Moving beyond the Standard Model for Actuarial Assessment for Sexual Offenders[1]

David Thornton
University of Bergen, Norway

R. Karl Hanson Leslie Helmus
Public Safety Canada

Available at http://ccoso.org/newsletter.php California Coalition on Sex Offending.

The Standard Model for Adjusted Actuarial Assessment

Up until relatively recently a common conception of good practice in sexual offender risk assessment might have gone something like this:

1) Determine the offender's Static-99 score.
2) Use the empirically-derived recidivism estimate associated with that particular score as a starting point in evaluating the risk presented by the offender.
3) Consider how similar the current case is to the offenders in the research samples used to derive the recidivism estimates. This consideration would affect the amount of weight that would be given to the Static-99 score in the overall risk evaluation.
4) Consider also whether the individual has other risk-relevant characteristics that differ markedly from what would be typical for a sexual offender with that Static-99 score. This would speak to whether the recidivism estimate should be adjusted up or down. Small adjustments have generally been preferred because it has been thought that (a) other risk-related factors would already have been largely taken into account in the Static-99 score (so their incremental predictive validity would be small), and (b) we can score Static-99 more reliably than we can assess other risk-related factors. There was also a concern that the subjective element in this process would degrade predictive accuracy.

This way of doing sexual offender evaluations will be referred to as the Standard Model. Research presented at the 2008 and 2009 ATSA conferences has called into question assumptions on which this evaluation model is based. This paper summarizes some of these research findings and considers their implications for the use of actuarial instruments in sexual offender risk assessment.

Base Rates

The Standard Model assumes that base rates are nearly constant. The rationale for this assumption was that they appeared to be constant across the samples described by Hanson and Thornton (2000). Research presented at the ATSA conferences and downloadable from www.static99.org demonstrates that this assumption is unsafe. There were two key findings regarding base rates.

First, the mean sexual recidivism base rate after 5-years and 10-years follow up was lower in a newly assembled collection of 28 Static-99/recidivism samples than it was in Hanson and Thornton (2000). Second, there was highly significant variation in sexual recidivism base rates among the 28 samples. A detailed account of the analyses of these data may be found in Helmus (2009) and specific normative recidivism estimates my be downloaded from

http://www.static99.org/pdfdocs/st-99rworkbookwithsamplesandsummaries.pdf

These findings of variable base rates have profound implications for actuarial risk assessment. They mean that in order to obtain a recidivism estimate, in addition to knowing how someone scores on Static-99, you also need to know the base rate for the population from which the offender comes. This step is not included within the Standard Model.

The Impact of Other Risk-Related Factors

[1] The opinions contained in this paper are those of the authors and not those of the organizations for which the authors work

The Standard Model assumes that other risk-related factors only make a small difference once you have taken an individual's Static-99 score into account. Three inter-related sets of findings call this assumption into question.

First, findings presented at the 2009 ATSA conference indicated that age on release has a substantive effect beyond Static-99 score. There have been individual findings supportive of this idea and it has been eloquently advocated for by Howard Barbaree and his colleagues for some time. The new results go beyond this in finding robust incremental effects of age in the analysis of 23 samples of Static-99/recidivism data.

These findings are sufficiently strong that the Static-99 authors responded by developing and validating a revised version of Static-99 designed to more fully incorporate the effects of age.

Second, even after variation in age on release is allowed for, there was still substantial variation in the sexual recidivism base rates in the new collection of Static-99/recidivism data sets. Analysis of this variation indicates that a major source of base rate variation is the degree that the sample as a whole had been selected for factors related to risk. More specifically, in samples which have been subject to strong pre-selection for risk, the recidivism rates associated with any given Static-99R score are about twice the rates associated with that same score in unselected samples. Pre-selection effects can only arise through the selection process being responsive to risk factors that are external to Static-99R. The magnitude of the base rate effect implies that these external risk factors can have a large effect.

Third, one of the symposia at the 2009 ATSA conference systematically examined the incremental predictive value of psychological risk factors, assessed either through questionnaires or ratings, relative to static actuarial assessment (Hanson and Helmus, 2009; Harkins, Thornton, & Beech, 2009; and Thornton & Knight, 2009). Strikingly similar findings in these three data sets show that the magnitude of the incremental validity of these ratings was sufficient to reproduce the base rate effects.

For example, the Dynamic Supervision Project examined predictors of sexual recidivism in a large sample of offenders under community supervision. STABLE-2007 is an empirically-derived scheme for rating psychological risk factors. For ratings made by conscientious staff, STABLE-2007 showed incremental validity relative to Static-99R. The table below shows the projected three-year sexual recidivism rates for those with average STABLE scores for this sample (scoring 5) and those whose scores are around the 90th percentile (scoring 14). It is based on further analysis of the data presented by Hanson and Helmus (2009).

