

# Predictive Accuracy of Static-99R and Static 2002R

Author: Terence W. Campbell, Ph.D., ABPP (forensic)  
Sterling Heights, Michigan  
[tcampbell3920@COMCAST.NET](mailto:tcampbell3920@COMCAST.NET)

## Abstract

Relying on the comparison groups recommended by the instruments' developers, this article reports the levels of predictive accuracy obtained by the Static-99R and Static-2002R. It identifies the Positive Predictive Values, Negative Predictive Values, and frequencies of true-positive, false-positive, true-negative, and false-negative outcomes for all Static-99R and Static-2002R scores. These data demonstrate that ruling-out recidivism risk is consistently more accurate than ruling-in recidivism risk for these two instruments.

**Keywords:** Static-99R, Static-2002R, predictive accuracy, positive predictive value, negative predictive value.

---

## Introduction

This article's title likely needs clarification and possibly some defense. Whether or not evaluators involved in Sexually Violent Predator (SVP) matters engage in risk prediction has been disputed (Doren, 2006). Nevertheless, Campbell (2007) has clearly demonstrated how professional associations such as the Association for the Treatment of Sexual Abusers (ATSA) recognize prediction as a necessary endeavor for SVP evaluators. Additionally, individuals such as Barbaree, Seto, Langton, and Peacock (2001), Epperson, Kaul, and Huot, (1995), Hanson (2000), Hanson and Harris (2000), Hanson and Morton-Bourgon (2004, 2009), and Quinsey, Rice, and Harris (1995) also recognize that prediction is centrally important in SVP proceedings. Therefore, this article addresses the predictive accuracy of the Static-99R and Static-2002R.

Though various methods are available for assessing the recidivism risk of previously convicted sex offenders, actuarial instruments are the most commonly used (Campbell & DeClue, 2010a; Hanson & Morton-Bourgon, 2009). Evaluators often supplement actuarial instruments with approaches such as adjusted actuarial assessment, or guided clinical risk assessment (Campbell & DeClue, 2010b). This paper, however, focuses on "pure" actuarial assessment, reporting the outcomes obtained when using only the Static-99R and/or Static-2002R.

## Identifying Predictive Accuracy

Identifying the predictive accuracy of an actuarial instrument such as the Static-99R or Static-2002R is challenging. It can leave attorneys, jurors, and many psychologists glassy eyed, wishing for some kind of a compass as a result of feeling hopelessly lost. Terms such as *areas under the curve* (AUC values), *d values*, *logistic regression*, *confidence intervals*, and *confidence intervals for an individual* are not easily understood.

Diagnostic testing terms used in medicine (Glaros & Kline, 1988) are relatively straightforward and more readily comprehended as a result. As is the case for diagnostic testing, there are four possible outcomes when assessing whether a previously convicted sex offender will reoffend

**True Positive (TP):** The actuarial instrument predicts the offender will reoffend and he does reoffend.

**False Positive (FP):** The actuarial instrument predicts the offender will reoffend, but the offender does not reoffend.

**True Negative (TN):** The actuarial instrument predicts the offender will not reoffend, and he does not reoffend.

**False Negative (FN):** The actuarial instrument predicts the offender will not reoffend, but he does reoffend.

As actuarial instruments, the Static-99R and Static-2002R are linear-additive models (Bani-Yaghoub, Federoff, Curry, & Amundsen, 2010). Linear-additive models, such as the Static-99R and Static-2002R, assume that higher scores are associated with an increased risk of recidivism. It is further assumed that lower scores are associated with a decreased risk of recidivism. Any Static-99R score or Static-2002R score can be used to rule-in or rule-out recidivism risk. Assume an offender scores 6 on the Static-99R or the Static-2002R. Ruling-in recidivism risk necessitates considering all outcomes (frequencies of TP, FP, TN, and FN) for scores of 6 and above. Ruling-out recidivism risk necessitates considering all outcomes (frequencies of TP, FP, TN, and FN) for scores of 6 and below. Therefore, depending on whether an evaluator is ruling-in or ruling-out recidivism risk, evaluators adopt one of the following decision-making rules.

1. If ruling-in recidivism risk for a particular offender, an evaluator has implicitly adopted the following decision-making rule: "Rule-in recidivism risk for all scores of X and above (where X is an offender's Static-99R or Static-2002R score)."
2. If ruling-out recidivism risk for a particular offender, an evaluator has implicitly adopted the following decision-making rule: "Rule-out recidivism risk for all scores of X and below (where X is an offender's Static-99R or Static-2002R score)."

## Positive Predictive Values and Negative Predictive Values

Hart, Webster, and Menzies (1993) expressed the following recommendations related to identifying the accuracy of violence risk assessments:

We further recommend that, when the focus of analyses is predictive accuracy, the following standard statistics should be reported: (a) positive predictive power, the accuracy of predictions that individuals will be violent, (b) negative predictive power, the accuracy of predictions that individuals will not be violent . . . (p. 698).

The eighth "Commandment" of Serin and Brown's (2000) "Ten Commandments of Risk Assessment" advises: "Thou shalt know thy false positive and false negative rates for specific cut offs" (p. 263).

