possible criminogenic need points. This is consistent with the level of criminogenic need that is commonly found in existing HRHN samples (see Supplemental Appendix). The score of 4 also approximately divided the present sample into approximately equal sized higher and lower Need groups. In the full sample with sufficient information to scores the SRA ( $n = 367$ ), the average SRA score was 4.0 ( $SD = 1.7$ ,  $Mdn = 4.2$ , range from zero to 7.7).

Importantly, the author who constructed the Need score and chose the threshold was blind to the recidivism data.

### ***Static-99R***

Static-99R (Hanson & Thornton, 2000; Helmus, Thornton, et al., 2012) is an empirical actuarial risk tool designed to assess risk of sexual recidivism among adult men charged or convicted of a sexual motivated offense. It contains 10 items related to demographic information (age, relationship history) and criminal history (e.g., prior sexual offences, prior nonsexual violence, any male victims). Total scores (range from -3 to 12) are calculated by summing all item points and can be used to place individuals in one of five risk levels: Level I – *very low risk* (scores of -3 to -2), Level II – *below average risk* (scores of -1 to 0), Level III- *average risk* (scores of 1 to 3), Level IVa – *above average risk* (scores of 4 to 5), and

Level IVb – *well above average risk* (scores of 6 or higher; Hanson, Babchishin et al., 2017). Total scores can be interpreted in terms of percentile ranks, relative risk, and absolute recidivism rates for 5-year, 10-year, and 20-year follow-up periods (Helmus, Lee et al., 2021). Meta-analyses based on large samples (> 15,000) indicate moderate ability to discriminate between sexual recidivists and nonrecidivists (average AUC values of .66 to .68; Helmus et al., 2022; Helmus, Hanson et al., 2021).

Rater reliability of the Static-99R total scores was not available in the current study because they were calculated from five variables already coded in the dataset (e.g., age, any male victim) and five criminal history items specifically coded for this study. The rater reliability for these five newly coded items ranged from .71 to .98 (median Kappa of .88 based on 59 cases coded by two raters). Previous studies have found overall good interrater reliability of the Static-99R total scores (intraclass correlation coefficient [ICC] ranging from .78 to .96; Gonçalves, 2020; Hanson, Lunetta et al., 2014; McGrath et al., 2012; Raymond et al., 2020; Stephens et al., 2017).

### ***Static-2002R***

Like Static-99R, Static-2002R (Hanson & Thornton, 2003; Hanson et al., 2010) is an empirical, actuarial risk tool for estimating the likelihood of sexual recidivism among adult males based on commonly available demographic and criminal history information. It was developed to improve conceptual consistency and increase predictive accuracy over Static-99. Subsequent research, however, has found Static-2002R and Static-99R to have similar predictive accuracy, and that each contributes incrementally to the likelihood of sexual recidivism (Babchishin et al., 2012a; Lehmann et al., 2013). Total scores (ranging from –2 to 13) are calculated by summing all item points and can be used to place individuals in one of five risk levels: Level I – *very low risk* (scores of –2 to –1), Level II – *below average risk* (scores of 0 to 1), Level III – *average risk* (scores of 2 to 4), Level IVa – *above average risk* (scores of 5 to 6), and Level IVb – *well above average risk* (scores of 7 or higher; Hanson, Babchishin et al., 2017).

Total scores can be interpreted in terms of percentile ranks, relative risk, and absolute recidivism rates for 5-year and 20-year follow-up periods (Helmus, Lee et al., 2021).

Rater reliability for Static-2002R total scores was not available because it calculated from variables already coded in the dataset (e.g., age, any male victim) and four items that were specifically coded for this study. The rater reliability for these four newly coded items ranged from .76 to .94 (median Kappa of .89 based on 59 cases coded by two raters). Previous studies have found good interrater reliability for Static-2002/R total scores (ICCs ranging from .87 to .98; Helmus & Hanson, 2007; Jung et al., 2017; Reeves et al., 2018; Smid et al., 2014).

### **Identifying Individuals at High Risk for Sexual Recidivism**

Two high risk subgroups were constructed from the 409 original cases. The first subgroup was intended to resemble the preselected high risk/high need (HRHN) samples in the STATIC norms. According to Hanson et al. (2016), HRHN samples have high average STATIC scores, and unusually high levels of criminogenic risk factors (i.e., riskiest 20%). In the current sample, the average Static-99R score was in the *above average* range ( $M = 4.3$ ,  $SD = 2.4$ ,  $n = 407$ ). Using a SRA Need score of 4 as the threshold for high levels of criminogenic needs (see above), the HRHN subgroup comprised approximately half of the original sample (200 out of 407 cases;  $n = 190$  with recidivism information). The average Static-99R score in the HRHN subgroup was 5.1 ( $SD = 2.3$ , range from -3 to 10). The syntax used to create the Need Score and the decision threshold for high risk were made blind to recidivism information.

The second high risk subgroup, the *well above average* group, not only had unusually high criminogenic needs, but also unusually high STATIC scores. Unusually high STATIC scores were defined as an average score on the Static-99R and Static-2002R in the well above average range (Level IVb). Given that the raw scores of Static-99R and Static-2002R are in different metrics, these scores were averaged using the metric of logged hazard ratios, consistent with previous studies (Babchishin et al.,

2012a; Babchishin et al., 2012b; Lehmann et al., 2013, see Appendix). Overall, 39.3% (160/407) individuals in the complete sample had average STATIC scores in the well above average range (Level IVb), of which 75% (105/140) also had high SRA scores (4+; SRA scores were not available for 20 of the 160 cases). For the recidivism analyses, the *well above average* group comprised individuals who had complete information on Static-99R, Static-2002R, and SRA, as well as recidivism information ( $n = 99$ ). The average Static-99R score in the *well above average* group was 6.8 ( $SD = 1.1$ , range from 5 to 10). The average SRA Need score was 5.4 ( $SD = 0.95$ , range of 4 to 7.67) in this group. Although high SRA scores and high STATIC scores were intended to be separate criteria, none of the individuals with high STATIC scores had an SRA score lower than four.

### **Recidivism**

Our primary source of recidivism information was the Canadian Police Information Centre (CPIC) criminal history records held by the Royal Canadian Mounted Police (RCMP). These records are intended to record all convictions across Canada, as well as some charges. In addition, news articles were searched via Google using personal identifiers (e.g., full name plus the province where most prior offences occurred). Forty-two relevant news articles were identified; in all cases, the individuals in the news articles had already been identified as a sexual recidivist based on RCMP records. The offence descriptions in the news reports, however, allowed the researchers to classify six of these events as contact sexual recidivism and two as non-contact sexual recidivism. It is not uncommon that publicly available media informs academic research on serious crime (Laporte et al., 2003; Tanguay, 2004), including the measurement of sexual recidivism (Hanson et al., 2015).

The minimum threshold for recidivism was a charge; however, given the primary source of information used (CPIC), the vast majority of recidivism events were convictions. Out of the initial sample of 409 cases, recidivism information was obtained for 377 (92.2%). Records could be unavailable in 2017 because the individual had been deceased for many years, pardoned, or for unknown reasons

(all 409 cases had CPIC records in 1997). The length of follow-up was sufficient (20+ years) that it is likely that some of the individuals with missing records had reoffended prior to their records going missing.

Follow-up time began a) at the date of assessment for the original non-recidivist group, or b) after release from the previously identified recidivism event for the recidivist group in the original study, i.e., the index offences of the recidivist group in the original study were replaced with the date of their previously identified recidivism events, creating new at-risk dates for the recidivist group. Follow-up ended with death ( $n = 34$ ), deportation ( $n = 1$ ), the date of the most recent CPIC record (2017), or the start of a period of incarceration that extended into one of these dates. Mortality was only noted when it appeared in the CPIC or news records (mortality information was not requested from Statistics Canada). The time at-risk prior to sexual recidivism did not include time incarcerated for nonsexual offences after the index offence. Overall, 34 individuals were known to have spent a month or more incarcerated prior to the survival end date for sexual recidivism (median of 4.7 months, range from 1 month to 7.8 years,  $n = 34$ ).

In the full sample with recidivism information ( $n = 377$ ), the average follow-up time was 19.6 years ( $Mdn = 20.5$ , range from 3 months to 35.6 years). Most of the sample was followed for at least 20 years (65.5%, 247/377); however, few individuals were followed for more than 25 years (3.2%, 12/377). The additional length of follow-up after 20 years was usually only a year or two (mean of 1.8 years, or 21.8 years total). Among the HRHN subgroup, 62% (118/190) were followed for longer than 20 years ( $Mdn = 21.5$ , range from 3 months to 31 years).

Sexual recidivism included both contact and noncontact sexual crimes. Violent recidivism included nonsexual violent offences (e.g., assault), arson, and contact sexual offences. Classification was based primarily on the name of the offence, although sexual motivation was attributed for certain

nonsexual offences when justified by associated information (e.g., news reports, sexual charge followed by a nonsexual violent conviction).

### **Procedure**

The original data were collected in 1996-1997 by four trained researchers working under the supervision of a project manager employed by the Solicitor General of Canada (now Public Safety Canada; see Hanson & Harris, 2000b). Probation and parole officers throughout Canada were asked to identify individuals who had recently sexually reoffended while on community supervision. These cases were then matched with individuals who had not reoffended. Case information was extracted from file review as well as from structured interviews with the supervising officers.

For the current study, the previously collected case information was linked to recidivism information collected in 2017. Two graduate-level research assistants working for Public Safety Canada coded the recidivism data along with items needed to score risk tools. The 2017-2018 update also resulted in some minor edits and corrections to the original data (e.g., revised at-risk dates).

The current study was conducted under a data sharing agreement with Public Safety Canada. This study received ethics approval from Carleton University Research Ethics Board-B (Project # 115736). All statistics in this manuscript were verified independently by the first and second authors using either Microsoft Excel, IBM/SPSS, or R Statistics.

### **Analyses**

#### ***Discrimination***

Two statistics were used to describe the extent to which recidivists were different from nonrecidivists (i.e., discrimination): a) the Area Under the Curve (*AUC*), and b) slope coefficients from logistic regression.

