*Journal of Sexual Offender Civil Commitment: Science and the Law*, *1*, 124-140. (2006).

# Evaluation of Civil Commitment Criteria in a High Risk Sample of Sexual Offenders

Jeffrey Abracen
Central District Parole, Correctional Service of Canada

Jan Looman[1]
Regional Treatment Centre (Ontario)

## ABSTRACT

The contribution of DSM-IV diagnoses of Paraphilia and Personality Disorder to predictions of sexual recidivism were examined for a group of high risk sexual offenders (i.e., scores of 5 or higher on the Satatic-99). It was found that none of the sub-groups of offenders re-offended at rates which approached 50% (approximately 5 year follow-up) and that only for a group that met both personality Disorder and Paraphilia diagnoses did the recidivism rate exceed 20%. In addition, Psychopathy did not add to the validity of predictions. Results are discussed in terms of the validity of the SVP criteria and clinicians ability to distinguish high-risk sexual offenders from other groups.

## Introduction

The issue of sexual offender commitment has received a great deal of attention in both the legal and mental health communities. A number of states and the District of Columbia have laws for the special commitment of convicted sexual offenders who are about to be released to the community. As of the summer of 2002 nearly 2500 sexual offenders were hospitalized, ostensibly for treatment, pursuant to this legislation (Fitch, 2003). As noted by Jackson, Rogers and Shuman (2004) typically four criteria exist on the basis of which an individual may be committed. The first is that the individual must have been convicted of a sexual offence (though in North Dakota even this criterion is apparently not necessary). The second criterion is the presence of a mental disorder or abnormality. In the majority of statutes there is also one phrase "or personality disorder" which helps to define the mental abnormality (Doren, 2002). It is important to note that the simple presence of a mental disorder is not sufficient; the key element

[1] Author to whom correspondence should be addressed. Regional Treatment Centre, 555 King St. W., P.O. Box 22, Kingston, Ontario, Canada, K7L 4V7. e-mail loomanja@csc-scc.gc.ca

is that the abnormality causes the lack of control (Jackson et al., 2004). Typically, diagnoses are based on the criteria listed in DSM-IV-TR (APA, 2000). However, as noted by Doren (2002) the legal phrase "mental disorder" is not synonymous with the DSM-IV concept. For example, Doren (2002) notes that the term personality disorder is left completely undefined in the majority of States whose commitment criteria discuss these disorders. The importance of the legal mental abnormality requirement is critical in that the legislation rests upon the presence of such a disorder.

Aside from the criteria listed above, all commitment statutes require that the individual is at risk for committing future acts of sexual violence. The extent of this likelihood and the degree of risk are less well defined. Doren (2002) notes that 12 of the statutes use the term "likely" to denote the required degree of risk necessary for commitment. At least two states (Iowa and Washington) have further refined this concept by using the phrase "more likely than not" (Doren, 2002). Evaluators typically interpret the phrase "more likely than not" as meaning 50% and above. With reference to how such risk is assessed, the evidence suggests that actuarially based assessments are the most accurate means currently available (Hanson, Morton & Harris, 2003). Recent advances in the assessment of dynamic or changeable risk using such measures as the STABLE 2000 (Hanson & Harris, 2001) offer promise and may be used to moderate the risk assessed by means of actuarial instruments. Among the best validated instruments for the assessment of sexual recidivism are the STATIC-99 (Hanson & Thornton, 2000) and the SORAG (Quinsey, Harris Rice & Cormier, 1998) which includes the offender's score on the Hare Psychopathy Checklist-Revised (PCL-R; Hare 1991, 2003). The Hare PCL-R, in turn, is one of the best validated measures for the assessment of risk for general or violent recidivism (see Hare 2003 for a review of the available research). As several articles have reviewed the relevant research related to these measures (e.g., Barbaree, Seto, Langton & Peacock, 2001; Hanson, Morton & Harris, 2003) this information will not be repeated here.

Doren (1998, 2002) has suggested that if sexual offenders are followed for extremely long follow-up periods (say 25 years) the rate at which sex offenders will possibly recidivate approaches or exceeds the "more likely than not" criterion. However, his perspective has been criticized by others (e.g., Wollert, 2001). In fact, Hanson, Morton & Harris (2003), in a recent meta-analysis of the available literature, note that the five-year sexual recidivism rate (involving new charges or convictions) was approximately 14% and the 20 year rate was 27%. Therefore, the average long term rate of sexual offence recidivism, for sexual offenders as a group, is considerably lower than that which is necessary for the basis of commitment. It should be noted that the argument made by Doren (1998) was based on only one sample which may have been at particularly high risk of sexual offence recidivism (see Wollert, 2001) whereas the meta-analytic data provided by Hanson et al. (2003) were based on 10 samples and comprised over 4000 ($n$=4724) sexual offenders, and thus likely more representative. Hanson and Morton-Bourgon (2005), in a more recent meta-analysis of the sex offender literature, have demonstrated that the average sexual recidivism rate of sexual offenders is quite low (i.e., 13.7% over a 5 to 6 year period) reinforcing the percentages listed above. The phenomenon of relatively low rates of sexual offence recidivism is commonly referred to as the base rate problem. The base rate problem is critical, as Janus & Meehl (1997) have demonstrated that with a base rate of under 30%, given the current state of the art in risk assessment, prediction will be wrong most of the time. Even if one disputes the assumptions made by Janus and Meehl (Doren & Epperson, 2001) the fact that there is a relatively low base rate of long term sexual recidivism suggests that evaluators may have a difficult time differentiating those offenders who are going to offend sexually in the future from those who are not.

A variety of recent research further reinforces these concerns. First, the impact of treatment on sexual recidivism rates needs to be considered (See Abracen & Looman, 2004 for a review). The results of the meta-analysis by Hanson, Gordon, Harris, Marques, Murphy, Quinsey and Seto (2002) clearly demonstrate the efficacy of contemporary approaches to sex offender treatment. More recently, Lösel & Schmucker (2005) have also demonstrated the efficacy of sexual offender treatment in a meta-analytic study which examined 69 investigations and represents one of the most exhaustive meta-analytic reviews of the literature to date. The authors noted that treated offenders showed six percentage points or 37% less sexual recidivism than controls.