Craig and Beech (2010) also advised:

One method of calculating change-corrected predictive accuracy is to calculate positive predictive accuracy (PPA) and negative predictive accuracy (NPA). PPA and NPA are different ways of conceptualizing predictive accuracy where PPA refers to the accuracy of predicting individuals that are dangerous while NPA refers to the accuracy of predicting individuals that are not dangerous (p. 282).

Positive Predictive Value(s) (PPV) identify the accuracy with which one rules-in recidivism risk.

1. PPVs are obtained by dividing the number of true-positive predictions by the number of true-positive predictions plus the number of false-positive predictions ( $PPV = TP/TP + FP$ ).

Negative Predictive Value(s) (NPV) identify the accuracy with which one rules-out recidivism risk.

2. NPVs are obtained by dividing the number of true-negative predictions by the number of true-negative predictions plus the number of false-negative predictions ( $NPV = TN/TN + FN$ ).

## Base Rates, Sensitivity and Specificity, and Comparison Groups

PPVs and NPVs are base-rate sensitive. As the base rate of some event (e.g., sex offender recidivism) increases, PPVs also increase while NPVs decrease. Conversely, as the base rate of some event decreases, PPVs decrease and NPVs increase.

AUC values identify how accurately an actuarial instrument orders recidivists and non-recidivists independent of the base rate. AUC values respond to this question: If randomly selecting a recidivist and a non-recidivist from a sample of sex offenders, what is

the likelihood that the actuarial score of the recidivist exceeds that of the non-recidivist? Base rates, however, influence applied assessments. As a result, PPVs and NPVs serve as the most appropriate indices of predictive accuracy when undertaking applied assessments.

Sensitivity values equal the percentage of recidivists accurately identified by a given cut-off score. The sensitivities of the Static-99R and Static-2002R can be increased by lowering the cut-off score. Specificity values equal the percentage of nonrecidivists accurately identified by a given cut-off score. The specificities of the Static-99R and Static-2002R can be increased by increasing the cut-off score. Thus, one can select cut-off scores to increase an instrument's sensitivity, or specificity; but simultaneously increasing both sensitivity and specificity is not possible. Hence, the term "sensitivity-specificity trade off" underscores the impossibility of selecting a cut-off score to both increase sensitivity and specificity.

The Static-99 website ([www.static99.org](http://www.static99.org)) reports outcome data for all four Static-99R comparison groups. These comparison groups can be considered as low base rate (Routine Sample), low-moderate base rate (Preselected for Treatment Need), moderate-high base rate (Non-Routine Sample), and high base rate sample (High Risk/need Group). Except for the Routine Sample (with only 5-year follow-up data available), there are 5- and 10-year follow-up data for the remaining groups. These data were taken from the "Norms" section of the Static-99 website, specifically the "Detailed recidivism tables Static-99R (October 2009)." For each comparison group and both follow-up periods, the data were taken from the "Fixed Follow-up" groups. These data can be found at [www.static99.org](http://www.static99.org), specifically in the section identified as "Detailed recidivism tables Static-99R (October 2009)." These data are not estimates, but instead correspond to known outcomes for actual offenders.

### Five-Year Follow-Up Data for Static-99R

Tables 1-4 report PPV, NPV, and the frequencies of TP, FP, TN, and FN outcomes, for 5-year follow-ups, at all scores for the Static-99R. That is to say, any Static-99R score can be used to rule-in, or rule-out, recidivism risk.

**Table 1:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-99R Routine sample: 5 year follow-up, N= 2406, Base Rate = .06 (145/2406), Recidivists = 145, Non-Recidivists = 2261.

99-R Score	PPV	NPV	TP	FP	TN	FN
-3 & above	.06		145	2261	0	0
-3 & below		1.00	145	2221	40	0
-2 & above	.06		145	2221	40	0
-2 & below		1.00	145	2156	105	0

-1 & above	.06		145	2156	105	0
-1 & below		.98	138	1903	358	7
0 & above	.07		138	1903	358	7
0 & below		.98	130	1617	644	15
1 & above	.07		130	1617	644	15
1 & below		.98	120	1277	984	25
2 & above	.09		120	1277	984	25
2 & below		.97	106	941	1320	39
3 & above	.10		106	941	1320	39
3 & below		.97	86	618	1643	59
4 & above	.12		86	618	1643	59
4 & below		.96	69	358	1903	76
5 & above	.16		69	358	1903	76
5 & below		.95	41	193	2068	104
6 & above	.18		41	193	2068	104
6 & below		.95	27	97	2164	118
7 & above	.22		27	97	2164	118
7 & below		.94	15	34	2227	130
8 & above	.31		15	34	2227	130
8 & below		.94	7	14	2247	138
9 & above	.33		7	14	2247	138
9 & below		.94	2	6	2255	143
10 & above	.25		2	6	2255	143
10 & below		.94	0	1	2260	145
11 & above	.00		0	1	2260	145
11 & below		.94	0	0	2261	145

**Table 2:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-99R Pre-Selected for Treatment sample: 5 year follow-up, N= 1782, Base Rate = .09 (163/1782), Recidivists = 163, Non-Recidivists = 1619.

