

## **The Role of Local Base Rate Information in Determining the Accuracy of Sexual Recidivism Actuarial Instruments**

**Brian R. Abbott, Ph.D.**  
**Private Practice, Sunnyvale, CA.**

Actuarial risk assessment is considered superior to clinical judgment when assessing the potential for sexual recidivism (see Janus & Prentky, 2003 for a comprehensive review of this issue related to sexual recidivism). For this reason, actuarial risk assessment has become the *sin quo non* in assessing an identified sexual offender's potential to engage in further sexually offending behavior. The actuarial method of risk assessment is appealing due to its mechanistic methods of completing the instrument, as well as its apparent scientifically derived accuracy and reliability. The aura of infallibility engendered by actuarial instruments has contributed to formulation of legislative mandates and administrative proscriptions for the use of actuarial instruments in assessing future risk for sexual reoffense among already identified sexual offenders.

For instance, California Senate Bill 1128 (aka: SARATSO) mandates the use of the actuarial instrument called the Static-99 in assessing risk of adult sex offenders in before release from prison and as part of presentencing investigations. The California Department of Mental Health (2008) requires contract and employed evaluators to administer the Static-99 in sexually violent predator evaluations. Use of the Static-99 in such ways presumes the measure has sufficient reliability and validity for these tasks when in actuality this has not been proven. We have no data in using the Static-99 with California sex offenders as to how many of them are determined to be likely recidivists when they actually never reoffend sexually (false positive error). Similarly, we have no idea how many sexual offenders are classified as unlikely to reoffend sexually based on their Static-99 score only to commit another sexual assault (false negative error).

These two forms of error have enormous implications. High rates of false positive decisions may subject individuals to unwarranted deprivation of liberty through imprisonment, civil commitment, or high control supervision practices. Such extreme measures to control sexual offenders waste precious resources in our cash-strapped state on possibly large numbers of individuals who will never reoffend sexually. High rates of false negative decisions mean that sexual offenders are released into the community to later victimize others. This outcome exacts large costs in terms of dealing with the sequelae of trauma to victims and investigating and prosecuting the sex crimes committed by reoffenders.

As a forensic psychologist who works primarily for the defense, I have become well acquainted with the deliberate or unintended misuse of actuarial instruments for various legal purposes. Debate abounds in the literature regarding the accuracy and reliability of actuarial risk assessment instruments to measure sexual recidivism (see Prentky et al., 2006 & Prentky & Janus, 2004 for detailed discussion of these issues). Due to space constraints in the newsletter and as a result of the complexity of this issue, I will focus on the singular topic related to the application of the Static-99 risk estimates to sex offenders in the U.S. in general and California in particular. Despite my narrow focus on one actuarial instrument, what I discuss here applies to the other major actuarial instruments developed to predict sexual recidivism including the Sex Offender Risk Appraisal Guide (Quinsey et al., 1998); Minnesota Sexual Offender Screening Tool-Revised (Epperson et al., 2003); and the Rapid Risk Assessment of Sex Offender Recidivism (Hanson, 1997).

The main argument in defending the use of the Static-99 when making high-stakes decisions rests on its "moderate predictive accuracy." Predictive accuracy is determined by using a specific statistical method called the Receiver Operator Characteristics Area Under the Curve ("AUC"). The AUC estimates predictive accuracy for the entire group of sexual offenders being measured without taking into account accuracy rate for a particular score or the amount of sexual recidivism that has occurred in the population. Using this statistic, various studies on the Static-99 (see Harris et al., 2003 page 72 for review of studies) have found an average AUC of 72%. What does this mean? Seventy-two

percent of randomly selected recidivists in a population of sexual offenders will have a higher risk bin score (from 0-6+) on the Static-99 than randomly selected nonrecidivists. On the other side of the AUC coin, 28% of the randomly selected nonrecidivists will have a higher score on the Static-99 than randomly selected recidivists. Thus, the Static-99 has a 28% false positive error rate.