Although there are those who remain sceptical regarding the efficacy of contemporary sexual offender treatment programs, the accumulating evidence suggests that contemporary treatment programs, particularly those with a cognitive-behavioral orientation, are successful at significantly reducing rates of sexual recidivism. There are those who have argued that there is little evidence as to the efficacy of contemporary approaches to sex offender treatment. Most notably, Quinsey and his colleagues have argued that, given the poor quality of the research in the area, there is no reason to believe the results of meta-analyses that have indicated that treatment reduces the risk of recidivism (e.g., Rice & Harris, 2003; Quinsey, Harris, Rice & Cormier, 2006). Marshall (2006), in an excellent critique of the arguments made by Rice & Harris (2003) has noted that when one considers effect size as an indication of treatment efficacy, the effect sizes which have been demonstrated for sex offender treatment programs are at or above acceptable levels and are, in many cases, higher than those available with reference to medical treatments (e.g., chemotherapy for breast cancer). Marshall (2006) also criticizes these authors for, on the one hand, claiming that studies that compare treated offenders to those not offered treatment are fundamentally flawed and, on the other hand, conducting studies that use exactly such a methodology themselves.

Recent research by Barbaree and his colleagues (Barbaree, Langton & Peacock, 2003), as well as some of our own research (Looman, 2006), has demonstrated that the STATIC-99 over predicts risk of sexual recidivism among treated sexual offenders. As well, as sexual offenders age their risk of recidivism appears to be reduced (Barbaree, Blanchard & Langton, 2003). Barbaree, Blanchard & Langton (2003) argue that therefore "it follows that the estimates for rates of recidivism given by actuarial methods will overestimate rates for older men" (p.70).

In a recent study by Jackson et al. (2004) it was observed that forensic psychologists only made accurate predictions of risk (based on data from actual case material where recidivism data for each defendant was known) approximately half the time. It may be that the base rate problem (i.e., the low rate of sexual offence recidivism among sexual offenders) combined with ambiguous definitions of commitment criteria may result in making it extremely difficult for even well trained clinicians to accurately assess risk of sexual recidivism.

The inclusion of DSM-based diagnoses may be particularly troubling in that actuarial assessments of risk alone are likely better indices of risk than are DSM based diagnoses. For example, with reference to antisocial personality disorder, Hare (1998) has noted that the base rates in prison samples are typically in the 50-75% range. If the majority of offenders meet the diagnostic criteria for antisocial personality disorder this suggests that the presence of such a diagnosis will be insufficient to determine who will be at a particularly elevated risk of recidivism. In spite of the fact that personality disorders (e.g., antisocial personality disorder) have been associated with sexual offence recidivism (Hanson & Morton-Bourgon, 2004) it is not clear whether the addition of such data to sexually violent offender assessments either adds to

the accuracy of prediction or obscures the prediction of sexual offence recidivism based on actuarial assessment alone. This is particularly true with reference to the assessment of high risk-high need sexual offenders who have received comparatively little attention in the empirical literature.

Given that sexually violent predator assessments presumably concern themselves with the assessment of high risk sexual offenders the utility of outcome research on specifically high risk sexual offenders is clearly relevant to the determination of the questions raised above. In order to investigate the rates of recidivism among a sample of high risk sexual offenders, clients treated at the Regional Treatment Centre Sex Offender Treatment program (RTCSOTP) were included in the present study. The RTCSOTP has been described elsewhere (Abracen & Looman, 2004, Looman, Abracen & Nicholaichuk, 2000). The reader is referred to these manuscripts for a description of the RTCSOTP. In brief, the RTCSOTP is designed for sexual offenders who are determined to be at elevated risk of sexual recidivism based on the assessment of actuarially determined risk or who present with significant treatment needs or both (see Mailloux, Abracen, Serin, Cousineau, Malcolm & Looman, 2003 for data comparing the RTCSOTP sample to other well known sex offender programs in the Ontario Region of Correctional Service of Canada).

Although there is relatively little outcome research in the area of civil commitment proceedings, several hypotheses were made based on the available literature related to treatment outcome with sexual offenders. First, it was predicted that actuarial risk assessment measures alone would provide reasonable estimates of actual risk of sexual recidivism. Second, it was predicted that the use of psychiatric diagnoses would not reliably distinguish sexual recidivists from non-recidivists. Last, it was predicted that the rate of sexual offence recidivism among the present sample of high risk sexual offenders would be relatively low, thus making the prediction of sexual offence recidivism very difficult. Although phallometric assessment data were included in the present study no predictions were made regarding these assessments given the inconsistent findings available in the literature.

## Method

### Participants

The offenders included in the study were taken from a larger sample of 495 consecutive admissions to the RTCSOTP for assessment and treatment between 1993 and 2004. Those retained for the study were offenders who were assigned scores of 5 or higher on the Static-99 and had been released at the time of follow-up ($n$=188). Analyses were restricted to this latter subsample of offenders, however in some analyses $n$'s vary due to missing data.

The mean age of the offenders included was 35.1 (SD = 9.2) with a range from 19 to 63 years. Offence history information, as well as scores on actuarial instruments, regarding the present sample can be found in Table 1. Based on their index offence, 101 (57.4%) of the sample were convicted of offences against victims 16-years of age or older (i.e., adult rapists), 20 (11.4%) had victims 13-15 years of age (i.e., hebephiles), 41 (23.3%) were extra-familial child molesters with victims 12 years of age or younger. Seven offenders (4.0%) had victims in more than one age group. None of the offenders had offended strictly within their own families. Over the course of their offending history 78 (51.0%) of the sample had only adult victims, while 75 (49.0%) had at least one child victim. Twenty-nine (20.0%) of the offender had only male victims, 111 (76.6%) had only female victims, and 5 (3.4%) had victims of both sexes.

