# Failure of Static-99 and SORAG to predict relevant reoffense categories in relevant sexual offender subtypes: A prospective study

**Reinhard Eher[1,2], Martin Rettenberger[1], Frank Schilling[1] & Friedemann Pfäfflin[2]**
[1]Evaluation Centre for Violent and Sexual Offenders, Vienna, Austria; [2]Forensic Psychotherapy, Ulm University, Ulm, Germany

## Abstract

*An Austrian sample of 275 sexual offenders was assessed clinically as well as with actuarial tools including the Static-99 and the Sexual Offender Risk Appraisal Guide (SORAG) before being released from prison. After an average time of 3.6 years at risk, their relapse rates were correlated with their previous scores on these two instruments.*
*The results demonstrate that the Static-99 and the SORAG, although predicting certain types of relapses quite well, are much better in predicting the non-relapse than the relapse: Less than 50% of the high scorers on both instruments relapsed during the follow-up time of 3.6 years at risk and were again convicted for a new sexual or violent offence. For practical purposes, e. g. for the clinician and the judge being in charge of deciding about the release of an individual offender, these overall results as well as more specific results for certain types of relapses with low base rates are discussed.*

*Key words: Static-99, SORAG, risk assessment, sex offenders, relapse, validity*

## Introduction

The prediction of relapses may be primarily based on (1) individual clinical assessments or on (2) group-statistics evaluating actuarial data drawn from random samples of re-offenders, or, as is usually the case, a combination of both (Dahle et al. 2007). Standardized actuarial risk assessment tools like the Static-99 (Hanson & Thornton 2000) and the SORAG (Sexual Offender Risk Appraisal Guide; Quinsey et al. 2006) reflect mostly static, unchangeable factors and do not take into account dynamic, potentially changing factors. Their advantage is, however, easy handling and producing a total score allowing calculating the risk of certain types of re-offenses. Presently, the Static-99 is the most widely used actuarial risk assessment tool for sex offenders. In spite of high validity, their predictions rates are, however, moderate (Helmus & Hanson 2007).

In their meta-analysis, Hanson & Morton-Bourgon (2007) demonstrated the superiority of actuarial risk assessments to clinical assessments in predicting sex offenders relapses with a new sex offense (Cohen s *d* 0.85 vs. 0.40). The same holds true for the prediction of a violent relapse (including sexually motivated violent offenses) (Cohen s *d* 0.85 vs. 0.31).

The Static-99 and the SORAG are empirically validated actuarial prognostic instruments (Hanson & Morton-Bourgon 2007). Their items were selected based on statistically validated correlations between individual items and relapse. Scoring is standardized. High reliability, constant observation periods, and few missing data increase the validity of such instruments (Harris & Rice 2003).

There is, however, also some controversy about actuarial risk assessments (Craig et al. 2006). Understanding and managing an offender by operating with just one total score may be of limited value (Hart 1998). Various subgroups of offenders might not have the same risk rates. Neither the Static-99 nor the SORAG predicted the sexual motivated relapse in a group of rapists (Harris et al. 2003b). Craig et al. (2003) examined 12 actuarial prognostic instruments and found that 10 out of 12 were better in predicting general than sexual recidivism.

Base rates play an outstanding role, and they vary between subgroups of sexual offenders especially so between the subgroups of rapists and the extra- and intra-familiar child abuser (Hanson 2002). Also, the longer the follow-up time, the higher are the base rates. On the other hand, the lower the base rates are, the higher is the risk to overestimate the relapse rates and to produce false positives. With a base rate of 4%, the Static-99 would overestimate the risk in 49 out of 50 cases (Hood et al. 2002).

## Recent results

Validation studies of the Static-99 (Harris et al. 2003a) and the SORAG (Quinsey et al. 1998) speak for the predictive validity of these instruments. Using ROC-analysis (**R**eceiver **O**perating **C**haracteristics), the results are quantified in AUC-values (**a**rea **u**nder the **c**urve) (Dahle et al. 2007, Nedopil 2005). AUC-values > .71 speak for a high predictive validity, AUC-values between .63 and .70 for a moderate predictive validity and AUC-values < .63 for a weak predictive validity. As measured by AUC-values, the predictive validity for the prediction of subsequent sexual (Table 1) and violent (Table 2) relapses was found to be moderate up to high. Referring to our own data the predictive validity, as judged from a clinical perspective, often is weak. Only the minority of studies distinguishes between contact/hands-on offenses and non-contact/hands-off offenses.

In a previous retrospective study based on 1208 follow-ups of sexual offenders from England and Canada the Static-99 reached an AUC of .71 for sexually motivated relapse (there was no differentiation between contact and non-contact offenses). For violent relapse (including sexual hands-on offenses) an AUC of .69 was found (Hanson & Thornton 2000). After 15 years follow-up even in the highest risk group of the Static-99 (score of 6 or 6+ = 12% of the sample) just about 50% of the offenders had relapsed. The authors, nevertheless, concluded that the Static-99 allowed the identification of a significant subgroup of sexual offenders with a relapse risk of more than 50% (Tables 1 and 2).

