Behavioral Sciences and the Law
Behav. Sci. Law (2013)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/bsl.2072

Sexually Violent Predators: Toward Reasonable Estimates of Recidivism Base Rates

Daniel J. Neller, Psy.D., ABPP (Forensic)*,† and Giovanni Petris, Ph.D.‡

The sexual recidivism rate of sex offenders is a controversial issue. Perhaps as controversial is the sexual recidivism rate of the select group of sex offenders who are examined pursuant to sexually violent predator (SVP) statutes. At present, reliable estimates of SVP recidivism are unavailable. We propose that reasonable estimates of SVP recidivism can be reached by considering three available pieces of data: (i) a likely recidivism rate of the general population of sex offenders; (ii) procedures typically followed by jurisdictions that civilly commit sex offenders; and (iii) classification accuracy of procedures. Although sexual recidivism rates vary across jurisdictions, the results of our analyses suggest sex offenders referred for examination pursuant to SVP statutes recidivate at substantially higher rates than typical sex offenders. Our results further suggest that sex offenders recommended for commitment as SVPs recidivate at even greater rates than SVP respondents who are not recommended for commitment. We discuss practice and policy implications of these findings. Copyright © 2013 John Wiley & Sons, Ltd.

Dr. Cy N. Tist, an academician who dedicated his career to researching sex offenders, has retired from his teaching position. He has chosen to apply his knowledge and skills to a clinical setting – assessment of sex offenders pursuant to sexually violent predator (SVP) statutes. Dr. Tist believes he will be able to identify a select group of high-risk[1] sex offenders by applying to a clinical endeavor his knowledge of risk factors, protective factors, and classification accuracy.

Dr. Tist fulfills two primary roles in his new clinical position. First, he assesses sex offenders for potential SVP commitment. Second, he assesses civilly committed SVP residents for potential release. After performing a number of SVP assessments, Dr. Tist begins to suspect that the men he examines represent an atypical class of sex offenders. He hypothesizes that the sexual recidivism risk posed by these sex offenders, as a group, is probably greater than the risk posed by the general sex offender population. Knowing that base rates vary across settings and populations, he also hypothesizes that the SVP respondents whom he examines for potential commitment probably recidivate at a lower rate than the SVP residents he examines for potential release from commitment.

Dr. Tist carefully reviews the literature. He finds limited data to support or refute his hypotheses. He realizes that efforts to collect local base rate data will be limited because

*Correspondence to: Daniel J. Neller, P.O. Box 1194, Wynne, AR 72396, U.S.A. E-mail: danieljneller@gmail.com
†Independent Practice, P.O. Box 1194, Wynne, AR 72396, U.S.A
‡University of Arkansas, Fayetteville AR
The opinions expressed in this manuscript do not necessarily reflect the views of any institutions, agencies, or departments with which the authors are affiliated.
[1] Although “risk” can be disaggregated into multiple elements (e.g., severity, imminence, etc.), in this paper we address exclusively probability of sexual recidivism. Similarly, we do not address other criteria that are potentially relevant to SVP commitment (e.g., mental abnormality, volitional control).

he practices in a jurisdiction where very few SVP residents are released from civil commitment (Jackson, Schneider, & Travia, 2011). Yet, as a scientist-practitioner, he recognizes that this issue – i.e., the base rate issue – represents perhaps "the most fundamental question about sex offender recidivism" (Furby, Weinrott, & Blackshaw, 1989, p. 22). Dr. Tist is astonished that an issue so fundamental to accurate assessment (Elwood, 1993a, 1993b) exists as little more than controversial discourse in the field of SVP assessment (e.g., Doren & Epperson, 2001; Janus & Meehl, 1997; for an exception, see Boccaccini, Murrie, Caperton, & Hawes, 2009). After considerable cogitation, he realizes that he can arrive at reasonable estimates of SVP recidivism by carefully considering three pieces of information.

First, Dr. Tist knows he can identify from extant literature a plausible rate of sexual recidivism for the general population of sex offenders; he refers to this as *a priori* probability (Mises, 1981). Second, Dr. Tist knows he can estimate classification accuracy of generic procedures by which many sex offenders are committed pursuant to SVP statutes, a process that typically consists of multiple stages (Lieb, 2003). Third, Dr. Tist knows he can estimate detection properties (i.e., sensitivity, specificity) of modestly accurate actuarial risk assessment instruments (ARAIs; Gottfredson & Moriarty, 2006; Hanson & Morton-Bourgon, 2009; Mossman, 2008) that often are used in SVP proceedings (Jackson & Hess, 2007). From these three sources of information, Dr. Tist knows he can generate algorithms (Bayes, 1763) that he believes will show plausible sexual recidivism rates at various stages of the SVP commitment and release process.

SVP RECIDIVISM RATES

Dr. Tist understands that base rates of sex offender recidivism vary widely across studies (Helmus, 2009). He attributes this variation to multiple factors (e.g., Blumstein & Larson, 1971), such as operational definitions of recidivism (e.g., re-arrest, re-conviction, re-incarceration), types of offenders sampled (e.g., ratio of child molesters to rapists), and jurisdictional reporting and apprehension rates. Nevertheless, two widely cited meta-analyses with which he is familiar (Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2005) show similar recidivism estimates for the general sex offender population. From these studies he concludes that approximately 13–14% of sex offenders sexually recidivate during 4- to 6-year follow-up periods. He also reviews literature indicating that longer follow-up periods (e.g., 20–25 years) typically yield greater recidivism rates, approximating 30–40% (Hanson, Morton, & Harris, 2003; for higher estimates, see Doren, 1998; Harris & Rice, 2007; Langevin et al., 2004). Dr. Tist presumes, based on these findings, that the general sex offender population recidivates at a rate of approximately 10–40% (Barbaree, 1997).

Some of these findings do not comport with Dr. Tist's clinical experience in an SVP program, however. Rather, he suspects that SVP respondents and residents probably present a greater risk of recidivism than sex offenders who are neither referred nor committed as such (Fabian, 2005). Dr. Tist's suspicion is corroborated by Levenson (2004). She found that sex offenders referred for SVP commitment scored higher on risk assessment instruments and showed a greater number of risk factors than sex offenders not referred for commitment. Dr. Tist finds additional support from research conducted by Milloy (2003), who reported recidivism rates of sex offenders recommended for commitment but for whom no petition was filed: They were

DOI: 10.1002/bsl

reconvicted of new felony sex offenses at a 6-year rate of approximately 30%, more than double the rates obtained in Hanson's meta-analyses (Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2005).