**Area Under the Curve (AUC).** AUC values range from zero to 1, and express the probability that randomly selected individual who has reoffended has a higher score than a randomly selected

nonrecidivist. Following Cohen's (1988) conventions for Cohen's  $d$ , AUC values of .56, .64, and .71 are considered small, moderate, and large, respectively (Rice & Harris, 2005). No association between the predictor and the outcome is indicated by an AUC value of .50.

**Logistic regression slope coefficients.** Slope coefficients indicate the amount of change in the likelihood of the outcome based on a one-unit increase in the risk tool. The coefficients ( $B1$ ) are logged odds ratios (logits), such that  $\exp(B1)$  are odds ratios. An odds ratio of 1.0 ( $B1$  of zero) indicate no relationship between the predictor and the outcome. Absolute values of the slope coefficients are difficult to interpret because they vary based on the scaling of the risk tool. In previous meta-analyses of preselected high risk/high need samples, the average odds ratio for Static-99R predicting sexual recidivism was 1.28 for 5-years, and 1.26 for 10-years; for Static-2002R, the average odds ratio was 1.24 at 5-years (Hanson, Thornton et al., 2016, Table 7).

### **Calibration**

Two statistics were used to measure the similarity between the observed and expected values (calibration): a) the  $E/O$  index and b) meta-analysis of logistic regression intercept coefficients (Hanson, 2022). The expected values were from the HRHN recidivism rate tables for Static-99R (5-year, 10-year, 20-year) and Static-2002R (5-year, 20-year; Helmus, Lee et al., 2021). Some STATIC scores do not have associated recidivism rates because they were weakly populated ( $n < 10$ ) in the normative samples. The expected rates for these scores were considered equal to the closest populated value. For example, the High Risk/High Needs norms provide rates for scores of -1 (e.g., 5.6% after 5 years) but not for scores of -2 or -3. Consequently, cases with scores of -2 or -3 were assigned an expected risk level of 5.6% after 5 years, the same as for the score of -1. The expected values for 15 years were calculated for this study using the methods described by Thornton et al. (2021). The observed rates for 5 years and 10 years were calculated using fixed follow-up. The observed rates for 15 years and 20 years were calculated using Kaplan-Meier survival analysis.

***E/O Index.*** The *E/O* index is a measure of calibration in which the expected number of recidivists is divided by an observed number of recidivists (Hanson, 2017; Vaillon et al., 2009). Perfect calibration is indicated by a value of 1.0. When the observed values were based on fixed follow-up, the 95% confidence intervals were estimates using the following equations (Equations 13.1 and 13.2 from Hanson, 2022; Rockhill et al., 2003):

$$\text{Upper Limit of 95\% CI of } \frac{E}{O} \text{ Index} = \left(\frac{E}{O}\right) e^{\left(1.96\sqrt{\frac{1}{O}}\right)}$$

$$\text{Lower Limit of 95\% CI of } \frac{E}{O} \text{ Index} = \left(\frac{E}{O}\right) e^{\left(-1.96\sqrt{\frac{1}{O}}\right)}$$

When the observed values were based on Kaplan-Meier survival analysis, the following equations were used for the 95% confidence intervals (Equations 13.3 and 13.4 from Hanson, 2022; Equation 14 in Viallon et al., 2009):

$$\text{Upper Limit of 95\% CI of } \frac{E}{O} \text{ Index} = \left(\frac{E}{O}\right) e^{\left(1.96\left[\frac{SE}{O}\right]\right)}$$

$$\text{Lower Limit of 95\% CI of } \frac{E}{O} \text{ Index} = \left(\frac{E}{O}\right) e^{\left(-1.96\left[\frac{SE}{O}\right]\right)}$$

In the above equations, *SE* is the Greenwood approximation of the standard error (Equation 10.8 in Singer & Willett, 2003), expressed in the units of cases (not proportions).

**Comparing Logistic Regression Parameters (*BO*).** Calibration was also examined by comparing the intercept values from logistic regression to those in the HRHN normative samples (Hanson, Thornton et al., 2016). For both Static-99R and Static-2002R, the intercepts were centered on the median value in routine/complete samples (score of 2 for Static-99R; 3 for Static-2002R); as such, the *BO*'s represents the expected recidivism rates for individuals with a score of 2 (Static-99R) or 3 (Static-2002R), expressed in logit units ( $\ln[p/\{1 - p\}]$ ). Centering on median values allows direct comparison with the *BO*'s reported in the STATIC norms.

Differences between the parameters in the current sample and those of the norms were tested

using fixed-effect meta-analysis (Borenstein et al., 2021) conducted with the package "metafor" (Version 3.0-2; Viechtbauer, 2010) for the statistical software R (Version 4.0.3; R Core Team, 2020).

### Results

In the full sample with recidivism information, 31.6% (119/377) were known to have committed a new sexual crime during the 20+ year follow-up period. Median follow-up time was 20.5 years and ranged from 3 months (one individual died shortly after release) to 35.6 years. Overall, 94.7% (357) were followed for at least 10 years, 89.1% (336) for at least 15 years, and 65.5% (247) for at least 20 years (all times exclusive of time in custody). The most serious sexual crime was a contact sexual crime in 109 cases and a non-contact sexual crime in 10 cases. There were 16 additional cases who violated their conditions of supervision for presumably sexual motivations (e.g., victim grooming, but no known offences); these cases were not included in the definition of sexual recidivism. In the full sample, 44.0% (166/377) were known to have committed a new violent offence (non-sexual violence, arson, or contact sexual offence) and 54.6% any new criminal offence (206/377). Among the 190 cases identified as High Risk/High Need (HRHN), the recidivism rate for sexual crimes was 40.0% (76/190), 56.3% (107/190) for violent crimes, and 68.9% (131/190) for any crime. Six of the HRHN cases received a Dangerous Offender designation for their recidivism event.

Both Static-99R and Static-2002R showed large relationships with sexual recidivism in the full dataset ( $n = 377$ ,  $AUC = .717$ ,  $SE = .028$ ; the same values for both Static-99R and Static-2002R, ragged follow-up). The average of Static-99R and Static-2002R scores ( $AUC = .725$ ,  $SE = .027$ ) was not significantly different from Static-99R ( $Z = 1.11$ ,  $p = .266$ ) or Static-2002R ( $Z = 0.735$ ,  $p = .462$ ; using the Delong test in IBM/SPSS version 27). In the HRHN subsample, all the STATIC measures showed moderate relationships with sexual recidivism ( $n = 190$ ; Static-99R,  $AUC = .671$ ,  $SE = .039$ ; Static-2002R,  $AUC = .685$ ,  $SE = .039$ ; Average of 99R + 02R,  $AUC = .681$ ,  $SE = .039$ ). The raw SRA need scores (without dichotomization) showed a moderate relationship to sexual recidivism in the full data set ( $n = 342$ ,  $AUC$

= .642,  $SE = .031$ ) and was incremental to the Static-99R (logistic regression incremental  $B$  of 0.170,  $SE = 0.086$ ,  $p = .048$ ) but not Static-2002R ( $B$  of 0.146,  $SE = .089$ ,  $p = .102$ ), or the Static-99R + Static-2002R average ( $B = 0.143$ ,  $SE = .089$ ,  $p = .106$ ).

When calibration was examined using the  $E/O$  index, the observed 5-year sexual recidivism rates for the HRHN subsample were lower than the expected HRHN values for both Static-99R ( $E/O = 1.44$ ) and Static-2002R ( $E/O = 1.72$ ; see Table 1). In contrast, the observed 10-year rates for Static-99R were close to the expected HRHN values ( $E/O = 1.13$ ), as were the 15-year rates ( $E/O = 0.96$ ) and 20-year rates ( $E/O = 1.00$ ; see Table 2). Similarly, the observed 15-year and 20-year rates for Static-2002R were close to the expected HRHN values ( $E/O$  index of 1.09 and 1.16, respectively; see Table 3).

When calibration was examined using logistic regression, none of the parameters were significantly different from the HRHN norms for either Static-99R (5-year, 10-year) or Static-2002R (5-year; see Table 4). Although there were no significant differences, there was a trend towards lower base rates (smaller  $BO$ 's) and better discrimination (higher  $B1$ 's) in the current HRHN sample compared to the HRHN norms. Contrary to expectation, the observed recidivism rates for our preselected subgroup were very close to the Routine/Complete norms (see Table 4). For five of the six comparisons with the Routine/Complete norms, the differences were less than expected by chance ( $Q < 1$ ,  $I^2 = 0$ ), and in the other comparison the variability was small ( $I^2 = 12.5\%$ ).

Figure 1 presents a visual representation of the calibration of Static-99R against the HRHN norms for 5 years, 10 years, 15 years, and 20 years. The Static-99R observed values are smoothed logistic regression curves (with 95% confidence intervals) plotted for each score that was populated by at least 10 cases. As can be seen from Figure 1, the Static-99R norms tended to overestimate the observed sexual recidivism rates at the shorter follow-up periods (5 years and 10 years); however, they were closely aligned with the observed values for the longer follow-up periods (15 years and 20 years). Calibration plots for Static-2002R for 5 years, 15 years, and 20 years are presented in Figure S1 (see

online supplemental materials). The Static-2002R HRHN norms overestimated the observed recidivism rates at 5 years, whereas the observed rates were close to the expected rates at 15 years and 20 years. We did not calculate calibration for Static-2002R at 10-year follow-up because there are no 10-year HRHN norms asserted by the test developers (see Helmus, Lee et al., 2021).

---

Insert Figure 1 about here

---

The next analyses examined the observed sexual and violent recidivism rates for the 99 individuals in the *well above average* subgroup (i.e., individuals with SRA scores > 4 and Static-99R and Static-2002R scores in the *well above average* range: Level IVb). In this subgroup, 52.5% (52/99) were known to have sexually reoffended during the follow-up period. Based on Kaplan-Meier survival analysis (see Figure 2), the sexual recidivism rates were 22.2% at five years, 43.7% at 10 years, and 52.1% at 20 years. There was only one new sexual recidivism event after 17 years (at 20.5 years). For violent recidivism, 73.7% (73/99) reoffended with a violent or contact sexual offence during the full follow-up period (ragged follow-up). Based on Kaplan-Meier survival analysis the violent recidivism rates were 44.4% at five years, 67.7% at 10 years, and 74.3% at 20 years. There was only one new violent recidivism event after 17 years (at 19 years). The sexual recidivism rates for the highest risk individuals (top 1%; Static-99R > 7; Static-2002R > 9) were in the range of 60% to 70%. Four received a Dangerous Offender designation for their recidivism event, and an additional 18 received Long Term Supervision Orders (a preventive detention provision with a determinant sentence).