Barbaree et al. (2001) compared a number of prognostic instruments by evaluating 215 sexual offenders with a time at risk of 4.5 years. They concluded that the SORAG came close to the Static-99 in predicting a sexual relapse (AUC .70 for both instruments) as well as in predicting other serious offenses (violent offense including sexual hands-on offense) (SORAG: AUC .73; Static-99: AUC .70). The study did not differentiate between sexual child abuse and rape (Tables 1 and 2).

Bartosh et al. (2003) examined re-imprisonment rates among 186 sexual offenders with a follow-up time of up to 5.5 years. They determined the validity indices for four prognostic instruments including the Static-99 and a modified form of the SORAG. They distinguished between rapists, intra- and extra-familiar child abusers and non-contact offenders. Taking the total sample, the Static-99 predicted re-imprisonment as a result of a sexual offense only moderately (AUC .64), the SORAG not at all (AUC .58). The predictive validity of the Static-99 for re-imprisonment due to a violent offense (sexual and non-sexual) was high (AUC .73); using the SORAG it was just moderate (AUC .64). In the group of the extra-familiar child abusers the Static-99 (AUC .65) and the SORAG (AUC .69) showed moderate validity. In predicting re-imprisonment due to a violent offense (sexual

and non-sexual), high AUC-values were found (AUC .94 and .93). For intra-familiar offenders findings were similar. For rapists there could not be shown any significant predictions (Tables 1 and 2).

Similarly, Nunes et al. (2002) did not find differences in the predictive validity between the Static-99 and the SORAG, but they did not distinguish between the various sexual offender subgroups. The AUC-values for a sexual relapse (no differentiation between contact and non-contact offenses) was reported with .70 (Static-99) and .65 (SORAG). As regards the prediction of a violent offenses (including a sexual offense) there was no difference between the two instruments (AUC .69) (Tables 1 and 2).

In a retrospective study Harris et al. (2003b) compared the Static-99, the SORAG and other risk assessment tools by using a sample of 396 male sexual offenders. The follow-up time at risk averaged approximately 5 years. In the total sample the Static-99 showed an AUC of .62 (low) for sexual relapse (only contact-offenses, sexual relapse was included as a subgroup of violent offenses). The values for the subgroups of rapists (AUC .59) and child abusers (AUC .65) also ranked low. Violent recidivism could be predicted only moderately or not at all (total sample AUC .63; subgroup of rapists AUC .58; subgroup of child abusers AUC .64). In terms of both relapse criterions (sexual and violent) the SORAG scored better than the Static-99. As regards the total sample, predictive validity for a sexual relapse was moderate (AUC .66), for a violent offense it was high (AUC .73). The SORAG was better than the Static-99 in predicting relapses in the subgroups of rapists and sexual child abusers, and in predicting non-specific violent relapses as compared to sexual relapses (sexual relapse: rapist AUC .62, child abusers AUC .70; violent relapse: rapists AUC .70, sexual child abusers AUC .72) (Tables 1 and 2).

In a Belgian comparative study with an average follow-up time of 4.2 years (Ducro & Pham 2006) better values were found for the Static-99 (AUC .70 and AUC .71) than for the SORAG (AUC .65 and AUC .64) regarding sexually motivated relapses in the groups of sexual child abusers and rapists. In the subgroup of rapists, the SORAG predicted the violent relapse better than the Static-99 (AUC .77 and AUC .65) (Tables 1 and 2).

**Table 1: Validity studies for the Static-99 and the SORAG for sexual relapses (AUC-values)**

| | Sexual relapse | | | | | |
| | Static-99 | | | SORAG | | |
| | Total N | Sexual child abusers | Rapists | Total N | Sexual child abusers | Rapists |
|---|---|---|---|---|---|---|
| HANSON & THORNTON, 2000+ | .71 | .72 | .71 | - | - | - |
| BARBAREE et al., 2001** | .70 | - | - | .70 | - | - |
| NUNES et al., 2002+ | .70 | - | - | .65 | - | - |
| BARTOSH et al., 2003+ | .64 | .65/.74* | .71 | .58 | .70/.72* | .71 |
| HARRIS et al., 2003b** | .62 | .65 | .59 | .66 | .70 | .62 |
| DUCRO & PHAM, 2006+ | .66 | .70 | .71 | .64 | .65 | .64 |

AUC-value > .71 = high, AUC-value > .63 = moderate, AUC-value < .63 = weak
* The authors differentiate between intra- and extra-familiar sexual child abuse
** hands-on offenses only
+ hands-on as well as hands-off