Finally, Dr. Tist finds support from research conducted by Knight and his colleagues (Knight & Thornton, 2007; Prentky, Lee, Knight, & Cerce, 1997), who studied sex offenders recommended for commitment pursuant to Massachusetts' Sexually Dangerous Person (SDP) statute. In that study, approximately 3,600 prisoners were recommended for initial consideration for commitment. Following a screening process, approximately 1,450 of the prisoners were transferred to a treatment center for full assessment. Approximately 350 of those 1,450 prisoners (i.e., 10% of the referred group, 25% of the evaluated subgroup) were committed to the SDP treatment program.

Knight and his colleagues (Knight & Thornton, 2007; Prentky et al., 1997) reported long-term recidivism rates of former SDP residents. The treated and released rapists considered *no longer sexually dangerous* recidivated at a rate of 39%. The treated and released child molesters considered *no longer sexually dangerous* recidivated at a rate of 52%. By contrast, the sex offenders who were evaluated but not committed to the SDP treatment program recidivated a rate of 13%.

Based on his review of the literature, Dr. Tist concludes that the studies by Levenson, Milloy, and Knight et al. represent the best available data pertaining to the recidivism risk posed by SVP respondents and residents. Dr. Tist believes that these data, when compared to the recidivism rates of general sex offenders, support his hypotheses that SVP respondents and residents probably recidivate at rates greater than sex offenders who are not referred for assessment pursuant to SVP commitment statutes.

STEP 1: A PLAUSIBLE *A PRIORI* PROBABILITY

After finding sufficient support for his belief that SVP respondents and residents probably recidivate at a high rate relative to general sex offenders, Dr. Tist has presented himself with an intriguing question: What are their most likely recidivism rates? Dr. Tist's first step toward answering this question involves the derivation of a plausible long-term recidivism rate for the general sex offender population. This base rate, or *a priori* probability, represents the rate of recidivism before Dr. Tist makes any attempts at classification.

Dr. Tist initially considers data from Hanson's meta-analyses (Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2005) to arrive at a reasonable *a priori* probability estimate. Because the follow-up periods of these meta-analyses are probably too short for application to SVP proceedings, and because recidivism rates are generally expected to rise with length of follow-up period (Hanson et al., 2003), Dr. Tist concludes that these rates are probably too low for his purposes. By contrast, he knows that findings of substantially greater long-term recidivism rates, such as the 60% rate reported in Langevin et al.'s (2004) 25-year follow-up, likely will be met with skepticism by many of his colleagues. Acknowledging that consensus may be impossible (Conroy, 2003), he eventually settles on 20% as his long-term recidivism rate. This *a priori* probability is considered a reasonable estimate for a 10-year follow-up period (Hanson et al., 2003), and Dr. Tist considers 10 years as reasonably long-term. This *a priori* probability also is commensurate with Knight and Thornton's (2007) "best compromise."

DOI: 10.1002/bsl

STEP 2: SUCCESSIVE HURDLES

To discern reasonably accurate SVP recidivism rates, as a second step Dr. Tist carefully examines the screening process by which sex offenders are referred to him. Dr. Tist knows that such processes, when effective, screen out lower-risk offenders and screen in higher-risk offenders, thereby increasing the base rate of the condition of interest in the population being studied – in this case, sexual recidivism rates of SVP respondents and residents.

In Dr. Tist's jurisdiction, the process of SVP commitment consists of three stages (cf. Lieb, 2003). At the first stage, a structured risk assessment instrument is used to screen in higher-risk offenders and screen out lower-risk offenders. Dr. Tist's jurisdiction uses as a screening tool an ARAI, the Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR; Hanson, 1997). At the second stage, a screening committee identifies sex offenders to refer to a prosecutor's office that routinely pursues commitment; this is a very small percentage of all sex offenders released from prison (see, e.g., Elwood, Doren, & Thornton, 2010; Janus & Walbek, 2000; McReynolds & Sandler, 2012). At the third stage, Dr. Tist himself examines all respondents whose commitment is pursued by the prosecutor's office. As do most SVP experts (Jackson & Hess, 2007), Dr. Tist assesses recidivism risk with the Static-99 (Hanson & Thornton, 1999, 2000), the most widely used instrument designed for assessing sexual recidivism risk (Archer, Buffington-Vollum, Stredny, & Handel, 2006). He refers to this three-stage process, in which presumptively lower-risk offenders are screened out and presumptively higher-risk offenders are screened in, as a "successive hurdles" approach (Meehl & Rosen, 1955) to identifying high-risk sex offenders.

STEP 3: CLASSIFICATION ACCURACY ESTIMATES

Dr. Tist considers the classification accuracy of the various procedures employed within his jurisdiction (i.e., ARAIs, screening committee judgment). For his purposes, Dr. Tist assumes the classification accuracy of these procedures remain generally stable at each hurdle and across base rates (Frederick & Bowden, 2009a, 2009b; Glaros & Kline, 1988; McFall, 2005; cf. Mossman, 2006; Singh, 2013). He is aware that a procedure's sensitivity provides the percentage of recidivists who are correctly classified as recidivists; that is, sensitivity provides the true positive rate. Dr. Tist also knows that a procedure's specificity provides the percentage of non-recidivists who are correctly classified as non-recidivists; that is, specificity provides the true negative rate.

Dr. Tist knows that the sensitivity and specificity estimates of contemporary ARAIs are contingent on the cut-off scores he selects (e.g., Mossman, 2013). He examines the average RRASOR and Static-99 scores of all sex offenders who are recommended for commitment in his jurisdiction. He finds an average RRASOR score of 4 and an average Static-99 score of 6. Other jurisdictions report similar findings (e.g., Elwood et al., 2010; Levenson, 2004).

Based on his examination of the actuarial tables for these instruments (Hanson, 1997; Hanson & Thornton, 1999, 2000; Neller & Frederick, 2013), Dr. Tist estimates the sensitivity and specificity of a RRASOR score of 4 as, respectively, 0.21 and 0.96.

