Advancing Fairness and Transparency:
National Guidelines for Post-Conviction Risk and Needs Assessment

Sarah L. Desmarais, David A. D'Amora, Lahiz P. Tavárez

CSG **Justice Center**
THE COUNCIL OF STATE GOVERNMENTS

BJA
Bureau of Justice Assistance
U.S. Department of Justice

Table of Contents

Dedication

We are pleased to dedicate this project to Dr. Ed Latessa. Dr. Latessa spent his entire career working to improve the criminal justice system, especially in the areas of risk and needs assessment and correctional programming. He was an invaluable advisor to The Council of State Governments (CSG) Justice Center and the Bureau of Justice Assistance within the U.S. Department of Justice's Office of Justice Programs for well over a decade. His advice and counsel improved our knowledge and strengthened the impact of our work in the field. Dr. Latessa transformed the fields of community supervision and correctional habilitation and treatment in the United States, and for that, we are eternally grateful.

Acknowledgments

This project was accomplished through the collaborative efforts of researchers, risk and needs assessment instrument developers, practitioners, and leaders in the field who gave generously of their time and expertise over the course of 2.5 years. It draws on an extensive review of literature and related research, observations from the field, feedback from national experts, several multidisciplinary forums and advisory group discussions, and a rigorous review process.

Thank you to this publication's group of expert advisors whose input greatly informed the development of these guidelines. The following collaborators provided insights and resources that were truly invaluable:

Project Credits

Writing: **Sarah L. Desmarais**, Policy Research Associates; **David A. D'Amora**, CSG Justice Center; **Lahiz P. Tavárez**, CSG Justice Center

Advising: **Elizabeth Lyon**, **Marshall Clement**, **Megan Quattlebaum**, CSG Justice Center

Editing: **Katy Albis**, **Leslie Griffin**, **Emily Morgan**, CSG Justice Center

Public Affairs: **Aisha Jamil**, CSG Justice Center

Design: **Michael Bierman**

Web Development: **Eleventy Group**

Advisory Group

Ashlee Barnes, Virginia Commonwealth University

Mark Bergstrom, Pennsylvania Commission on Sentencing

Guy Bourgon, Public Safety Canada (ret.), Clinical and Consulting Psychology Practice

Kristofer "Bret" Bucklen, Pennsylvania Department of Corrections

Christina Campbell, University of Cincinnati

Bree Derrick, Idaho Department of Correction

Kevin Douglas, Simon Fraser University

Grant Duwe, Minnesota Department of Corrections

Joel Dvoskin, University of Arizona College of Medicine

Carey Green, Criminal Justice Assistance Division, Texas Department of Criminal Justice

Stephen Hart, Protect International Risk and Safety Services

Kirk Heilbrun, Drexel University

Natalie Jennifer Jones, Evaluation Division of the Social Sciences and Humanities Research Council of Canada

KiDeuk Kim, Urban Institute

Daryl Kroner, Southern Illinois University

Edward Latessa, University of Cincinnati

Brian Lovins, Justice System Partners

Eyitayo Onifade, Clark Atlanta University

Osonde Osoba, Center for Scalable Computing and Analysis, Pardee RAND Graduate School

Sarah Picard, Center for Court Innovation

Charles Robinson, Administrative Office of the United States Courts

Ralph Serin, Criminal Justice Decision Making Laboratory, Carleton University

Christopher Slobogin, Vanderbilt University Law School

Raymond "Chip" Tafrate, Central Connecticut State University

Faye Taxman, George Mason University

Gina Vincent, University of Massachusetts Medical School, Center for Mental Health Services Research

Introduction

Risk and needs assessment instruments are widely used in criminal justice settings across the United States to inform decisionmaking at various stages in case processing. These guidelines were developed to promote accuracy, fairness, transparency, and effective communication and use of risk and needs assessment instruments to inform decisionmaking following conviction.

Methodology

The guidelines are drawn from empirical research and reflect the perspectives of an advisory group comprising nearly 30 researchers, practitioners, and policymakers. Advisory group members participated in several discussions on key concerns in the development, validation, and implementation of post-conviction risk and needs assessment instruments. These guidelines also consider existing statements and guidance on the use of risk and needs assessment instruments written by diverse groups and for diverse contexts.

Usage

These guidelines pertain to the use of post-conviction risk and needs assessment instruments to inform decisions and case planning that occur after court disposition—specifically, after conviction and sentencing. They may also be used in the application of assessment results to inform decisionmaking and case planning in the context of alternative forms of criminal justice processing, such as after a decision has been made to offer a diversion program.

