Predictive Accuracy of Static-99R in Three Prospective Field Validity Studies

Symposium Chair: L. Maaike Helmus, Ph.D., Global Institute of Forensic Research

Although Static-99/R is the most widely researched sex offender risk assessment scale (Hanson & Morton-Bourgon, 2009), virtually all studies of the scale have been scored for research purposes and not by field staff as part of routine practice. It is not clear how well the findings from research studies will generalize when the scale is used as part of routine correctional practice (though generalizability may be strongly influenced by factors such as quality of training and institutional support for the scale, as well as access to appropriate records for scoring). Preliminary research in this area has been conflicting, with discouraging results for the scale first published in Texas (Boccaccini, Murrie, Caperton, & Hawes, 2009), followed by very encouraging results from Canada (Hanson, Helmus, & Harris, 2015) and California (Hanson, Lunetta, Phenix, Neeley, & Epperson, 2014).

The purpose of this symposium is to highlight results from three new studies of the field validity of Static-99/R using large sample sizes. The first presentation will include more recent findings from Texas, with a much increased sample size ($n > 34,000$). The second presentation will describe a new Canadian field study from British Columbia ($n > 4,000$). The third presentation will present findings from a new sample from California ($n > 1,000$). Together, these findings will shed new light on the utility of Static-99R in jurisdiction-wide implementations of the scale. Importantly, each presentation will also discuss predictive accuracy in terms of both discrimination and calibration.

References

**Financial Disclosure:**
L. Maaike Helmus is a certified Static-99R trainer. The copyright for Static-99R is held by the Government of Canada.

Revisiting the Field Validity of Static-99/R Scores in Texas

Marcus T. Boccaccini, Ph.D., Sam Houston State University
L. Maaike Helmus, Ph.D., Global Institute of Forensic Research
Amanda K. Rice, M.A., Sam Houston State University

In 2009, we published the first field validity study of Static-99 scores in the United States (Boccaccini, Murrie, Caperton, & Hawes, 2009). We found that Static-99 scores were weaker predictors of future sexual offending than expected (AUC = .60, $d = .36$), well outside the 95% confidence interval of the predictive effect from Hanson and Morton-Bourgon's meta-analysis ($d = .67$, 95% CI [.62, .72]). The purpose of the current study was to revisit the field validity of Static-99 and Static-99R scores in Texas, using a large and representative sample of 34,687 sexual offenders released from custody after being scored on the Static-99. Although we only had access to offenders' total scores, we were able to convert Static-99 scores to Static-99R scores using offender birth and release dates. The mean follow-up time was 5.23 years ($SD = 3.18$), and the base rate of sexual recidivism was 3.6% ($n = 1,241$).

The overall predictive effects for Static-99 and Static-99R scores tended to be larger than those from the 2009 study. In terms of calibration analyses, the routine recidivism estimates currently available for Static-99R significantly and meaningfully overestimated recidivism rates compared to the observed data in Texas, with overestimation of recidivism most pronounced for higher risk scores. Because there is some evidence that the field reliability of Texas Static-99 scores improved after the release of the 2003 Static-99 scoring manual (see Rice, Boccaccini, Harris, & Hawes, 2014), we compared effects from offenders scored before 2004 ($n = 15,680$) to those scored after 2004 ($n = 19,007$) and found significantly stronger effects for post-2004 Static-99 scores than pre-2004 Static-99 scores.
Learning Goals and Objectives
At the conclusion of the presentation, attendees will be able to:

- Understand the difference between field validity and non-field validity studies in risk assessment.
- Understand the range of predictive effects indicated by prior Static-99/R field validity research.
- Understand why predictive validity effects in field settings may be stronger for some cohorts of offenders than others.
- Understand how Static-99/R scores perform in the largest ever Static-99R field study.
- Use findings from the large field study to better inform their assessment practices.

Predictive Accuracy of the Static-99R in British Columbia:
Results of a Prospective Field Study with an Average 5-Year Follow-Up

L. Maaike Helmus, Ph.D., Global Institute of Forensic Research
Carmen L. Zabarauckas, Ph.D., BC Corrections

Starting in 2005, the Canadian province of British Columbia began assessing all sex offenders on community supervision using the Static-99R. The current study examined all Static-99R scores entered between 2005 and 2012, with an average follow-up of approximately five years ($n > 4,000$). The distribution of total scores in British Columbia will be compared to existing Static-99R percentile norms. Analyses of the scale’s discrimination accuracy will include AUCs and results from Cox and logistic regression. Analyses of the scale’s calibration accuracy will include logistic regression and the E/O effect size index to compare how the recidivism rates at each score compare to the rates predicted for routine, correctional samples. The findings from the current study will be compared to previous field and research studies of Static-99/R, with implications for applied assessments discussed.

Learning Goals and Objectives
After attending this presentation, audience members will be able to:

- Compare the distribution of risk scores in British Columbia to the current normative data on Static-99R percentiles
- Discuss the discrimination accuracy of Static-99R in British Columbia
- Compare the discrimination accuracy of Static-99R in British Columbia to other field and research studies of the scale
- Discuss the calibration accuracy of Static-99R in British Columbia
- Compare the calibration accuracy of Static-99R in British Columbia to other field and research studies of the scale
The Field Validity of Static-99/R among Sex Offenders in California

Seung C. Lee, M.A., Carleton University

Since 2007 in California, Static-99/R has been designated as the official static risk assessment instrument for adult male sex offenders. Static-99/R is currently used in California for diverse decision-making purposes with sexual offenders (e.g., sentencing, sex offender registry, levels of supervision, or GPS monitoring device). Given the important role of Static-99/R in the criminal justice system of California, the validity of Static-99/R in this jurisdiction should be empirically investigated. Hanson et al. (2014) found good predictive accuracy of Static-99/R (AUC = .824 and .817, respectively) and acceptable fit between expected and observed recidivism rates among 475 adult male sex offenders in the State of California. The purpose of this study is to replicate the previous study with a larger sample size (n = 1,198 from prison and n = 428 from probation samples). First, we examined the ability of Static-99/R to discriminate between recidivists and nonrecidivists with a fixed five-year period by using odds ratios from logistic regression and the area under the curve (AUC) from receiver operating characteristics (ROC) analysis. Furthermore, we conducted calibration analyses by using E/O index as well as fixed-effect meta-analysis by comparing logistic regression parameters of current study with new norms for routine samples of Static-99R from Hanson et al. (2015). The current findings support previous research by Hanson and colleagues (2014) demonstrating good discrimination and adequate calibration for Static-99/R as used in the State of California.

Learning Goals and Objectives
After attending this presentation, audience members will be able to:

- Discuss how Static-99/R is implemented in the criminal justice system with diverse decision-making purposes in California
- Discuss the appropriate methods to evaluate the predictive accuracy of risk assessment instruments (e.g., Static-99/R) and how to interpret them
- Discuss the discrimination and calibration accuracy of Static-99/R in California
- Compare the discrimination and calibration accuracy of Static-99R in California with other filed studies and new norms of Static-99R
- Discuss the overall field validity of Static-99/R