DOI: 10.1002/bsl

This means that, if he uses this cut-off score, he can expect 21% of recidivists to be accurately detected as recidivists and 4% of non-recidivists to be falsely classified as recidivists. Based on the original data published on the Static-99, the sensitivity and specificity estimates associated with a Static-99 score of 6 are 0.25 and 0.92, , respectively. This means that, if he uses this cut-off score, he can expect 25% of recidivists to be accurately detected as recidivists and 8% of non-recidivists to be falsely classified as recidivists. More recently published data offer sensitivity and specificity estimates of a Static-99 score of 6 as, respectively, .43 and .84 (retrieved from www.static99.org on 17 June 2012).[2] This means that, if he uses this cut-off score, he can expect 43% of recidivists to be accurately detected as recidivists and 16% of non-recidivists to be falsely identified as recidivists. Dr. Tist chooses the more recent data for his analyses because he believes they are more likely to apply to a contemporary sex offender population (Boccaccini et al., 2009).

To estimate the accuracy of the screening team, Dr. Tist uses published estimates of clinical judgment reported by Wollert (2006) – sensitivity of 0.42 and specificity of 0.64. That is, the screening team is expected to correctly identify 42% of recidivists as recidivists, and to incorrectly classify 36% of non-recidivists as recidivists.

To summarize, in his effort to discern reasonable estimates of SVP recidivism rates, Dr. Tist assumes a long-term *a priori* probability of detected sexual recidivism of 20%. He considers the possible effects of a commitment process comprising at least three hurdles: (i) the scoring of a brief ARAI; (ii) the clinical judgment of a committee; and (iii) an assessment by a forensic examiner who relies upon a second, presumably more comprehensive, ARAI. Dr. Tist considers estimates of classification accuracy based on relevant empirical data. He applies these estimates to each stage of the SVP commitment process.

TOWARD PROPOSED BASE RATES

Dr. Tist is prepared to estimate recidivism base rates from the information he gathered. He does this by computing Bayesian (Bayes, 1763) posterior probabilities. Like many people (Gigerenzer, 2002; Gigerenzer & Hoffrage, 1995), Dr. Tist finds computations based on algebraic formulas somewhat difficult to follow. By contrast, he finds graphically depicted computations relatively simple to understand, so he uses them for his analyses.

As illustrated in Figure 1, Dr. Tist selects a large but otherwise arbitrary number to represent the relevant population of sex offenders ($N = 10{,}000$). He uses 20% as his *a priori* probability of sexual recidivism; therefore, he expects 2,000/10,000 to be recidivists and 8,000/10,000 to be non-recidivists before any formal procedures are applied. Based on sensitivity and specificity estimates of a score of 4 on the RRASOR, he expects to classify 21% of the 2,000 recidivists as recidivists and 4% of the 8,000 non-recidivists as recidivists. Thus, based on RRASOR scores, he expects 420/2,000 recidivists and 320/8,000 non-recidivists to be screened in. The remaining sex offenders are screened out and are not considered for commitment. After the first hurdle, therefore, the probability that a screened-in sex offender will recidivate is 57% [i.e., $420/(420 + 320) = 0.57$].

[2] The sensitivity estimate is 0.435; we rounded down to offer a conservative analysis. We considered data from the entire sample (rather than the “high-risk, high-needs” sample) because of its large size.

DOI: 10.1002/bsl

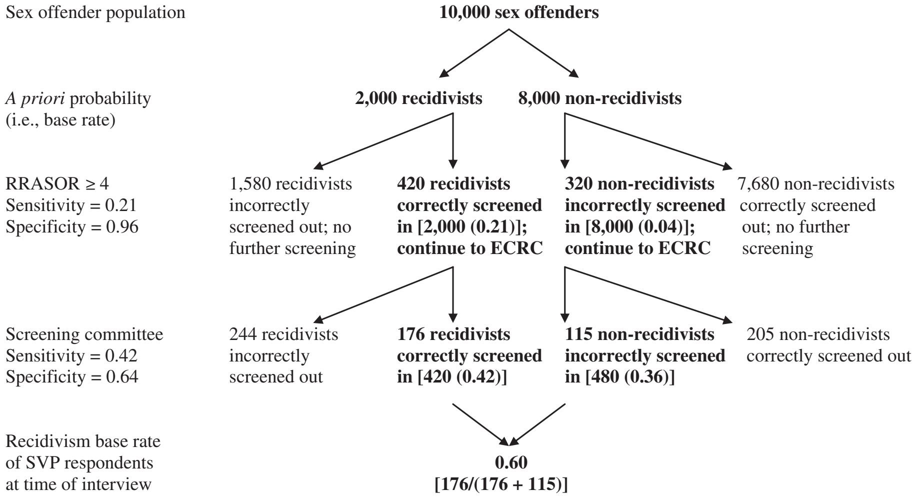

Figure 1. Estimated probability of sexual recidivism of sexually violent predator (SVP) respondents based on screening process. ECRC, End of Confinement Review Committee; RRASOR, Rapid Risk Assessment for Sexual Offense Recidivism.

At the second hurdle, the screening committee is expected to accurately detect 42% of the remaining 420 recidivists as recidivists (i.e., 176) and to mistakenly classify 36% of the 320 non-recidivists as recidivists (i.e., 115). The remaining sex offenders are screened out. Therefore, when a sex offender is referred to Dr. Tist for potential SVP commitment, and before Dr. Tist conducts any examination, he is prepared to assume the respondent's likelihood of sexual recidivism approximates 60% [i.e., 176/ (176 + 115) = 0.60].

So far, the sex offender has crossed the first two hurdles (i.e., RRASOR and screening committee). Assuming the prosecutor chooses to pursue commitment, the SVP respondent next is examined by Dr. Tist, the third hurdle. Using a Static-99 cut-off score of 6, Dr. Tist expects to accurately identify 43% of the remaining 176 recidivists (i.e., 76) as recidivists; he expects to inaccurately classify 16% of the remaining 115 non-recidivists (i.e., 18) as recidivists. Dr. Tist notes that the probability of recidivism among sex offenders recommended for SVP commitment appears to be 81% [i.e., 76/ (76 + 18) = 0.81]; however, Dr. Tist recognizes the possibility that this probability estimate could be inflated because scores from two of the assessment methods – the RRASOR and Static-99 – correlate highly with each other, roughly at $r = 0.76$ (Barbaree, Seto, Langton, & Peacock, 2001; Langton et al., 2007).[3]

To account for the correlation between RRASOR and Static-99 scores, Dr. Tist simultaneously considers all data (i.e., *a priori* probability, successive hurdles, classification accuracy, correlations between scores) in a Monte Carlo simulation. This is a computer-generated simulation of all RRASOR and Static-99 scores for "synthetic" populations of recidivists and non-recidivists. It is produced by repeated random

[3] Dr. Tist does not expect the screening committee's judgment to artificially inflate the risk estimate because research shows that clinical judgments of future violence do not significantly correlate with actuarial risk scores (e.g., Hilton & Simmons, 2001).