Timing

Many of the guidelines describe processes or steps that should ideally be taken prior to implementation of a post-conviction risk and needs assessment instrument. But many corrections and community supervision agencies are already using a post-conviction risk and needs assessment instrument. For those agencies, we recommend:

(1) Conducting an informal review of the extent to which current policies and practices adhere to the action items under each guideline.

(2) Setting specific goals and identifying priorities for implementing the remaining guidelines over a realistic time period.

Whatever the setting, the guidelines presume that the intended use of post-conviction risk and needs assessment instruments is to support accurate, fair, and transparent decisions regarding a person's risk of recidivism. These guidelines also presume that the purpose of post-conviction risk and needs assessment instruments is, ultimately, to promote public safety and positive outcomes for people in the criminal justice system through the least restrictive means possible.

Purpose

The purpose of this publication is to provide the information that criminal justice agencies need to ensure that the implementation of post-conviction risk and needs assessment instruments promotes accuracy, fairness, transparency, and effective communication and use. The sections of the publication are organized by these four principles. Each section provides the rationale for the related guideline, recommends actions that should be taken or requirements that must be met to follow the guideline, and reviews practical considerations for planning, implementation, and continuous quality improvement (CQI).

Intended Audience

The intended audience for this publication includes people who support agency administrators, supervisors, and other stakeholders involved in selecting or implementing post-conviction risk and needs assessment instruments, the development of related policy, and decisions regarding their ongoing use. These individuals may include trainers, quality assurance personnel, research partners, or other consultants. The content is also relevant for developers of post-conviction risk and needs assessment instruments and researchers or evaluators who may evaluate the performance of assessment results in studies or in practice. Assessors and their supervisors additionally may find utility in the information presented here to support training and CQI-related efforts.

Other stakeholders may find that some of the additional information provided herein supports a deeper understanding of the guidelines. For instance, system actors (e.g., judges, attorneys, service providers, or probation/parole officers) or people in contact with the criminal justice system may find that this publication helps them understand what informs the application of assessment results in individual case processing.

Principle I: Accuracy

The first four guidelines speak to strategies that agencies can use to promote accuracy in the use of post-conviction risk and needs assessment instruments. Accuracy refers to the degree to which the assessment results predict the recidivism outcomes they were designed to predict, as measured in relation to the observed rate and severity of criminal behavior. Promoting accuracy also involves considering whether the post-conviction risk and needs assessment instruments are completed and used as intended to inform case decisions and planning within facilities and in the community.

❝

Criminal justice is a domain where it is imperative to exercise maximal caution and humility in the deployment of statistical tools. (Partnership on AI, 2019, 33)

❞

We recommend the following guidelines to promote accuracy of post-conviction risk and needs assessment instruments:

1. Conduct a local evaluation of the post-conviction risk and needs assessment instrument to ensure that the instrument is suitable for the agency's population.
2. Meet the minimum performance thresholds of post-conviction risk and needs assessments completed in the field according to statistical standards.
3. Use a continuous quality improvement (CQI) process to ensure successful implementation of the post-conviction risk and needs assessment instrument.
4. Use a multistep approach to assess risk and needs over time.

Guideline 1: Conduct a local evaluation of the post-conviction risk and needs assessment instrument to ensure that the instrument is suitable for the agency's population.

Overview

This guideline establishes that the post-conviction risk and needs assessment instrument can be completed reliably and with acceptable levels of accuracy in predicting the outcome(s) of interest in practice. Many different factors can affect the reliability and validity of the assessment results, including assessors, information used to complete the assessment, resources, policies, and practices. A local evaluation may produce information that not only bears on the reliability and validity of the assessments but also elucidates potential issues or concerns in local practices and policies that should be addressed.

Action Items

Establish inter-rater reliability prior to using the instrument.

Inter-rater reliability is relevant to any assessment instrument to ensure that the assessment is completed consistently and accurately.[1] Indeed, establishing that assessments can be completed with consistency across independent assessors is a necessary criterion for establishing validity.[2] Even if the assessment does not involve an interview and most—or even all—of the items are completed using official records, there may be differences in how that information is extracted and interpreted by assessors or errors may occur in the coding process. To that end, all individuals who will be conducting the assessments should complete a minimum of three practice cases to consensus after training and prior to using the instrument in the field.[3]

There are different ways to ensure that these practice cases occur. For instance, the trainers hired to conduct the pre-service training (see Guideline 3) may provide additional case materials and ratings that agencies may use for these practice cases.[4] If not, agency representatives may ask the expert trainer or other qualified professional to help them develop practice cases and ratings. Alternatively, agencies may develop case studies that experienced in-house assessors have coded and use them for consensus ratings as new assessors are trained. Another strategy may be to have new assessors review case materials and complete the assessment for an individual who is currently being assessed—or has recently been assessed—by a more experienced assessor in the agency. Assessors should have these opportunities to practice *after training* but *before use in practice* to increase their understanding of the assessment process, get feedback on their ratings, and gain experience. Information on specific metrics for determining whether inter-rater reliability is acceptable is provided later on (see Guideline 2).