**Table 2: Validity studies for the Static-99 and the SORAG for violent relapses (AUC-values)**

| | Violent relapse | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Static-99 | | | SORAG | | |
| | Total N | Sexual child abusers | Rapists | Total N | Sexual child abusers | Rapists |
| HANSON & THORNTON, 2000 | .69 | .71 | .69 | - | - | - |
| BARBAREE et al., 2001 | .70 | - | - | .73 | - | - |
| NUNES et al., 2002 | .69 | - | - | .69 | - | - |
| BARTOSH et al., 2003 | .73 | .94/.86* | .53 | .64 | .93/.82* | .57 |
| HARRIS et al., 2003b | .63 | .64 | .58 | .73 | .72 | .70 |
| DUCRO & PHAM, 2006 | .68 | .72 | .65 | .72 | .70 | .77 |

AUC-value > .71 = high, AUC-value > .63 = moderate, AUC-value < .63 = weak
* The authors differentiate between intra- and extra-familiar sexual child abuse

There are also some validation studies of the Static-99 from German-speaking countries based on samples of sexual offenders. Stadtland et al. (2006) found an AUC of .72 for the violent relapse; Rettenberger & Eher (2006b) an AUC of .74 for the sexually motivated and an AUC of .76 for the violent relapse. The validation of the German translation of the SORAG also resulted in AUC-values of .71 or higher for the sexual and violent relapses (Rettenberger & Eher 2007a,b), except for the sexually motivated relapse of rapists.

# Aims of the Study

Once validations of the German translations of the Static-99 and the SORAG had been established (Rettenberger & Eher 2006 a,b) one could proceed to the next step: the quantitative analysis of actual relapses. Even though the statistical analysis is very important for the scientist, and even though the statistical values of the Static-99 and the SORAG are moderate or even high, the judge and the clinician want to know how to apply them to the individual case and how well they predict relapses individually. Against the background of limited resources the answer to these questions is of vital importance for the allocation of resources within the risk management of prisons. One does not want to produce false positives (Hood et al. 2002), and one wants to apply therapeutic activities of the staff as effectively as possible.

# Methods

## The prognostic tools Static-99 and SORAG

The Static-99 is a synthesis of two former prognostic instruments, the Rapid Risk Assessment for Sex Offender Recidivism (RRASOR, Hanson 1997) and the Structured Anchored Clinical Judgement　Minimum (SACJ　Min; Grubin 1998). All criteria were extracted from earlier empirical relapse studies (Hanson & Bussiere 1998). The short form of the SACJ and the RRASOR were combined into the 10 items of the Static-99. A complex revision was published in 2003 (Harris et al. 2003a). The SORAG is a modification of the Violence Risk Appraisal Guide (VRAG; Quinsey et al. 1998) developed for the purpose to predict violent (including sexually motivated violent) relapses of released offenders. The SORAG had been validated on a small sample of only 178 released sexual offenders. 14 of the most predictive variables were included into the SORAG in a weighted manner.

In the years 2002　2007, German versions of Static-99 and the SORAG were used in the Austrian prison system in the context of expert opinions at the Evaluation Centre for Violent and Sexual Offenders　(BEST) an institution of the Federal Department of Justice for the assessment of violent and sexual offenders. During this period, 714 convicted sexual offenders were assessed, 629 completed the Static-99 and 614 the SORAG. Good inter-rater reliability and predictive validity had been found for both instruments in earlier studies (Rettenberger & Eher 2006a,b, 2007a,b).

# Sample

Since 2002, all Austrian sexual offenders with a prison sentence have to be reported to the BEST. From the beginning of 2002 until September 2007 about 60% of all sex offenders imprisoned in Austria were thoroughly assessed clinically and under criminological aspects (Eher et al. 2006). For 275 of the 714 sex offenders information was collected about relapses in terms of criminal record advice of the Federal Department of the Interior. The average follow-up time at risk was 3.6 years (2.3　5.9 years). The average age at the time of release was 40.9 years (16.1　71.4). 136 offenders showed an index offense committed on minors and 132 an index offense committed on adults. 7 offenders had not committed a hands-on offense. The overall relapse rate was 30.2%. 17.5% were re-incarcerated. 6.2% had relapsed with a sexual offense, but only 3.3% with a hands-on offense. New violent offenses including sexual hands-on offenses were found in 14.6% of the cases, incarceration as a result of suchlike offenses was imposed in 8.36% (see Table 3).