The observed sexual recidivism rates per score on the Static-99R and Static-2002R risk tools are provided in the online supplemental materials (see Table S1, Table S2, Table S3, and Table S4). The 5-year and 10-year rates are based on fixed follow-up periods; the 15-year and 20-year rates are based on Kaplan-Meier survival analysis. The supplemental tables also provide a line of best fit for these observed

values, smoothed using logistic regression. For the purpose of the logistic regression fit, the observed values from the Kaplan-Meier survival analyses were rounded to integers.

### Discussion

This study examined the extent to which it is possible to identify, in advance, individuals who present a high risk for sexual recidivism. Using a structured approach to risk assessment and long-term follow-up, we identified a group whose observed sexual recidivism rates were above 50%. Although there is no universal threshold for high risk, sexual recidivism rates in this range are clearly problematic and would presumptively meet the “more likely than not” threshold enshrined in certain civil commitment statutes in the USA. Furthermore, three out of four individuals in this *well above average* group reoffended with a nonsexual violent or contact sexual offence, indicating substantial risk of a serious personal injury offence (a criteria of Canada’s Dangerous Offender provision). Given that a certain proportion of the convictions for nonsexual violence will be for crimes with a sexual motivation (e.g., kidnapping), the actual rate of sexual recidivism will be higher than the rate estimated when only counting offences that are explicitly sexual.

The observed sexual recidivism rates would also underestimate the true rates because not all offences are detected. Although the size of the gap between detected and undetected reoffenders remains a topic of debate (Abbott, 2020; Lave et al., 2021; Scurich & John, 2019), there is no question that nonreporting obscures the true recidivism rate. We believe, however, that the gap between detected and undetected offenders should decrease as the follow-up time increase. Even if the likelihood of detection per offence is low, individuals who repeatedly offend are likely to eventually get caught.

When the decision threshold concerns lifetime risk, most recidivism rate estimates are limited because they are based on shorter follow-up periods (e.g., 5 years, 10 years). Consequently, evaluators must extrapolate from the reported rates to lifetime rates. Hanson, Thornton and colleagues (Hanson et

al., 2018; Thornton et al., 2021) have argued that for sexual recidivism the 20-year rates closely approximate the lifetime rates because the hazards of sexual recidivism become vanishingly small for individuals who remain 20 years sexual offence free in the community. This assumption was supported in the current study. The residual hazard of sexual recidivism was small after 15 years sexual offence free. In the full sample of 377 individuals, the residual hazard was only 3.4% after 15 years, compared to 29.2% between time of release and 15 years. Among the 181 individuals who remained sexual offence free for more than 20 years, there were only two subsequent sexual recidivism events (at 20.5 and 21.4 years). Even though there were few individuals followed for more than 25 years, the consistency of the current findings with previous research (Hanson et al., 2018; Moore, 2018) suggests that the observed 20-year sexual recidivism rates in the current study reasonably approximate the rates that would have been observed if all the individuals had been followed to the end of their lives.

A related goal was to evaluate the predictive accuracy (calibration) of the Static-99R and Static-2002R sexual recidivism prediction tools for high risk/high need (HRHN) samples. The results were mixed. The HRHN norms substantially overestimate the observed sexual recidivism rates for the shorter follow-up periods (5 years) for both Static-99R and Static-2002R, but were close to the observed values at 15 years and 20 years. This pattern was opposite to our expectations. We had greater *a priori* confidence in the 5-year and 10-year rates given that they required fewer inferences than the 15-year and 20-year norms. Nevertheless, the good calibration of the STATIC risk tools at 20-year follow-up should be reassuring to evaluators wishing to use these measures to estimate long-term (e.g., lifetime) sexual recidivism rates.

Even though the current study broadly supported the high risk/high needs norms for the STATIC measures, the observed values also fit the norms for routine/complete samples. This pattern could be explained, in retrospect, by the unusually high STATIC scores in this sample. In routine/complete samples, the median Static-99R value is usually 2 ( $SD = 2.1$ ; Hanson, Lloyd et al., 2012), whereas the

median value in the current study was 5; furthermore, 76% of the sample (144/190) had a Static-99R score in the *above average* range (4 or higher). Conversely, there were almost no individuals in the *very low* (1/190) or *below average* (6/190) risk levels. The logistic regression equations that define the STATIC norms assume that Static-99R and Static-2002R show greater discrimination in routine/complete samples than in high risk/high need samples, which results in routine/complete samples having similar expected rates for high scores (Hanson et al., 2016). Consequently, the expected values are substantially similar for both sets of recidivism rate norms (routine/complete and high risk/high need) for the Static-99R values populated in this study.

Most studies find a consistent decline in sexual recidivism risk for each year that the individual remains offence-free in the community (Hanson et al., 2018). In previous meta-analyses, the sexual recidivism rates between 5 years and 10 years were approximately half the sexual recidivism rates observed between time of release and 5 years (Hanson et al., 2014; Hanson et al., 2018; Harris & Hanson, 2004). In contrast, the current study found similar sexual recidivism hazard rates during the first 10 years, followed by a sharp decline in risk for subsequent years (see Figure 2). This pattern may be a random feature of this dataset, or a consequence of delayed processing of sexual assault cases in the criminal justice system (i.e., conviction dates are often years after offence dates). It is possible, however, that it is related to protective effects of community supervision. All the individuals in the original study were on community supervision at that time, and many were nominated for the study by their supervising officers because of concerns about sexual recidivism risk. Although the risk reduction effects of community supervision remain an open research topic, the current study is not the first to show relatively flat hazard functions during the first 10 years for individuals who started then exited community supervision (Ducro et al., 2020; Zgoba et al., 2012).

When the sexual recidivism outcome was expanded to include sexually motivated community supervision violations, the recidivism rate during the first five years (24.4%,  $n = 190$ ) was larger (as

expected) than the rate during years 5 to 10 (13.9%). Given that the circumstances of the recidivism events were not always known, some of the community supervision violations could have been for sexual crimes. Consequently, in addition to suppressing actual recidivism, community supervision could be suppressing the apparent reconviction rate (i.e., some sexual crimes could be sanctioned by return to prison rather than new charges). Including violations in the definition of sexual recidivism would move the observed rates away from the routine/complete norms and closer to the HRHN norms.

Nevertheless, we recommend against including community supervision violations in the definition of sexual recidivism because the actual behaviours were unknown.

### **Limitations**

Missing records is a common challenge for long term follow-up studies, and the current study was no exception. In 1997, official criminal history records were available for all 409 cases identified for the original study, whereas in 2017 recidivism information was only available for 377 cases (92.2%). The reasons for missing records are not known, and would vary over the course of the 20+ year follow-up time. Nevertheless, our previous experience with the RCMP CPIC records is that they are unlikely to go missing for individuals who remain in the criminal justice system (i.e., the recidivists). In the *well above average* subgroup there were 99 individuals with codable recidivism information, of which 52 were known to have reoffended with a sexual crime (raw recidivism rate of 52.5% = 52/99). As well, there were an additional 6 individuals without codable recidivism information (four without FPS records plus two with FPS records that lacked the necessary released dates). If all six individuals in this group were considered nonrecidivists, the observed estimate would decrease to 49.5% (52/105); if all six were recidivists, the estimate would increase to 55.2% (58/105). Given that both values are close to unadjusted value of 52.5%, it is unlikely that missing records introduced substantial bias.

Another limitation was that the SRA scheme used to identify a high risk/high need subsample was applied retrospectively using variables previously coded. Although the dataset was sufficiently rich

that the SRA implementation was plausible, it was not ideal. In particular, as compared to typical implementations of the SRA Need Framework, the present implementation was based heavily on observations made by supervising agents, and so is more oriented to recent functioning than is desirable where the intention is to assess long-term vulnerabilities. Additionally, for the most part, the variables being integrated were not from established measures with established reliability and validity.

A related limitation is that the threshold for HRHN (SRA score of 4) was unique to this study. Although plausibly aligned with the STATIC definition, there were no established norms to justify this specific score for this particular operationalization of the SRA system. In order to explore the potential influence of the choice of threshold, it was varied from 4 (approximately the 50<sup>th</sup> percentile), to 3 (30<sup>th</sup> percentile), and to 5 (70<sup>th</sup> percentile). The raw sexual recidivism rate of the *well above average* group remained unchanged: 4 (52.5%, 52/99), 3 (52.5%, 52/99) and 5 (52.5%, 32/61). Consequently, it is unlikely that minor changes to the SRA threshold would substantially influence the overall findings.

### **Implications for Practice**

The study provided general support for the High Risk/High Need (HRHN) recidivism rate norms of the Static-99R and Static-2002R sexual recidivism risk tools. Although these norms overestimated the observed recidivism rates at 5-years, they were well calibrated for the longer follow-up periods. The poor calibration at 5 years, on the other hand, suggests that evaluators should be cautious when making inferences about the likelihood of recidivism during this time period, particularly if the individual in question has external protective factors, such as ongoing treatment and a credible supervision plan.

The use of the HRHN norms requires an initial judgement that the individual has unusually high levels of criminogenic needs (Hanson et al., 2016). The specific method used in the current study, however, cannot be directly implemented in applied practice because it was based on archival research data. Instead, prudent evaluators should make the determination that the individual is a member of a presumptively high-risk group based on a validated measure of criminogenic needs, such as the STABLE-

2007 (Brankley et al., 2021), and the VRS-SO (Olver et al., 2007), or the forensic version of the SRA (e.g., SRA-FV; Thornton & Knight, 2015),

### **Implications for Research**

More studies of HRHN samples are needed. Static-99R studies of routine/complete samples are sufficiently common to support regular updates of the recidivism rate norms for this reference group (Lee & Hanson, 2021). To our knowledge, this is the first HRHN validity study of either Static-99R or Static-2002R since 2008 (Hanson et al., 2016). The total number of HRHN samples remains small ( $k = 6$  for Static-99R;  $k = 3$  for Static-2002R, including the current study).