Demographic Variables for Sample (N = 356)

| Demographic Variable | Mean (SD) |
|---|---|
| Age at testing | 35.1 (9.2) |
| Age at release | 36.4 (9.2) |
| Age at first arrest | 17.9 (3.5) |
| Number of nonsexual violent convictions | 1.9 (2.2) |
| Number of sexual convictions | 4.0 (5.2) |
| Number of non-violent convictions | 15.9 (13.9) |
| Total number of convictions | 21.7 (14.8) |
| Sentence length in years | 5.6 (3.7) |
| PCL-R total score | 23.7(7.5) |
| PCL-R Factor 1 | 9.0 (3.4) |
| PCL-R Factor 2 | 11.6 (4.1) |
| RRASOR score | 2.6 (1.3) |
| SORAG Score | 21.5 (10.5) |

**Table 1**

The sample scored, on average, in the 6th risk bin of the SORAG and the top of the moderate psychopathy range (Hare, 2003) on the PCL-R. Seventy-three (48.0%) of the sample scored 25 or higher, while 30 (19.7%) score 30 or higher on the PCL-R. The average RRASOR score was in the moderate risk range.

DSM diagnostic status was determined based on review of files. The sexual offenders treated at the RTCSOTP have all been assessed by psychiatrists for sentencing or parole purposes, thus this information is readily available. The most recent diagnosis was coded. In terms of diagnosis 140 (74.5%) of the offenders had a DSM diagnosis of Personality Disorder (PD); typically Antisocial (*n*=127, 67.6%) or Borderline (*n*=11, 5.9%) PD, while 92 (43.6%) met the criteria for one or more Paraphilias. Thirty-six (19.1%) met the criteria for Pedophilia, 12 (6.4%) for Sadism, and 31 (16.5%) for paraphilia NOS (Hebephilia or preferential rape). Sixty-three (33.5%) of the offenders met the criteria for both a Personality Disorder and a Paraphilia. Only 21 (11.2%) did not meet the criteria for at least one DSM based diagnosis. Other diagnoses present in the sample included psychotic disorders (*n*=13, 6.9%), mood disorders (*n*=6, 3.2%) and developmental disability (*n*=8, 4.3%). Seventy (42.7%) offenders report substantial alcohol problems and 33 (20%) report substantial drug abuse problems.

**Procedure**

As noted above, the RTCSOTP has been described elsewhere (e.g., Abracen & Looman, 2004; Looman, Abracen, Serin & Marquis, 2005; Looman et al., 2000). On intake to the RTCSOTP all offenders complete a thorough pre-treatment assessment, conducted by a Masters level or higher psychologist, including phallometric testing, as well as a risk assessment. This assessment includes the scoring of the Psychopathy Checklist - Revised (Hare, 1991, 2003).

*Coding of Risk Assessment Tools*

For offenders who entered the RTCSOTP after 1999 ($n$=53) the Static-99 was completed during the pre-treatment assessment. For the remainder of the current sample ($n$=135) the Static-99 was completed based on an archival case review, through files and the Correction Service of Canada computerized Offender Management System (OMS). The persons collecting the data had no knowledge of the recidivism rates of the offenders within the sample while completing the assessments.

Inter-rater reliability data was obtained for 26 of the Static-99 scores, using a trained, independent rater. The correlation between total scores was $r$=.90, indicating a high level of agreement. Similarly, inter-rater reliability data was obtained for 20 of the SORAG scores by the same trained rater. These were the same offenders chosen for the Static-99, however six of the offenders chosen were missing PCL-R scores. Inter-rater reliability for these 20 scores was $r$=.90.

*Static 99.* The Static 99 (Hanson and Thornton, 2000) is a ten item assessment tool, made up of the four RRASOR factors and items from the Structured Anchored Clinical Judgement, and was created for the prediction of violent sexual recidivism. The six non-RRASOR items of the Static 99 are single, index non-sexual violence, prior non-sexual violence, prior sentencing dates, convictions for non-contact sex offences, and having stranger victims.

The reported average AUC given in Beech's et al (2003) review was .71. Harris and Rice (2003) report similar ROC areas .71 and .84 with a two year, and .69 and .89 with a five year follow up period. Looman (2005) using a subset of the current sample, reported an AUC of .63 for sexual recidivism and .56 for violent (including sexual) recidivism.

*Sex Offender Risk Appraisal Guide (SORAG).* The SORAG (Quinsey, et al., 1998) was developed in Canada, from the Violent Risk Assessment Guide (VRAG), to predict violent recidivism of sex offenders. It is made up of 14 items including: lived with both natural parents until age of 16, maladjustment at school, history of alcoholism, marital status, total non-violent history score using the Cormier-Lang classification scale, total violent history score using the Cormier-Lang classification scale, number of previous sexual offences, were sex offences committed exclusively against female victims under 14 years of age, breach of previous conditional release, age at time of current offence, personality disorder diagnosis according to DSM-III criteria, Schizophrenia diagnosis according to DSM-III criteria, Plethysmograph test results, and Raw score of the Psychopathy checklist (PCL-R) (Quinsey et al, 1998).

Hanson, Morton and Harris (2003) found the SORAG to have a predictive validity slightly higher than the RRASOR and lower than the Static 99. Harris and Rice (2003) reported high ROC areas of .85 and .90 with a two year, and .95 and .83 with a 5 year follow up period. Looman (2005) using a subset of the current sample, reported an AUC of .69 for both strictly sexual and for violent (including sexual) recidivism.