**Table 3: Relapse rates of total N and differentiated according to the subgroups of offenses against minors and adults after 3.6 years at risk**

|  | Total N (N=275) | Index offense with minors (n=136) | Index offence with adults (n=132) |
|---|---|---|---|
| Relapse in general | 83 (30.2%) | 30 (22.1%) | 48 (36.4%) |
| Relapse in general and incarceration | 48 (17.5%) | 18 (13.2%) | 28 (21.2%) |
| Relapse "violent" (not sexually motivated) | 40 (14.6% | 8 (5.9%) | 31 (23.5%) |
| Relapse "violent" (not sexually motivated) and incarceration | 24 (8.7%) | 3 (2.2%) | 20 (15.2%) |

| | | | |
|---|---|---|---|
| Relapse "sexual" | 17 (6.2%) | 12 (8.8%) | 3 (2.3%) |
| Relapse "sexual" and incarceration | 16 (5.8) | 12 (8.8%) | 3 (2.3%) |
| Relapse "sexual" with hands-on | 9 (3.3%) | 5 (3.7%) | 3 (2.3%) |
| Relapse "sexual" with hands-on and incarceration | 9 (3.3%) | 5 (3.7%) | 3 (2.3%) |
| Relapse "sexual" with hands-off | 8 (2.9%) | 7 (5.2%) | 0 |
| Relapse "sexual" with hands-off and incarceration | 7 (2.6%) | 7 (5.2%) | 0 |
| Relapse „violent" ("sexual" with hands-on + not sexually motivated violence) | 40 (14.6%) | 12 (8.8%) | 26 (19.7%) |
| Relapse „violent" and incarceration | 23 (8.4%) | 6 (4.4%) | 16 (12.1%) |

7 of the total of N = 275 could not be assigned to one of the subgroups

# Statistics

ROC-analyses were performed as validity indices regarding the prognostic value of the instruments. ROC-analysis is a procedure of logistic regression with the advantage of being largely independent of base rates. AUC-values correspond both, conceptually as well as numerically, with effect sizes (Harris et al. 2003b). Basically, the AUC-value indicates the probability to which a re-offender shows a higher score as compared to someone not re-offending. AUC-values above .71 are considered as high, values above .63 as moderate, and those below .63 as low (Dahle et al. 2007).

# Results

The individual results in terms of AUC-values are summarized in Table 4. Taking the sample as a whole, both instruments demonstrated a high prognostic validity for almost all relapse categories. Values were mostly above those previously described in the literature (see Tables 1 and 2). Only in the category of sexual hands-on offenses both instruments showed only a moderate prognostic validity: for the SORAG, AUC-values were not significant from the very beginning, and neither were they significant for the Static-99 (after correction of type-I-error). In none of the relapse categories significant differences were found between the Static-99 and the SORAG (paired comparison of AUC-values).

In the group of sexual child abusers high prediction values were found, especially so as regards sexual relapses (AUC > 0.8). The SORAG even reached high values in predicting sexual hands-on offenses of sexual child abusers. Both instruments proved to be good in predicting violent relapses. Not satisfactorily predictable was the sexual relapse for rapists. Basically higher prognostic values were found in the group of child abusers (see Table 4).

**Table 4: Validity indices of the Static-99 and the SORAG (AUC-values)**

| | AUC-values for the total N = 275 | AUC-values for the subgroup of | AUC-values for the subgroup of |
|---|---|---|---|

|  | Static-99 | SORAG | child abusers | | rapists | |
|---|---|---|---|---|---|---|
|  |  |  | Static-99 | SORAG | Static-99 | SORAG |
| Relapse in general | .70*** | .77*** | .69** | .78*** | .67** | .71*** |
| Relapse in general plus incarceration | .73*** | .78*** | .75** | .83*** | .69** | .71** |
| Relapse violent | .69** | .75*** | .64 | .66 | .67** | .71*** |
| Relapse violent plus incarceration | .70** | .78*** | .60 | .74 | .67* | .72*** |
| Relapse sexual | .74*** | .71** | .81*** | .82*** | .54 | .65 |
| Relapse sexual plus incarceration | .72*** | .71* | .81*** | .82*** | .54 | .65 |
| Relapse sexual with hands-on | .68* | .69 | .74 | .80* | .54 | .65 |
| Relapse sexual with hands-on plus incarceration | .68* | .69 | .74 | .80* | .54 | .65 |
| Relapse violent (incl. sexual with hands-on) | .71*** | .74*** | .73* | .74** | .65* | .67** |
| Relapse violent plus incarceration | .71** | .76*** | .74 | .83* | .66* | .67* |

AUC-value > .71 = high, AUC-value >.63 = moderate, AUC-value > .63 = weak

Tables 5  8 list the relapses according to the risk categories of the instruments.  Indicated are the criminologically relevant relapse categories ( relapse with sex offence ,  relapse with sexual hands-on offense ,  relapse with violent offense ,  re-imprisment because of a violent offense ).

# Static-99 risk categories and relapse incidents in sexual child abusers

Table 5 shows the respective Static-99 values for the group of child abusers. The highest risk category shows 33% relapses with new sexual offenses whereas less relapses were found in the first three risk categories. Only 14% in the highest risk group relapsed with a sexual hands-on offense. Relapses in the category  violent  increased steadily with Static-99 risk categories, whereas even in the highest risk group  only  19% were re-convicted and  only  about 10% were re-incarcerated.

