Preselection Effects Can Explain Group Differences in Sexual Recidivism Base Rates in Static-99R and Static-2002R Validation Studies

R. Karl Hanson & David Thornton

Association for the Treatment of Sexual Abusers
October 19, 2012
Denver, Colorado
Karl.Hanson@ps.gc.ca; David.Thornton@wi.gov
Base Rates and Risk Prediction

- A base rate is the proportion of a pre-defined group (or category) that will experience the adverse outcome.
- Risk prediction concerns future events; consequently, it can never been known with certainty.
- Future base rates estimated from past experience with similar cases.
The Evolution of Static-99

• The absolute recidivism rates for the original Static-99 was tested on 3 distinct samples
• No significant variability was observed
• Consequently, all samples were combined into one (and only one) recidivism rate table
STATIC-99
Sexual Reconviction Rates

Years after release

NOTE: Untreated Samples
STATIC-99
Sexual Reconviction Rates

Years after release

- Low
- Medium-Low
- Medium-High
- High
Life Gets More Complicated

- Updated the norms in 2008 (Harris et al., ATSA presentation) and 2009 (Helmus’ MA Thesis)
- Significant variability was found
- The differences in recidivism rates across samples was large enough to matter
Leslie Helmus, R. Karl Hanson, David Thornton, Kelly M. Babchishin and Andrew J. R. Harris
*Criminal Justice and Behavior* 2012 39: 1148 originally published online 21 May 2012
DOI: 10.1177/0093854812443648

The online version of this article can be found at:
http://cjb.sagepub.com/content/39/9/1148
Questions?

• Is the variability random?
• Different jurisdictions?
• Different definitions of recidivism?
• Preselection effects on risk relevant variables?
Degree of Preselection

• None
  – Routine samples, not obviously selected on risk relevant variables

• Some
  – mostly treatment samples, demonstrating some criminogenic needs

• Lots
  – High Risk/High Needs samples, explicitly selected on risk relevant variables

Rater Reliability: Kappa = .92 (95% agreement; 19/20)
5 Year Sexual Recidivism: Static-99R

![Graph showing 5 Year Sexual Recidivism: Static-99R with different treatment needs and risk levels.](image)

- **Static-99R score**
- **(%)**

- **Routine**
- **Treatment Need**
- **High Risk/Need**
Part 1

• Evidence of preselection effects on risk relevant variables
Samples: Study 1

- 20 samples from the Static-99 re-norming project
  - N ranged from 133 to 1,278 (N = 7,778)
  - Majority from Canada (k = 9) or the U.S. (k = 5)
- Studies excluded if they:
  - (1) Could not be classified into a sample type
    - pre-selected low risk offenders (Cortoni & Nunes, 2007)
    - only sexual murderers (Hill et al., 2008)
  - (2) Were identified as a statistical outlier
    - Saum (2007)
- None of the samples were preselected based on Static-99/R scores
Sample Types Score Differently on Sexual and General Criminality

- No Preselection
- Some Preselection
- Preselected High Risk/Need
Hypothetical Pattern: No Significant Variability on Base Rates

- No Preselection
- Some Preselection
- Preselected High Risk/Need
Observed Pattern: Sample Types Differ on Sexual Recidivism Rates

- Mean Static-99R Score
  - No Preselection
  - Some Preselection
  - Preselected High Risk/Need

5-Year Sexual Recidivism (%)

- No Preselection
- Some Preselection
- Preselected High Risk/Need
Part 1 Conclusions

• Selection process resulted in meaningful differences in Static-99R scores
• Incremental effect of sample type indicates that they were also preselected on risk factors external to Static-99R
Part 2: Evidence of Preselection Effects on Risk Factors External to Static-99R

• Are different samples different in the density of external risk factors?
• And if so, by how much?
Samples for Study 2

• Studies included if reported:
  – (1) On measure that was found to provide incremental validity over Static-99 or Static-99R
  – (2) Could be classified in one of the three sample types

• Measures included if:
  – Means and standard deviations for at least two of the sample types

• Three measures selected:
  – VRS-SO (Olver et al., 2007)
  – SRA-FV (Thornton)
  – STABLE-2007 (Hanson et al., 2007)
Sample: Study 2

• $K = 19$; $n$ ranged from 15 to 792 (Total $N = 3,976$)
  – Most from Canada ($k = 8$) and the U.S. ($k = 7$)
  – Measures:
    • VRS-SO ($N = 991$; $k = 5$), SRA-FV ($N = 952$; $k = 5$), and STABLE-2007 ($N = 2,033$; $k = 9$)
  – Sample types:
    • Routine ($N = 1,198$; $k = 2$), Treatment ($N = 1,566$; $k = 12$), and High risk/high need ($N = 1,212$; $k = 5$)
Transforming to a Common Metric (Z scores)

\[ Z = \frac{X_i - \bar{X}}{SD} \]

Ideal:

\[ Z = \frac{(Score_i) - (Population\ Mean)}{SD\ for\ population} \]

Available:

\[ Z = \frac{(Score_i) - (Treatment\ Mean)}{Total\ SD\ for\ treatment\ groups} \]
Sample Type Differ in their Degree of Risk Factors External to Static-99/R

- High risk samples
- Treatment samples
- Routine samples
Part 2 Conclusions

• Consistent differences in external risk factors based on sample type
• High Risk/High Need samples 1 SD above Treatment Samples
• Routine samples 1 SD below Treatment Sample
# Average Scores on External Risk Factors based on Degree of Preselection