*Psychopathy Checklist-Revised:* The PCL-R (Hare, 1991; 2003) was scored for each of the participants, as part of a broader risk/treatment needs assessment. As described in the manual

(Hare, 1991, 2003) the PCL-R is comprised of 20 items, thought to describe the core features of psychopathy. Participants are assigned scores of 0, indicting an absence of that feature; 1, indicating the feature may be present, or that criteria are partially met; or 2, indicating that the feature is definitely present. Because the ratings ere completed as part of the pre-treatment assessment, rather than as part of a research program, inter-rater reliability is not available for these ratings. However, as reported in the latest edition of the PCL-R manual (Hare 2003), average inter-rater reliabilities, using trained raters for the PCL-R total score for male offenders is $\alpha$=.85, for Factor 1 $\alpha$=.80 and for Factor 2 $\alpha$=.75. For Facet 1a, $\alpha$=.71, for Facet 1b, $\alpha$=.71; for Facet 2a $\alpha$=.67 and for Facet 2b $\alpha$=.64. Previous research with sexual offenders has resulted in an average total PCL-R score of 24.2, with the Factor 1 score being 9.8 and the Factor 2 score being 13.5 (Hare, 2003).

*Phallometric assessment*

Every offender who enters the RTCSOTP, for either assessment or treatment, completes a phallometric assessment. The assessment battery includes the Quinsey Female Sexual Violence assessment (Quinsey, Chaplin & Varney, 1981), the Barbaree sexual violence assessment (Barbaree, Marshall & Lanthier, 1979), the Quinsey Child Sexual Violence (Quinsey & Chaplin, 1988) and an age/gender slide assessment. Due to changes in the administration of the battery over time, while all subjects were administered the Quinsey Sexual Violence assessment and all men with offences against children were administered the Quinsey Child Sexual Violence assessment, the Child Sexual Violence assessment has only been routinely administered to rapists since 1996. Thus, data on this stimulus set is available for only 98 (52.1%) of the participants. Similarly, while the age/gender assessment was administered to all subjects while it was in use, this stimulus set was dropped from the battery in 2001 due to ethical concerns regarding the use of child pornography. Thus data for this set is available for 146 (77.7%) of the participants. Finally, the Barbaree sexual violence assessment was inconsistently used during the study period, thus data is only avialble for 89 (47.3%).

*Quinsey Female Sexual Violence assessment* (Quinsey, Chaplin & Upfold, 1984): The Quinsey Female Sexual Violence assessment consists of 14 audiotaped stimili taken from the assessment set described in Quinsey et al. (1984). The stimuli were of up to approximately one minute in duration all recorded in the same male voice: four each of consensual, sexual assault and non-sexual violence, and two neutral interactions. There are two consenting stimuli in which the woman is the initiator, two consenting stimuli in which the man is the initiator, two rape stimuli with a sexual motivation, two rape stimuli with an anger motivation, two non-sexual physical assault with an anger motivation, two non-sexual physical assault with a robbery motivation and two neutral (i.e. non-sexual, nonviolent) interactions between a man and a woman.

*Barbaree sexual violence assessment* (Barbaree, Marshall & Lanthier, 1979): The Barbaree stimulus set consisted of six audiotaped stimuli (three consenting sex and three rape), each of two minutes in duration and recorded in the same male voice. The consensual stimuli consisted of one depiction in which the woman is the initiator, one in which the man is the initiator while in the third depiction the woman participated reluctantly, but willingly. The first rape depiction involved verbal coercion, the second verbal coercion plus physical restraint, while the third rape involved a physical assault.

*Age/gender assessment:* The age/gender assessment is a slide assessment developed for the Ontario Region of the Correctional Service of Canada. This assessment set was in use at the

RTCSOTP from 1992 until 2001. This assessment consisted of 21 colored slides of single, nude individuals: three adults, three pubescent and three pre-pubescent individuals of each sex, as well as three neutral (scenery) slides.

*The Quinsey Child Sexual Violence assessment* (Quinsey & Chaplin, 1988). The child sexual violence assessment consisted of 22 audiotaped stimuli: 11 depicted an adult male interacting with a female adult or child, and 11 depicted an adult male interacting with a male adult or child. Within each group of 11 stimuli, two depicted consensual sexual interactions with an adult, two depicted sexual interactions with a passively cooperating child, two with a child who is verbally coerced into sex, two with a child who is sexually assaulted with physical violence, two with a child who is non-sexually assaulted, and one neutral (non-sexual, non-assaultive) interaction.

Phallometric testing was conducted under conditions we have described elsewhere (Looman & Marshall, 2001). Penile circumference was measured using a mercury-in-rubber strain gauge (DM Davis). Changes in circumference were detected by a Parks Electronic, Model 270 Strain Gauge Plethysmograph that outputs reading in mm stretch to a Hewlett Packard 9000 series computer. Care was taken to ensure that changes in the output of the plethysmograph were linearly related to changes in the circumference of the penis (Davidson, Malcolm, Lanthier, Barbaree & Ho, 1981). Readings were taken every 100 millisecs beginning 30 seconds before the onset of the stimulus, throughout the stimulus presentation, and ending 30 seconds after stimulus termination. Low responding subjects (i.e., less than 15% full erection) with otherwise valid profiles were retained for analyses following the recommendation by Harris, et al. (1992).

Deviance indices were calculated by dividing the averaged responses to deviant stimuli by the averaged responses to appropriate stimuli. When more than one deviant stimulus category was available, the most deviant index was used.

*Recidivism:* Recidivism data were collected from official criminal records maintained by the Royal Canadian Mounted Police (RCMP). The official Finger Print Service (FPS) sheets for each offender were obtained electronically and new convictions were coded according to the Cormier-Lang system (Quinsey, Harris, Rice & Cormier, 1998). Violent convictions were convictions listed as Group 1 offences according to the Cormier-Lang system (e.g., assault, robbery with violence). Non-violent offences were those listed under group 2 in this system (e.g., theft, possession of stolen property). New sexual offences were those offences clearly of a sexual nature according to the recorded conviction (e.g. sexual assault, gross indecency, invitation to sexual touching). Since the SVP legislation is concerned specifically with risk for future sexual re-offence, this outcome was treated as the outcome of interest.

Since all offenders in the sample had scores on the Static-99 of 5 or higher, Static-99 scores were consider a constant in the analyses. For this reason, the majority of the following analyses did not employ Static-99 scores.