99-R Score	PPV	NPV	TP	FP	TN	FN
-3 & above	.09		163	1619	0	0
-3 & below		1.00	163	1587	32	0
-2 & above	.09		163	1587	32	0
-2 & below		.97	161	1557	62	2
-1 & above	.09		161	1557	62	2
-1 & below		.98	159	1385	234	4
0 & above	.10		159	1385	234	4
0 & below		.97	150	1148	471	13
1 & above	.12		150	1148	471	13
1 & below		.95	128	905	714	35
2 & above	.12		128	905	714	35
2 & below		.95	111	641	978	52
3 & above	.15		111	641	978	52
3 & below		.94	91	429	1190	72
4 & above	.18		91	429	1190	72
4 & below		.94	70	246	1373	93
5 & above	.22		70	246	1373	93
5 & below		.93	46	129	1490	117
6 & above	.26		46	129	1490	117
6 & below		.92	25	63	1556	138
7 & above	.28		25	63	1556	138
7 & below		.91	14	32	1587	149
8 & above	.30		14	32	1587	149
8 & below		.91	4	9	1610	159
9 & above	.31		4	9	1610	159
9 & below		.91	2	3	1616	161

10 & above	.40		2	3	1616	161
10 & below		.91	0	1	1618	163
11 & above	.00		0	1	1618	163
11 & below		.91	0	0	1619	163

**Table 3:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-99R Non-Routine sample: 5 year follow-up, N= 3353, Base Rate = .15 (497/3353), Recidivists = 497, Non-Recidivists = 2856.

99-R Score	PPV	NPV	TP	FP	TN	FN
-3 & above	.15		497	2856	0	0
-3 & below		1.00	497	2818	38	0
-2 & above	.15		497	2818	38	0
-2 & below		.95	493	2774	82	4
-1 & above	.15		493	2774	82	4
-1 & below		.97	486	2544	312	11
0 & above	.16		486	2544	312	11
0 & below		.95	467	2238	618	30
1 & above	.17		467	2238	618	30
1 & below		.94	435	1885	971	62
2 & above	.19		435	1885	971	62
2 & below		.93	398	1529	1327	99
3 & above	.21		398	1529	1327	99
3 & below		.92	343	1157	1699	154
4 & above	.23		343	1157	1699	154
4 & below		.90	272	776	2080	225
5 & above	.26		272	776	2080	225
5 & below		.89	193	455	5401	304
6 & above	.30		193	455	2401	304
6 & below		.87	114	240	2616	383
7 & above	.32		114	240	2616	383
7 & below		.86	60	110	2746	437

8 & above	.35		60	110	2746	437
8 & below		.86	24	37	2819	473
9 & above	.39		24	37	2819	473
9 & below		.85	8	9	2847	489
10 & above	.47		8	9	2847	489
10 & below		.85	0	2	2854	497
11 & above	.00		0	2	2854	497
11 & below		.85	0	0	2856	497

**Table 4:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-99R High-Risk sample: 5 year follow-up, N= 1313, Base Rate = .21 (276/1313), Recidivists = 276, Non-Recidivists = 1037

99-R Score	PPV	NPV	TP	FP	TN	FN
-3 & above	.21		276	1037	0	0
-3 & below		1.00	276	1036	1	0
-2 & above	.21		276	1036	1	0
-2 & below		1.00	276	1030	7	0
-1 & above	.21		276	1030	7	0
-1 & below		.95	274	1000	37	2
0 & above	.22		274	1000	37	2
0 & below		.94	271	959	78	5
1 & above	.22		271	959	78	5
1 & below		.94	266	873	164	10
2 & above	.23		266	873	164	10
2 & below		.90	250	801	236	26
3 & above	.24		250	801	236	26
3 & below		.88	227	663	374	49
4 & above	.26		227	663	374	49
4 & below		.86	185	485	552	91
5 & above	.28		185	485	552	91
5 & below		.84	139	301	736	137



6 & above	.32		139	301	736	137
6 & below		.82	86	165	872	190
7 & above	.34		86	165	872	190
7 & below		.81	44	76	961	232
8 & above	.37		44	76	961	232
8 & below		.80	20	28	1009	256
9 & above	.42		20	28	1009	256
9 & below		.79	6	7	1030	270
10 & above	.46		6	7	1030	270
10 & below		.79	0	2	1035	276
11 & above	.00		0	2	1035	276
11 & below		.79	0	0	1037	276

### Discussion of Static-99R 5-Year Follow-Up Data

Without exception, the NPVs for any Static-99R score, over a 5-year follow-up, exceed the PPVs for the same score. Consequently, using the Static-99R to rule-out recidivism risk is inevitably more accurate than ruling-in recidivism risk. Static-99 scores of 6 and higher have often prompted recommendations for civil commitment. Comparing the PPV and NPV for Static-99R scores of 6 and above provides an interesting contrast.

Using Routine Sample Static-99R scores of 6 and above, Table 5 demonstrates how PPVs are computed. Note that there are a total of 234 rule-in predictions; 41 of those decisions are TPs, and 193 of those decisions are FPs. If ruling-in recidivism risk for all Routine Sample Static-99R scores of 6 and above, the data indicate that an evaluator would be correct in 18% of cases (41 of a total of 234 rule-in decisions).

