Washington State Institute for Public Policy

110 Fifth Avenue SE, Suite 214 • PO Box 40999 • Olympia, WA 98504 • 360.586.2677 • www.wsipp.wa.gov

February 2014

Predicting Criminal Recidivism:
A Systematic Review of Offender Risk Assessments in Washington State

Summary

The 2009 Legislature required the Department of Corrections (DOC) to use a risk assessment, recommended by the Washington State Institute for Public Policy (WSIPP), which has the highest predictive accuracy for recidivism.

To complete this task, WSIPP employed a systematic research approach. We reviewed the research literature on risk assessments and found five that have been tested on adult offenders in Washington. Among the five options, our review indicates that, to date, the Static Risk and Offender Needs Guide-Revised (STRONG-R) has the highest predictive accuracy of criminal recidivism.

Under Washington State's sentencing laws, an adult convicted of a felony in superior court receives a sentence as prescribed within the ranges of the state's sentencing guidelines. Depending on the seriousness of the crime and a person's criminal history, some sentences may result in confinement in prison, community supervision, or both.[1] The Department of Corrections (DOC) has jurisdiction over offenders sentenced to more than one year of confinement as well as those who receive a sentence of supervision in the community.[2]

In 1999, the Legislature enacted the Offender Accountability Act (OAA) that set state policy regarding the intensity of community supervision.[3] The law requires DOC to classify offenders according to their future risk for re-offense and the harm they have caused society in the past. DOC must deploy more staff and rehabilitative resources to higher-risk offenders. Since the passage of the OAA, DOC has implemented two different risk assessments to assist with the classification of offenders.[4]

The 2009 Legislature required DOC to use a risk assessment "recommended to the department by the Washington State Institute for Public Policy as having the highest degree of predictive accuracy for assessing an offender's risk of re-offense."[5] We focus our systematic review on assessments that have been tested on offender populations in Washington State.

The Washington State Institute for Public Policy (WSIPP) was approached in 2012 by DOC to determine if a new risk assessment under consideration by DOC has the highest degree of predictive accuracy of future recidivism.

To fulfill this legislative requirement, WSIPP systematically reviewed the literature on risk assessments that have been statistically "validated." That is, we examined tools developed and tested on offenders in Washington to determine the degree of accuracy of predicting recidivism.

Suggested citation: Drake, E. (2014). *Predicting criminal recidivism: A systematic review of offender risk assessments in Washington State* (Doc. No. 14-02-1901). Olympia: Washington State Institute for Public Policy.

[1] RCW 9.94A, RCW 9.94A.501, and RCW 9.94A.701 & 702.
[2] See, for example, RCW 9.94A.190 and RCW 9.94A.501.
[3] Engrossed Second Substitute Senate Bill 5421, Chapter 196, Laws of 1999.
[4] Barnoski, R. & Drake, E. (2007). *Washington's Offender Accountability Act: Department of Corrections' static risk assessment* (Doc. No. 07-03-1201). Olympia: Washington State Institute for Public Policy.
[5] Engrossed Substitute Senate Bill 5288, Chapter 375, Laws of 2009.

Background

Risk assessments have been an essential function of correctional agencies since punishments were first handed down to criminals.[6] To date, criminologists acknowledge four "generations" of assessments.[7]

First generation (1G) assessments are unstructured clinical judgments, typically by psychologists, on the likelihood that an offender will commit a new crime. Second generation (2G) assessments are empirical assessments based on "static" risk factors (see key terms in sidebar box) that estimate the probability that an offender will commit a new crime. Third generation (3G) assessments are also empirically based on static risk factors, but include "dynamic" risk factors or criminogenic needs.

Fourth generation (4G) assessments, the latest generation, include static and dynamic risk factors but have two important differences compared with 3G instruments. First, 4G instruments incorporate the "responsivity" principle, which is the concept that interventions must be aligned with the offender's motivation and abilities. The second difference is that 4G instruments allow for reassessment during the case management process enabling corrections staff to measure changes in the offender's behavior over time.

DOC currently uses the Static Risk Assessment (SRA), a 2G assessment based on static risk factors (e.g., criminal history).[8] A separate needs assessment is used in conjunction with the SRA to refer offenders into appropriate interventions.