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Z-score</th>
<th>Weighted $M$</th>
<th>$N$ ($k$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>VRS-SO</td>
<td>SRA-FV</td>
</tr>
<tr>
<td>Routine</td>
<td>-1.06</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Preselected for Treatment</td>
<td>0.0</td>
<td>20.74</td>
<td>2.22</td>
</tr>
<tr>
<td>Preselected High Risk/Need</td>
<td>1.04</td>
<td>27.24</td>
<td>3.26</td>
</tr>
</tbody>
</table>

Note: Z-scores based on M(SD) of Treatment Groups
Part 3

• How Well Can Individual Differences on External Risk Factors Estimate Sample Type Recidivism Rates Norms?
Samples: Study 3

- Four datasets that included the Static-99/R AND the VRS-SO, SRA-FV, and/or Stable-2007
- Fixed 5-year follow-up

<table>
<thead>
<tr>
<th>Sample</th>
<th>Country</th>
<th>Measure</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eher et al. (2012)</td>
<td>Germany</td>
<td>Stable-2007</td>
<td>259</td>
</tr>
<tr>
<td>Hanson et al. (2012)</td>
<td>Canada</td>
<td>Stable-2007</td>
<td>262</td>
</tr>
<tr>
<td>Olver et al. (unpublished data)</td>
<td>Canada and New Zealand</td>
<td>VRS-SO</td>
<td>538</td>
</tr>
<tr>
<td>Thornton (unpublished data)</td>
<td>U.S.</td>
<td>SRA-FV</td>
<td>418</td>
</tr>
</tbody>
</table>
Calculating Estimates 1: Study Specific Estimates

• The basic equation is:

Logit = $B_0 + B_1 \times (\text{STATIC}-99R) + B_2$ (other measure)

Logit = $(B0 \pm [1\text{SD adjustment}]) + B_1 \times (\text{STATIC}-99R)$
Calculating Estimates 2
Averaged Across Studies

- $B_0$ is the logit of the recidivism rate of the treatment group (7.2% or $B_2 = -2.5563$) for a Static-99R score of 2, used as a constant reference point.
- $B_1$ for Static-99R is the relative risk estimate from Helmus, Hanson et al. (2012)($B_1 = .293$).
Calculating Estimates 3

- $B_2$ is the average incremental effect of the external risk measures across the 4 samples in Z units ($B_2 = 0.5273$)
- Therefore, the $1 \text{ SD}$ adjustment can take on three values:
  - $-1 \times 0.5273 = -0.5273$
  - $0 \times 0.5273 = 0$
  - $1 \times 0.5273 = 0.5273$
Calculating Estimates - 4

- For routine:
  \[ \text{logit} = (-2.5563 - 0.523) + .293 \text{ (Static-99R scores)} \]

- For HR/HN:
  \[ \text{logit} = (-2.5563 + 0.523) + .293 \text{ (Static-99R scores)} \]

- For Treatment:
  \[ \text{logit} = (-2.5563 + 0) + .293 \text{ (Static-99R scores)} \]
Expected Sexual Recidivism Rates after Adjusting for Presence of External Risk Factors

![Graph showing expected sexual recidivism rates with different risk scores and treatment samples. The x-axis represents Static-99R Score, ranging from -3 to 12. The y-axis represents Expected Sexual Recidivism (%), ranging from 0 to 70. The graph includes three lines: 1 SD above treatment samples (solid dark line), Treatment samples (dashed dark line), and 1 SD below treatment samples (solid orange line).]
How Well Does Adjusting for External Risk Factors Match the Static-99R Sample Type Norms?

![Graph showing expected sexual recidivism (%) against Static-99R score for different sample types: High risk samples, 1 SD above treatment samples, Treatment samples, 1 SD below treatment samples, and Routine samples. The graph illustrates the relationship between Static-99R score and expected sexual recidivism for each sample type.]
Have We Explained It All?

• Is there residual variability left over after accounting for Static-99/R and the density of external risk factors?
Part 3 Conclusions

• The Static-99R sample type recidivism rates closely match the recidivism rates expected for offenders who have different density of criminogenic needs

• Considering external risk factors with Static-99R improve the calibration of risk predictions

• Further research is needed on other sources of residual variability across samples
General Conclusion

• There are strong preselection effects on risk relevant variables across samples
• The Static-99R sample types can be interpreted as corresponding to groups that are 1 SD above (High Risk/High Need) or 1 SD below (Routine) the density of criminogenic needs found in treatment samples
• Several different measures can be used to assess the density of criminogenic needs
Acknowledgements

• Co-authors
  – Kelly Babchishin, Leslie Helmus

• Research assistants
  – Chelsea Sheahan, Heather VanZuylen, Rebecca Mugford
Acknowledgements

• Data Sharing
  – Alfred Allan, Tony Beech, Susanne Bengtson, Jacques Bigras, Sasha Boer, Jim Bonta, Sébastien Brouillette-Alarie, Franca Cortoni, Jackie Craissati, Margretta Dwyer, Reinhard Eher, Doug Epperson, Tina Garby, Randolph Grace, Steve Gray, Andy Haag, Leigh Harkins, Grant Harris, Andreas Hill, Steve Johansen, Ray Knight, Niklas Långström, Terry Nicholaichuk, Kevin Nunes, Mark Olver, Jean Proulx, Martin Rettenberger, Marnie Rice, Steve Saum, Rebecca Swinburne Romine, Daryl Ternowski, Robin Wilson, and Annie Yessine.
Contact Information

Karl.Hanson@ps.gc.ca
David.Thornton@wi.gov

www.Static99.org