The AUC statistic fails to account for the rate of sexual recidivism in a given population of sexual offenders, which I will refer to hereinafter as the base rate. The base rate of sexual recidivism plays a crucial role in predictive accuracy. Users who interpret the Static-99 risk information blind to the base rate of sexual recidivism in the population of interest will have no idea as to the rate of erroneous and correct decisions. As the base rate of sexual recidivism falls below 50% in a population of sexual offenders, a sexual recidivism actuarial measure, like the Static-99, performs poorly as seen by the fact that the proportion of detecting true recidivists is less than the proportion of nonrecidivists who are misclassified as likely reoffenders (Prentky & Burgess, 2000; Prentky & Janus, 2003; Prentky et al., 2006; & Vrieze & Grove, 2007). This problem can be illustrated with the Static-99 developmental sample that has an 18% base rate at the five-year follow up period.

At a score of 6+, only 26.5% of actual recidivists in the Static-99 developmental sample (Hanson & Thornton, 2000) were correctly classified at the expense of incorrectly classifying 8.8% of nonrecidivists. At first blush, this seems like acceptable odds but one has to consider that as the base rate diminishes the number of nonrecidivists grows increasingly larger than the recidivists. In the Static-99 developmental sample, nonrecidivists outnumber the recidivists on the order of almost five to one. Turning to the actual numbers in the Static-99 developmental sample, we find that at a score of 6+ 50 recidivists are correctly classified at the expense of misclassifying 79 nonrecidivists as likely reoffenders. As seen in Table 1, when considering the combination the base rate in the Static-99

**Table 1: Proportion of Correct and Incorrect Classification of Sexual Recidivists Using Different Accuracy Measures**

Accuracy Method	Proportion of Correct Predictions	Proportion of Incorrect Predictions
AUC Statistic	72%	28%
Cut-off Score of 6+ & base rate of 18%	38.5%	61.2%

developmental sample at the five-year follow up period and the proportion recidivists and nonrecidivists are a risk level score of 6+, the actual proportion of correct classification of recidivists to incorrect classification of nonrecidivists as likely reoffenders is the opposite of what the AUC statistic reflects. Consequently, the AUC statistic will overestimate predictive accuracy when rendering risk assessment decisions for sex offenders in low base rate populations at every categorical risk level.

To further illustrate this point with a practical example: Let's say a probation officer was using the Static-99 score of 6+ to assign sex offenders to a high control case load providing intensive probation supervision. In this circumstance, the probation officer would assign three sex offenders who would never sexually reoffend to the high control supervision caseload to capture the two offenders who actually meet the criteria for high risk. Now let's say this enlightened probation department decides to study the effectiveness of their high supervision case load. Not surprisingly, the researcher finds a low recidivism rate for the sex offenders under the high control caseload. If this researcher is not cognizant of the error rate of assigning more nonreoffenders to the caseload than reoffenders by using the Static-99, we really don't know if the high supervision methods worked or whether the preponderance of individuals in this group who would never reoffend contributed to the low recidivism rate.