**Table 5**

Positive	True 41	False 193	Totals 234
Negative	True Positive 104	False Positive 2,068	2,172
Totals	False Negative 145	True Negative 2,261	2,406

Using Routine Sample Static-99R scores of 6 and below, Table 6 demonstrates how NPVs are computed. Note that there are a total of 2,172 rule-out predictions, and 2,068

of those decisions are TNs, and 104 of those 2,172 decisions are FNs. If ruling-out recidivism risk for all Routine Sample Static-99R scores of 6 and below, the data indicate that an evaluator would be correct in 95% of cases (2,068 of a total of 2,172 rule-out decisions).

**Table 6**

	True	False	Totals
Positive	27	97	124
	True Positive	False Positive	
Negative	1,118	2,164	2,282
	False Negative	True Negative	
Totals	145	2,261	2,406

**Table 7:** PPV and NPV for a Static-99R score of 6 and above, and 6 and below, for all four comparison groups for 5-year follow-ups.

	Routine	Pre-selected	Non-routine	High-risk
PPV	.18	.25	.30	.32
NPV	.95	.92	.87	.82

Table 7 demonstrates how PPVS gradually increase as base rates increase. Correspondingly, NPVs gradually decrease as base rates increase.

### Static-99R - 10-Year Follow-Up Data

Tables 8-10 report PPV, NPV, and the frequencies of TP, FP, TN, and FN outcomes, for 10-year follow-ups, at all scores for the Static-99R. Ten-year follow-up data are available for the Preselected for Treatment sample, the Non-Routine sample, and the High-Risk sample. There are no 10-year follow-up data available for the Routine sample.

**Table 8:** PPV, NPV and frequencies of TP, FP, TN, and FN for the Static-99R Pre-selected for Treatment sample: 10 year follow-up, N= 866, Base Rate = .14 (118/866), Recidivists = 118, Non-Recidivists = 748

99-R Score	PPV	NPV	TP	FP	TN	FN
-3 & above	.14		118	748	0	0
-3 & below		1.00	118	732	16	0
-2 & above	.14		118	732	16	0
-2 & below		.93	116	722	26	2

-1 & above	.14		116	722	26	2
-1 & below		.95	112	641	107	6
0 & above	.15		112	641	107	6
0 & below		.93	102	537	211	16
1 & above	.16		102	537	211	16
1 & below		.93	93	415	333	25
2 & above	.18		93	415	333	25
2 & below		.91	73	293	455	45
3 & above	.20		73	293	455	45
3 & below		.90	60	196	552	58
4 & above	.23		60	196	552	58
4 & below		.90	46	99	649	72
5 & above	.32		46	99	649	72
5 & below		.88	27	51	697	91
6 & above	.35		27	51	697	91
6 & below		.88	15	26	722	103
7 & above	.37		15	26	722	103
7 & below		.87	8	16	732	110
8 & above	.33		8	16	732	110
8 & below		.86	1	5	743	117
9 & above	.17		1	5	743	117
9 & below		.86	0	2	746	118
10 & above	.00		0	2	746	118
10 & below		.86	0	1	747	118
11 & above	.00		0	1	747	118
11 & below		.86	0	0	748	118

**Table 9:** PPV, NPV, and frequencies of TP, FP, TN, and FN for the Static-99R Non-Routine sample: 10 year follow-up, N=1626, Base Rate = .20 (332/1626), Recidivists = 332, Non-Recidivists = 1294.

99-R Score	PPV	NPV	TP	FP	TN	FN
-3 & above	.20		332	1294	0	0
-3 & below		1.00	332	1278	16	0
-2 & above	.21		332	1278	16	0
-2 & below		.93	330	1266	28	2
-1 & above	.21		330	1266	28	2
-1 & below		.94	324	1163	131	8
0 & above	.22		324	1163	131	8
0 & below		.93	311	1030	264	21
1 & above	.23		311	1030	264	21
1 & below		.93	297	857	437	35
2 & above	.26		297	857	437	35
2 & below		.90	263	703	591	69
3 & above	.27		263	703	591	69
3 & below		.88	227	540	754	105
4 & above	.30		227	540	754	105
4 & below		.86	178	339	955	154
5 & above	.34		178	339	955	154
5 & below		.84	118	185	1109	214
6 & above	.39		118	185	1109	214
6 & below		.82	70	103	1191	262
7 & above	.40		70	103	1191	262
7 & below		.81	33	54	1240	299
8 & above	.38		33	54	1240	299
8 & below		.80	13	21	1273	319
9 & above	.38		13	21	1273	319
9 & below		.80	4	8	1286	328

10 & above	.33		4	8	1286	328
10 & below		.80	0	2	1292	332
11 & above	.00		0	2	1292	332
11 & below		.80	0	0	1294	332

**Table 10:** PPV, NPV, and frequencies of TP, FP, TN, and FN for the Static-99R High Risk/need Group: 10 year follow-up, N= 703, Base Rate = .29 (204/703), Recidivists = 204, Non-Recidivists = 499.