We hypothesized that community supervision suppressed recidivism during the first 5 years following release (i.e., flattened the hazard rates). This speculation, however, was made without knowledge of the actual interventions received. Future research could advance understanding of effective interventions by monitoring the type and intensity of treatment and supervision efforts, and by monitoring changes in psychological and community adjustment (e.g., Babchishin & Hanson, 2020).

Further research is also needed on the effect of illness and mortality on long term recidivism rates. There is anecdotal evidence that exceptionally good health may mitigate the protective effect of advanced age on sexual recidivism risk (Mattek & Hanson, 2018). Much knowledge could be gained by data linkage studies that combined criminal justice, health, and mortality records. In particular, it would inform case management decisions about individuals previously identified as high risk who are now in their advanced old age.

### **Conclusion**

Preventative detention measures, such as Sexually Violent Person's laws in the US and Dangerous Offender provisions in Canada, rest on the assumption that there are certain individuals who present a high risk for sexual recidivism. For such provisions to work, evaluators and decision-makers must be able to identify these high-risk individuals in advance. The current findings support these

assumptions. Although the ethical and legal implications of preventive detention need to be carefully considered, there are individuals who present a significant, enduring risk to public safety. Furthermore, evaluators using existing technologies can identify this problematic subgroup with reasonable accuracy.

## References

- Abbott, B. R. (2020). Illuminating the dark figure of sexual recidivism. *Behavioral Sciences and the Law*, 38(6), 543-558. <https://doi.org/10.1002/bsl.2494>
- Aelick, C. A., Babchishin, K. M., & Harris, A. J. R. (2020). Severe mental illness diagnoses and their association with reoffending in a sample of men adjudicated for sexual offences. *Sexual Offending: Theory, Research, and Prevention*, 15(1), Article e3123. <https://doi.org/10.5964/sotrap.3123>
- Alper, M., & Durose, M. R. (2019). *Recidivism of sex offenders released from state prison: A 9-year follow-up (2005-14)*. Special Report NCJ 251773. Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice.
- Babchishin, K. M., & Hanson, R. K. (2020). Monitoring changes in risk of reoffending: A prospective study of 632 men on community supervision. *Journal of Consulting and Clinical Psychology*, 88(10), 886-898. <https://doi.org/10.1037/ccp0000601>
- Babchishin, K. M., Hanson, R. K., & Helmus, L. (2012a). Even highly correlated measures can add incrementally to predicting recidivism among sex offenders. *Assessment*, 19(4), 442-461. <https://doi.org/10.1177/1073191112458312>
- Babchishin, K. M., Hanson, R. K., & Helmus, L. (2012b). Communicating risk for sex offenders: Risk ratios for Static-2002R. *Sexual Offender Treatment*, 12(2), 1-12.
- Baudin, C., Nilsson, T., Sturup, J., Wallinius, M., & Andiné, P. (2021). A Static-99R validation study on Individuals with mental disorders: 5 to 20 years of fixed follow-up after sexual offenses. *Frontiers in Psychology*, 12, 140. <https://doi.org/10.3389/fpsyg.2021.625996>
- Blais, J., & Bonta, J. (2015). Tracking and managing high risk offenders: A Canadian initiative. *Law and Human Behavior*, 39(3), 253-265. <https://doi.org/10.1037/lhb0000109>

- 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. <http://dx.doi.org/10.1037/pas0000377>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2<sup>nd</sup> ed.). Wiley.
- Bourgon, G., Mugford, R., Hanson, R. K., & Coligado, M. (2018). Offender risk assessment practices vary across Canada. *Canadian Journal of Criminology and Criminal Justice, 60*(2), 167-205. <https://doi.org/10.3138/cjccj.2016-0024>
- 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>
- Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge Academic.
- Criminal Code of Canada. RSC (1985). c.C-46.
- Ducro, C., Telle, E., & Pham, T. (2020). Recidivism rates of sex offenders released from the French Belgian Judicial System. *Acta Psychiatrica Belgica, 120*(1), 43-49.
- Gonçalves, 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), 203–219. <https://doi.org/10.1177/1079063218821117>
- 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. (2022). *Prediction statistics for psychological assessment*. American Psychological Association.
- Hanson, R. K., Babchishin, K. M., Helmus, L. M., Thornton, D., & Phenix, A. (2017). Communicating the results of criterion referenced prediction measures: Risk categories for the Static-99R and Static-

2002R sexual offender risk assessment tools. *Psychological Assessment*, 29(5), 582-597.

<https://doi.org/10.1037/pas0000371>

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66(2), 348-362.

<https://doi.org/10.1037/0022-006X.66.2.348>

Hanson, R. K., & Harris, A. J. R. (2000a). *The Sex Offender Need Assessment Rating (SONAR): A method for measuring change in risk levels*. User Report 2000-01. Department of the Solicitor General of Canada.

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

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.

<https://doi.org/10.1177/0886260514526062>

Hanson, R. K., Harris, A. J. R., Letourneau, E., Helmus, L. M., & Thornton, D. (2018). Reductions in risk based on time offense free in the community: Once a sexual offender, not always a sexual offender. *Psychology, Public Policy and Law*, 24(1), 48-63. <https://doi.org/10.1037/law0000135>

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., Helmus, L., & Thornton, D. (2010). Predicting recidivism among sexual offenders: A multi-site study of Static-2002. *Law and Human Behavior*, 34(3), 198-211.

<https://doi.org/10.1007/s10979-009-9180-1>

Hanson, R. K., Lloyd, C. D., Helmus, L., & Thornton, D. (2012). Developing non-arbitrary metrics for risk communication: Percentile ranks for the Static-99/R and Static-2002/R sexual offender risk

scales. *International Journal of Forensic Mental Health*, 11(1), 9-23.

<https://doi.org/10.1080/14999013.2012.667511>

Hanson, R.K., Lunetta, A., Phenix, A., Neely, J., & Epperson, D. (2014). The field validity of Static-99R Sex Offender Risk Assessment Tool in California. *Journal of Threat Assessment and Management*, 1(2), 102-117. <https://doi.org/10.1037/tam0000014>

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>

Hanson, R. K., & Thornton, D. (2003). *Notes on the development of Static-2002*. (Corrections Research User Report No. 2003-01). Department of the Solicitor General of Canada. <https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/nts-dvlpmnt-sttc/nts-dvlpmnt-sttc-eng.pdf>

Hanson, R. K., Thornton, D., Helmus, L.-M., & Babchishin, K. M. (2016). What sexual recidivism rates should be associated with Static-99R and Static-2002R scores? *Sexual Abuse: A Journal of Research and Treatment*, 28(3), 218-252. <https://doi.org/10.1177/107906321557410>

Harris, A. J. R. & Hanson, R. K. (2004). *Sex offender recidivism: A simple question*. Corrections Users Report No. 2004-03: Public Safety and Emergency Preparedness Canada.

Helmus, L. M., & Babchishin, K. M. (2017). Primer on risk assessment and the statistics used to evaluate its accuracy. *Criminal Justice and Behavior*, 44(1), 8-25. <https://doi.org/10.1177/0093854816678898>

Helmus, L., & Hanson, K. (2007). Predictive validity of the Static-99 and Static-2002 for sex offenders on community supervision. *Sexual Offender Treatment*, 2, 1-14.

Helmus, L. M., Hanson, R. K., Murrie, D. C., & Zabarauckas, C. L. (2021). Field validity of Static-99R and STABLE-2007 with 4,433 men serving sentences for sexual offences in British Columbia: New findings and meta-analysis. *Psychological Assessment, 33*(7), 581-595.

<https://doi.org/10.1037/pas0001010>

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.

<https://doi.org/10.1177/0093854812443648>

Helmus, L. M., Kelley, S. M., Frazier, A., Fernandez, Y. M., Lee, S. C., Rettenberger, M., & Boccaccini, M. T. (2022). Static-99R: Strengths, limitations, predictive accuracy meta-analysis, and legal admissibility review. *Psychology, Public Policy, and Law*. Advance online publication.

<https://doi.org/10.1037/law0000351>

Helmus, L. M., Lee, S. C., Phenix, A., Hanson, R. K., & Thornton, D. (2021). *Static-99R & Static-2002R evaluators' workbook*. SAARNA: The Society for the Advancement of Actuarial Risk Need Assessment. Available at [www.saarna.org](http://www.saarna.org)

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. <https://doi.org/10.1177/1079063211409951>

Jung, S., Ennis, L., Hermann, C. A., Pham, A. T., Choy, A. L., Corabian, G., & Hook, T. (2017). An evaluation of the reliability, construct validity, and factor structure of the Static-2002R. *International Journal of Offender Therapy and Comparative Criminology, 61*(4), 464–487. <https://doi.org/10.1177/0306624X15595228>

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. Little, Brown Spark.

- Kelley, S. M., Ambroziak, G., Thornton, D., & Barahal, R. M. (2020). How do professionals assess sexual recidivism risk? An updated survey of practices. *Sexual Abuse, 32*(1), 3-29.  
<https://doi.org/10.1177/1079063218800474>
- Knighton, J. C., Murrie, D. C., Boccaccini, M. T., & Turner, D. B. (2014). How likely is “likely to reoffend” in sex offender civil commitment trials? *Law and Human Behavior, 38*(3), 293–304.  
<https://doi.org/10.1037/lhb0000079>
- Laporte, L., Poulin, B., Marleau, J., Roy, R., & Webanck, T. (2003). Filicidal women: jail or psychiatric ward? *The Canadian Journal of Psychiatry, 48*(2), 94-98.  
<https://doi.org/10.1177/070674370304800205>
- Lave, T. R., Prescott, J. J., & Bridges, G. (2021). The problem with assumptions: Revisiting “The dark figure of sexual recidivism”. *Behavioral Sciences & the Law, 39*(3), 279-306.  
<https://doi.org/10.1002/bsl.2508>
- Lee, S. C., & Hanson, R. K. (2021). Updated 5-year and new 10-year sexual recidivism rate norms for Static-99R with routine/complete samples. *Law and Human Behavior, 45*(1), 24–38.  
<https://doi.org/10.1037/lhb0000436>
- Lee, S. C., Hanson, R. K., & Blais, J. (2020). Predictive accuracy of the Static-99R and Static-2002R risk tools for identifying Indigenous and White individuals at high risk for sexual recidivism in Canada. *Canadian Psychology/Psychologie canadienne, 61*(1), 42-57.  
<https://doi.org/10.1037/cap0000182>
- Lee, S. C., Hanson, R. K., Fullmer, N., Neeley, J, & Ramos, K. (2018). *The Predictive validity of Static-99R over 10 years for sexual offenders in California: 2018 update*. State Authorized Risk Assessment Tools for Sex Offenders (SARATSO).  
[http://saratso.org/pdf/Lee\\_Hanson\\_Fullmer\\_Neeley\\_Ramos\\_2018\\_The\\_Predictive\\_Validity\\_of\\_S\\_.pdf](http://saratso.org/pdf/Lee_Hanson_Fullmer_Neeley_Ramos_2018_The_Predictive_Validity_of_S_.pdf)

Lee, S. C., Restrepo, A., Satariano, A, & Hanson, R. K. (2016). *The predictive validity of Static-99R for sex offenders in California: 2016 update*. State Authorized Risk Assessment Tools for Sex Offenders (SARATSO).