---

1. Douglas et al., "Clinical Forensic Psychology."
2. Douglas, Skeem, and Nicholson, "Research Methods in Violence Risk Assessment," 325–46; Gottfredson and Moriarty, "Statistical Risk Assessment," 178–200.
3. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice.*
4. Ibid.

Complete a local validation, ideally prior to using the instrument, to confirm that the assessment results are predicting recidivism using local data and in the context of current and local practices.

Agencies should complete a local validation to demonstrate predictive validity. Predictive validity is not a property of the post-conviction risk and needs assessment instrument itself but rather a property of the assessment results.[5] So, while we can be confident that an instrument already validated in large research studies or in other jurisdictions will produce reliable and valid assessment results, we cannot assume that the same level of reliability and predictive validity will be achieved locally. Factors that can affect the validity and reliability of assessment results include local record keeping practices; assessor attitudes, training, and knowledge; and variations in penal codes and base rates of recidivism.[6] Indeed, absent the necessary information and time, implementation with fidelity—and, consequently, reliability and validity—is not possible, even with highly motivated, knowledgeable, and well-trained staff.

Agencies should conduct a local validation study to establish performance in relation to jurisdiction-specific rates of recidivism, ideally prior to using the instrument.[7] This may be possible to achieve through a retrospective study design, for example, if assessors can extract the information needed to complete the assessment and document recidivism from existing records (e.g., jail/prison records, court records, etc.).

If the assessment requires information that is not available in local records or requires an interview, then an alternative study design would be to conduct a pilot implementation with a subset of cases and test the validity of these assessments prospectively (i.e., looking forward) prior to a full-scale implementation. A sample size of 500 people would likely be sufficient for the local validation.[8] Further discussion of the research methods for conducting local validations are available elsewhere (see, for example, the **Public Safety Risk Assessment Clearinghouse** ⧉). We provide information on specific metrics for determining whether predictive validity is acceptable later on (see Guideline 2).

Revalidate assessment results at least every 5 years—or sooner if there are major policy or population changes—to verify that the assessment results continue to meet minimum performance thresholds.

Regular revalidation will ensure that the assessment instruments continue to be completed as intended and that the results continue to demonstrate acceptable reliability and validity. There may be changes over time in the reliability and validity of assessment results for both expected and unexpected reasons. In particular, there may be meaningful changes in the makeup of the criminal justice population over time because of reforms in policing, charging, or prosecution, for example. Such changes will call for a re-examination of the assessment results to ensure that they continue to meet minimum performance thresholds (see Guideline 2).

---

5. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, *Standards for Educational and Psychological Testing*.
6. Mills, Jones, and Kroner, "An Examination of the Generalizability of the LSI-R and VRAG Probability Bins," 565–85; de Vogel and de Vries Robbé, "Adapting Risk Assessment Tools to New Jurisdictions," 26–39.
7. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice*.
8. For more information on this sample size estimation, see Hanson et al., *A Five-Level Risk and Needs System*.

Even in the absence of policy and practice changes, we recommend revalidation of assessment results at routine intervals. There will inevitably be drift from coding and administration protocols as time since training elapses or with staffing changes and turnover. As part of the planning process—ideally, prior to using the instrument— agencies should establish a timeline for revalidation and identify and allocate resources and staffing to support the revalidation. We recommend that revalidation occur at least every 5 years to balance system demands and resources and to allow sufficient time from implementation to evaluation for recidivism outcomes to be observed and documented.

Consult with experts such as university partners or other experienced evaluators, as needed, to ensure that local evaluations adhere as much as possible to best practices in risk and needs assessment research and standards in test validation.

There are many factors related to the design and methods of an evaluation that affect the reliability and validity of its findings.[9] To that end, there are established standards that should be applied, to the extent possible, to ensure that local validations are conducted in a sufficiently rigorous manner. These standards are found in psychological and educational testing,[10] accepted practices in risk assessment research methods,[11] and guidelines for reporting risk assessment research methods and findings.[12] Applying such standards will also promote the likelihood that the evaluation's findings are an accurate reflection of the performance of the post-conviction risk and needs assessment results and cannot be attributed to other factors.