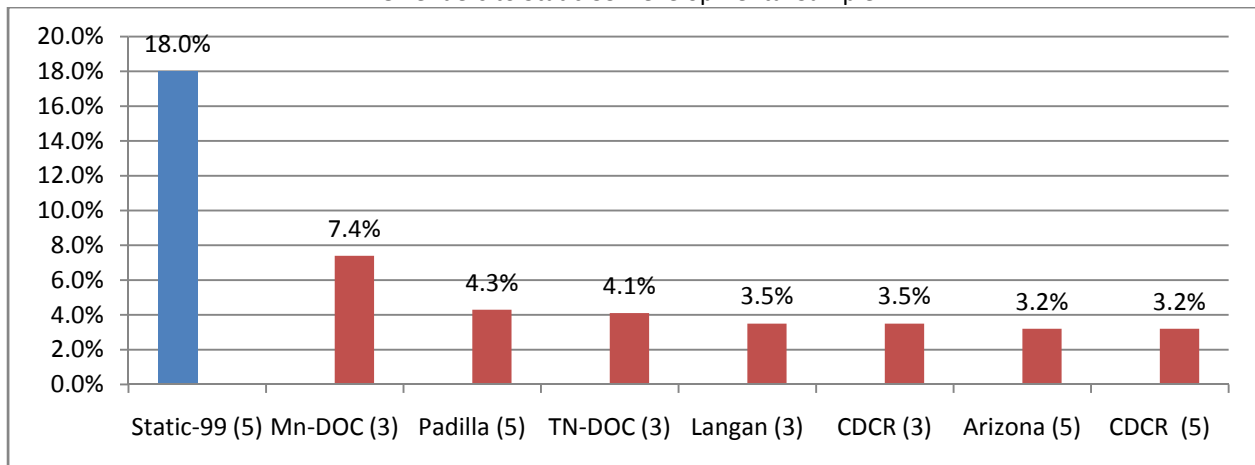
As my example shows, base rates of sexual recidivism matter when interpreting the categorical risk ranking of 0-6+ or low to high as determined by offenders raw scores on the Static-99. Research

has found that the Static-99 categorical risk levels, from 0-6+ or low to high, appear to be fairly reliable in ranking relative risk (Harris et al., 2003). Existing research has found the risk percentages reported by the Static-99 (Hanson & Thornton, 2000) are unstable or unreliable when applied to other groups of sexual offenders (Helmus, 2008; Looman, 2006; Abracen & Looman, 2006; Hanson, 2006; & Doren, 2004). These studies tell us that a score of 6+ on the Static-99 yields five-year recidivism rates between 11% and 27% compared to 39% as found with the Static-99 developmental sample; a rate of overprediction ranging between 30% and 72%. Doren (2004) reported one sample out of eight where the corresponding risk estimate for the Static-99 score of 6+ at five years underpredicted sexual recidivism by 10%.

The data related to the Static-99 overprediction rate tells us that a high categorical risk level (low to high) does not inevitably translate into an actual high percentage risk for sexual reoffense. If a score of 6+ on the Static-99 is associated with an 11% chance for sexual reoffense, this means that 89% of the sexual offenders at this high risk level did not reoffend sexually. On practical level, decision makers relying on the Static-99 risk bin of 6+ to correctly classify likely recidivists for whatever purpose will incorrectly classify eight nonrecidivists as likely recidivists in order to capture the one true recidivist (assuming the base rate is 11% in a local population for a score of 6+).

The instability of the Static-99 risk percentages across various populations of sexual offender highlights the need to consider base rate data on local populations of sexual offenders (Janus & Prentky, 2003 & Prentky et al., 2006) rather than to rely on the risk estimates propounded by risk estimate tables contained in actuarial risk assessment instruments. Until recently, little contemporary data has been published on sexual recidivism rates for local populations of sexual offenders. Figure 1 provides an overview of sexual recidivism rates in contemporary samples of United States sex offenders as compared to the five-year base rates of sexual recidivism found in the Static-99 developmental sample unadjusted for risk level. I made the comparison based on risk unadjusted numbers because the studies did not have risk assessment data from which to classify offenders by actuarial risk bins. The numbers listed in parenthesis after the title for each sample corresponds to the follow-up period. The reference for each study is listed at the conclusion of this article.<sup>1</sup> As can be seen in these comparisons, the contemporary base rates for U.S. sexual offenders are substantially lower than what is reported for the Static-99 developmental sample. At the five year risk interval, U.S. base rates for sexual recidivism are

Figure 1: Comparison of 5-Year Recidivism Rates: U.S. Sex Offenders to Static-99 Developmental Sample



