Examining the Validity and Reliability of the Ohio Risk Assessment System Community Supervision Tool and Community Supervision Screening Tool

Executive Summary

Prepared for the Commonwealth of Massachusetts
Office of the Commissioner of Probation

Edward Latessa, Ph.D.
Principal Investigator

Jennifer Lux, Ph.D.
Project Director

Melissa Lugo, M.S.
Research Assistant

Joshua Long, M.S.
Research Assistant

University of Cincinnati Corrections Institute

Submitted on: October 14, 2017
OVERVIEW

The Massachusetts Probation Service (MPS), adopted the Ohio Risk Assessment System Community Supervision Screening Tool (ORAS-CSST) and the Ohio Risk Assessment System Community Supervision Tool (ORAS-CST) in 2013 to promote law-abiding behavior in the community and to reduce recidivism. Although the reliability and validity of the ORAS-CSST and ORAS-CST have been examined in other states, their efficacy has not been assessed specifically for Massachusetts. As such, MPS contracted with the University of Cincinnati Corrections Institute (UCCI) in January 2016 to evaluate the validity and reliability of both instruments. The current report presents the results of this study, which was designed to meet the following goals:

1. Assess the inter-rater reliability of the ORAS-CST;

2. Examine normative information for the ORAS-CSST and ORAS-CST for MPS clients and evaluate the tools’ internal reliability;

3. Evaluate the predictive validity of the ORAS-CSST and ORAS-CST, as well as examine the validity of the tools for different sub-groups (e.g., gender, race); and

4. Develop recommendations for revision, as appropriate, to include changes to scoring guides and adjust cutoff scores.

METHODS

Inter-Rater Reliability

In order to assess inter-rater reliability, ORAS-CST vignettes and an officer characteristics survey were administered to staff using Qualtrics, a web-based survey platform. Staff received a total of five vignettes (three male and two female) and scored out an ORAS-CST instrument for each vignette based on the information provided. The officer characteristics survey collected data concerning staff’s demographic information (e.g., position in the agency, gender, education level), work experience, experience conducting risk assessments with offenders, and whether officers received training to administer ORAS assessments. The inter-rater reliability sample consisted of 120 participants.

Norming Data, Internal Reliability, & Predictive Validity

This portion of the study included a sample of adult offenders under MPS supervision in the state of Massachusetts. The sample consisted of all offenders who received an ORAS-CSST and/or ORAS-CST assessment in 2014. These assessments were limited to the initial ORAS assessment for the offender and only one assessment per offender if he/she had multiple assessments during the time period. The resulting sample was 11,640 offenders for the ORAS-CSST and 10,548 for the ORAS-CST.¹

¹ Due to the large sample size for both tools, the initial samples for the ORAS-CSST and the ORAS-CST were randomly split in half to create a development sample and a validation sample. Results conducted in the development sample were then replicated in the validation sample to ensure findings were not due to chance. Findings were nearly identical across the two samples, therefore, the results presented in the current report are based on the full sample.
RESULTS

The following section briefly summarizes the findings and conclusions that can be drawn from the validation study:

Inter-Rater Reliability

- High levels of agreement among raters were observed for the majority of average ORAS-CST domain scores. On average, the Education, Employment, and Financial Situation and Neighborhood Problems domains produced the highest rates of reliability. Conversely, the Peer Associations and Criminal Attitudes and Behavioral Patterns sections produced the lowest rates of reliability, on average.

- Average inter-rater agreement overall fell just below 80 percent. A substantial number of individual items produced less than 80 percent agreement levels. Specifically, 18 items yielded rates of inter-rater reliability below the 80 percent threshold.

- Three of the four ORAS-CSST items produced inter-rater agreement rates below 80 percent.

- Kalpha analyses revealed that the Criminal History, Neighborhood Problems, and Peer Associations domains were the only areas that met the “acceptable reliability” threshold.

- Kalpha analyses also indicated that both the total risk score and overall risk level across raters was reliable with values meeting the threshold of “acceptable reliability”.

- No significant differences were found between officer characteristics (e.g., gender, role, region) and rates of agreement.

- Agreement rates were lower for two of the five vignettes—Jane Smith and Alfredo Diaz.

Predictive Validity

- For both new arraignments and new convictions, offenders recidivated at an increasingly greater rate as levels of risk increased. These relationships were found to be significant for both ORAS-CSST and ORAS-CST instruments for both outcomes of interest.