**Table 5: Relapses in the subgroup of sexual child abusers (N = 133) according to the Static-99-categories in absolute and relative numbers (missing data of 3 individuals)**

| Static-99-value | Relapse sexual | Relapse sexual, hands-on | Relapse violent (incl. sexual, hands-on) | Relapse violent and incarceration |
|---|---|---|---|---|
| static 0,1 (n=53) | 0 | 0 | 1 (1.89%) | 0 |

| static 2,3 (n=43) | 5 (11.63%) | 2 (4.65%) | 4 (9.30%) | 3 (6.98%) |
| static 4,5 (n=16) | 0 | 0 | 2 (12.5%) | 0 |
| static 6,6+ (n=21) | 7 (33.33%) | 3 (14.29%) | 4 (19.05%) | 2 (9.52%) |

## Static-99-risk categories and relapse incidents in rapists

There was no correlation between the risk category according to the Static-99 and sexual relapses. However, in terms of a violent offenses, a continuous rise in relapse incidents parallel to higher Static-99-risk categories was revealed. In the highest Static-99 risk group the relapse rate leading to incarceration was 23% (see Table 6).

**Table 6: Relapses in the subgroup of rapists (N = 128) according to the Static-99-categories in absolute and relative numbers**
**(missing data of 4 individuals)**

| Static-99-value | Relapse sexual | Relapse sexual, hands-on | Relapse violent (incl. sexual, hands-on) | Relapse violent and incarceration |
|---|---|---|---|---|
| static 0,1 (n=12) | 0 | 0 | 1 (8.33%) | 0 |
| static 2,3 (n=43) | 1 (2.33%) | 1 (2.33%) | 5 (11.62%) | 4 (9.3%) |
| static 4,5 (n=47) | 2 (4.26%) | 2 (4.26%) | 12 (25.53%) | 6 (12.77%) |
| static 6,6+ (n=26) | 0 | 0 | 8 (30.77%) | 6 (23.08%) |

## SORAG-risk categories and relapse incidents in sexual child abusers

The SORAG revealed similar results as the Static-99 in the group of child abusers (see Table 7), i.e. an increase in the number of relapses in correlation to the risk group. The highest risk group in the SORAG confirmed after all a probability of 30-40% regarding a sexual relapse. This correlation was less distinct in the group of sexual hands-on offenses. In the second highest risk category only one out of five offenders actually re-offended. No single one met the criteria of the highest risk group. The correlation between risk category and relapse with a violent offense was again similar to the findings with the Static-99, i.e. a steady relapse incident increase in correlation to the risk category. In the highest risk category, however, far more offenders did not re-offend than vice versa.

**Table 7: Relapses in the subgroup of child abusers (N = 130) according to SORAG-categories in absolute and relative numbers**
**(missing data of 6 individuals)**

|  | Relapse sexual | Relapse sexual, hands-on | Relapse violent (incl. sexual, hands-on) | Relapse violent and incarceration |
|---|---|---|---|---|
| SORAG 1 (n=22) | 0 | 0 | 1 (4.55%) | 0 |
| SORAG 2 (n=33) | 1 (3.03%) | 0 | 1 (3.03%) | 0 |
| SORAG 3 (n=22) | 1 (4.55%) | 1 (4.55%) | 0 | 0 |
| SORAG 4 (n=15) | 1 (6.67%) | 1 (6.67%) | 1 (6.67%) | 1 (6.67%) |
| SORAG 5 (n=13) | 2 (15,38%) | 1 (7.69%) | 2 (15.38%) | 1 (7.69%) |
| SORAG 6 (n=11) | 2 (18,18%) | 0 | 1 (9.09%) | 1 (9.09%) |
| SORAG 7 (n=9) | 3 (33,33%) | 1 (11.11%) | 3 (33.33%) | 1 (11.11%) |
| SORAG 8 (n=5) | 2 (40%) | 1 (20%) | 2 (40%) | 1 (20%) |
| SORAG 9 (n=0) | - | - | - | - |

# SORAG-risk categories and relapse incidents in rapists

In the group of rapists the SORAG showed similar results as the Static-99. No correlation was found between the risk category and the actual sexually motivated relapse, however a clear correlation was shown between risk category and violent relapse: more than 40% in the two highest risk categories of the SORAG relapsed with a violent offence, and in both categories nearly 30% of the offenders were again incarcerated (see Table 8).