## Results

Of the sample included in the analyses (*N*=188) 25 (13.3%) of the offenders re-offended sexually, while 69 (36.9%) re-offended in a violent (including sexual) manner. The follow-up period was 4.8 (sd=2.9) years (range 1 month to 11.6 years). Table 2 displays the sexual reoffence rates for each of the levels of score on the Static-99 and the corresponding expected recidivism rates based on Harris, Phenix and Hanson (2003). Note that in Harris et al., none of the offenders scored above 9, thus no expected recidivism rates for scores of 9 or 10 were available.

Observed and expected recidivism rates for Static-99 scoring groups

| Static-99 score | Observed 5-year sexual recidivism rate (*n*,%) | Expected 5-year sexual recidivism rate (%) |
|---|---|---|
| 5 (*n*=55) | 7 (12.7) | 33.0 |
| 6 (*n*=65) | 7 (10.8) | 36 |
| 7 (*n*=36) | 7 (19.4) | 43 |
| 8 (*n*=21) | 4 (19.0) | 33 |
| 9 (*n*=7) | 0 | 33 |
| 10 (*n*=4) | 0 | ---- |

**Table 2**

Since the designation of SVP requires a high level of actuarially determined risk as well as a diagnosis of a mental disorder which contributes to the offender's risk for sexual offending, recidivism rates were examined based on combinations of diagnoses. Comparing sexual recidivism rates of those offenders with and without diagnoses (see Table 3) those offenders with diagnoses of paraphilia do not re-offend at significantly higher rates than those without a paraphilia diagnosis, $X^2$(*N*=188, 1)=1.80, *p* ns. Similarly those with a personality disorder diagnosis do not re-offend at a higher rate than those without a PD diagnosis, $X^2$ (*N*=188, 1)=1.38, *p* ns. Significant differences were obtained, however, when comparing those who carry both a PD diagnosis and a diagnosis of paraphilia, with those with both diagnoses sexually re-offending at about twice the rate of other offenders (9.6% vs. 20.6%; $X^2$ (*N*=188, 1)=4.42, *p*<.035).

Offenders with only adult victims were examined separately based on diagnosis of personality disorder. There was no significant difference in terms of sexual recidivism between those with a diagnosis of Antisocial Personality Disorder (*n*=10; 15.6%) and those who had no Personality disorder (*n*=2; 16.7%), $X^2$ (*N*=78, 2)=0.38, *p* ns. None of the offenders with Borderline PD diagnoses re-offended. Offenders with only child vicims were not examined separately due to low *n*'s.

Recidivism rates by diagnosis

| Diagnosis | recidivism |
|---|---|
| No Paraphilia (*n*=106) | 11 (10.4%) |
| Paraphilia (*n*=82) | 14 (17.1%) |
| No Personality Disorder (*n*=48) | 4 (8.3%) |
| Personality Disorder (*n*=140) | 21 (15%) |
| PD + Paraphilia (*n*=63) | 13 (20.6%)[a] |
| Others (*n*=125) | 12 (9.6%)[a] |

Note: values with the same superscript are significantly different $p<.03$

**Table 3**

In order to control for the influence of time, sexual recidivism rates of the offenders with both Paraphilia and PD diagnoses were compared to the rest of the sample by means of survival analysis. Results indicated a significant difference, Wilcoxan 6.12(1), $p<.013$, (see Figure1) in terms of the proportion of men who carried both types of diagnoses re-offended compared to the remainder of the sample.

Since the PCL-R is also widely used in SVP determination, groups were formed based on PCL-R scores, with offenders scoring less than 20 (*n*=38) forming a low PCL group, offenders scoring between 20 and 30 (*n*=84) forming a moderate PCL group and offenders scoring 30 or higher (*n*=30) forming a high PCL group. The reader is reminded that all of these men scored 5 or higher on the Static-99. These three groups were compared by means of survival analysis (see Figure 2). No differences were found in group sexual recidivism rates. Five (13.2%) of the Low PCL group sexually recidivated, while 8 (9.5%) of the moderate PCL group sexually re-offended. Finally, 5 (16.7%) of the high PCL group sexually re-offended.