DOI: 10.1002/bsl

sampling of RRASOR and Static-99 scores from their joint distribution in the populations of recidivists and non-recidivists, and it enables him to build an entirely new probability distribution based on these data. Dr. Tist uses large synthetic populations in his simulation – 200,000 recidivists and 800,000 non-recidivists (i.e., reflecting the *a priori* probability of 20%) – so he can safely assume that the variability, inherent in any simulation procedure, remains negligible for practical purposes. Additional details regarding the simulation procedure are reported in the Appendix.

Table 1 reports in percentages the cross-tabulation of the scores on the two ARAIs for non-recidivists and recidivists. The strong correlation among the scores is readily apparent for both populations. From these cell counts, Dr. Tist computes the true- and false-positive rates corresponding to each cut-off score for the two ARAIs (see Table 2). From this, he discerns that less than 4% of non-recidivists will be incorrectly identified as recidivists if they score 4 or higher on the RRASOR *and* 6 or higher on the Static-99. Next, Dr. Tist computes the positive predictive value (PPV) for each combination of RRASOR and Static-99 scores; this enables him to estimate probabilities of recidivism based on various combinations of scores on the two measures without the influence of the screening committee. The full set

Table 1. Cross-tabulation of non-recidivist and recidivist scores on the RRASOR and Static-99 (in %)

	RRASOR						
	0	1	2	3	4	5	Total
Non-recidivists							
Static-99							
0	10.42	2.36	0.14	0.00	0.00	0.00	12.93
1	6.63	6.14	0.98	0.03	0.00	0.00	13.77
2	4.00	8.48	2.89	0.16	0.01	0.00	15.53
3	1.79	7.67	4.95	0.51	0.02	0.00	14.94
4	0.74	6.20	7.91	1.64	0.14	0.00	16.62
5	0.12	2.18	5.38	2.05	0.32	0.01	10.06
6	0.02	0.83	3.74	2.53	0.63	0.05	7.79
7	0.00	0.18	1.66	2.01	0.87	0.11	4.84
8	0.00	0.02	0.43	1.00	0.79	0.20	2.45
9	0.00	0.00	0.05	0.18	0.24	0.12	0.60
10	0.00	0.00	0.01	0.08	0.18	0.19	0.47
Total	23.73	34.06	28.13	10.18	3.20	0.69	100.00
Recidivists							
Static-99							
0	1.25	0.39	0.05	0.00	0.00	0.00	1.69
1	1.80	1.87	0.45	0.02	0.00	0.00	4.14
2	1.63	3.90	1.88	0.18	0.01	0.00	7.59
3	1.01	4.65	4.29	0.80	0.08	0.00	10.84
4	0.52	4.77	8.74	3.16	0.58	0.04	17.81
5	0.10	2.01	6.92	4.46	1.41	0.15	15.05
6	0.03	0.84	5.35	5.97	2.99	0.49	15.67
7	0.00	0.21	2.52	4.94	4.11	1.32	13.10
8	0.00	0.03	0.61	2.33	3.56	2.25	8.78
9	0.00	0.00	0.06	0.38	0.95	1.27	2.65
10	0.00	0.00	0.01	0.14	0.59	1.94	2.68
Total	6.35	18.67	30.87	22.39	14.27	7.46	100.00

RRASOR, Rapid Risk Assessment for Sexual Offense Recidivism.

DOI: 10.1002/bsl

Table 2. False-positive rates for non-recidivists and true-positive rates for recidivists

	RRASOR					
	0	1	2	3	4	5
(a) False positive rates for non-recidivists						
Static-99						
0	100.00	76.27	42.21	14.08	3.90	0.69
1	87.07	73.76	42.06	14.08	3.90	0.69
2	73.30	66.62	41.06	14.05	3.89	0.69
3	57.76	55.09	38.01	13.89	3.89	0.69
4	42.83	41.94	32.53	13.35	3.86	0.69
5	26.20	26.05	22.84	11.57	3.73	0.69
6	16.14	16.12	15.08	9.20	3.40	0.68
7	8.36	8.36	8.15	6.00	2.72	0.63
8	3.52	3.52	3.49	3.00	1.74	0.52
9	1.06	1.06	1.06	1.01	0.75	0.32
10	0.47	0.47	0.47	0.46	0.38	0.19
(b) True-positive rates for recidivists						
Static-99						
0	100.00	93.65	74.99	44.12	21.73	7.46
1	98.31	93.22	74.94	44.12	21.73	7.46
2	94.17	90.87	74.46	44.09	21.73	7.46
3	86.57	84.91	72.40	43.90	21.72	7.46
4	75.73	75.08	67.22	43.02	21.64	7.46
5	57.92	57.79	54.70	39.24	21.02	7.42
6	42.87	42.84	41.77	33.22	19.46	7.27
7	27.21	27.20	26.97	23.77	15.98	6.78
8	14.11	14.11	14.08	13.40	10.55	5.46
9	5.33	5.33	5.33	5.26	4.75	3.21
10	2.68	2.68	2.68	2.67	2.53	1.94

RRASOR, Rapid Risk Assessment for Sexual Offense Recidivism.

of these conditional probabilities is provided in Table 3.[4] Finally, Dr. Tist uses a graphical depiction to estimate the probability that sex offenders will recidivate if they: (a) are drawn from a population of sex offenders with a 10-year sexual recidivism rate of 20%; (b) score 4 or higher on the RRASOR; (c) are screened in by a committee whose judgment is unrelated to the RRASOR score; and (d) score 6 or higher on the Static-99. As illustrated in Figure 2, Dr. Tist expects such sex offenders to recidivate at a rate of 65% during a 10-year follow-up period.