[http://saratso.org/pdf/ThePredictiveValidity\\_of\\_Static\\_99R\\_forSexualOffenders\\_inCalifornia\\_2016v1.pdf](http://saratso.org/pdf/ThePredictiveValidity_of_Static_99R_forSexualOffenders_inCalifornia_2016v1.pdf)

Lehmann, R. J. B, Hanson, R. K., Babchishin, K., Gallasch-Nemitz, F., Biedermann, J., & Dahle, K.-P. (2013). Interpreting multiple risk scales for sex offenders: Evidence for averaging. *Psychological Assessment, 25*(3), 1019-1024. <https://doi.org/10.1037/a0033098>

McGrath, R. J., Lasher, M. P., & Cumming, G. F. (2012). The Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Psychometric properties and incremental predictive validity With Static-99R. *Sexual Abuse, 24*(5), 431–458. <https://doi.org/10.1177/1079063211432475>

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. <https://doi.org/10.1177/1079063210366039>

Mattek, R., & Hanson, R. K. (2018). Committed as a violent sexual predator in his 10th decade: A case study. *Archives of Sexual Behavior, 47*(2), 543-550. <https://doi.org/10.1007/s10508-017-1041-2>

Moore, L. (2018). *Static risk assessment of sexual offenders in New Zealand: Predictive accuracy, classification of risk, and the moderating effect of time offence-free in the community*. Doctoral Dissertation, Psychology, University of Canterbury, Christchurch, NZ.

Neal, T. M., & Grisso, T. (2014). Assessment practices and expert judgment methods in forensic psychology and psychiatry: An international snapshot. *Criminal Justice and Behavior, 41*(12), 1406-1421. <https://doi.org/10.1177/0093854814548449>

Olver, M. E., Kelley, S. M., Johnson, L., & Wong, S. C. P. (2020). *Violence Risk Scale-Sexual Offense Version (VRS-SO): Users' Workbook*. Retrieved from <https://psynergy.ca/vrs-so>

Olver, M. E., Kelley, S. M., Kingston, D. A., Beggs Christofferson, S. M., Thornton, D., & Wong, S. C. P.

(2021). Incremental contributions of static and dynamic sexual violence risk assessment integrating Static-99R and VRS-SO common language risk levels. *Criminal Justice and Behavior*, 48(8), 1091-1110. <https://doi.org/10.1177/0093854820974400>

Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the

Violence Risk Scale-Sexual Offender version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment*, 19(3), 318–329. <https://doi.org/10.1037/1040-3590.19.3.318>

Prentky, R. A., Barbaree, H. E., & Janus, E. S. (2015). *Sexual predators: Society, risk, and the law*.

Routledge.

Public Safety Canada. (2020). *2019 Corrections and Conditional Release Statistical Overview*.

<https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/ccrso-2019/index-en.aspx#e3>

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for

Statistical Computing. <https://www.R-project.org/>

R. v Currie. (1997). 2. S.C.R. 260.

Raymond, B. C., McEwan, T. E., Davis, M. R., Reeves, S. G., & Ogloff, J. (2020). Investigating the predictive

validity of Static-99/99R scores in a sample of older sexual offenders. *Psychiatry, Psychology and Law*, 28(1), 120–134. <https://doi.org/10.1080/13218719.2020.1767714>

Rockhill, B., Byrne, C., Rosner, B., Louie, M. M., & Colditz, G. (2003). Breast cancer risk prediction with a

log-incidence model: Evaluation of accuracy. *Journal of Clinical Epidemiology*, 56(9), 856-861. [https://doi.org/10.1016/S0895-4356\(03\)00124-0](https://doi.org/10.1016/S0895-4356(03)00124-0)

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*, 30(8), 887–907.

<https://doi.org/10.1177/1079063217712216>

- 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. <https://doi.org/10.1007/s10979-005-6832-7>
- Scurich, N., & John, R. S. (2019). The dark figure of sexual recidivism. *Behavioral Sciences and the Law, 37*(2), 158-175. <https://doi.org/10.1002/bsl.2400>
- Singer, J., & Willet, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.
- Smid, W. J., Kamphuis, J. H., Wever, E. C., & Van Beek, D. J. (2014). A comparison of the predictive properties of nine sex offender risk assessment instruments. *Psychological Assessment, 26*(3), 691-703. <https://doi.org/10.1037/a0036616>
- Stephens, S., Cantor, J. M., Goodwill, A. M., & Seto, M. C. (2017). Multiple indicators of sexual interest in prepubescent or pubescent children as predictors of sexual recidivism. *Journal of Consulting and Clinical Psychology, 85*(6), 585–595. <https://doi.org/10.1037/ccp0000194>
- Tanguay, D. (2004). *Récits motards: Examen d'un conflit en milieu criminel*. [Biker stories: Study of a conflict in the criminal underworld]. Unpublished Master's thesis. École de Criminologie, Université de Montréal, Québec, Canada.
- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14*(2), 137-151. <https://doi.org/10.1177/107906320201400205>
- Thornton, D. (2016). Structured risk assessment. In A. Phenix & H.M. Hoberman (Eds.), *Sexual offending: Predisposing antecedents, assessments and management* (pp. 503-520). Springer.
- Thornton, D., Hanson, R. K., Kelley, S. M., & Mundt, J. C. (2021). Estimating lifetime and residual risk for individuals who remain sexual offense free in the community: Practical applications. *Sexual Abuse, 33*(1), 3-33. <https://doi.org/10.1177/1079063219871573>

Thornton, D., & Knight, R. A. (2015). Construction and validation of SRA-FV need assessment. *Sex Abuse: A Journal of Research and Treatment*, 27(4), 360-375.

<https://doi.org/10.1177/1079063213511120>

Viallon, V., Ragusa, S., Clavel-Chapelon, F., & Bénichou, J. (2009). How to evaluate the calibration of a disease risk prediction tool. *Statistics in Medicine*, 28(6), 901–916.

<https://doi.org/10.1002/sim.3517>

Viechtbauer, W. (2010). Conducting meta-analyses in R with the *metafor* package. *Journal of Statistical Software*, 36(3), 1-48. <https://doi.org/10.18637/jss.v036.i03>

Zgoba, K., Miner, M., Knight, R., Letourneau, E., Levenson, J., & Thornton, D. (2012). *A multi-state recidivism study using Static-99R and Static-2002R risk scores and tier guidelines from the Adam Walsh Act*. U.S. Department of Justice Document No. 240099.

<https://www.ojp.gov/ncjrs/virtual-library/abstracts/multi-state-recidivism-study-using-static-99r-and-static-2002-risk>

**Table 1**

*Calibration Analyses (E/O Index) with High Risk/High Need Sexual Recidivism Norms for Static-99R and Static-2002R*

		N	Expected recidivism		Observed recidivism		E/O	95% CI
			n of recidivists (E)	Percent (E)	n of recidivists (O)	Percent (O)		
<b>Static-99R</b>								
5-year								
I	Very low	1	0.056	5.6	0	0	-	-
II	Below average	6	0.42	7.0	0	0	-	-
III	Average	39	4.82	12.4	3	7.7	1.61	0.52, 4.98
IVa	Above average	51	9.95	19.5	7	13.7	1.42	0.68, 2.98
IVb	Well above average	87	26.51	30.5	19	21.8	1.40	0.89, 2.19
	All	184	41.76	22.7	29	15.8	<b>1.44</b>	<b>1.001, 2.07</b>
10- year								
I	Very low	1	0.11	10.6	0	0	-	-
II	Below average	6	0.76	12.6	0	0	-	-
III	Average	39	8.03	20.6	5	12.8	1.61	0.67, 3.86
IVa	Above average	49	14.72	30.0	10	20.4	1.47	0.79, 2.74
IVb	Well above average	87	36.46	41.9	38	43.7	0.96	0.70, 1.32
	All	182	60.07	33.0	53	29.1	1.13	0.87, 1.48

Table continues

Table 1 continued

		<i>N</i>	Expected recidivism		Observed recidivism		E/O	95% CI
			<i>n</i> of recidivists (E)	Percent (E)	<i>n</i> of recidivists (O)	Proportion (O)		
<b>Static-2002R</b>								
5-year								
I	Very low	1	0.074	7.4	0	0	-	-
II	Below average	2	0.18	9.0	0	0	-	-
III	Average	23	3.14	13.7	2	8.7	1.57	0.39, 6.28
IVa	Above average	53	11.3	21.3	5	9.4	2.26	0.94, 5.44
IVb	Well above average	105	35.1	33.4	22	21.0	<b>1.60</b>	<b>1.05, 2.42</b>
	All	184	49.8	27.1	29	15.8	<b>1.72</b>	<b>1.19, 2.47</b>

Note: Bolded values are statistically significant at  $p < .05$ .