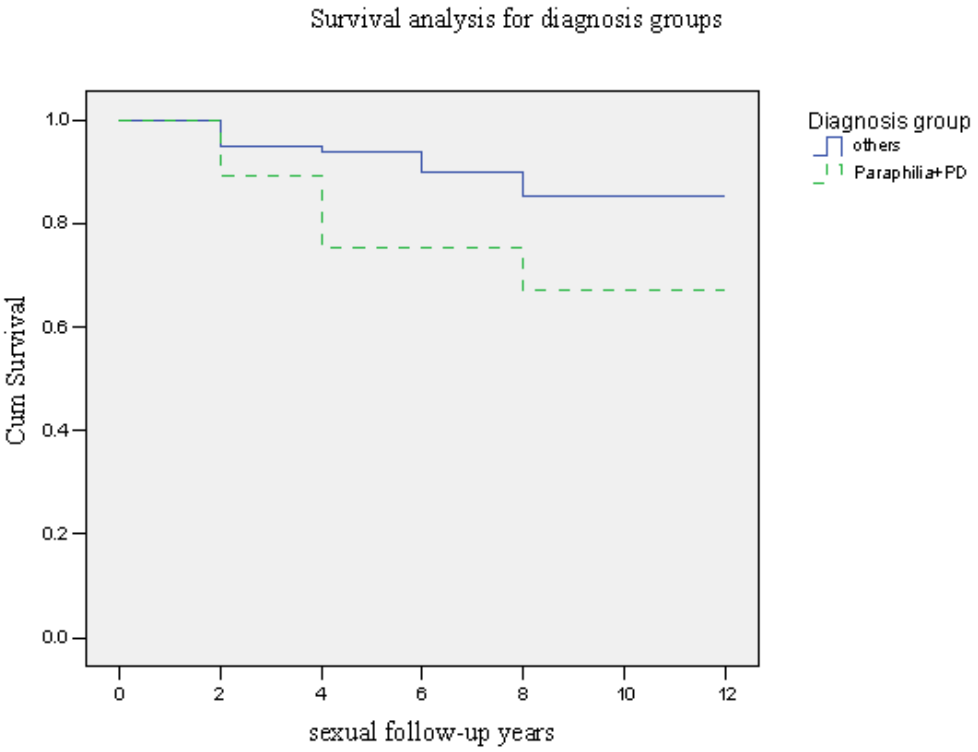


**Figure 1**

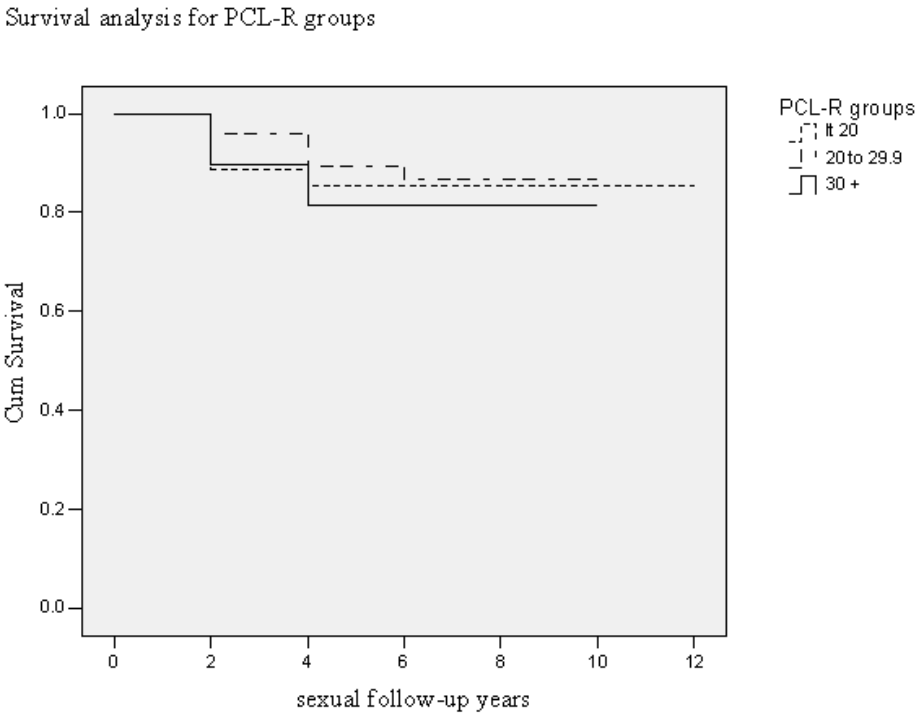


**Figure 2**

Groups were also formed based on their phallometric test results. Offenders whose overall phallometric profile indicated a preference for deviant stimuli were classified as deviant (*n*=115, 68.9% of the sample) while those with a preference for appropriate stimuli were classified

as non-deviant (n=52, 31.1% of the sample). When comparing these groups on sexual recidivism, no differences were found X(*N*=167, 1)=0.83, *p* ns. Twelve (10.4%) of the deviant offenders re-offended compared to 8 (15.4%) of the non-deviant offenders.

## Discussion

The results of the present study fail to support the validity of the criteria commonly used in SVP assessments. Although it is difficult to comment authoritatively regarding which sexual offenders are typically referred for commitment it seems plausible to suggest that many of the offenders included in the present investigation would be among those referred for commitment. Aside from the elevated STATIC-99 scores in the current sample, the majority met diagnostic criteria for one or more mental disorders that might increase their risk of offending. As well, on average the current sample had approximately 20 convictions on their record and an average of 4 sexual convictions. Nonetheless, recidivism rates were well below the standard set by SVP commitment criteria.

It might be argued that this statement is unfair in that many States (e.g., Washington) include charges or even the possibility of having committed offences which have not come to the attention of the authorities under the rubric of "more likely than not." Leaving aside the fact that the current state of the art in risk assessment appears to be the prediction of charges or convictions rather than events which are hard or impossible to detect in large scale investigations, Hanson, Morton, & Harris (2003) have suggested that the actual rates of recidivism are at least 10-15% higher than the observed rates. The authors note that this statement is based on the assumption that 60% or less of recidivists commit 5 or fewer new offences over a 20-year period and that the probability of detection is 15% per offence. Even if we were to double the observed rate of sexual recidivism in the current study (approximately 13%) the standard established by SVP commitment legislation would not, on average, be met. It is, once again, important to keep in mind that these finings do not apply to a randomly selected group of incarcerated sexual offenders. These offenders were chosen because they represented a high risk sub-sample of sexual offenders who were deemed to require a high intensity sex offender treatment program.

The current sample was selected based on their status on the Static-99 indicating a high risk to re-offend. Groups were formed within this high risk sample based on the common practice of using DSM diagnoses or psychopathy in declaring a person to meet the SVP criteria. The key finding in this study is that, despite the high actuarial risk assessment scores, none of the groupings resulted in a sub-sample of offenders who are "more likely than not to re-offend" (i.e., re-offence rates of 50% or higher).

The lower than expected recidivism rates are of interest. Given that this sample was selected based on high Static-99 scores, the fact that only 13.3% re-offended sexually over an average 5 year follow-up was not expected. As noted in the introduction, however, similarly lower than expected recidivism rates were obtained by Barbaree et al. (2003) and Looman (2006). Given that the majority of the current sample (as well as the sample reported on by Looman, 2006) was drawn from men who participated in the Ontario Region's High Intensity Sexual Offender Treatment Program, it is possible that the completion of treatment may have moderated the recidivism rate in this sample. However, 17% of the current sample did not complete the treatment program (i.e., either withdrew or were discharged). The recidivism rate of this sub-group was not different from those who completed the program (12.9% vs. 14.0% respectively). Both of these recidivism rates are well below that expected based on the sample average score on

the Static-99 (6.3; 39% over 5 years follow-up). Given the low rate of sexual offence recidivism observed in the current study one would not expect to see significant differences between treated and untreated groups unless a much longer follow-up period was employed (see Barbaree, 1997 for a discussion).