Agencies may not have the in-house expertise, resources, and knowledge to design and field a validation study that would stand up to public and peer review. We recommend that agencies consult with experts such as university partners or other experienced evaluators to inform the methods of their local evaluation efforts.

9. For instance, validation studies should account for time at risk and length of follow-up. Time at risk refers to the amount of time for which an individual may actually be able to engage in criminal behavior and length of follow-up refers to the period from assessment to the end of the follow-up. Time at risk and follow-up periods are critical for understanding the base rates of criminal behavior. For some individuals in the study, these values may be the same; for others they may be different. For example, if someone is assessed at the point of admission to a prison, incarcerated for 2 years, and then followed for another 2 years in the community, the follow-up period would be 4 years, but actual time at risk is only 2 years.
10. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, *Standards for Educational and Psychological Testing*.
11. Douglas, Skeem, and Nicholson, "Research Methods in Violence Risk Assessment," 325–46.
12. Singh et al., "Reporting Guidance for Violence Risk Assessment Predictive Validity Studies," 15–22.

Guideline 2: Meet the minimum performance thresholds of post-conviction risk and needs assessments completed in the field according to statistical standards.

Overview

Establishing that post-conviction risk and needs assessments completed in the field meet the minimum performance thresholds according to statistical standards is critical to evaluating accuracy. The performance metrics provided reflect well-established statistical standards for measuring the strength or degree of agreement not only among assessors but also between the assessment results and recidivism. To be clear, these are minimum performance thresholds. Agencies may elect to require more—but not less—stringent performance thresholds than the minimums provided here as a matter of policy for all cases or for specific contexts.

Action Items

Demonstrate good agreement or better among assessors for post-conviction risk and needs assessments completed in the field.

Agreement among assessors may be evaluated using different statistical approaches, depending on the rating or scoring of interest, and each approach may have advantages and disadvantages. For categorical ratings such as yes/no or low/moderate/high, the level of observed agreement is the most straightforward and easiest to calculate (number of agreements/number of agreements and disagreements), but it does not account for expected agreement. Kappa considers both observed and expected agreement, but it can produce erroneous results when there is little to no response variability.[13] Intra-class coefficient (ICC) looks at whether assessors rate individuals similarly on a continuous scale as opposed to across categories.

The guidelines for interpreting the strength or practical significance of assessor agreement reflect well-established standards in social and epidemiological sciences.[14] Based upon these standards, good agreement is indicated by:

• Observed agreement among assessors of 80 percent or greater

• Kappa = .60–.74

• ICC = .60–.74

For agencies wishing to adopt more stringent criteria, excellent agreement is indicated by:

• Observed agreement among assessors of 90 percent or greater

• Kappa = .75–1.00

• ICC = .75–1.00

13. Feinstein and Cicchetti, "High Agreement but Low Kappa," 543–49.
14. Cicchetti, "The Precision of Reliability and Validity Estimates Re-Visited," 695–700.

Demonstrate good validity or better in predicting the likelihood of recidivism with post-conviction risk and needs assessments completed in the field.

To evaluate the performance of post-conviction risk and needs assessments in predicting the likelihood of recidivism, consider two different metrics:[15] (1) the observed rates of criminal behavior at each risk level and (2) an overall index of predictive validity.

The observed rate of criminal behavior at each risk level represents a simple calculation examining the proportion of individuals who went on to recidivate within each risk level. To demonstrate, if 75 people were classified as Risk Level 3 and 5 of them recidivated, then the observed rate of criminal behavior at Risk Level 3 is 6.7 percent (i.e., 5/75 x 100). There are no set performance standards or established benchmarks for what would be "good or better" for this metric. However, as risk levels increase, so too should the observed rates of criminal behavior. In other words, at higher levels, we would expect to see higher rates of criminal behavior than at lower levels. Additionally, observed rates of criminal behavior should increase at each subsequent risk level. For example, if the observed rate of criminal behavior at Risk Level 3 is 6.7 percent, then the observed rate of criminal behavior at Risk Level 2 should be less than 6.7 percent, and at Risk Level 4 greater than 6.7 percent.[16]