**Table 2***Calibration (E/O index) of 15-year and 20-year sexual recidivism rates for Static-99R*

		<i>N</i>	Expected recidivism		Observed recidivism		E/O	95% CI
			<i>n</i> of recidivists (E)	Proportion (E)	<i>n</i> of recidivists (O, SE)	Proportion (O)		
15-year								
I	Very low	1	0.11	11.0	0	0	-	-
II	Below average	6	0.80	13.3	0	0	-	-
III	Average	39	8.60	22.1	9.09 (2.65)	23.3	0.95	0.53, 1.68
IVa	Above average	54	17.4	32.3	14.5 (3.32)	26.9	1.20	0.77, 1.89
IVb	Well above average	90	41.2	45.8	47.5 (4.77)	52.8	0.87	0.71, 1.06
	All	190	68.1	35.9	71.2 (6.74)	37.5	0.96	0.80, 1.15
20- year								
I	Very low	1	0.12	12.0	0	0	-	-
II	Below average	6	0.86	14.4	0	0	-	-
III	Average	39	9.36	24.0	9.09 (2.65)	23.3	1.03	0.58, 1.82
IVa	Above average	54	19.2	35.6	16.7 (3.48)	31.0	1.15	0.76, 1.73
IVb	Well above average	90	45.3	50.3	48.6 (4.77)	54.0	0.93	0.77, 1.13
	All	190	74.8	39.4	74.5 (6.82)	39.2	1.00	0.84, 1.20

Note: Median follow-up time of 20.5 years, with 89% followed for at least 15 years.

**Table 3***Calibration (E/O index) of 15-year and 20-year sexual recidivism rates for Static-2002R*

		<i>N/n</i>	Expected recidivism		Observed recidivism		E/O	95% CI
			<i>n</i> of recidivists (E)	Proportion (E)	<i>n</i> of recidivists (O, SE)	Proportion (O)		
15-year								
I	Very low	1	0.13	12.8	0	0	-	-
II	Below average	2	0.31	15.5	0	0	-	-
III	Average	23	5.22	22.7	5.00 (1.98)	21.7	1.04	0.48, 2.27
IVa	Above average	54	18.4	34.0	11.2 (3.01)	20.8	1.64	0.97, 2.77
IVb	Well above average	110	53.8	48.9	54.9 (5.30)	49.9	0.98	0.81, 1.19
	All	190	77.9	41.0	71.2 (6.74)	37.5	1.09	0.91, 1.32
20- year								
I	Very low	1	0.139	13.9	0	0	-	-
II	Below average	2	0.336	16.8	0	0	-	-
III	Average	23	5.71	24.8	5.00 (1.98)	21.7	1.14	0.53, 2.48
IVa	Above average	54	20.2	37.4	13.5 (3.26)	25.0	1.49	0.93, 2.40
IVb	Well above average	110	60.3	54.8	55.9 (5.30)	50.9	1.08	0.89, 1.30
	All	190	86.7	45.6	74.5 (6.82)	39.2	1.16	0.97, 1.39

Note: Median follow-up time of 20.5 years, with 89% followed for at least 15 years.

**Table 4**

*Comparison of Logistic Regression Parameters in the current sample to the High Risk/High Needs (HRHN) and Routine/Complete Norms for Static-99R and Static-2002R.*

	<i>B (SE)</i>	% or odds ratio	Comparison with Current		
			<i>Q<sub>between</sub></i> ( <i>df</i> = 1)	<i>p</i>	<i>f</i> <sup>2</sup>
<b>Static-99R</b>					
5-year					
Intercept ( <i>B</i> <sub>0</sub> )					
Current	-2.725678 (0.474215)	6.1%			
HRHN	-2.063626 (0.153444)	11.3%	1.764	.184	43.3
Routine/Complete	-3.035568 (0.066871)	4.6%	0.419	.518	0.00
Slope ( <i>B</i> <sub>1</sub> )					
Current	0.300292 (0.108799)	1.35			
HRHN	0.250091 (0.042447)	1.28	0.185	.667	0.00
Routine/Complete	0.372803 (0.021437)	1.45	0.428	.513	0.00
10-year					
Intercept ( <i>B</i> <sub>0</sub> )					
Current	-2.155873 (0.386044)	10.4%			
HRHN	-1.442304 (0.186037)	19.1%	2.773	.096	63.9
Routine/Complete	-2.556535 (0.118530)	7.2%	0.984	.321	0.00
Slope ( <i>B</i> <sub>1</sub> )					
Current	0.372137 (0.092412)	1.45			
HRHN	0.230673 (.056260)	1.26	1.710	.191	41.5
Routine/Complete	0.364911 (0.037367)	1.44	0.005	.942	0.00

Table continues

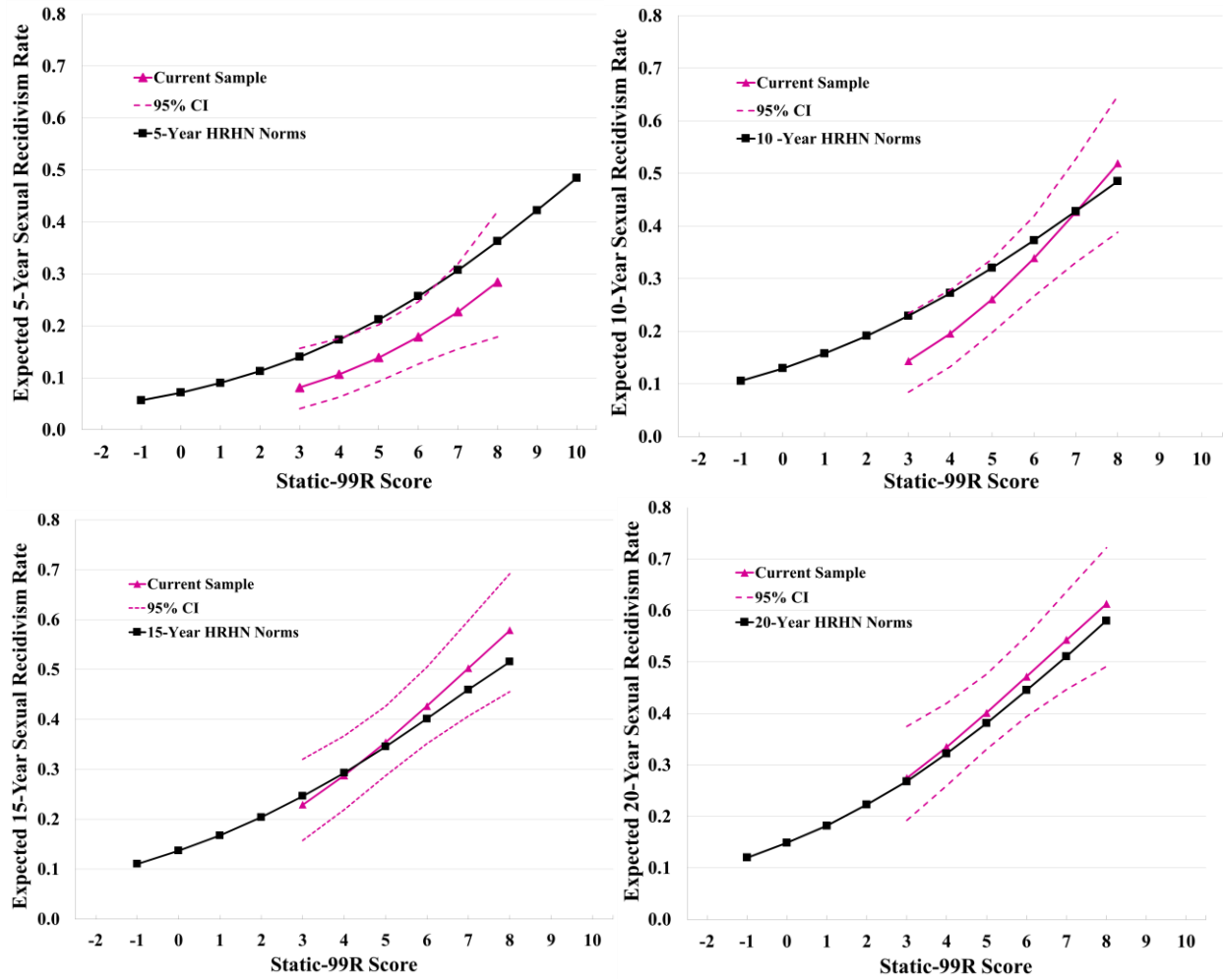
**Table 4** continued

	<i>B</i> ( <i>SE</i> )	% or odds ratio	Comparison with Current		
			<i>Q</i> <sub>between</sub> ( <i>df</i> = 1)	<i>p</i>	<i>f</i> <sup>2</sup>
<b>Static-2002R</b>					
5-year					
Intercept ( <i>B</i> <sub>0</sub> )					
Current	-2.865787 (0.538280)	5.4%			
HRHN	-1.877664 (0.173236)	13.3%	3.054	.081	67.2
Routine/Complete	-2.623861 (0.109268)	6.8%	0.194	.660	0.00
Slope ( <i>B</i> <sub>1</sub> )					
Current	0.276160 (0.105218)	1.32			
HRHN	0.217648 (0.050133)	1.24	0.252	.616	0.00
Routine/Complete	0.395056 (0.036062)	1.48	1.143	.285	12.5

Note: The sample size for current analyses were as follows: Static-99R/5-year 29/184, Static-99R/10-year 53/182, and for Static-2002R/5-year 29/184 (number of recidivists/*N*). Static-99R norms from Lee and Hanson (2021); Static-2002R norms from Hanson et al. (2016).