99-R Score	PPV	NPV	TP	FP	TN	FN
-3 & above	.29		204	499	0	0
-3 & below		.00	204	499	0	0
-2 & above	.29		204	499	0	0
-2 & below		1.00	204	497	2	0
-1 & above	.29		204	497	2	0
-1 & below		.92	202	475	24	2
0 & above	.30		202	475	24	2
0 & below		.91	199	447	52	5
1 & above	.31		199	447	52	5
1 & below		.91	194	396	103	10
2 & above	.33		194	396	103	10
2 & below		.85	180	368	131	24
3 & above	.33		180	368	131	24
3 & below		.81	158	308	191	46
4 & above	.34		158	308	191	46
4 & below		.78	124	212	287	80
5 & above	.37		124	212	287	80
5 & below		.76	85	118	381	119
6 & above	.42		85	118	381	119
6 & below		.74	53	67	432	151
7 & above	.44		53	67	432	151
7 & below		.72	24	35	464	180

8 & above	.41		24	35	464	180
8 & below		.72	12	15	484	192
9 & above	.44		12	15	484	192
9 & below		.71	4	6	493	200
10 & above	.40		4	6	493	200
10 & below		.71	0	1	498	204
11 & above	.00		0	1	498	204
11 & below		.71	0	0	499	204

The recidivism base rates, and sample sizes in parentheses, for each Static-99R comparison group can be seen in Table 11.

**Table 11:** Recidivism base rates, and sample sizes, for each Static-99R comparison group.

	Routine	Pre-selected	Non-routine	High-risk
5 year	.06 (2406)	.09 (1782)	.15 (3353)	.21 (1313)
10 year	N/A	.14 (866)	.20 (1626)	.29 (703)

### Predictive Accuracy of the Static-2002R

The Static-99 website ([www.static99.org](http://www.static99.org)) reports outcome data for all three Static-2002R comparison groups. These comparison groups can be considered as low-base-rate sample (Routine), moderate-base-rate sample (Non-Routine), and high-base-rate sample (High Risk). Except for the Routine Sample (with only 5 year follow-up data available), there are 5- and 10-year follow-up data for the remaining groups. These data were taken from the "Static-2002" section of the Static-99 website, specifically the "Detailed Static-2002R Recidivism Tables." For each group and both follow-up periods, the data were taken from the "Fixed Follow-up" groups.

Tables 12-16 report PPV, NPV, and the frequencies of TP, FP, TN, and FN outcomes -- for 5-year follow-ups -- at any and all scores for the Static-2002R.

**Table 12:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-2002R Routine sample: 5 year follow-up, N= 526, Base Rate = .05 (28/526), Recidivists = 28, Non-Recidivists = 498.

99-R Score	PPV	NPV	TP	FP	TN	FN
-2 & above	.05		28	498	0	0
-2 & below		1.00	28	490	8	0

-1 & above	.05		28	490	8	0
-1 & below		1.00	28	474	24	0
0 & above	.06		28	474	24	0
0 & below		1.00	28	438	60	0
1 & above	.06		28	438	60	0
1 & below		.99	27	391	107	1
2 & above	.06		27	391	107	1
2 & below		.99	26	335	163	2
3 & above	.07		26	335	163	2
3 & below		.98	24	267	231	4
4 & above	.08		24	267	231	4
4 & below		.98	21	191	307	7
5 & above	.10		21	191	307	7
5 & below		.97	16	127	371	12
6 & above	.11		16	127	371	12
6 & below		.96	12	64	434	16
7 & above	.16		12	64	434	16
7 & below		.96	7	30	468	21
8 & above	.19		7	30	468	21
8 & below		.95	4	12	486	24
9 & above	.25		4	12	486	24
9 & below		.95	1	3	495	27
10 & above	.25		1	3	495	27
10 & below		.95	1	1	497	27
11 & above	.50		1	1	497	27
11 & below		.95	0	0	498	28
12 & above	.00		0	0	498	28
12 & below		.95	0	0	498	28
13 & above	.00		0	0	498	28
13 & below		.95	0	0	498	28

**Table 13:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-2002R Non-Routine sample: 5 year follow-up, N= 1121, Base Rate = .20 (222/1121), Recidivists = 222, Non-Recidivists = 899

99-R Score	PPV	NPV	TP	FP	TN	FN
-2 & above	.20		222	899	0	0
-2 & below		1.00	222	897	2	0
-1 & above	.20		222	897	2	0
-1 & below		.89	221	891	8	1
0 & above	.20		221	891	8	1
0 & below		.92	219	864	35	3
1 & above	.20		219	864	35	3
1 & below		.95	217	803	96	5
2 & above	.21		217	803	96	5
2 & below		.93	209	724	175	13
3 & above	.22		209	724	175	13
3 & below		.94	204	627	272	18
4 & above	.25		204	627	272	18
4 & below		.89	172	489	410	50
5 & above	.26		172	489	410	50
5 & below		.87	144	370	529	78
6 & above	.28		144	370	529	78
6 & below		.86	117	239	660	105
7 & above	.33		117	239	660	105
7 & below		.85	87	160	739	135
8 & above	.35		87	160	739	135
8 & below		.83	50	80	819	172
9 & above	.38		50	80	819	172
9 & below		.82	28	34	865	194
10 & above	.44		28	34	865	194
10 & below		.80	7	16	883	215



11 & above	.30		7	16	883	215
11 & below		.80	3	4	895	219
12 & above	.43		3	4	895	219
12 & below		.80	0	1	898	222
13 & above	.00		0	1	898	222
13 & below		.80	0	0	899	222

**Table 14:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-2002R High-Risk sample: 5 year follow-up, N= 931, Base Rate = .22 (204/931), Recidivists = 204, Non-Recidivists = 727.