Because the observed rate of criminal behavior at each risk level is a purely descriptive metric,[17] we recommend using the areas under receiver operating characteristic curves (AUCs) to provide an overall index of predictive validity. AUCs represent the likelihood that a randomly selected individual who recidivated during the follow-up period received a higher risk score than a randomly selected individual who did not recidivate during the follow-up period. AUC is a preferred metric because its values are not affected by rates of recidivism to the same degree that other metrics are influenced by rates of recidivism.[18] The guidelines for interpreting the strength or practical significance of AUCs, again, reflect well-established research standards.[19] Based upon these standards, good validity will be indicated by AUC values of .65–.70. For agencies wishing to adopt more stringent criteria, excellent validity will be indicated by AUC values of .71–1.00.[20]

---

15. While there are additional metrics that agencies and others may wish to examine, such as the false positive and false negative rates or the positive predictive value (PPV) and negative predictive value (NPV), we do not recommend their use. We highlight two issues of particular importance here. First, post-conviction risk and needs assessment instruments do not make binary predictions about future criminal behavior (e.g., yes or no), nor do they make binary decisions about an individual (e.g., detain or release). Rather, they estimate the likelihood of recidivism using multiple categories or levels to provide decisionmakers with the information necessary to make such decisions. Consequently, these metrics do not reflect how the instruments are designed or intended to be used. Second, these metrics are dependent upon sample size and recidivism rates. As a result, values that may be interpreted as reflecting poor validity could instead represent errors in a small number of cases or successful mitigation of recidivism. More specifically, PPV and NPV are based upon a single threshold or cutoff, but there are no post-conviction risk and needs assessment instruments that use a single threshold or cutoff. Instead, they typically use at least three risk levels or categories. To calculate the PPV and NPV, then, requires selecting a threshold; this may include the use of a single numerical score or risk level as the threshold, or the PPV and NPV may be calculated for each risk level. The former is what is done most frequently; however, this does not reflect how the assessment results are used in practice. The latter is more akin to how the instruments are used in practice, but the calculated values will be affected by the relatively small number of cases at each level. Specifically, even a small number of "errors" may dramatically affect the observed PPV or NPV. Because fewer individuals are typically assessed at higher relative to lower risk levels, this means that even within a single validation study, the estimates of PPV and NPV will be less stable for higher than lower risk levels. Further, the base rate of recidivism in a given jurisdiction puts boundaries on the possible range of values: PPV will increase with increases in the prevalence of recidivism, while NPV will decrease with increases in recidivism. This means that in jurisdictions with relatively low rates of recidivism, it is not possible to observe high PPVs. Only with higher rates of recidivism will higher PPVs be observed. The converse is true for NPV.

16. Ideally, the increase in the observed rate of criminal behavior would be statistically significant from one level to the next; however, this may not be realistic if there are small numbers of people assessed at each level and low base rates of criminal behavior. Consequently, a substantive increase in the observed rate of criminal behavior from one level to the next is sufficient.

17. Singh, "Predictive Validity Performance Indicators in Violence Risk Assessment," 8–22; Helmus and Babchishin, "Primer on Risk Assessment," 8–25.

18. Smith, "The Effects of Base Rate and Cutoff Point Choice," 83–11; Rice and Harris, "Comparing Effect Sizes in Follow-up Studies," 615–20.

19. Cohen's d is a measure of the difference between the averages of two groups. It is the most commonly used measure against which to interpret the strength of association in the social and epidemiological sciences.

20. An AUC value of 1.00 indicates perfect discrimination between those who went on to recidivate from those who did not recidivate during follow-up period; .50 indicates discrimination at chance levels; and 0.00 indicates completely incorrect discrimination (i.e., all those who did not recidivate were identified as higher risk for recidivism while those who did recidivate were identified as lower risk and vice versa).

Guideline 3: Use a CQI process to ensure successful implementation of the post-conviction risk and needs assessment instrument.

Overview

Even the most well-established, vetted, and validated post-conviction risk and needs assessment instrument may fail to produce the desired results if not implemented with fidelity. Successful implementation will require significant planning and resources at the outset, as well as the establishment and deployment of strategies to monitor the implementation and assessment processes over time. Deliberate, pre-planned CQI efforts will allow for prompt identification of issues that may interfere with the effectiveness of post-conviction risk and needs assessment instruments and enable the deployment of strategies to address those issues, thereby promoting the accuracy of assessment results.

Action Items

Document the protocols for applying the post-conviction risk and needs assessment instrument.