**Table 8: Relapses in the subgroup of rapists (N = 123) according to SORAG-category in absolute and relative numbers (missing data of 9 individuals)**

|  | Relapse sexual | Relapse sexual, hands-on | Relapse violent (incl. sexual, hands-on) | Relapse violent and incarceration |
|---|---|---|---|---|
| SORAG 1 (n=9) | 0 | 0 | 0 | 0 |
| SORAG 2 (n=13) | 0 | 0 | 0 | 0 |
| SORAG 3 (n=20) | 0 | 0 | 3 (15%) | 2 (10%) |
| SORAG 4 (n=16) | 1 (6.25%) | 1 (6.25%) | 4 (25%) | 2 (12.50%) |
| SORAG 5 (n=15) | 2 (13.33%) | 1 (13.33%) | 5 (33.33%) | 3 (20.0%) |
| SORAG 6 (n=16) | 0 | 0 | 1 (6.25%) | 1 (6.25%) |
| SORAG 7 (n=16) | 0 | 0 | 3 (18.75%) | 3 (18.75%) |
| SORAG 8 (n=11) | 0 | 0 | 5 (45.45%) | 3 (27.27%) |
| SORAG 9 (n=7) | 0 | 0 | 3 (42.86%) | 2 (28.57%) |

# Discussion

The most important conclusion is the obvious discrepancy between the high validity indices of the Static-99 and the SORAG on the one hand which, by definition, suggest that relapses can be predicted very well with these instruments, and, on the other hand, their limited practical benefit when it comes to identifying recidivists with high enough certainty. This can be concluded from the fact that in the highest risk category often far below 50% of the offenders finally were re-incarcerated.

Although the (statistically verified) validity of actuarial prognostic instruments has been demonstrated by meta-analysis (Hanson et al. 2007), their practical value has to be discussed when the clinician or the judge have to decide in individual cases. In the primary sample of the Static-99 even in the high risk group (6,6+) a relapse risk for another sexual offense of only just above 50% was found    and this only after 15 years at risk. Taking the primary sample of the SORAG for comparison, the results of this instrument seem to allow better predictions at first glance, at least for violent  relapses (sexual hands-on and not sexually motivated violent offenses) of sexual offenders: in the highest risk group 100% relapsed after 7    10 years. It has, however, to be taken into account that the primary sample of the SORAG was very small, and only just 1%    3 offenders    was represented in its highest risk category. In a replication study including 258 offenders (Nunes et al. 2002) as well as in our study not a single offender was found in the highest risk category. Even in the second highest risk category of the SORAG only 8%    2 offenders    were found, and one (50%) of them relapsed with a violent offence (Nunes et al. 2002). Taking this into account, one has to conclude that even high and highest validity scores of such instruments do not predict who will actually re-offend and who will not.

Depending on the definition of cut-offs for dangerousness, prognostic instruments produce many false positives. Uncritical use of these instruments causes a systematic overestimation of the real (verifiable) risk. For example, taking 5 as the cut-off in the Static-99 after an average follow-up time of about 8 years at risk, 46% will re-offend in the 5, 5+ group with a sexual and/or non-sexual violent offense, but 54% will not (Stadtland et al. 2006). In the original sample (Harris et al. 2003a) 61% out of the highest risk group (6,6+) had not relapsed after 5 years; after 10 years 55%, and after 15 years 48% had not relapsed with a new sexual offense (Harris et al. 2003a). In the study of Nunes et al. (2002) the rate of false positives for the same cut-off (6,6+) and the same relapse category was 70% after approximately 7 years at risk. Comparably differentiated data do not exist for the SORAG, but Nunes et al. (2002) reported a 50% relapse rate for sexual offenses in the second highest risk category and only a 25% relapse rate in the third highest risk category.

In our study satisfactory validity indices were found, especially so for certain subgroups and relapse categories. Some of them were higher than previously published ones, e.g. for sexual relapses of child abusers. The values in the category hands-on sexual offenses were, however, much less impressing. The predictive performance of the Static-99 for this category was only moderate with an AUC of .68; and with the SORAG it could not be predicted at all in a significant manner. The poor or missing validity of the instruments for hands-on sexual offenses is due to the exceptional low relapse rate of 3,3% in our sample. On the other hand, these data lead to the conclusion that the accuracy of the instruments in this specifically relevant relapse category is of low level    even more so as there are other offense categories with much less relapses but still better indices. Apparently this relapse category is likely to be forgotten in spite of all its relevance. In a number of validation studies the respective differentiation is missing (see Table 1). Yet it is obvious from a victimological as well as from a more general sociopolitical point of view that it makes a difference if the offender directly commits victim abuse with physical harm or for example   only consumes child pornography. If one analyses the data accurately one must conclude that these

instruments are not valid for relevant subgroups, e.g. sexual-motivated hands-on relapses, or at least, that the proof of validity is still missing. The two studies, that examined the relapse category of sexual hands-on offenses either found only moderate AUCs in the total group (.70    a further differentiation in subgroups was not performed (Barbaree et al. 2001)   , or the values proved to be poor in the first place for the total group (.62 for the Static-99 and .66 for the SORAG) and further differentiation showed no validity for rapists (.59 for the Static-99, and .62 for the SORAG) and moderate validity for child abusers (.65 for the Static-99 and .70 for the SORAG) (Harris et al. 2003b).