99-R Score	PPV	NPV	TP	FP	TN	FN
-2 & above	.22		204	727	0	0
-2 & below		.00	204	727	0	0
-1 & above	.22		204	727	0	0
-1 & below		1.00	204	726	1	0
0 & above	.22		204	726	1	0
0 & below		.92	202	705	22	2
1 & above	.22		202	705	22	2
1 & below		.94	201	676	51	3
2 & above	.23		201	676	51	3
2 & below		.90	193	628	99	11
3 & above	.24		193	628	99	11
3 & below		.92	188	552	175	16
4 & above	.25		188	552	175	16
4 & below		.86	158	436	291	46
5 & above	.27		158	436	291	46
5 & below		.85	133	335	392	71
6 & above	.28		133	335	392	71
6 & below		.84	107	215	512	97
7 & above	.33		107	215	512	97
7 & below		.82	79	142	585	125

8 & above	.36		79	142	585	125
8 & below		.80	45	72	655	159
9 & above	.38		45	72	655	159
9 & below		.80	28	31	696	176
10 & above	.47		28	31	696	176
10 & below		.78	7	15	712	197
11 & above	.32		7	15	712	197
11 & below		.78	3	4	723	201
12 & above	.43		3	4	723	201
12 & below		.78	0	1	726	204
13 & above	.00		0	1	726	204
13 & below		.78	0	0	727	204

**Table 15:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-2002R Non-Routine sample: 10 year follow-up, N= 766, Base Rate = . 27 (209/766), Recidivists = 209, Non-Recidivists = 557.

99-R Score	PPV	NPV	TP	FP	TN	FN
-2 & above	.27		209	557	0	0
-2 & below		1.00	209	556	1	0
-1 & above	.27		209	556	1	0
-1 & below		.80	208	553	4	1
0 & above	.27		208	553	4	1
0 & below		.88	206	534	23	3
1 & above	.28		206	534	23	3
1 & below		.93	204	491	66	5
2 & above	.29		204	491	66	5
2 & below		.91	197	436	121	12
3 & above	.31		197	436	121	12
3 & below		.90	189	367	190	20
4 & above	.34		189	367	190	20
4 & below		.83	153	282	275	56

5 & above	.35		153	282	275	56
5 & below		.80	121	213	344	88
6 & above	.36		121	213	344	88
6 & below		.79	97	142	415	112
7 & above	.41		97	142	415	112
7 & below		.77	74	95	462	135
8 & above	.44		74	95	462	135
8 & below		.75	41	47	510	168
9 & above	.47		41	47	510	168
9 & below		.74	26	23	534	183
10 & above	.53		26	23	534	183
10 & below		.73	5	12	545	204
11 & above	.29		5	12	545	204
11 & below		.73	3	3	554	206
12 & above	.50		3	3	554	206
12 & below		.73	0	1	556	209
13 & above	.00		0	1	556	209
13 & below		.73	0	0	557	209

**Table 16:** PPV, NPV, and frequencies of TP, FP, TN, and FN, for the Static-2002R High-Risk sample: 10 year follow-up, N= 642, Base Rate = .29 (189/642), Recidivists = 189, Non-Recidivists = 453.

99-R Score	PPV	NPV	TP	FP	TN	FN
-2 & above	.29		189	453	0	0
-2 & below		.00	189	453	0	0
-1 & above	.29		189	453	0	0
-1 & below		.00	189	453	0	0
0 & above	.29		189	453	0	0
0 & below		.89	187	437	16	2
1 & above	.30		187	437	16	2
1 & below		.93	186	414	39	3

2 & above	.31		186	414	39	3
2 & below		.88	179	377	76	10
3 & above	.32		179	377	76	10
3 & below		.88	171	323	130	18
4 & above	.35		171	323	130	18
4 & below		.80	138	250	203	51
5 & above	.36		138	250	203	51
5 & below		.76	109	195	258	80
6 & above	.36		109	195	258	80
6 & below		.76	87	129	324	102
7 & above	.40		87	129	324	102
7 & below		.75	66	87	366	123
8 & above	.43		66	87	366	123
8 & below		.73	38	43	410	151
9 & above	.47		38	43	410	151
9 & below		.73	26	21	432	163
10 & above	.55		26	21	432	163
10 & below		.71	5	11	442	184
11 & above	.31		5	11	442	184
11 & below		.71	3	3	450	186
12 & above	.50		3	3	450	186
12 & below		.71	0	1	452	189
13 & above	.00		0	1	452	189
13 & below		.71	0	0	453	189

**Table 17:** PPV and NPV for a Static-2002R score of 6 and above, and 6 and below, for all three comparison groups for 5-year follow-ups.

	Routine	Non-routine	High-risk
PPV	.11	.28	.28
NPV	.96	.86	.84

**Discussion of Static-2002R Data**

As was the case for the Static-99R, NPVs far exceed PPVs for the Static-2002R. In other words, ruling-out recidivism risk is always more accurate than ruling it in when relying on the Static-2002R.