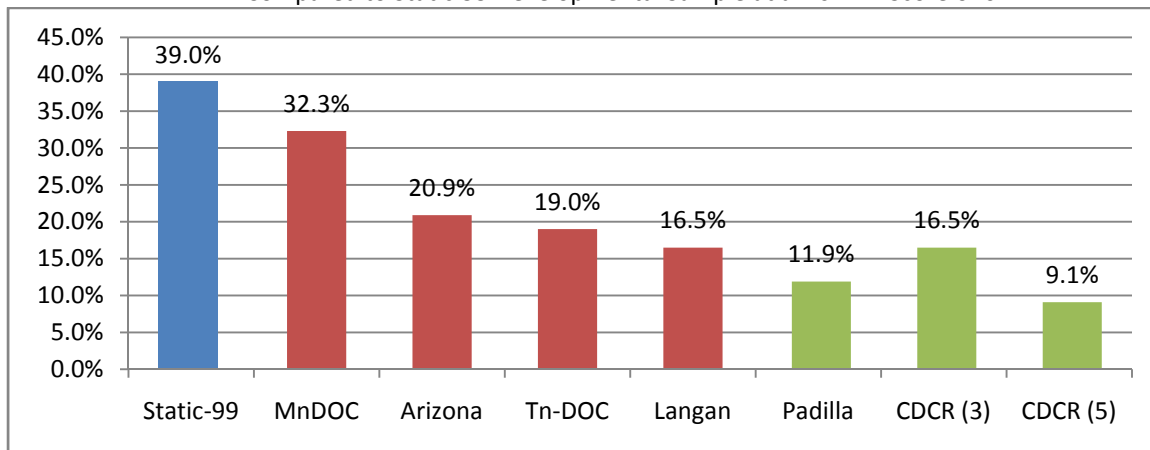
between 58%-82% lower what the Static-99 reports. Similar trends are seen at the ten year risk interval. Due to space limitations, I did not include the graph of these data. Nevertheless, U.S. samples show rates of sexual reoffense between 44% and 81% less than the Static-99 at the ten year interval. The reader will notice that some of the follow up periods at five years for the U.S. studies are slightly

shorter than the Static-99 follow up period. The Static-99 will still over-predicts by a large margin even when adjusting U.S. studies upward to compensate for the shorter follow up periods.

The reader critical of my argument will fault the above analysis based on the fact that the rates of sexual recidivism are not adjusted by actuarial risk level. I find this criticism unconvincing because it can be logically reasoned that when controlling for actuarial risk the recidivism rates in low base rate populations will still be lower than reported by the Static-99. If the 6+ risk estimate at five years for the Static-99 developmental sample is 2.167 times higher than the risk unadjusted base rate of 18% for this same sample, it is reasonable to expect populations with lower base rates will exhibit the same proportional increase in recidivism rates when adjusting for risk level. But don't take my word for it.

A recent article (Donaldson & Wollert, 2008) provides a mathematical calculation to adjust the actuarial risk estimates using known base rates in local populations (e.g., sexual offenders in California) using a probability theory referred to as Bayes Theorem. Figure 2 provides the base rate adjusted risk estimates for a Static-99 score of 6+ with for sex offenders from California (green bars) and several other states (red bars) using Bayes Theorem. I first normalized the base rate for each study to correspond to a five year risk interval. This was accomplished by taking the base rate of the study dividing it by the duration of the follow up period and multiplying this result by five (e.g., base rate of 3.5% with follow up of three years-  $3.5\% / 3 = 1.166\% \times 5 = 5.8\%$  five-year recidivism rate). The 5-year normalized base rate for each study was used to apply the base rate adjustment at a Static-99 risk bin score of 6+, as explicated by Donaldson and Wollert (2008).