Protocols for administration should document how, when, and for whom and by whom assessments will be completed. Agencies should develop and document these administration protocols as part of the planning process, ideally before using the instrument. For agencies that have already implemented a post-conviction risk and needs assessment instrument, documentation of current administration protocols should be prioritized and completed over a short, but feasible, timeframe. Doing so will not only help promote accuracy in the assessments, ensuring that they are completed as intended, but also will provide clarity and transparency on the appropriate—and inappropriate—use of the post-conviction risk and needs assessment instrument.

The protocols should describe the required and recommended sources of information to use to complete the ratings. They also should describe which individuals should be assessed, when assessments should be completed for them, and what decisions and processes the results should inform.[21] Additionally, protocols should describe when re-assessments should be conducted; this may include specification of the timeframe for routine re-assessment (e.g., every 6 months)[22] or certain conditions that would prompt re-assessment (e.g., change in relationships, employment, housing, health, legal status, etc.).[23] For more on re-assessment, see Guideline 4.

---

21. For example, post-conviction risk and needs assessments may be required for individuals in specific programs or charged with certain offenses. In terms of timing and decisions, there may be a requirement to complete an initial assessment within 2 weeks of intake to a new program or agency to inform case planning or within 4 weeks of release from a program or setting to inform release planning.
22. Barnes et al., "Validity of Initial, Exit, and Dynamic Juvenile Risk Assessment," 21–38; Lloyd et al., "Reassessment Improves Prediction of Criminal Recidivism," 568–81.
23. Buchanan et al., "Psychiatric Violence Risk Assessment," 340.

In addition to documenting the administration protocols, we recommend that agencies document protocols for how, when, and by whom CQI will be conducted. These CQI protocols should ideally be developed and documented in consultation with diverse stakeholders—including instrument developers or other experts, staff, supervisors, and administrators—before using the instrument. However, they should also be revisited periodically as there may be emergent issues that call for changes in the frequency or focal points of CQI-related efforts. Thinking through and planning for CQI before using the instrument will ensure that the necessary data, resources, and staffing are available to support CQI over time.

Prior to their use of the instrument in practice, provide all assessors with training on the rating procedures and protocols for applying assessment results to inform case plans.

Assessors should complete all required training prior to using a post-conviction risk and needs assessment instrument in practice. The minimum training requirements for administration may be specified by the instrument developers, but at the very least should include training on the strategies to gather and interpret information, procedures for rating items, and how to interpret results. Typically, trainings on these fundamentals are provided by experts, including the instrument developers or others who are well trained and qualified in using the instrument. They may be offered live in person, online, or available on demand via a licensed provider or organization. These training options and modalities will be determined largely as a function of the instrument selected, as well as agency resources, needs, and practical considerations (e.g., staff schedules, onboarding processes, certification or credentialing requirements, etc.). Assessors additionally should complete one (or more) practice cases. Either during or after training, most trainers will provide the practice case materials and the experts' "gold standard" ratings against which to compare trainees' assessments.[24] In total, assessors should complete and receive feedback on at least four to five practice cases before they begin to use the instrument in the field (i.e., one or two in the context of the training and three after completing the training; see Guideline 1).

Assessors also should receive training on the site-specific policies and protocols for applying assessment results to inform case plans before they begin using the instrument in the field. This training can be completed in conjunction with or after the training on the administration of the instrument. It should cover local policies regarding for whom and when post-conviction risk and needs assessments will be conducted and for what purpose(s), ensuring that these uses match the tasks for which the instrument was developed. This training should be provided by local experts, supervisors, or other administrators involved in developing and overseeing the implementation of the post-conviction risk and needs assessment instrument, as appropriate.

Supervisors and others who will be involved in or affected by the implementation should receive some level of training on the procedures and protocols for the use of the post-conviction risk and needs assessment instrument. This may include a short, overview presentation on the basic approach and use of the post-conviction risk and

---

24. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice.*

needs assessment instrument for other agency staff, decisionmakers, or community representatives. Supervisors will need sufficient knowledge of how to conduct the assessment and use the assessment results in decisions and case planning to be able to conduct CQI, including case reviews and booster sessions, as described below. It may not be necessary for them to attend all the trainings or to complete all the practice cases. However, the more supervisors and other local leaders demonstrate commitment to the use of the post-conviction risk and needs assessment instrument, the more buy-in there may be from staff and other stakeholders.

Agencies that already use an instrument should develop a strategy to provide this training within a 6-month period.

Complete case reviews at least twice yearly during implementation to identify problems to correct through individual coaching or booster training.

At least twice per year, supervisors should conduct case reviews that examine:

- Fidelity to the rating and scoring guidelines.
- Adherence to the implementation protocols.
- Concordance among assessment results and case decisions, resource allocation, and service provision.