IMPLICATIONS

The results of these analyses admittedly are based on informed speculation. And, for multiple reasons, the recidivism estimates are less precise than they might initially appear (e.g., Cooke & Michie, 2010, 2011; Hart, Michie, & Cooke, 2007; but see responses by Hanson & Howard, 2010; Harris, Rice, & Quinsey, 2008; Mossman &

[4] Some of these values are not only improbable but also, based on the coding rules of each instrument, impossible to obtain if the instruments are scored correctly. For instance, because the instruments share items that are scored in the same manner, a sex offender could not accurately obtain a 10 on the Static-99 and a 0 on the RRASOR. The proportion of these cases is so low that they are not expected to meaningfully influence the results of the simulation.

DOI: 10.1002/bsl

Table 3. Probability of recidivism based on combinations of RRASOR and Static-99 scores

	RRASOR					
	0	1	2	3	4	5
Static-99						
0	0.20	0.23	0.31	0.44	0.58	0.73
1	0.22	0.24	0.31	0.44	0.58	0.73
2	0.24	0.25	0.31	0.44	0.58	0.73
3	0.27	0.28	0.32	0.44	0.58	0.73
4	0.31	0.31	0.34	0.45	0.58	0.73
5	0.36	0.36	0.37	0.46	0.59	0.73
6	0.40	0.40	0.41	0.47	0.59	0.73
7	0.45	0.45	0.45	0.50	0.59	0.73
8	0.50	0.50	0.50	0.53	0.60	0.72
9	0.56	0.56	0.56	0.57	0.61	0.72
10	0.59	0.59	0.59	0.59	0.63	0.72

RRASOR, Rapid Risk Assessment for Sexual Offense Recidivism.

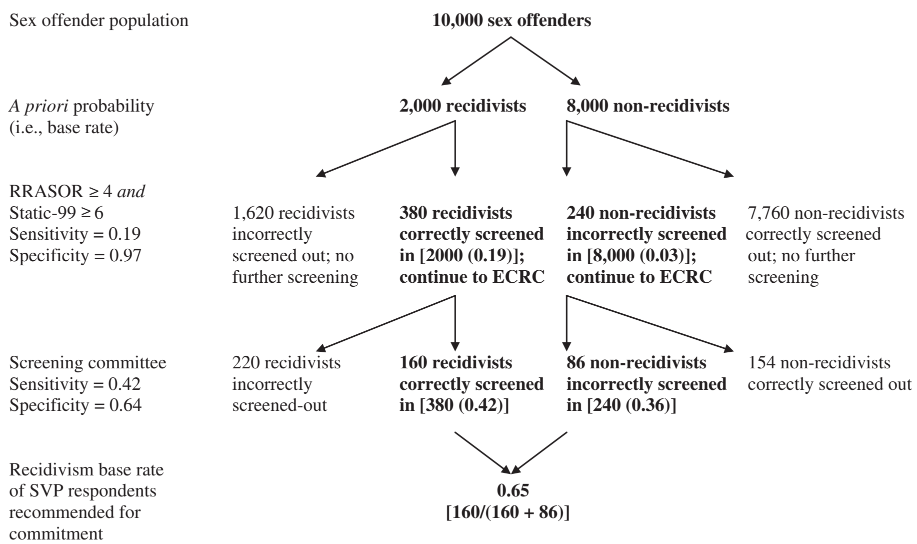

Figure 2. Estimated probability of sexual recidivism of sexually violent predator (SVP) respondents based on entire commitment process. ECRC, End of Confinement Review Committee; RRASOR, Rapid Risk Assessment for Sexual Offense Recidivism.

Sellke, 2007). As stated by Faust (2003, p. 438), however, "Even if the base rate information we seek is limited or far from ideal, we still may do better by using this imperfect information as opposed to disregarding it."

These findings have implications for practitioners who conduct SVP assessments in jurisdictions that employ screening processes similar to the one described here. First, the results show that forensic examiners and fact-finders can assume that SVP respondents probably recidivate at a rate that is higher than that of typical sex offenders – perhaps 60% within 10 years. Although some (e.g., Vrieze & Grove, 2008; but see Mossman, 2008) contend that modestly accurate instruments are too inaccurate to use in certain contexts, it is noted that moderately accurate instruments can be quite helpful when base rates approach 50% (e.g., Meehl & Rosen, 1955). Therefore, when

DOI: 10.1002/bsl

successive hurdles are incorporated into an SVP process, it appears that ARAIs can prove quite helpful at identifying high-risk sex offenders.

Second, the results of these analyses suggest that practitioners are able to identify a small group of sex offenders whose likelihood of recidivism substantially exceeds the risk posed by general sex offenders. Indeed, the present findings show that roughly two out of three sex offenders recommended for civil commitment – a figure more than threefold that of general sex offenders – can be expected to sexually recidivate within 10 years if certain assumptions are met. One such assumption is a high score on more than one ARAI, an infrequent occurrence among general sex offenders (Barbaree, Langton, & Peacock, 2006). Accordingly, practitioners might be appropriately reluctant to recommend release of SVP residents who were recommended for commitment because: (i) the residents scored relatively high on actuarial measures, which are superior to unstructured judgments (Ægisdotter et al., 2006; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Grove & Meehl, 1996; Hanson & Morton-Bourgon, 2009; Hilton, Harris, & Rice, 2006; Sawyer, 1966; Swets, Dawes, & Monahan, 2000; but see Litwack, 2001); (ii) predictions that go against base rates likely will be wrong (Arkes, 1981; Borum, Otto, & Golding, 1993; Meehl & Rosen, 1955; Vrieze & Grove, 2008; Wollert, 2006); (iii) pre-treatment scores appear to be more accurate at predicting recidivism than post-treatment scores (Barnett, Wakeling, Mandeville-Norden, & Rakestrow, 2012); and (iv) high-quality research fails to clearly support the effectiveness of sex offender-specific treatment (Hanson, Bourgon, Helmus, & Hodgson, 2009; Hanson, Broom, & Stephenson, 2004; Harris & Rice, 2003; Marques, Wiederanders, Day, Nelson, & van Ommeren, 2005; cf. Hanson et al., 2002).