**Table 18:** Recidivism base rates and sample sizes for each Static-2002R comparison group.

	Routine	Non-routine	High-risk
5 year	.05 (526)	.20 (1121)	.22 (931)
10 year	N/A	.27 (766)	.29 (642)

**Which Comparison Group and Which Follow-Up Period?**

Static-99R scores can be compared to any one of four comparison groups. Static-2002R scores can be compared to any one of three comparison groups. These circumstances lead to the obvious question of how to select the most appropriate comparison group? In her thesis that prompted wholesale changes in the Static-99, Helmus (2009) recommended correctional systems developing their own "local" norms. Though local norms amount to an ideal option, they remain unavailable for SVP jurisdictions in the United States. Helmus also advised using the Routine Sample as they are most representative of the entire population of sex offenders. Helmus additionally identified using all available comparison groups as another option.

Using all the comparison groups is clearly the most transparent option. Evaluators doing so can acknowledge that there are no generally recognized and accepted decision-making criteria available for selecting one comparison group in particular. Using all the comparison groups allows identifying a range of PPVs and NPVs. In turn, evaluators can explain that the true PPVs and NPVs for a given offender most likely fall within that range.

Spada, Perillo, Mercado, and Jeglic (2011) followed 2,756 previously incarcerated, New Jersey, male sex offenders for an average of 6.5 years (range of 2-13 years). Of those offenders who reoffended, more than 50% reoffended before the third year at large. For offenders who remain offense-free for five years, their risk is much less than it was at the time of their release five years previously, and what risk remains is mostly influ-

enced by dynamic variables. For example, Harris, Phenix, Hanson, and Thornton (2003) have advised: "In general, the expected sexual offence recidivism rate should be reduced by about half if the offender has five to ten years of offence-free behaviour in the community" (p. 59). Bani-Yaghoub et al. (2010) detailed how linear additive instruments, such as the Static-99R and Static-2002R, too often overlook instances of "parabolic decline."

Parabolic decline ensues when an offender's heightened risk of recidivism markedly declines in a few years. The effects of dynamic variables such as supportive release environment, stable interpersonal relationships, and positive treatment participation, exercise a greater influence on recidivism outcomes than static variables. Assessing these variables prior to an offender's release, however, borders on the impossible. Attempting to predict recidivism beyond a five-year window of opportunity, relying entirely on static variables, is therefore ill-advised. As a result, five-year follow-up data are most appropriately used for the Static-99R and Static-2002R.

**Discussion of Static-99R and Static-2002R Data**

Hard core empiricists are wont to say, "The data speak for themselves." The data reported herein speak words of caution for any SVP evaluator inclined to rule-in recidivism risk. Vars (in press) has reported standards of SVP commitment for the federal government, and for the 19 states with SVP statutes. Vars further distinguished between "commitment standards" and "proof standards." Proof standards correspond to either a "clear-and-convincing" standard, or a "beyond-a-reasonable-doubt standard." Vars additionally identified whether any jurisdiction's commitment standard specified a likelihood of recidivism as greater than 50%, at 50%, less than 50%, or unspecified.

In Vars' opinion, satisfying a clear and convincing standard of proof necessitates at least a 75% likelihood of recidivism. Satisfying a beyond a reasonable doubt standard necessitates at least 90% likelihood of recidivism. The commitment standards and proof standards for each SVP jurisdiction are summarized in Table 19. Vars mistakenly identified a "beyond-a-reasonable-doubt" standard for the federal government. In fact, the federal proof standard is "clear-and-convincing" evidence.

**Table 19:** Standards of Commitment and Proof by Jurisdiction

Commitment Standard		Proof Standard	
Likelihood of recidivism	Clear & Convincing (75%)	Reasonable Doubt (90%)	
>50%	MN, NJ	AZ, IL	
50%	FL, MO, NB	IA, WA, WI	
<50%	Fed	CA, MA	
Unspecified	NH, NY, ND, VA	KS, SC, TX	

Except for a Static-2002R score of 10 and above, followed over a 10-year period, the PPVs for the Static-99R and Static-2002R never exceed .50 for the available comparison data. As a result, these PPVs fall far below the .75 and .90 likelihood standards advocated by Vars. In view of these data one might ask: "What amounts to 'best practice' when assessing the reoffending risk of previously convicted sex offenders?"

The data reported herein emphatically underscore the importance of recognizing and responding to base rates. Daniel Montaldi (2011), affiliated with Florida's SVP program, advises: "Given decreased base rates [of sexual recidivism] over the past 20 years, the most accurate method now may be to just use the overall (low) reconviction base rate ... We would have false negative errors but perhaps fewer errors overall" (p. 2). If predicting that no one will reoffend, Montaldi's assessment is entirely accurate when applied to these Static-99R and Static-2002R data. The greatest level of accuracy is obtained by predicting that no one will reoffend. The PPVs for the Static-99R and Static-2002R never exceed the NPVs for the same score. These data, obtained from observed recidivism rates, demonstrate that the Static-99R and Static-2002R lead evaluators into false-positive predictions much more often than false-negative predictions.