Figure 2: Base Rate Adjusted Actuarial Risk Estimates for U.S. Sex Offenders Compared to Static-99 Developmental Sample at a Risk Bin Score of 6+



The Bayes Theorem procedure yields a base rate correction for the actuarial risk estimates contained in the Static-99 developmental sample at each risk bin level. Applying the correction to the risk bin level 6+ reveals that sex offenders from all reported states have recidivism rate from between 17% and 77% lower than what the Static-99 would suggest. Most notable is the fact that the three lowest base-rate adjusted actuarial risk estimates are among California sex offender populations (Padilla, CDCR (3), & CDCR (5)). Langan (2003) contains more than 3,000 California sexual offenders. The base rates reflected in these studies show that the Static-99 overpredicts by a magnitude of between 57% and 77% when applied to California sex offenders who fall within the 6+ risk bin. In practical terms, an evaluator who relied on the Static-99 risk bin score of 6+, when assessing California sex offenders, would falsely classify between five and ten nonrecidivists as likely recidivists in order to correctly classify the one true recidivist.

## **In Conclusion**

Using peer reviewed published literature I have shown how high categorical risk levels as defined by the Static-99 do not necessarily correspond with high rates of sexual reoffense. While I have shown serious deficiencies when relying on the risk information contained in the Static-99 developmental sample, my intent in writing this article is not to advocate abandoning the use of the Static-99. Rather, I argue for the responsible use of the Static-99. Toward this end, I close with the following recommendations.

Professionals who use the Static-99 with California sex offenders should be trained not only in its proper administration but also have the knowledge about common psychometric terms that will assist in responsibly interpreting the risk information. This issue becomes of increasing urgency as legislation has resulted in correctional officials using the Static-99. These users rarely have proper training to understand psychometric principles necessary to understand and interpret findings from actuarial risk measures. Users of sexual recidivism actuarial instruments, like the Static-99, should know the extent to which actuarial results are erroneous. Without knowing and understanding the error rate in decision making, users risk misinforming others about the risk potential of an individual.

Professionals who administer the Static-99 should not rely on the risk information contained in the Static-99 developmental sample to interpret the risk potential for California sex offenders. As I have documented in this article, ample literature exists that allows us to use local base rate data to derive more accurate estimates of risk for California sex offenders. Users of the Static-99 should apply the base rate correction procedure explained by Donaldson and Wollert (2008) when reporting risk information about California sex offenders.

If the government requires the use of actuarial risk assessment methods that may result in erroneous decisions that jeopardize citizens' liberty interests or results in increasing potential harm to members of the community, the government should be required to ensure these decisions minimize false positive and false negative errors. Consequently, state legislators or the Department of Mental Health who require the use of the Static-99 for various legal and administrative purposes should conduct validation studies on contemporary groups of sexual offenders in California. The goal of this research would be the development of reliable actuarial risk estimates that can be confidently applied to sex offenders in California and to reduce the potential of false positive and false negative decisions. This research would need to be conducted periodically as base rates may fluctuate over time.

## **REFERENCES**

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, 124-139.

Arizona Department of Corrections. (1998). Sex Offender Recidivism: Fact Sheet. Available at: <http://www.adcprisoninfo.az.gov/FACTSHEETS/Fact%20Sheet%2098-06.htm>

California Department of Mental Health (2008, August). Clinical Evaluator Handbook and Standardized Assessment Protocol. Unpublished. Available upon request from author at: dr.abbott@comcast.net

California Sexual Offender Management Board. (2008, January). An Assessment of Current Management Practices of Adult Sex Offenders in California: Initial Report. Available at: <http://www.calcasapublicpolicy.org/wp-content/uploads/2008/02/somb-report.pdf>

California Sexual Offender Management Board. Recidivism of paroled sex offenders- A five (5) year study. Unpublished. Available from author upon request at dr.abbott@comcast.net

California Sexual Offender Management Board. Recidivism of paroled sex offenders- A ten (10) year study. Unpublished. Available from author upon request at [dr.abbott@comcast.net](mailto:dr.abbott@comcast.net)

Donaldson, T. & Wollert, R. (2008). A Mathematical Proof and Example That Bayes's Theorem Is Fundamental to Actuarial Estimates of Sexual Recidivism Risk. Sexual Abuse: A Journal of Research and Treatment, 20(2), 206-217.