The present results further show that, in the context of assessing SVPs, the PPV of a high score on a simple, four-item ARAI is only modestly increased by a high score on a second, presumably more comprehensive, ARAI (also see Seto, 2005; cf. Babchishin, Hanson, & Helmus, 2011).[5] In examining Table 3, we see that the PPV ranges for RRASOR scores of 4 and 5 vary by no more than five percentage points when the Static-99 score is taken into consideration; it is difficult to conceive of a fact-finder who, all else being equal, would be persuaded to civilly commit a sex offender whose likelihood of recidivism were 63% but would release him if his likelihood of recidivism were 58%.

By contrast, a high score on the Static-99 might contribute to decision-making when the RRASOR score is more moderate. For instance, the PPV associated with a score of 3 on the RRASOR is 0.44, and the PPV associated with a score of 8 on the Static-99 is 0.42. Considered independently, these probability estimates might be insufficient for

[5] The successive hurdles approach and predictive value estimates are distinguished from the three decision rules and accuracy indices examined by Seto (2005). It is worth highlighting that, all other things being equal, the hurdles will probably contribute greatest to accuracy when the instruments' scores are uncorrelated with one another. Of the instruments included in their study, Langton et al. (2007) found that scores on the RRASOR and Minnesota Sex Offender Screening Tool – Revised (MnSOST-R; Epperson, Kaul, & Hesselton, 1999) correlate least with each other ($r = 0.34$), followed closely by scores on the RRASOR and Sex Offender Risk Appraisal Guide (SORAG; Quinsey, Harris, Rice, & Cormier, 1998, 2006), which correlated $r = 0.40$. However, the benefit of employing the MnSOST-R or SORAG might be lost on their weaknesses with respect to SVP assessments. First, fewer studies have been conducted on the MnSOST-R and SORAG than on the Static-99 and RRASOR (Hanson & Morton-Bourgon, 2009). Second, scores on the MnSOST-R are less reliable than scores on other available instruments (Barbaree et al., 2001), a property that will limit validity. Third, the follow-up period for the MnSOST-R developmental sample was 6 years, which may be too short for SVP commitment. Fourth, instead of sexual-specific recidivism, the SORAG uses as its criterion all forms of violent recidivism; although all violent acts may be concerning, most jurisdictions with SVP statutes identify sexual-specific recidivism as the outcome of interest.

DOI: 10.1002/bsl

recommending commitment in jurisdictions that utilize a "more likely than not" standard. When these scores are considered jointly, however, the PPV rises to 0.53; this new estimate might cross a legally established threshold (i.e., 50%) in some jurisdictions, making otherwise ineligible offenders eligible for commitment. Along the same lines, SVP residents who were committed after scoring high on one but not both of the instruments might be considered for release in these same jurisdictions, as their joint probability of recidivism might fall below a legally established threshold. In this way, the conditional probabilities reported in Table 3 might prove helpful to examiners and fact-finders by offering conceptually sound estimates of sexual recidivism risk based on various combinations of RRASOR and Static-99 scores. We are familiar with no other conceptually sound, published risk estimates derived from the combination of these two ARAIs.

In addition to implications for practice, the results of these analyses have implications for policy. Perhaps most of all, they show the level of effort necessary to identify a very small group of sex offenders whose risk for recidivism is substantially greater than that of the general population of sex offenders. This is particularly striking in light of the number of recidivists who escape detection. Consequently, jurisdictions with SVP statutes must be willing to expend considerable resources to identify an extremely high-risk group of sex offenders. Even so, they must expect to miss many more sexual recidivists than they can expect to commit. These findings indicate that jurisdictions willing to expend a high level of effort, while missing many recidivists in the process, can expect to identify a small group of high-risk sex offenders if they employ actuarially based, successive hurdles approaches to SVP commitment.

ACKNOWLEDGEMENTS

We are grateful to Richard Frederick for his comments on an earlier version of this manuscript, and to Douglas Mossman for his contributions.

REFERENCES

Ægisdotter, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., & Rush J. D. (2006). The meta-analysis of clinical judgment project: Fifty six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, *34*, 341–382.

Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, *87*(1), 84–94.

Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology*, *49*(3), 323–330.

Babchishin, K. M., Hanson, R. K., & Helmus, L. (2011). The RRASOR, Static-99R and Static-2002R all add incrementally to the prediction of recidivism among sex offenders. (User Report No. 11-02). Ottawa: Public Safety Canada.

Barbaree, H. E. (1997). Evaluating treatment efficacy with sexual offenders: The insensitivity of recidivism studies to treatment effects. *Sexual Abuse: A Journal of Research and Treatment*, *9*, 111–128.

Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006). Different actuarial risk measures produce different risk rankings for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, *18*, 423–440.

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

Barnett, G. D., Wakeling, H. C., Mandeville-Norden, R., & Rakestrow, J. (2012). How useful are psychometric scores in predicting recidivism for treated sex offenders? *International Journal of Offender Therapy and Comparative Criminology*, *56*(3), 420–446.