As part of the planning process, ideally before using the instrument, we recommend that agencies develop a case review checklist that includes observable indicators of fidelity issues or errors in the assessments themselves (e.g., missing ratings, inconsistencies between item scores and risk levels) as well as the steps in the assessment process (e.g., collecting information from records and interviews, if appropriate). The checklist should also include items for documenting the population(s) that should be assessed, if appropriate, and the timing of the assessments as detailed in the local administration protocols. Finally, the checklist should include items that speak to whether case decisions, resource allocation, and service provision are in line with the assessment results. We recommend the application of the Risk-Need-Responsivity model[25] as a framework for examining concordance between assessment results and practices; see Guideline 6 for more on this.

The **Risk Assessment Quality Improvement (RAQI) protocol** ⧉ provides a starting point for structuring the case review checklist and CQI process. However, agency leaders may wish to consult with the developers of the specific instrument they use or other experts to ensure that they have adequately captured the relevant issues in their local case review checklist. When case reviews reveal issues that need to be addressed, supervisors may wish to address them at the individual or group level, as they come up, or during the annual booster training sessions. However, it is imperative that supervisors and other stakeholders consider whether the identified issues stem uniquely from assessors themselves (e.g., knowledge or motivational concerns) or whether they rise to the broader system or interagency level (e.g., lack of specificity in the protocols, unavailability of required documents, etc.).

---

25. Bonta and Andrews, *Risk-Need-Responsivity Model*.

Conduct booster training at least annually for all assessors during implementation.

At least once per year, assessors using the instrument should complete a booster training session to prevent drift and promote assessment accuracy. These booster sessions should:

- Review rating procedures and protocols for using assessment results.
- Require completion of at least one—but ideally two—practice cases to good agreement or better with a "gold standard" or expert rating.
- Address any other issues identified in the case reviews.

Booster sessions may be conducted by experienced in-house assessors or by outside experts (e.g., the instrument developers or other experts qualified to train on the instrument).[26] As with the initial training sessions, these sessions could be held for all assessors in a group—whether in person or online—or through an on-demand format that could be accessed by individual assessors as needed. There is no one best or recommended approach. The booster session training format and modality may be informed by the instrument that is selected for implementation, the needs and resources of the agency, and other practical considerations (e.g., the number of staff to be trained, specific training needs identified in the CQI reviews).

Practice cases can be completed during or after the booster session in various ways, such as having a staff member present a case to the group for assessment and review or using practice cases developed in collaboration with the expert trainer or other qualified professional. The goal of these practice cases is to offer real-time feedback on the accuracy of the assessment ratings and the connection between assessment results and case planning. Finally, while booster sessions are an opportunity to discuss and address issues identified via case reviews, it may be necessary to provide individual assessors or teams with more specific and targeted feedback through team meetings or one-on-one supervision, as appropriate.[27]

26. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice.*
27. Ibid.

Guideline 4: Use a multistep approach to assess risk and needs over time.

Overview

While not required, a multistep approach to reassess risk and needs over time may contribute to greater accuracy and efficiency in the post-conviction risk and needs assessment process. Agencies may wish to implement a multistep approach for various reasons. In particular, the use of a risk screening instrument as an initial step in a comprehensive post-conviction risk and needs assessment process may help expedite initial decisionmaking and case processing. Additionally, the routine re-administration of post-conviction risk and needs assessment instruments that include dynamic factors and needs will afford the detection of changes in risk and needs that can be used to amend risk management strategies and case plans.

Action Items

Follow post-conviction risk screening instruments, if used, with a comprehensive risk and needs assessment only for those identified as being at potentially heightened risk of recidivism.

Although the terms are often used interchangeably, screening and assessment refer to two different, but related, processes. In the post-conviction context, screening refers to the universal implementation of a short, easily administered set of items to quickly identify individuals who are potentially at heightened risk of recidivism and should receive a more in-depth, comprehensive risk and needs assessment. In other words, screening instruments can be used as a first step to identify and "screen out" individuals who pose limited risk of recidivism and, thus, do not need to be evaluated further and to identify and "screen in" those who are at *potentially* heightened risk of recidivism and therefore warrant further, more comprehensive evaluation of their risk and needs.