In our study the SORAG as well as the Static-99 could not predict the sexually motivated relapse in the subgroup of rapists. This was certainly also due to the low relapse rate. On the other hand, Harris et al. (2003b) as well as Ducro & Pham (2006) had similar problems in predicting rapists sexually motivated relapses, whereas non sexual violent offenses could be predicted better with the SORAG as compared to the Static-99. This is no surprise since the SORAG had been designed for the prediction of violent relapse.  On the other hand, in the group of rapists    with all its heterogeneity    the violent element is the more important problem in comparison to the sexual element (Prentky et al. 1995).

In our study the violent relapse (sexual hands-on offense and not sexually motivated violent offense) could be predicted with satisfactory (rapists) and high (child abusers) performance using both instruments in both subgroups.

If analyzing the actual relapse incidents in the different risk categories the value of the prognostic instruments is being put into perspective once more for the user: Regarding the highly relevant relapse categories sexually motivated hands-on offense and violent offense we found in the highest risk group (6,6+) of the Static-99 in the sample of child abusers only 14% and 19%, respectivel,y actual relapses (and 9.5% re-incarcerations). In the group of rapists there were only 0%, respectively 31% actual relapses (and 23% re-incarcerations). The respective numbers for the two highest risk categories of the SORAG in the group of child abusers are 20% and 40% (with 20% re-incarcerations), in the group of rapists 0% and 44%, respectively (and 28% re-incarcerations).

In other words: At the average follow-up time of 3.6 years at risk, the highest risk category correctly identified only every 5th to every 10th recidivist with a sexually motivated hands-on relapse and only every 2nd to every 10th violent recidivist (with the range depending on risk category and prognostic instrument). The prediction of a new sentence for imprisonment due to such relapse incidents was just as imprecise: the highest risk category identified only every 4th to every 10th recidivist correctly.

When relying on these cut-offs for relapse prediction, the absolute numbers of true positives and false negatives are the same. Although the percentage of false negatives is lower due to the absolute number of those who are predicted not to re-offend, it is exactly those false negatives that play an outstanding role in the public discussion, nourishing doubts as to the qualification and the professionalism of the forensic prognostic expert.

Laws enforcing civil commitment or prison sentences for sexual offenders due to their   dangerousness   demand at least a high probability as regards prospective offences, although in most of these laws neither   high probability   nor the evil that should be prevented are properly defined and quantified (Abracen & Looman 2006). Analyzing our data and the results of the studies quoted it is obvious that even in those subgroups with the highest risk levels we are often far from being able to establish at least a probability of 50% for a negative prognostic outcome.

The prediction of relapses and their severity with standardized prognostic instruments is still

insufficient in many respects. For sexual offenders there are the two major categories   sexual and   violent   but these are much too imprecise for the legal practitioner or the clinician. As shown in our study as well as in other studies, sexual relapses frequently are hands-off offenses. Nevertheless, the legal reaction is mostly a prison sentence. Under the perspective of the damage to society caused by such offenses, many of them are minor offenses rarely justifying severe sanctions. On the other hand, the category   violent   in our study included relapses not resulting in a prison sentence, which indicates that the courts did not evaluate them as severe enough to justify such a sanction.

In conclusion we think that the Static-99 and the SORAG are not yet sufficiently validated as regards their predictive power for the most important relapse categories, i.e. sexually motivated hands-on offenses and those violent offenses resulting in re-incarcerations. Although both tools have high validity indices, their practical usefulness is restricted and they certainly do not allow the prediction of the behavior of individuals. This, however, has never been the claim, and one should not throw out the baby with the bathwater. They are most useful in establishing base rates of recidivism for various categories of offenders and offences on which the clinician and the legal practitioner may build their individual assessments and decisions.

# References

1. Abracen, J., & Looman, J. (2006). Evaluation of civil commitment criteria in a high risk sample of sexual offenders. Journal of Sexual Offender Civil Commitment: Science and the Law, 1, 124-140.
2. Barbaree, H.E., Seto, M.C., Lanton, 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.
3. Bartosh, D.L., Garby, T., Lewis, D., & Gary, S. (2003). Differences in the predictive validity of actuarial risk assessments in relation to sex offender type. International Journal of Offender Therapy and Comparative Criminology, 47, 422-438.
4. Craig, L.A., Beech, A., & Browne, K.D. (2006). Cross-validation of the risk matrix 2000 sexual and violent scales. Journal of Interpersonal Violence, 21, 612-633.
5. Craig, L.A., Browne, K.D., & Stringer, I. (2003). Risk scales and actors predictive of sexual offenses recidivism. Tauma, Violence & Abuse: A Review Journal, 4, 45-68.
6. Dahle, K.P., Schneider, V., & Ziethen, F. (2007). Standardisierte Instrumente zur Kriminalprognose. Forensische Psychiatrie, Psychologie, Kriminologie, 1, 15-26.
7. Ducro, C., & Pham, T. (2006). Evaluation of the SORAG and the Static-99 on Belgian sex offenders committed to a forensic facility. Sex Abuse, 18, 15-26.
8. Eher, R., Fruehwald, S., & Frottier, P. (2006). Die standardisierte Begutachtung von Sexualstraftätern im Österreichischen Strafvollzug - erste Ergebnisse und Ausblicke. Neuropsychiatrie, 20, 50-55.
9. Grubin, D. (1998). Sex offending against children: Understanding the risk. Police Research Series, Paper 99. London: Home Office.
10. Hanson, R.K. (1997). The Development of a Brief Actuarial Risk Scale for Sexual Offense Recidivism. User Report 1997-04. Ottawa: Department of the Solicitor General of Canada.
11. Hanson, R.K. (2002). Recidivism and age follow-up data from 4,673 sexual offenders. Journal of Interpersonal Violence, 17, 1046-1062.
12. 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.
13. Hanson, R.K., & Morton-Bourgon, K.E. (2007). The accuracy of recidivism risk assessments for sexual offenders : A meta-analysis. (Rep. No. Corrections User Report No 2007-01.). Ottawa: Public Safety and Emergency Preparedness Canada.