DOI: 10.1002/bsl

Bayes, T. (1763). An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418.
Blumstein, A., & Larson, R. C. (1971). Problems in modeling and measuring recidivism. *Journal of Research in Crime and Delinquency*, *8*(2), 124–132.
Boccaccini, M., Murrie, D., Caperton, J., & Hawes, S. (2009). Field validity of the STATIC-99 and MnSOST-R among sex offenders evaluated for civil commitment as Sexually Violent Predators. *Psychology, Public Policy, and Law*, *15*(4), 278–314.
Borum, R., Otto, R., & Golding, S. (1993). Improving clinical judgment and decision making in forensic evaluation. *The Journal of Psychiatry and Law*, *2*, 35–76.
Conroy, M. A. (2003). Evaluation of sexual predators. In: I. B. Weiner, & A. M. Goldstein (Ed.), Handbook of Psychology: Vol. 11. Forensic Psychology Series ed. (pp. 463–484). Hoboken, NJ: John Wiley & Sons, Inc.
Cooke, D., & Michie, C. (2010). Limitations of diagnostic precision and predictive utility in the individual case: A challenge for forensic practice. *Law and Human Behavior*, *34*, 259–274.
Cooke, D., & Michie, C. (2011). Violence risk assessment: Challenging the illusion of certainty. In: B. McSherry, & P. Keyzer (Eds.), Dangerous People: Policy, Prediction, and Practice (pp. 147–161). New York, NY: Routledge.
Doren, D. M. (1998). Recidivism base rates, predictions of sex offender recidivism, and the "Sexual Predator" commitment laws. *Behavioral Sciences & the Law*, *16*, 97–114.
Doren, D. M., & Epperson, D. L. (2001). Great analysis, but problematic assumptions: A critique of Janus and Meehl (1997). *Sexual Abuse: A Journal of Research and Treatment*, *13*, 45–51.
Elwood, R. W. (1993a). Clinical determinations and neuropsychological tests: An appeal to Bayes' Theorem. *The Clinical Neuropsychologist*, *7*(2), 224–233.
Elwood, R. W. (1993b). Psychological tests and clinical determinations: Beginning to address the base rate problem. *Clinical Psychology Review*, *13*, 409–419.
Elwood, R. W., Doren, D. M., & Thornton, D. (2010). Diagnostic and risk profiles of men detained under Wisconsin's Sexually Violent Person law. *International Journal of Offender Therapy and Comparative Criminology*, *54*(2), 187–196.
Epperson, D., Kaul, J., & Hesselton, D. (1999). Minnesota Sex Offender Screening Tool – Revised (MNSOST-R): Development, performance, and recommended risk level cut scores. Unpublished manuscript.
Fabian, J. M. (2005). The risky business of conducting risk assessments for those already civilly committed as Sexually Violent Predators. *William Mitchell Law Review*, *32*(1), 81–159.
Faust, D. (2003). Holistic thinking is not the whole story: Alternative or adjunct approaches for increasing the accuracy of legal evaluations. *Assessment*, *10*(4), 428–441.
Frederick, R. I., & Bowden, S. (2009a). The test validation summary. *Assessment*, *16*, 215–236.
Frederick, R. I., & Bowden, S. (2009b). Using the test validation summary to explore constructs in symptom validity tests. *The Journal of Head Trauma Rehabilitation*, *24*, 105–122.
Furby, L., Weinrott, M., & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin*, *105*(1), 3–30.
Gigerenzer, G. (2002). Calculated risks: How to know when numbers deceive you. New York: Simon & Schuster.
Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*(4), 684–704.
Glaros, A. G., & Kline, R. B. (1988). Understanding the accuracy of tests with cutting scores: The sensitivity, specificity, and predictive value model. *Journal of Clinical Psychology*, *44*(6), 1013–1023.
Gottfredson, S. D., & Moriarty, L. J. (2006). Statistical risk assessment: Old problems and new applications. *Crime & Delinquency*, *52*(1), 178–200.
Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law*, *2*, 293–323.
Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, *12*(1), 19–30.
Hanson, R. K. (1997). The development of a brief actuarial risk scale for sexual offense recidivism. (User Report 97-04). Ottawa: Department of the Solicitor General of Canada.
Hanson, R. K., & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66*, 348–362.
Hanson, R. K., & Howard, P. D. (2010). Individual confidence intervals to not inform decision-makers about the accuracy of risk assessment evaluations. *Law and Human Behavior*, *34*(4), 275–281.
Hanson, R. K., & Morton-Bourgon, K. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, *73*, 1154–1163.
Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, *21*(1), 1–21.
Hanson, R. K., & Thornton, D. (1999). Improving actuarial risk assessments for sex offenders. Unpublished manuscript.

DOI: 10.1002/bsl

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, *24*, 119–136.
Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). A meta-analysis of the effectiveness of treatment for sexual offenders: Risk, need, and responsivity. Ottawa, Ontario, Canada: Public Safety Canada.
Hanson, R. K., Broom, I., & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Science*, *36*(2), 87–96.
Hanson, R. K., Gordon, A., Harris, A. J. R., Marques, J. K., Murphy, W., Quinsey, V. L. & Seto, M. C. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, *14*, 169–194.
Hanson, R. K., Morton, K. E., & Harris, A. J. R. (2003). Sexual offender recidivism risk: What we know and what we need to know. In: R. Prentky, E. Janus, & M. Seto (Eds.), Vol. 989: Sexually Coercive Behavior: Understanding and Management (pp. 154–166). New York: Annals of the New York Academy of Sciences.
Harris, G. T., & Rice, M. E. (2003). The size and sign of treatment effects in sex offender therapy. In: R. Prentky, E. Janus, & M. Seto (Eds.), Vol. 989 – Sexually Coercive Behavior: Understanding and Management (pp. 428–440). New York: Annals of the New York Academy of Sciences.
Harris, G. T., & Rice, M. E. (2007). Characterizing the value of actuarial violence risk assessments. *Criminal Justice and Behavior*, *34*(12), 1638–1658.
Harris, G. T., Rice, M. E., & Quinsey, V. L. (2008). Shall evidence-based risk assessment be abandoned? *The British Journal of Psychiatry*, *192*, 154.
Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments: Evaluating the 'margins of error' of group v. individual predictions of violence. *The British Journal of Psychiatry*, *190*, s60–s65.
Helmus, L. (2009). Re-norming Static-99 recidivism estimates: Exploring base rate variability across sex offender samples. Unpublished master's thesis, Carleton University.
Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior*, *25*(4), 393–408.
Hilton, N. Z., Harris, G. T., & Rice, M. E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *The Counseling Psychologist*, *34*(3), 400–409.
Jackson, R. L., & Hess, D. T. (2007). Evaluation for civil commitment of sex offenders: A survey of experts. *Sexual Abuse: A Journal of Research and Treatment*, *19*, 425–448.
Jackson, R. L., Schneider, J. E., & Travia, T. (2011). SOCCPN annual survey of sex offender civil commitment programs 2011. Toronto, Canada: Presented at Association for the Treatment of Sexual Abusers Annual Convention.
Janus, E. S., & Meehl, P. E. (1997). Assessing the legal standard for predictions of dangerousness in sex offender commitment proceedings. *Psychology, Public Policy, and Law*, *3*(1), 33–64.
Janus, E. S., & Walbek, N. H. (2000). Sex offender commitments in Minnesota: A descriptive study of second generation commitments. *Behavioral Sciences & the Law*, *18*, 343–374.
Knight, R. A. & Thornton, D. (2007). Evaluating and improving risk assessment schemes for sexual recidivism: A long-term follow-up of convicted sexual offenders. (Document No. 21768). Unpublished manuscript.
Langevin, R., Curnoe, S., Federoff, P., Bennett, R., Langevin, M., Peever, C. & Sandhu, S. (2004). Lifetime sex offender recidivism: A 25-year follow-up study. *Canadian Journal of Criminology and Criminal Justice*, *46*(5), 531–552.
Langton, C. M., Barbaree, H. E., Seto, M. C., Peacock, E. J., Harkins, L., & Hansen, K. T. (2007). Actuarial assessment of risk for reoffense among adult sex offenders: Evaluating the predictive accuracy of the Static-2002 and five other instruments. *Criminal Justice and Behavior*, *34*(1), 37–59.
Levenson, J.S. (2004). Sexual predator civil commitment: A comparison of selected and released offenders. *International Journal of Offender Therapy and Comparative Criminology*, *48*, 638–648.
Lieb, R. (2003). State policy perspectives on sexual predator laws. In: B. Winick, & J. LaFond (Eds.), Protecting Society from Sexually Dangerous Offenders (pp. 41–59). Washington, DC: American Psychological Association.
Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy, and Law*, *7*(2), 409–433.
Marques, J. K., Wiederanders, M., Day, D. M., Nelson, C., & van Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment*, *17*(1), 79–107.
McFall, R. M. (2005). Theory and utility – Key themes in evidence-based assessment: Comment on the special section. *Psychological Assessment*, *17*(3), 312–323.
McReynolds, L. S., & Sandler, J. C. (2012). Evaluating New York State's Sex Offender Management and Treatment Act: A matched historical cohort analysis. *Criminal Justice Policy Review*, *23*(2), 164–185.