Doren, D. (2004). Stability of the interpretative risk percentages for the RRASOR and Static-99. Sexual Abuse: A Journal of Treatment and Research. 16(1), 25-36.

Epperson, D.L., Kaul, J.D., Huot, S.J., Hesselton, D., Alexander, W., & Goldman, R. (2003). Minnesota sex offender screening tool- revised (Mn-SOST-R): Development, performance, and recommended risk level cut scores.

Hanson, R.K. (2006). Does the static-99 predict recidivism among older sexual offenders? Sexual Abuse: A Journal of Research and Treatment, 18, 343-355.

Hanson, R.K. & Thornton, D. (2000). Improving risk assessment for sex offenders: A comparison of three actuarial scales. Law and Human Behavior, 24(1), 119-136.

Hanson, R.K. (1997). The Development of a Brief Actuarial Risk Scale for Sexual Offense Recidivism. Available at: [http://ww2.ps-sp.gc.ca/publications/corrections/199704\\_e.pdf](http://ww2.ps-sp.gc.ca/publications/corrections/199704_e.pdf)

Harris, A., Phenix, A., Hanson, R.K., & Thornton, D. (2003). Static-99 Coding Rules Revised- 2003. Canada: Dept of Solicitor General.

Helmus, L. (2008). Static-99 recidivism percentages by risk level. Supplemental data provided available from author at [leslie.helmus@ps.gc.ca](mailto:leslie.helmus@ps.gc.ca).

Janus, E.S. & Prentky, R.A. (2003). Forensic use of actuarial risk assessment with sex offenders: Accuracy, admissibility, and accountability. American Criminal Law Review, 40(1143), 1-59.

Langan, P.A., Schmitt, E.L., & Durose, M.R. (2003). Recidivism of sex offenders released from prison in 1994. Washington, D.C.: U.S. Dept of Justice.

Looman, J. (2006). Comparison of two risk assessment instruments for sexual offenders. Sexual Abuse: A Journal of Research and Treatment. 18(2), 193-206.

Minnesota Department of Corrections (2007). Sexual offender recidivism in Minnesota: April 2007. Available from [www.doc.state.mn.us](http://www.doc.state.mn.us).

Padilla, J. (2006, October). Personal communication to Jim McEntee, Public Defender. Available from author upon request at [dr.abbott@comcast.net](mailto:dr.abbott@comcast.net)

Prentky, R.A. & Burgess, A.W. (2000). Forensic Management of Sexual Offenders. New York: Kluwer Academic.

Prentky, R.A., Janus, E., Barbaree, H., Schwartz, B.K., & Kafka, M.P. (2006a). Sexually violent predators in the courtroom: Science on trial. Psychology, Public Policy, and the Law, 12(4), 357-393.

Tennessee Bureau of Investigation. (2007). Recidivism Study. Available at: [http://www.tbi.state.tn.us/Info%20Systems%20Div/TIBRS\\_unit/Publications/Sex%20Offender%20Recidivism%202007%208-14-07.pdf](http://www.tbi.state.tn.us/Info%20Systems%20Div/TIBRS_unit/Publications/Sex%20Offender%20Recidivism%202007%208-14-07.pdf)

Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C.A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Vrieze, S.L. & Grove, W.M. (2008). Predicting sex offenders recidivism. I. Correcting for item overselection and accuracy overestimation in scale development. II. Sampling error-induced attenuation of predictive validity over base rate information. Law and Human Behavior, 32(3), 266-278.

Wollert, R. (2006). Low base rates limit expert certainty when current actuarials are used to identify sexually violent predators. Psychology, public policy and the law, 12(1), 56-85.

---

<sup>1</sup> Mn-DOC- Minnesota Department of Corrections (2007)

Padilla- Padilla (2006)

TN-DOC- Tennessee Bureau of Investigations (2007)

Langan- Langan, Schmitt, & Durose (2003)

Arizona- Arizona Department of Corrections (1998)

CDCR- California Department of Corrections and Rehabilitation- see references for California Sexual Offender Management Board