In 2013, DOC issued a request for information on 4G risk and needs assessments. The goal was to procure an assessment to be the centerpiece of its case management system.[9] The assessment would align offenders with programs based on their risk factors (e.g., education or chemical dependency treatment). In addition, DOC sought a risk assessment that would allow for reassessment throughout the case management process to determine if desired changes were taking place.

DOC awarded the contract to Dr. Zach Hamilton, Director of the Institute for Criminal Justice Research at the Washington State University. After Dr. Hamilton developed the new risk assessment, DOC approached WSIPP to determine if WSIPP would recommend the tool as "having the highest degree of predictive accuracy for assessing an offender's risk of re-offense."[10] As we do with other research topics, WSIPP takes a scientific and systematic approach in this report to review the research literature on validated risk assessments.[11]

Key Terms

What Are "Static" and "Dynamic" Risk Factors?[a]
Risk factors that cannot decrease, such as criminal history, are static. Once a criminal record is obtained, it will always be a part of an offender's history. Dynamic risk factors, such as drug dependency, can change through treatment.

What Is "Risk Need & Responsivity (RNR)"?[b]
This term was developed by Canadian researchers in 1990 and is defined as follows:

- ✓ Risk principle: use interventions commensurate with risk for re-offense.
- ✓ Need principle: target criminogenic risk factors such as anti-social attitudes or substance use.
- ✓ Responsivity principle: use interventions aligned with the offender's abilities and motivation (focusing on cognitive behavioral or social learning interventions).

[a] D.A. Andrews & J. Bonta. (1998). *The psychology of criminal conduct.* Cincinnati, Ohio: Anderson Publishing Co.
[b] Andrews, D., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19–52.

[6] Bonta, J. Risk-needs assessment and treatment. In Harland, A. (1996) *Choosing correctional options that work.* (pg. 18-32). Thousand Oaks, CA: Sage.
[7] Andrews, D.A., Bonta, J., & Wormith, S. (2006). *The recent past and near future of risk and/or need assessment. Crime & Delinquency, 52*(1).
[8] Barnoski & Drake, (2007).
[9] Communication with DOC, October 31, 2012.
[10] Engrossed Substitute Senate Bill 5288, Chapter 375, Laws of 2009.
[11] For example, see: Drake, E. (2013). *Inventory of evidence-based and research-based programs for adult corrections* (Doc. No. 13-12-1901). Olympia: Washington State Institute for Public Policy.

Research Approach

The research question at hand is: what tool has the highest predictive accuracy for assessing an offender's risk of re-offense in Washington State? To empirically assess this policy question, we systematically reviewed the research literature. We established the following relevant research criteria and included all studies that met these criteria in our review.

First, to be included in our systematic review, the risk assessment had to measure recidivism as an outcome. Thus, we excluded studies of risk assessments with outcomes such as pre-trial risk of flight or pre-trail threats to public safety.

Second, we included risk assessments designed for general offender populations as opposed to special populations (e.g., sex offenders or domestic violence offenders). We employed this criterion because DOC is interested in a risk instrument for classification purposes throughout their population.

Third, to be included in our systematic review, the risk assessment must be a "validated" instrument. Empirical risk assessments are "constructed" using a sample of offenders to determine which characteristics are most predictive of recidivism. Then, the instrument is "validated" or tested on a different sample of offenders to determine how well the risk assessment performs.

Fourth, the risk assessment must have been tested (validated) on a Washington State DOC offender population. Some experts in the field believe that risk assessments are not transportable across offender populations since the assessment has been designed based on the offender characteristics of that specific jurisdiction.[12] Thus, we do not directly compare the predictive accuracy of risk assessment results across different states or jurisdictions.

Fifth, studies must report a statistical measure called the Area Under the Curve (AUC). The AUC is a commonly used statistic that measures the strength of association between risk classification and recidivism.[13] AUCs range from 0.500 to 1.000 with higher AUCs demonstrating higher predictive accuracy for assessing an offender's risk.