14. 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.
15. Harris, A.J., Phenix, A., Hanson, R.K., & Thornton, D. (2003a). Static-99 Coding Rules Revised - 2003. Correctional Service of Canada.
16. Harris, G.T., & Rice, M.E. (2003). Actuarial assessment of risk among sex offenders. Annuals of the New York Academy of Science, 989, 198-210.
17. Harris, G.T., Rice, M.E., Quinsey, V.L., Lalumiere, M.L., Boer, D., & Lang, C. (2003b). A multisite comparison of actuarial risk instruments for sex offenders. Psychological Assessment, 15, 413-425.
18. Hart, S.D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. Legal and Criminological Psychology, 3, 121-137.
19. Helmus, L.M.D., & Hanson, R.K. (2007). Predictive validity of the Static-99 and Static-2002 for sex offenders on community supervision. Sexual Offender Treatment, 2, 1-14 (www.sexual-offender-treatment.org/60.0.html)
20. Hood, R., Shute, S., Feilzer, M., & Wilcox, A. (2002). Sex offenders emerging from long-term imprisonment: A study of their long-term reconviction rates and of parole board members' judgements of their risk. British Journal of Criminology, 42, 371-394.
21. Nedopil, N. (2005). Prognosen in der Forensischen Psychiatrie. Ein Handbuch für die Praxis. Lengerich: Pabst.
22. Nunes, K.L., Firestone, P., Bradford, J.M., Greenberg, D.M., & Broom, I. (2002). A comparison of modified versions of the Static-99 and the Sex Offender Risk Appraisal Guide. Sex Abuse, 14, 253-269.
23. Prentky, R.A., Knight, R.A., Lee, A.S., Cerce, D.D. (1995). Predictive validity of lifestyle impulsivity for rapists. Criminal Justice and Behavior, 22, 106-128.
24. Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C. (1998). Violent Offenders - Appraising and Managing Risk. (1st ed.) Washington DC: American Psychological Association.
25. Quinsey, V.L., Harris, A.J., Rice, M.E., & Cormier, C. (2006). Violent Offenders: Appraising and Managing Risk. (2nd ed.) Washington DC: American Psychological Association.
26. Rettenberger, M., & Eher, R. (2006a). Actuarial assessement of sex offender recidivism risk: A validation of the German version of the Static-99. Sexual Offender Treatment, 1[3], 1-11(www.sexual-offender-treatment.org/51.0.html).
27. Rettenberger, M., & Eher, R. (2006b). Die deutsche Übersetzung und Adaptierung des Static-99 zur aktuarischen Kriminalprognose verurteilter Sexualstraftäter. Monatsschrift für Kriminologie und Strafrechtsreform, 89, 352-365.
28. Rettenberger, M., & Eher, R. (2007a). Aktuarische Kriminalprognosemethoden und Sexualdelinquenz: Die deutsche Version des SORAG. Monatsschrift für Kriminologie und Strafrechtsreform, 90, 484-497.
29. Rettenberger, M., & Eher, R. (2007b). Predicting reoffence in sexual offender subtypes: A prospective validation study of the German version of the Sexual Offender Risk Appraisal Guide (SORAG). Sexual Offender Treatment, 2, 1-12 (www.sexual-offender-treatment.org/60.1.html).
30. Stadtland, C., Hollweg, M., Kleindienst, N., Dietl, J., Reich, U., & Nedopil, N. (2006). Rückfallprognosen bei Sexualstraftätern - Vergleich der prädiktiven Validität von Prognoseinstrumenten. Nervenarzt, 77, 587-595.

# Author address

**Reinhard Eher**, *MD, Associate Professor*
*Evaluation Centre for Violent and Sexual Offenders*
*Gerichtsgasse 6*
*A-1210 Wien*
*Austria*
*reinhard.eher@justiz.gv.at*