Another possible explanation for this unexpected low recidivism rate is provided by Wollert (2006), who used a Bayes's theorem analysis of Static-99 age-based recidivism rates (Hanson, 2002) to illustrate that offenders in the 35-39 age range re-offended at an average rate of 14.2%, quite similar to the current 13.3% rate for the current sample, with an average age at release of 36.4 (9.2) years.

Although Janus & Meehl (1997) have been criticized for their comments regarding accuracy of prediction with base rates below 30%, (Doren & Epperson, 2001) a sexual recidivism rate below 15% it would be unlikely that even well trained clinicians could accurately predict who might represent a lifetime risk of sexual offence recidivism of over 50% except under the most extreme of circumstances. Given that such extreme circumstances are very rare (e.g., a score of 7 or higher on the STATIC-99 perhaps with a clear statement that the offender intends to reoffend sexually) one has to question the reliability of predictions where the necessary criterion (i.e., over 50%) is so many times higher than the observed rates of sexual recidivism. It is interesting to note that the recidivism rates reported in the current study are very similar to the meta-analytic data reported by Hanson, Morton & Harris (2003) over a five-year follow-up period.

With reference to the hypothesis regarding whether psychiatric diagnoses increase the risk of recidivism the present study produced mixed results. For both rapists and child molesters DSM based diagnoses were not related to increased risk of recidivism. However, when the entire sample was used and an offender had both a diagnosis of personality disorder and a paraphilia their risk was, in fact, significantly higher (20.6% vs. 9.6%). However, again the recidivism rate did not approach that which would warrant a civil commitment decision. It must be pointed out, however, that the low recidivism rate reduced the power of the analyses, as a result of small cell sizes. However, the fact remains that in now case did the recidivism rate exceed 25%. Langstrom, Sjostedt, & Grann (2004) reported a sexual recidivism rate of 5.9% (the average follow-up time reported in the study was 5.7 years) among their sample of offenders. These authors observed that a diagnosis of personality disorder increased the odds of criminal recidivism by a magnitude of 10 times. As well, there is research suggesting that the combination of psychopathy and deviant arousal may significantly increase the risk of sexual recidivism (See Roberts, Doren & Thornton, 2002 for a review of some of this research). As well, a recent study by Hildebrand, de Ruiter and de Vogel (2004) observed that the sexual offence failure rate for psychopathic deviant offenders was 82%. Nonetheless, according to the authors many of these offenders likely received psychodynamic based treatment which would not be considered to be a current treatment methodology (Hanson, Gordon, Harris, Marques, Murphy, et al., 2002). Given that current methodologies have been shown in recent meta-analytic work (Hanson, Gordon, et al, 2002) to be more effective in the reduction of sexual recidivism this likely had an effect on the observed recidivism rates Hildebrand et al. (2004). As well, in the current study sexual deviance was not associated with sexual recidivism. Clearly, more work is needed in this area. Perhaps the interaction between psychopathy and sexual deviance offers some hope of identifying a group at particularly high risk of sexual recidivism. At present, however, it would appear too early to say that this combination of features alone would typically place a sexual offender at or above the criteria established for SVP commitment. As well, Marshall & Fernandez (2000) have noted several reservations regarding the psychometric properties of phallometric assessment which have not been addressed to this point in time.

Over the last 15 years there have been significant advances in our ability to reliably distinguish high risk groups of offenders from lower risk groups (see Quinsey, Harris, Rice & Cormier, 1998 for a review of some of these measures). However, even with these advances it appears that the specificity (i.e., true positive rates) of the best of the actuarial instruments designed to assess risk for sexual recidivism is only about 30% for offenders in the age range of 35-50 years; and for those over 50 years of age it is even lower. Thus, to expect mental health professionals to reliably identify individuals who are at several times the expected rate of sexual recidivism (as is the expectation in SVP assessments) may be doing a disservice to the profession. Inadvertently, forensic health professionals may be returning to a point in history where predictions of risk are at chance levels. It's not that the technologies haven't improved, they clearly have. However, the requirement that a very small sub-group of high risk sexual offenders be reliably distinguished from "typical" high-risk offenders may be beyond the current technologies to predict. This may be especially true in the case of sex offenders.

The data included in the present study as well as in Looman (2006) and Barbaree et al. (2003) have shown that the STATIC-99 significantly over-predicts risk among treated sexual offenders. The question as to whether individuals should be committed for extended periods of time, following the completion of their sentence, based on untested scientific principles needs to be raised on ethical grounds. At present, there does not appear to be scientific merit in assuming that psychiatric diagnoses reliably add to the prediction of risk among sexual offenders whereby those offenders who are "more likely than not" can be reliably distinguished from those who do not meet this criterion. Given that commitment legislation appears to rest on this assumption, one has to question the legitimacy of such legislation from a scientific perspective.

## References

Barbaree, H. (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., Blanchard, R, & Langton, C. M. (2003). The development of sexual aggression through the life span: The effect of age on sexual arousal and recidivism among sex offenders. *Annals of the New York Academy of Sciences, 989*, 59-71.

Barbaree, H. E., Marshall, W. L. and Lanthier, R. D. (1979). Deviant sexual arousal in rapists. *Behaviour Research and Therapy, 17,* 215-222.

Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2003; October). The evaluation of sex offender treatment efficacy using samples stratified by levels of actuarial risk. Paper presented at the annual meeting of the Association for the Treatment of Sexual Abusers (ATSA), St. Louis, Missouri.

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*, 490-521.

Beech, Anthony R., Fisher, Dawn D., & Thornton, David. (2003). Risk assessment of sex offenders. *Professional Psychology: Research and Practice, 34*, 339 – 352.

Davidson, P. R., Malcolm, P. B., Lanthier, R. D., Barbaree, H. E. & Ho, T. P. (1981). Penile response measurement: Operating characteristics of the Parks plethysmograph. *Behavioral Assessment, 3,* 137-143.