After systematically reviewing the literature and employing our criteria, we found five risk assessments that have been tested on the DOC offender population in Washington:[14]

1) Level of Service Inventory-Revised (LSI-R)
2) Static Risk and Offender Needs Guide-Revised (STRONG-R)
3) Static Risk Assessment (SRA)[15]
4) Static Risk Assessment, revised (SRA2)[16]
5) Ohio Risk Assessment System (ORAS)

[12] See, for example, Duwe, G. (2013). The development, validity, and reliability of the Minnesota screening tool assessing recidivism risk (MnSTARR). *Criminal Justice Policy Review*, XX, 1-35. DOI: 10.1177/0887403413478821. Some experts also question how comparable validated risk assessments are for offenders in the same jurisdiction, but at different time periods.

[13] Some studies report correlation coefficients, but the AUC is a better measure of association since recidivism is a dichotomous outcome.

[14] The studies in our review were (1) Barnoski, R. & Aos, S. (2003). *Washington's Offender Accountability Act: An analysis of the Department of Corrections' risk assessment* (Doc. No. 03-12-1202). Olympia: Washington State Institute for Public Policy. (2) Hamilton, Z., Kigerl, A., Campagna, M., Barnoski, R., Block, L. & Lee, S. (2014). The development and validation of recidivism risk assessment: Introducing the STRONG-R. Working document.

[15] The AUC reported in Barnoski & Drake, 2007 (0.742) was validated on a population 8 years prior to the population used to validate the SRA in Hamilton et al., 2013. We use SRA validation results reported in Hamilton et al., 2013 in lieu of the results reported in Barnoski & Drake, 2007 because Hamilton uses the same offender population (N=35,788) making the performance of all four assessments directly comparable.

[16] Information on this assessment can be found in: Barnoski, R. (2010) *Washington State static risk assessment—version 2.0.* [Modified to improve reliability and validity, requested by Washington State Center for Court Research].

Findings & Recommendations

We were able to compare the results of five risk instruments to determine which has the highest predictive accuracy for assessing an offender's risk for re-offense in Washington State. The results are displayed in Exhibit 1.

As previously mentioned, the AUC is used to measure the strength of association between risk classification and recidivism. Risk assessments with higher AUCs demonstrate higher predictive accuracy. Among the five assessments we were able to review, DOC's proposed tool—the STRONG-R—has the highest AUC at 0.720. Thus, WSIPP recommends the STRONG-R as the instrument with the highest predictive accuracy of risk for recidivism.

Considerations

The goal of determining which risk assessment has the highest predictive accuracy is to find an assessment that has the greatest ability to minimize error. It is important to note, however, that no risk assessment is 100% accurate.

Two types of errors are relevant to risk assessment prediction:

1) Type I errors, or false positives, occur when an offender is classified as a potential recidivist, but does not commit a new crime.
2) Type II errors, or false negatives, occur when an offender is not classified as a potential recidivist, but does commit a new crime.

The AUC measures how accurately the instrument classifies an offender (recidivist or not) compared to actual (observed) recidivism. While the STRONG-R has the highest AUC of the five tools we were able to review, it is certainly not 100% accurate.

As offender populations change in DOC over time, so will the risk characteristics that are predictive of recidivism. Thus, in order to maintain the highest level of predictive accuracy, it is recommended that the instrument used by DOC be tested and re-weighted periodically.

Exhibit 1

Studies that Test Risk Assessments on a DOC Offender Population in Washington State

Study/risk assessment	AUC*	Number in validation sample
Level of Service Inventory-Revised (LSI-R)	0.660	22,533
Ohio Risk Assessment System-WA (ORAS-WA)	0.660	35,788
Static Risk and Offender Needs Guide-Revised (STRONG-R)*	0.720	35,788
Static Risk Assessment (SRA)	0.689	35,788
Static Risk Assessment, revised (SRA2)	0.660	35,788

Notes: The area under the curve (AUC) is a commonly used statistic that measures the strength of association between risk classification and recidivism. AUCs range from 0.500 to 1.000 and larger AUCs demonstrate higher predictive accuracy.

* For comparison purposes, the AUCs reported here are for male and female populations combined. The STRONG-R has different weights, thus different AUCs for male and female populations (0.720 and 0.700, respectively).

For further information, contact:
Elizabeth Drake at 360.586.2767, Elizabeth.Drake@wsipp.wa.gov

Document No. 14-02-1901