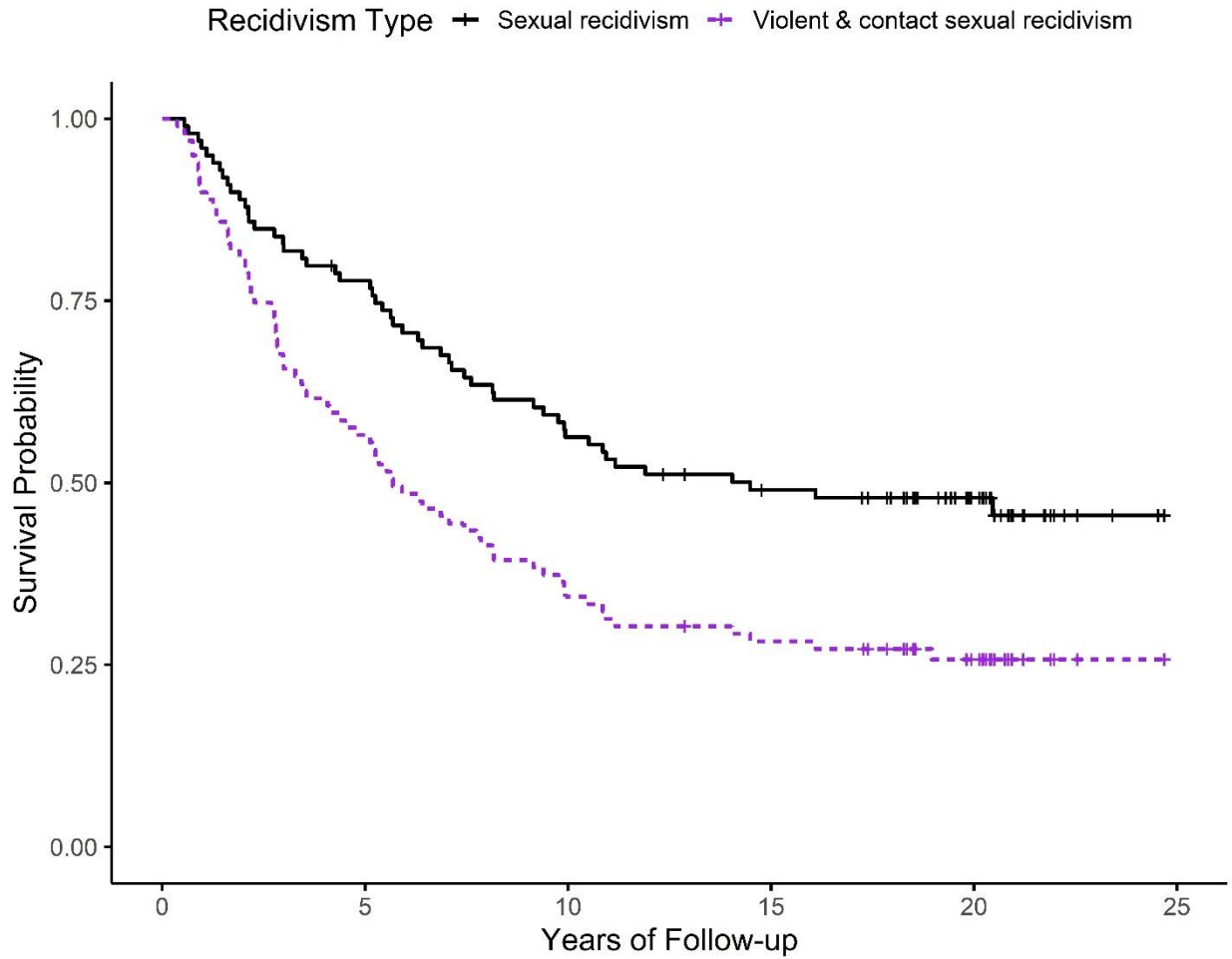
**Figure 1**

*Logistic curves for 5, 10, 15, and 20-year sexual recidivism rates of the current sample with the Static-99R High Risk/High Need (HRHN) norms*



**Figure 2**

*Sexual and violent (including contact sexual) recidivism rates for individuals identified as well above average risk to reoffend (Kaplan-Meier survival curves; n = 99).*



## Appendix

### Assigning standardized risk levels based on averaging Static-99R and Static-2002R scores.

Previous research has indicated that the average of Static-99R and Static-2002R scores has greater predictive accuracy than either measure on its own (Babchishin et al., 2012a; Lehmann et al., 2013). It is not obvious, however, how evaluators should calculate the average for specific cases because the raw scores of these measures are in different metrics. Consistent with previous research, in the current study the scores were averaged in the metric of logged hazard ratios (Babchishin et al., 2012a; Babchishin et al., 2012b; Lehmann et al., 2013). Hazard ratios were privileged because relative risk inferences are stable features of these prediction tools (Helmus, Hanson et al., 2012). The hazard ratios were log transformed because the effect of a one-point increase in STATIC scores is linear for logged hazard ratios (not for raw hazard ratios).

In the STATIC norms, logged hazard ratios are presented from Cox regression survival analysis, centered on the median values ( $\beta = 0.3318$ ; median of 2 for Static-99R;  $\beta = .322$ , median of 3 for Static-2002R; Helmus, Lee et al., 2021). In order to translate the averages into standardized risk levels, the boundaries between the risk levels were calculated as the midpoint between the hazards for adjacent risk levels. Specifically, when identifying the boundary between Level IVa and IVb for Static-99R, the logged hazard ratio for a score of 5 (the top of Level IVa) was 0.9954 and the logged hazard ratio for a score of 6 (the bottom of Level IVb) was 1.3272; consequently, the boundary was  $1.1613 = (0.9954 + 1.3272)/2$ . For Static-2002R, the boundary between Level IVa and Level IVb was similar (1.127) but not identical. Consequently, for the purpose of assigning risk levels, the boundary was calculated as the average of the boundaries for Static-99R and Static-2002R (i.e., logged hazard ratio of 1.14415; hazard ratio of  $e^{1.14415} = 3.140$ ).

## Online Supplemental Material

Table S1

Sexual recidivism rates associated with Static-99R scores at 5 years and 10 years based on fixed follow-up

	5-year				10-year			
	Observed		Smoothed Logistic Regression		Observed		Smoothed Logistic Regression	
	%	n / N	%	95%CI	%	n / N	%	95%CI
<b>Scores</b>								
-3	0.0	0/1	1.44	[0.21, 9.26]	0.0	0/1	1.77	[0.36, 8.33]
-2	-	0/0	1.93	[0.35, 10.1]	-	0/0	2.55	[0.61, 9.95]
-1	0.0	0/1	2.59	[0.57, 10.9]	0.0	0/1	3.65	[1.06, 11.9]
0	0.0	0/5	3.47	[0.94, 11.9]	0.0	0/5	5.21	[1.81, 14.1]
1	11.1	1/9	4.63	[1.55, 13.0]	11.1	1/9	7.39	[3.08, 16.7]
2	0.0	0/7	6.15	[2.52, 14.2]	0.0	0/7	10.4	[5.15, 19.8]
3	8.7	2/23	8.13	[4.04, 15.7]	17.4	4/23	14.4	[8.44, 23.4]
4	18.2	4/22	10.7	[6.29, 17.5]	19.0	4/21	19.6	[13.3, 27.9]
5	10.3	3/29	13.9	[9.31, 20.2]	21.4	6/28	26.1	[19.7, 33.7]
6	13.2	5/38	17.9	[12.7, 24.6]	39.5	15/38	33.9	[26.7, 41.9]
7	19.2	5/26	22.7	[15.6, 31.9]	42.3	11/26	42.7	[33.1, 52.8]
8	41.2	7/17	28.4	[17.9, 42.0]	52.9	9/17	51.9	[38.8, 64.8]
9	40.0	2/5	34.9	[19.9, 53.7]	60.0	3/5	61.0	[44.2, 75.6]
10	0.0	0/1	42.0	[21.7, 65.4]	0.0	0/1	69.4	[49.5, 84.1]
Total	15.8	29/184			29.1	53/182		

Logit (5-year) =  $-2.725678[0.474215] + 0.300292[0.108799](\text{Static-99R} - 2)$ ; correlation of estimates =  $-0.899396$

Logit (10-year) =  $-2.155873[0.386044] + 0.372137[0.092412](\text{Static-99R} - 2)$ ; correlation of estimates =  $-0.894705$

Note: In the above equations, the parameters are followed by their standard errors in brackets.

**Table S2**

*Sexual recidivism rates associated with Static-99R score at 15 years and 20 years based on Kaplan-Meier survival analysis*

	15-year					20-year				
	N	Observed		Smoothed Logistic Regression		Observed		Smoothed Logistic Regression		
		%	n (n)	%	95%CI	%	n / N	%	95%CI	
<b>Scores</b>										
-3	1	0.0	0 (0)	3.12	[0.78, 11.6]	0.0	0(0)	3.44	[0.89, 1.24]	
-2	0	-	0 (0)	4.38	[1.29, 13.8]	-	0(0)	4.82	[1.46, 14.7]	
-1	1	0.0	0 (0)	6.11	[2.13, 16.3]	0.0	0(0)	6.71	[2.40, 17.4]	
0	5	0.0	0 (0)	8.46	[3.48, 19.2]	0.0	0(0)	9.27	[3.90, 20.5]	
1	9	11.1	1.0 (1)	11.6	[5.61, 22.5]	11.1	1.0 (1)	12.7	[6.27, 24.0]	
2	7	14.3	1.0 (1)	15.7	[8.90, 26.3]	14.3	1.0 (1)	17.1	[9.88, 28.0]	
3	23	30.7	7.1 (7)	21.0	[13.7, 30.7]	30.7	7.1 (7)	22.7	[15.1, 32.6]	
4	23	22.4	5.2 (5)	27.4	[20.2, 35.9]	22.4	5.2 (5)	29.4	[22.1, 38.0]	
5	31	30.0	9.3 (9)	34.9	[27.9, 42.5]	37.2	11.5 (12)	37.2	[30.1, 44.9]	
6	38	50.1	19.0 (19)	43.2	[35.6, 51.2]	50.1	19.0 (19)	45.7	[37.9, 53.7]	
7	26	50.0	13.0 (13)	52.0	[42.2, 61.5]	54.5	14.2 (14)	54.5	[44.7, 63.9]	
8	18	55.6	10.0 (10)	60.6	[48.1, 71.8]	55.6	10.0 (10)	63.0	[50.6, 73.8]	
9	7	81.0	5.7 (6)	68.6	[53.5, 80.6]	81.0	5.7 (6)	70.7	[56.0, 82.1]	
10	1	0.0	0 (0)	75.6	[58.6, 87.2]	0.0	0 (0)	77.4	[61.1, 88.2]	
Total	190	37.5				39.2				

Note: In the column *n(n)*, *n* indicates the observed number of recidivists estimated by Kaplan-Meier survival analysis and (*n*) is *n* rounded to an integer (for logistic regression).

Logit (15 year) = -1.678214 [0.330643] + 0.351399 [0.082134] (Static-99R – 2). Correlation of estimates = -0.87679797

Logit (20 year) = -1.577884 [0.322587] + 0.351363 [0.081044] (Static-99R – 2). Correlation of estimates = -0.87252089

Note: In the above equations, the parameters are followed by their standard errors in brackets.