Doren, D. (1998). Recidivism base rates, predictions of sex offender recidivism, and the "sexual predator" commitment laws. *Behavioral Sciences and the Law, 16,* 97-114.

Doren. D. M. (2002). *Evaluating sex offenders: A manual for civil commitments and beyond.* London: Sage Publications

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.

Fitch, W. L. (2003). Sexual offender commitment in the United States: Legislative and policy concerns. *Annals of the New York Academy of Sciences, 989,* 489-501.

Hanson, R. K. (2002). Recidism and age. *Journal of Interpersonal Violence, 17,* 1046-1062.

Hanson, R. Karl, 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. Karl, & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis. Ottawa, Ontario: Public safety and Emergency Preparedness Canada (Cat. No.: PS3-1/2004-2E-PDF).

Hanson, R. Karl., & 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. Karl, Morton, K. E., & Harris, A. J. R. (2003).Sexual offender recidivism risk: What we know and what we need to know. *Annals of the New York Academy of Sciences, 989,* 154-166.

Hanson, R. Karl. & Thornton, D. (2000). Improving risk assessment for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24,* 119-136.

Hare, R (1991). *Manual for the Revised Psychopathy Checklist.* Toronto, Canada: Multi-Health Systems.

Hare, R (2003). *Manual for the Revised Psychopathy Checklist: 2nd edition.* Toronto, Canada: Multi-Health Systems.

Harris, A. J. R., Phenix, A., Hanson, R. K., & Thornton, D. (2003). Static-99 coding rules: Revised 2003. Ottawa : Department of the Solicitor General of Canada .

Harris, G. T. & Rice, M. E. (2003). Actuarial assessment of risk among sex offenders. *Annals of the New York Academy of Science, 989,* 198-210.

Hildebrand, M., de Ruiter, C., & de Vogel, V. (2004). Psychopathy and sexual deviance in treated rapists : Association with (sexual) recidivism. *Sexual abuse: A Journal of Research and Treatment, 16,* 1-24.

Jackson, R. L., Rogers, R., & Shuman, D. W. (2004). The adequacy and accuracy of sexually violent predator evaluations: Contextualized risk assessment in clinical practice. *International Journal of Forensic Mental Health, 3,* 115-129.

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,* 33-64

Långström, N; Sjöstedt, G. & Grann, M. (2004) Psychiatric Disorders and Recidivism in Sexual Offenders. Sexual Abuse: *Journal of Research and Treatment. 16,* 139-150.

Looman, J. (2006). Comparison of two risk assessment instruments for sexual offenders. In press, *Sexual abuse: A Journal of research and Treatment.*

Looman, J., Abracen, J., & Nicholaichuk, T. P. (2000). Recidivism among treated sexual offenders and matched controls: Data from the Regional Treatment Centre (Ontario). *Journal of Interpersonal Violence, 15,* 279-290.

Looman, J., Abracen, J., Serin, R., Marquis, P. (2005). Psychopathy, Treatment Change and Recidivism in High Risk High Need Sexual Offenders. *Journal of Interpersonal Violence, 20,* 549-568.

Looman, J. & Marshall, W. L. (2001). Phallometric assessments designed to detect arousal to children: the responses of rapists and child molesters. *Sexual Abuse: A Journal of Research and Treatment, 13,* 3-13.

Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology, 1,* 117-146.

Mailloux, D., L., Abracen, J., Serin, R., Cousineau, C., Malcolm, B., & Looman, J. (2003). Dosage of treatment to sexual offenders: Are we overprescribing? *International Journal of Offender Therapy and Comparative Criminology, 47,* 171-184.

Marshall, W. L., & Fernandez, Y. M. (2000). Phallometric testing with sexual offenders: Limits to its value. *Clinical psychology Review, 20,* 807-822.

Marshall, W. L. (2006). Appraising treatment outcome with sexual offenders. In W. L. Marshall, Y. M. Fernandez, L. E. Marshall & G. A. Serran (Eds.), *Sexual offender treatment: Controversial issues.* England: John Wiley and Sons, Ltd.

Rice, M. E., & Harris, G. T. (2003). The size and sign of treatment effects in sex offender therapy. *Annals of the New York Academy of Sciences, 989,* 428-440.

Roberts, Caton F., Doren, D. M., & Thornton, D. (2002). Dimensions associated with assessments of sex offender recidivism risk. *Criminal Justice and Behaviour, 29,* 569 – 589.

Skeem, J. L., Miller, J. D., Mulvey, E., Tieman, J., & Monahan, J. (2005). Using a five-factor lens to explore the relation between personality traits and violence in psychiatric patients. *Journal of Consulting and Clinical Psychology, 73,* 454-465.

Quinsey, V. L. and Chaplin, T. C. (1988). Penile response of child molesters and normals to descriptions of encounters with children involving sex and violence. *Journal of Interpersonal Violence, 3,* 259-274.

Quinsey, V. L., Chaplin, T. C. and Varney, G. (1981). A comparison of rapists and non-sex offenders sexual preferences for mutually consenting sex, rape and physical abuse of women. *Behavioral Assessment, 3,* 137-135.

Quinsey, V. L., Chaplin, T. C. & Varney, G. (1981). A comparison of rapists and non-sex offenders sexual preferences for mutually consenting sex, rape and physical abuse of women. *Behavioral Assessment, 3,* 137-135.

Quinsey, V. L., Harris, G. T., Rice, M. E. & Cormier, C. A. (1998). *Violent Offenders: Appraising and Managing Risk.* Washington: American Psychological Association.

Quinsey, V. L., Harris, G. T., Rice, M. E. & Cormier, C. A. (2006). *Violent Offenders: Appraising and Managing Risk (2nd Ed.).* Washington: American Psychological Association.

Wollert, R. (2001). Adversarial forum: An analysis of the argument that clinicians under-predict sexual violence in civil commitment cases. *Behavioral Sciences and The Law, 19,* 171-184.

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,* 56-85

Accepted: March 22, 2006