Clinically adjusted actuarial assessment cannot overcome the shortcomings of the Static-99R and Static-2002R. Adjusted actuarial assessment (AAA) raises the question of "incremental validity." Walters (2011) defines incremental validity as "... a measure's ability to contribute to the prediction of a criterion above and beyond what could be achieved with more conventional, less expensive, or simpler schemes" (p. 227). Applied to AAA, the question becomes whether AAA increases the predictive accuracy obtained by the Static-99R and/or Static-2002R alone. In other words, if AAA allows incremental validity beyond the Static-99R and Static-2002R, the PPV and/or NPV would have to increase as a result of AAA. That is to say, the frequency of TPs or TNs would have to increase, while the frequency of FPs and FNs correspondingly decrease. The risk factors, which many SVP evaluators use to rule-in recidivism risk, are found more frequently in samples of offenders who do not reoffend compared to samples of offenders who do reoffend (Campbell & DeClue, 2010b). Consequently, there is little likelihood of AAA providing incremental validity above and beyond the Static-99R and Static-2002R.

When assessing the recidivism risk of sex offenders, base rates amount to an "inconvenient truth." Although we might want to wish them away via AUC values obtained via ROC methods, base rates are a persistent nuisance when attempting to predict infrequently occurring events. SVP evaluators who disregard the base rate problem too often commit false-positive errors.

Received March 30, 2011; revision submitted August 18, 2011; accepted August 20, 2011

## References

- Bani-Yaghoub, M., Federoff, J. P., Curry, S., & Amundsen, D. E. (2010). A time series modeling approach in risk appraisal of violent and sexual recidivism. *Law and Human Behavior, 34*(5), 349-366.
- Barbaree, H. E., Seto, M., 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*, 490-521.
- Campbell, T. W. (2007). When prophecy fails: Retreating from prediction. *Journal of Sexual Offender Civil Commitment: Science and the Law, 2*, 1-11.
- Campbell, T. W., & DeClue, G. (2010a). Maximizing predictive accuracy in sexually violent predator evaluations. *Open Access Journal of Forensic Psychology, 2*, 148-232.
- Campbell, T. W., & DeClue, G. (2010b). Flying blind with naked factors: Problems and pitfalls in adjusted actuarial sex offender risk assessment. *Open Access Journal of Forensic Psychology, 2*, 75-101.
- Craig, L. A., & Beech, A. R. (2010). Towards a guide to best practice in conducting actuarial risk assessments with sex offenders. *Aggression and Violent Behavior, 15*, 278-293.
- Doren, D. M. (2006). Inaccurate arguments in sex offender civil commitment proceedings. In A. Schlank (Ed.), *The sexual predator: Law and Public Policy - Clinical Practice*, Vol. III. Kingston, NJ: Civic Research Institute.
- Epperson, D. L., Kaul, J. D., & Huot, S. J. (1995 October). Predicting risk of recidivism for incarcerated sex offenders: Updated development of the Sex Offender Screening Tool (SOST). Paper presented at the 14th annual research and treatment conference of the Association for the Treatment of Sexual Abusers, New Orleans, LA.
- 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.
- Hanson, R. K. (2000). Will they do it again: Predicting sex-offense recidivism. *Current Directions in Psychological Science, 9*, 106-109.
- Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sex offense recidivism. *Criminal Justice and Behavior, 27*, 6-35.



- Hanson, R. K., & Morton-Bourgon, K. E. (2004). Predictors of sexual recidivism: An updated meta-analysis. (Research Rep No. 2004-02). Ottawa, Canada: Public Safety and Emergency Preparedness, Canada.
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, *21*(1), 1-21.
- Harris, A., Phenix, A., Hanson, R. K., & Thornton, D. (2003). Static-99 coding rules: Revised 2003. [www.sgc.ca](http://www.sgc.ca)
- Hart, S. D., Webster, C. D., & Menzies, R. J. (1993). A note on portraying the accuracy of violence predictions. *Law and Human Behavior*, *17*, 695-700.
- Helmus, L. (2009 Sept). Re-norming Static-99 recidivism estimates: Exploring base rate variability across sex offender samples. Thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment for the degree, Master of Arts. Department of Psychology, Carleton University, Ottawa, Ontario.
- Montaldi, D. F. (2011). Probability and sexual offenders: Some preliminary remarks. *ATSA Forum*, XXIII (2), Spring 2011.
- Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence*, *10*, 85-105.
- Serin, R. C., & Brown, S. L. (2000). The clinical use of the Hare Psychopathy Checklist-Revised in contemporary risk assessment. In G. C. Gacono (Ed.), *The clinical and forensic assessment of psychopathy*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Spada, A., Perillo, A. D., Mercado, C. C., & Jeglic, E. (March 2011). Predicting sex offender recidivism: An examination of factors associated with early failure among sexual offenders. Presented at the Annual Conference of the American Psychology-Law Society, Miami, FL.
- Vars, F. E. (in press). Rethinking the indefinite detention of sex offenders. *Connecticut Law Review*, Volume 44.
- Walters, G. D. (2011). Taking the next step: Combining incrementally valid indicators to improve recidivism prediction. *Assessment*, *18*(2), 227-233.