**Table S3***Sexual recidivism rates associated with Static-2002R scores at 5 years and 10 years based on fixed follow-up*

	5-year				10-year			
	Observed		Logistic Regression		Observed		Logistic Regression	
	%	n / N	%	95%CI	%	n / N	%	95%CI
<b>Scores</b>								
-1	0.0	0/1	1.85	[0.30, 10.7]	0.0	0/1	1.62	[0.33, 7.46]
0	-	0/0	2.43	[0.48, 11.4]	-	0/0	2.37	[0.59, 9.08]
1	0.0	0/2	3.17	[0.77, 12.2]	0.0	0/2	3.46	[1.03, 11.0]
2	0.0	0/7	4.14	[1.22, 13.1]	0.0	0/7	5.03	[1.79, 13.3]
3	14.3	1/7	5.39	[1.94, 14.1]	14.3	1/7	7.25	[3.09, 16.1]
4	11.1	1/9	6.98	[3.05, 15.2]	22.2	2/9	10.4	[5.25, 19.4]
5	5.0	1/20	9.00	[4.72, 16.5]	10.0	2/20	14.6	[8.70, 23.4]
6	12.1	4/33	11.5	[7.07, 18.3]	18.2	6/33	20.1	[13.9, 28.3]
7	14.8	4/27	14.7	[10.0, 20.9]	28.0	7/25	27.1	[20.6, 34.7]
8	22.2	6/27	18.5	[13.1, 25.4]	44.4	12/27	35.4	[27.9, 43.7]
9	9.7	3/31	23.0	[15.6, 32.6]	29.0	9/31	44.8	[34.7, 55.3]
10	60.0	9/15	28.2	[17.5, 42.3]	73.3	11/15	54.5	[40.9, 67.4]
11	0.0	0/5	34.2	[19.1, 53.3]	60.0	3/5	63.8	[46.8, 78.0]
Total	15.8	29/184			29.1	53/182		

Logit (5-year) = -2.865787 [0.538280] + 0.276160 [0.105218] (Static-2002R – 3). Correlation of estimates = -0.923317.

Logit (10-year) = -2.548537 [0.458494] + 0.389657 [0.091938] (Static-2002R – 3). Correlation of estimates = -0.925513.

Note: In the above equations, the parameters are followed by their standard errors in brackets.

**Table S4**

*Sexual recidivism rates associated with Static-2002R scores at 15 years and 20 years based on Kaplan-Meier survival analysis*

	N	15-year				20-year			
		Observed		Smoothed Logistic Regression		Observed		Smoothed Logistic Regression	
		%	n (n)	%	95%CI	%	n / N	%	95%CI
<b>Scores</b>									
-1	1	0.0	0 (0)	2.48	[0.61, 9.54]	0.0	0 (0)	3.18	[0.83, 11.5]
0	0	-	0 (0)	3.61	[1.05, 11.7]	-	0 (0)	4.52	[1.39, 13.8]
1	2	0.0	0 (0)	5.23	[1.80, 14.2]	0.0	0 (0)	6.39	[2.31, 16.5]
2	7	0.0	0 (0)	7.51	[3.06, 17.3]	0.0	0 (0)	8.95	[3.81, 19.6]
3	7	28.6	2.0 (2)	10.7	[5.13, 20.9]	28.6	2.0 (2)	12.4	[6.21, 23.3]
4	9	33.3	3.0 (3)	15.0	[8.45, 25.1]	33.3	3.0 (3)	16.9	[9.90, 27/5]
5	20	15.0	3.0 (3)	20.6	[13.5, 30.0]	15.0	3.0 (3)	22.7	[15.3, 32.3]
6	34	24.2	8.2 (8)	27.6	[20.5, 36.0]	30.7	10.4 (10)	29.7	[22.5, 38.2]
7	28	36.9	10.3 (10)	35.9	[28.9, 43.6]	40.8	11.4 (11)	37.9	[30.8, 45.5]
8	28	50.0	14.0 (14)	45.2	[37.2, 53.3]	50.0	14.0 (14)	46.8	[38.7, 55.0]
9	33	42.4	14.0 (14)	54.8	[44.6, 64.6]	42.4	14.0 (14)	55.9	[45.7, 65.6]
10	15	80.0	12.0 (12)	64.1	[51.2, 75.2]	80.0	12.0 (12)	64.6	[51.8, 75.5]
11	6	80.0	4.8 (5)	72.4	[57.3, 83.7]	80.0	4.8 (5)	72.4	[57.5, 83.6]
Total	190	37.5				39.2			

Logit (15-year) = -2.124851 [0.403935] + 0.386419 [0.084062] (Static-2002R – 3); Correlation of estimates = -0.91735935.

Logit (20-year) = -1.954604 [0.388157] + 0.365002 [0.081592] (Static-2002R – 3); Correlation of estimates = -0.91256369.

Note: In the above equations, the parameters are followed by their standard errors in brackets.

### Supplemental Appendix

#### Constructing a Need Score based on the SRA Framework

The SRA Need framework proposes that psychological risk factors related to sexual recidivism can be understood as long-term vulnerabilities (LTVs) falling into four domains comprising Sexual Interests, Distorted Attitudes, Relational Style, and Self-Management (Thornton, 2016). To be regarded as long-term vulnerability, a problem needs to have been present in an enduring and pervasive way such that it is not manifested solely in one context or solely over a limited period of time. An SRA Need assessment should provide a balanced assessment of a range of long-term vulnerabilities from at least three of the four domains. Indicators for these needs should demonstrate appropriate validity and reliability.

The table below summarizes the long-term vulnerabilities falling within each domain.

Domains	Sub-Domains	Long-term vulnerabilities (LTV)
Sexual Interests	Sexual Preoccupation	Intense, Impersonal Sexual Interests Sexual Coping Diverse Outlets
	Offense-Related Sexual Interests	Sexual Preference for Children Sexualized Violence
Distorted Attitudes	Victim Schema	Women as Deceitful/Malicious Adversarial Sexual Attitudes Child Abuse Supportive Beliefs
	Rights & Means Schema	Excessive Entitlement Machiavellianism
Relational Style	Inadequate Relational Style	Dysfunctional Self-Evaluation Emotional Congruence with Children
	Lack of Emotionally-Intimate Relationships with Adults	Lack of Sustained Marital-Type Relationships Marital Relationships marred by repeated Violence/Infidelity
	Aggressive Relational Style	Callousness Grievance Thinking Poorly Managed Anger
Self-Management	Social Deviance	Lifestyle Impulsiveness Resistance to Rules and Supervision
	Dysfunctional Coping	Poor Problem-Solving Poor Emotional-Control

To construct an implementation for the current study the data dictionary was reviewed and variables classified according to their potential relevance to the constructs in the above table. Where alternative

descriptions were available variables which implied some persistence of the construct were chosen. Within each domain selected variables were subjected to exploratory factor analysis to group them into more homogeneous categories, better corresponding to the long-term vulnerabilities. For each domain at least two summary scores and levels of these were recoded so that they fell between 0 and 2. The summary scores for each domain were averaged to create a summary score for each domain and then these domain summary scores were summed to give a Need score running between 0 and 8.

The following table shows the within domain summary scores that resulted from this process.

<b>Domains</b>	<b>Within Domain Summary Scores</b>
Sexual Interests	Sexual Preoccupation Sexual Interest in Children
Distorted Attitudes	Attitude to Children Attitude to Women
Relational Style	Lack of Emotionally Intimate Relationships with Adults Isolation Angry Aggression
Self-Management	PCL-SV Factor Two (Social Deviance) Resistance to Supervision

It was not possible to find an operationalization of Sexualized Violence as a long-term vulnerability with the available variables, nor were there appropriate indicators for Dysfunctional Coping, but otherwise coverage of the domains was fairly good.

This implementation has two substantive limitations. First, with the exception of the PCL-SV, none of the indicators are established measures with known reliability and validity. Secondly, with some exceptions, most ratings are heavily based on the conduct of individuals as seen through the eyes of supervising agents who saw them for a limited period of time. This latter limitation means they likely overemphasize how long-term vulnerabilities manifested themselves during recent periods of supervision.

### **Determining a Threshold Level of Need to Identify HRHN Individuals**

Three considerations informed the selection of this threshold.

Conceptually we understand “high need” individuals as having multiple intense criminogenic needs. This can be contrasted with individuals whose criminogenic needs are less intense, or more circumscribed, or those whose offending resulted from transitory, situational, factors, and who therefore do not have sustained criminogenic needs. In SRA terms this concept suggests that “high need” individuals should have serious criminogenic needs from more than one domain and that if only two domains are involved then there should be serious problems in more than one criminogenic need within each domain.

For clinical practice, Thornton & Helmus (2021) have recommended using the HRHN norms when an individual’s score on a dynamic instrument is equal to or exceeds the mean score typically found in HRHN samples. Hanson and Thornton (2012) reported a review of mean scores in HRHN samples of

27.24, 3.26. and 14.70 for the VRS-SO Initial Dynamic Risk scale, the SRA-FV, and the STABLE-2007. Although these means vary substantially, resulting from how ratings are scaled in the different instruments, in each case the HRHN sample mean is close to half the maximum possible score on the instrument. The maximum on the VRS-SO Initial Dynamic Risk scale is 51, for SRA-FV it is 6, and for STABLE-2007 it is 26. Thus, in each case the means in HRHN samples are close to but just a little above half the maximum possible score on the instrument. This suggests that the threshold adopted on our instrument should correspond approximately to half the maximum possible score.

The threshold adopted should be consistent with selecting a higher need group while also assigning sufficient numbers to that group to permit statistical analyses. This suggests that the threshold adopted should be close to the sample median.

These considerations do not fully constrain what the threshold should be but they do provide guidance. The actual threshold adopted was a score of at least 4. Such a score can be obtained by having serious, multi-need, problems within two domains; having a serious single-need problem in all four domains; or some mixture of serious, single and multi-need, problems in three domains. It therefore is consistent with the first consideration. It corresponds to half the maximum possible score, so it is consistent with the second consideration. And it is close to the median found in the complete sample which was 4.2. It is therefore consistent with the third principle.

Importantly, both construction of the present SRA Need score and selection of the threshold were done while blind to the recidivism data.

**Figure S1**

*Logistic curves for 5, 15, and 20-year sexual recidivism rates of a current sample with the Static-2002R High Risk/High Need norms*