- The “not low-risk” ORAS-CSST subgroup showed higher rates of new arraignments and convictions compared to the low-risk ORAS-CSST subgroup. However, the low-risk group failed at a higher rate than would be expected.

- ORAS-CSST and ORAS-CST total scores were significantly correlated with new arraignment and new conviction outcome measures across the overall sample, as well as for each subgroup.
• Overall, the ORAS-CSST and ORAS-CST appear to predict future arraignments with slightly more accuracy compared to convictions. The ORAS-CSST is able to predict new arraignments and convictions for both gender and racial subgroups. The ORAS-CST was predictive of new arraignments and convictions for white and non-white males.

• The ORAS-CST does not predict recidivism outcomes equally well for females. The results suggest it predicts new arraignments and convictions among non-white females. Among white females, the ORAS-CST is predictive of new convictions; however, high-risk white females are slightly less likely to have a new arraignment compared to the moderate-risk group. These findings may be due to the small sample size for high-risk females. Out of the 1,382 females in the ORAS-CST sample, only 63 were categorized as high-risk (4.5%). Further, high-risk females represented 0.6 percent of the 10,548 individuals in the entire ORAS-CST sample.

• When examining individual items, both tools appear to predict future arraignments best. This was also true for domain scores on the ORAS-CST. All correlations were positive and the majority were significant for new arraignments and new convictions.

• ROC area under the curve analyses indicated that the ORAS-CSST and ORAS-CST instruments predicted offenders’ recidivism better than chance for two recidivism measures—new arraignment and new conviction.

RECOMMENDATIONS

Based on these results, several recommendations can be made:

1. Policies and procedures concerning the risk assessment process should be updated based on the results of the study. Staff should be made aware of these policy/procedure changes in writing. MPS may also wish to develop additional resources to enhance the efficacy of the risk assessment. For example, an “assessment checklist” could be created that outlines the steps that should be taken when conducting/scoring an ORAS.

2. The results from the current study should be examined to identify which items are areas of concern for staff. These items include those that received less than 80.0 percent agreement and a Kalpha score lower than 0.67. MPS may want to have a group discussion regarding these items and take time to review the scoring criteria for these tools.

3. MPS should continue maintaining its current ORAS Committee and its ongoing strategic planning efforts in the area of risk assessment. Additionally, MPS should hire a continuous quality improvement team charged with the advancement of risk assessment practices and procedures. The team would be tasked with developing strategies to improve inter-rater reliability and assessment implementation, including, but not limited to, providing training, coaching, auditing, and other capacity building tasks to field staff in the area of risk assessment. The existing ORAS Committee, in conjunction with the continuous quality improvement team, should consider developing structured feedback forms outlining the strengths and areas of improvement for end-users’ interviews. MPS should continue their
inter-rater reliability efforts as well as improve these efforts by randomly auditing selected cases and recording end-users’ interviews to allow for multiple assessors of the same assessment.

4. The department may want to consider implementing a new screener. This tool may be created based on the data from the current study or MPS may want to consider using the ORAS-Misdemeanor Assessment Tool (ORAS-MAT).

5. The ORAS-CST did not predict the rate of recidivism as expected for female offenders with high-risk females failing at a slightly lower rate than moderate-risk females for both new arraignments and convictions. MPS may want to consider two approaches to improving the predictive accuracy of the tool. These options include: 1) reevaluating the predictive validity of the ORAS-CST with a larger sample of female offenders or 2) adjusting cut-off scores for females so more cases are categorized as high-risk.

**CONCLUSION**

The overall results of the current study suggest that MPS should focus its risk assessment related efforts on improving inter-rater agreement. In particular, the Department may wish to incorporate the present findings into their current policies and procedures manual discussing problematic areas on the tools. Additionally, MPS should hire a continuous quality improvement team tasked with the advancement of risk assessment practices and procedures. The team would be responsible for developing strategies and implementing these strategies to improve inter-rater reliability and assessment implementation. These strategies may include, but are not limited to, providing training, coaching, auditing, and other capacity building tasks to properly equip staff in the area of risk assessment. In order to sustain the efforts in improving assessment implementation, MPS should continue to engage in quality assurance checks to monitor consistency across end-users. Other considerations include adopting a new screener tool and reevaluating the predictive validity of the ORAS-CST with a larger sample of female probationers or modifying the ORAS-CST cut-off scores for females.