Static-99R Score	Three-Year Recidivism		Five-Year Recidivism	
	STABLE=5	STABLE=14	Routine	Selected for Risk/Need
2	3%	7%	5%	12%
5	7%	18%	11%	25%
7	14%	32%	19%	38%

The Dynamic Supervision Project data shown in the left side of the table is for a three-year follow up. Three Static-99R scores were considered: 2 - corresponding to the median value, 5 – moderate-high (10th to 19th percentile), and 7 – high (2nd to 5th percentile), For each of the three Static-99R scores shown, a STABLE-2007 score of 14 implies at least twice the sexual recidivism found for an average STABLE score. For comparison the right side of the table shows the five-year sexual recidivism rates for unselected samples and samples that had been intensely pre-selected for risk (from the Static-99R norms presented in 2009). Samples intensely pre-selected for risk show on average twice the recidivism rate of unselected samples. Taking into account that the Dynamic Supervision Project data are based only two-thirds of the follow up period (3 years versus 5 years) for the normative samples, it is clear that intense selection on STABLE-2007 essentially reproduces within a Routine sample the kind of recidivism rates that are found in samples categorized as pre-selected for risk.

The other two studies cited showed similar effects using the Structured Risk Assessment framework (Thornton, 2002) for a pre-selected for treatment sample (Harkins, Thornton, & Beech, 2009) and a pre-selected for risk sample (Knight & Thornton, 2009). Other instruments that map into the Structured Risk

Assessment framework have also shown substantial incremental predictive validity relative to Static-99 (e.g. Olver, Wong, Nicholaichuk, & Gordon, 2007; Allan, Grace, Rutherford, & Hudson, 2007).

Taken together these results imply that the incremental predictive effect of risk factors external to Static-99 (or even Static-99R) are substantial, and certainly of sufficient magnitude that you cannot assume that all offenders with the same Static-99 score present the same risk. Rather than a single recidivism estimate properly indicating the risk implied by a given base rate and Static-99R score, a wide variety of risk levels may apply depending on the level of psychological risk factors that are also involved.

Practice Implications

Where the intention is to make broad resource allocation decisions for groups of offenders, Static-99 (or Static-99R) provides an efficient way of creating groups that differ in terms of relative risk. A much more challenging situation applies in high stakes risk assessment where legal criteria or public policy concerns require both an estimate of absolute recidivism risk and our best efforts at asserting the relevance of this recidivism estimate to the individual being evaluated. This is particularly so in the context of SVP civil commitment hearings where the potential consequences for both the individual being evaluated and for society are particularly high.

The research reviewed above implies that in addition to high scores on static actuarial instruments, it is only when unusually high levels of psychological risk factors are also present that the levels of risk required for civil commitment can reasonably be attributed. Prior selection processes that implicitly attribute raised levels of treatment need / psychological risk are one kind of cue which may indicate the presence of psychological risk factors but they are not sufficient to confidently infer that this applies to the individual. Prudent evaluators would only make this kind of attribution if they can also see direct evidence of high levels of psychological risk factors.

Finding direct evidence for psychological risk factors is a challenge both because offenders will seek to conceal risk in adversarial settings and because instruments for assessing it have been less well researched than static risk scales for sexual offenders. The new findings do, however, make one thing easier. The combination of high static risk and high psychological risk seems to imply levels of risk consistent with civil commitment in all three base rate groups identified in the recent Static-99 research. Consequently, evaluators are less dependent on uncertain inferences about which base rate group an offender belongs to so long as they are able to do a good assessment of long-term psychological risk factors.

Assessment of long-term psychological risk factors is best done using some empirically-based, structured instrument. Relevant instruments include the VRS-SO (Olver et al., 2007), STABLE-2007 (Hanson, Harris, Scott & Helmus, 2007), and the forensic version of SRA (Thornton & Knight, 2009). All of the existing instruments would benefit from considerably more statistical research exploring their properties. Nevertheless the obvious overlap in the factors assessed and the consistency of the findings across studies of different instruments provide a basis for having some confidence in them. The three instruments have different strengths and weaknesses. STABLE-2007 may be the instrument of choice for offenders who have had substantial amounts of time in the community in the last four years so long as evaluators have access to credible data regarding the offender's community functioning. The VRS-SO and SRA may be preferable for offenders who have not had significant recent time in the community and both also have the advantage of demonstrated incremental predictive value over long-term follow ups.

References

Allan, M., Grace, R. C., Rutherford, B. and Hudson, S. M. (2007). Psychometric assessment of dynamic risk factors for child molesters. *Sexual Abuse: A Journal of Research and Treatment, 19*, 347-367.

Hanson, R.K., Harris, A. J. R., Scott, T-L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project.* User Report, Corrections Research, Ottawa: Public Safety Canada. Available at www.ps-sp.gc.ca/res/cor/rep

Hanson, R. K. & Helmus, L. (2009, October) *Methods for combining historical and psychological risk factors: An example using Static-2002 and STABLE-2007.* Presentation at the 28th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Dallas, Texas.

Harkins, L., Thornton, D., & Beech, A. (2009, October). *The use of dynamic risk domains assessed using psychometric measures to revise relative risk assessment using RM 2000 and Static 2002.* Presentation at the 28th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Dallas, Texas.

Helmus, L. (2009). *Re-norming Static-99 recidivism estimates: Exploring base rate variability across sex offender samples.* Unpublished M.A. thesis, Carleton University, Ottawa, Ontario, Canada.

Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale – Sexual Offender Version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19*, 318-329.

Thornton, D. (2002) Constructing and Testing a Framework for Dynamic Risk Assessment. Sexual Abuse: *A Journal of Research and Treatment, 14*, 137-151.

Thornton, D. & Knight, R. (2009, October). *Using SRA Need domains based on structured judgment to revise relative risk assessments based on Static-2002 and Risk Matrix 2000.* Presentation at the 28th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Dallas, Texas.