DOI: 10.1002/bsl

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*, 194–216.
Milloy, C. (2003). Six-year follow-up of released sex offenders recommended for commitment under Washington's Sexually Violent Predator law, where no petition was filed. (Document No. 03-12-1101). Olympia: Washington State Institute for Public Policy.
Mises, R. v. (1981). Probability, statistics and truth. Mineola, NY: Dover.
Mossman, D. (2006). Another look at interpreting risk categories. *Sexual Abuse: A Journal of Research and Treatment*, *18*(1), 41–63.
Mossman, D. (2008). Analyzing the performance of risk assessment instruments: A response to Vrieze and Grove (2007). *Law and Human Behavior*, *32*(3), 279–291.
Mossman, D. (2013). Evaluating risk assessments using receiver operating characteristic analysis: rationale, advantages, insights, and limitations. *Behavioral Sciences & the Law*. DOI: 10.1002/bsl.2050.
Mossman, D., & Sellke, T. (2007). Avoiding errors about 'margins of error. *The British Journal of Psychiatry*, *191*, 561.
Neller, D. J., & Frederick, R. I. (2013). Classification accuracy of actuarial risk assessment instruments. *Behavioral Sciences & the Law*. DOI: 10.1002/bsl.2047.
Prentky, R. A., Lee, A. F. S., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior*, *21*(6), 635–658.
Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). Violent offenders: Appraising and managing risk. Washington, DC: American Psychological Association.
Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). Violent offenders: Appraising and managing risk, second *ed*. Washington, DC: American Psychological Association.
R Development Core Team. (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
Sawyer, J. (1966). Measurement *and* prediction, clinical *and* statistical. *Psychological Bulletin*, *66*(3), 178–200.
Seto, M. C. (2005). Is more better? Combining actuarial risk scales to predict recidivism among adult sex offenders. *Psychological Assessment*, *17*(2), 156–167.
Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences & the Law*. DOI: 10.1002/bsl.2052.
Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*(1), 1–26.
Vrieze, S. I., & Grove, W. M. (2008). I. Correcting for item overselection and accuracy overestimation in scale development. II. Sampling error-induced attenuation of predictive validity over base rate information. *Law and Human Behavior*, *32*(3), 266–278.
Wollert, R. (2006). Low base rates limit expert certainty when current actuarials are used to identify sexually violent predators: An application of Bayes's Theorem. *Psychology, Public Policy, and Law*, *12*(1), 56–85.

APPENDIX

In this appendix we describe how we fitted a model to RRASOR and Static-99 data and used it to simulate synthetic scores for two populations, one of recidivists and one of non-recidivists.

The binormal model is a popular way to examine diagnostic tool scores of people who might or might not have a condition of interest – in our case, recidivists and non-recidivists. For a tool with n possible values, the model assumes that there exist $n - 1$ cut-points $\tau_1 < \tau_2 < \ldots < \tau_{n-1}$ and that each individual has an associated unobservable continuous "decision variable" X such that his or her score on the tool is defined to be $Y = k$ if $\tau_{k-1} < X < \tau_k$. The model also assumes that the distribution of X is standard normal for those individuals not having the condition, while it is normal with mean a and standard deviation b^{-1} for those individuals having the condition. The unknown model parameters are a, b, τ_1, τ_2, ..., τ_{n-1}.

From two samples, one consisting of individuals with the condition, the other of individuals without the condition, the unknown parameters can be estimated using the method of maximum likelihood. In our case, this can be done separately for RRASOR and Static-99, leading to two sets of estimated parameters, one for each diagnostic tool. Accordingly, two decision variables, X_R and X_S, can be identified for

DOI: 10.1002/bsl

any individual who is assessed on both instruments. If this individual is a non-recidivist, then both X_R and X_S have a standard normal distribution. So, in order to simulate the scores of a non-recidivist on RRASOR and Static-99, we simply simulate the two standard normal random variables X_R and X_S. However, if we want the scores on the RRASOR and Static-99 to be correlated, then X_R and X_S cannot be independent. The solution is to simulate independent standard normal random variables X_R and Z and set $X_S = \alpha X_R + (1 - \alpha^2)^{1/2} Z$. It is not hard to see that for any α, where $-1 < \alpha < 1$, X_S has also a standard normal distribution. Using the two sets of estimated cut-points, the correlation between the scores on the two instruments, Y_R and Y_S, can be numerically evaluated for any specific value of α. Let $r(\alpha)$ be this correlation. If we want the correlation between Y_R and Y_S to be a specific value ρ, 0.76 in our case, we can numerically solve for α the equation $r(\alpha) = \rho$. A similar approach was used to simulate correlated RRASOR and Static-99 scores for recidivists.

All the calculations and simulations were performed using the R software environment (R Development Core Team, 2012). The R code used in the analysis is available from the authors upon request.

DOI: 10.1002/bsl