Screening, by definition, is not a standalone process. Instead, the addition of screening to a comprehensive post-conviction risk and needs assessment process, while not necessary, may prove useful for agencies with large caseloads that are seeking to prioritize resources. Screening instruments are designed to cast a wide net; they are calibrated during the development and validation process to over- (as opposed to under-) estimate risk. That is, they are intentionally designed to reduce the likelihood of false negatives—individuals who are misidentified as low risk for recidivism. However, in doing so, the number of false positives—individuals who are misidentified as being at heightened risk of recidivism—will be high. If used, screening instruments must be followed by a comprehensive risk and needs assessment for those "screened in."

To be clear, post-conviction risk and needs assessment instruments can be implemented in the absence of risk screening instruments. Agencies that are seeking to adopt an evidence-based assessment approach do not need to implement a universal risk screening protocol. However, the opposite is not true. Do not use risk screening instruments in lieu of comprehensive post-conviction risk and needs assessment instruments. This is a misapplication of screening instruments and will overestimate risk, which, in turn, will contribute to unnecessary

individual, assessor, and system costs and can potentially contribute to increases in recidivism. For these reasons, using risk screening instruments in lieu of comprehensive post-conviction risk and needs assessment instruments also will threaten fairness.

Re-administer post-conviction risk and needs assessment instruments that include dynamic factors and needs at routine intervals to monitor individual progress and inform amendments to case planning, as needed.

Re-administering post-conviction risk and needs assessment instruments that include dynamic factors and needs will improve assessment accuracy.[28] Dynamic risk factors and needs, by definition, are capable of change. For that reason, post-conviction risk and needs assessment instruments that include dynamic factors and needs require re-administration over time. Doing so will provide not only a measure of change—if any—in an individual's risk level overall but also an opportunity to review the appropriateness—and effectiveness—of the current risk management and intervention strategies. An overall decrease in risk level across repeated assessments may prompt consideration of a reduction in the level of supervision and services. An overall increase in risk level may suggest the need for greater supervision and services. Alternatively, a lack of change in risk level may prompt consideration of whether the appropriate factors are being targeted at the appropriate level via the intervention and, if so, in such a way as to promote individual responsivity.[29]

The timeframe for re-administering the post-conviction risk and needs assessment may depend on the instrument selected but also should account for the assessment's purpose, population, context, and local resources.[30] It may be useful to consult with the instrument developers or other experts to ascertain a timeframe for re-administration that balances resources with utility.

The emphasis here has been on the re-administration of post-conviction risk and needs assessment instruments that include dynamic factors and needs. We anticipate that static risk factors will change little, if at all, by definition. Change in static factors, if any, will typically be in the direction of increased rather than decreased risk (e.g., new charges or convictions that contribute to a higher criminal history rating). However, we may also see some reductions in risk level over time—even on risk factors thought to be static in nature—if items specify behaviors in a certain timeframe (e.g., convictions in the prior 2 years). So, agencies are encouraged to document circumstances in which re-administration of any post-conviction risk and needs assessment instruments, even those composed of static risk factors, may be needed to promote assessment accuracy.

---

28. Barnes et al., "Validity of Initial, Exit, and Dynamic Juvenile Risk Assessment," 21–38; Lloyd et al., "Reassessment Improves Prediction of Criminal Recidivism," 568–81.
29. See Bonta and Andrews, *Risk-Need-Responsivity Model*.
30. For example, some instruments may include more acute dynamic factors in which we might expect more frequent or rapid change, while others may include more stable dynamic factors that might change more slowly over months or years, if at all. As another example, some settings may confer more stability, expose individuals to fewer changes in their environment, or afford fewer opportunities for intervention, resulting in relatively limited change in functioning and risk. Alternatively, some agencies may implement post-conviction risk and needs assessments to support periods of transition, whether in or out of a particular setting or program. We may anticipate periods of transition to be times during which there will be considerable fluctuation in risk and needs. As a final example, some populations may show more or less change; we may anticipate greater change in risk and needs among some people convicted for first-time offenses or younger people in the criminal justice system, but less change in risk and needs among those who have had longer or more chronic justice system involvement.

Principle II: Fairness

The next three guidelines speak to strategies that agencies can use to promote fairness in the use of post-conviction risk and needs assessment instruments. Broadly speaking, fairness refers to the equitable use of the results of post-conviction risk and needs assessment instruments to inform case decisions, resource allocation, and services overall. However, fairness as it relates to disparities in racial, ethnic, gender, or other characteristics such as mental illness in post-conviction processes should consider, more specifically, the degree to which assessment results have the same meanings and applications across groups defined by these characteristics. Fairness should be considered in the development, validation, and implementation of post-conviction risk and needs assessment instruments.

"

[O]ne cannot expect any risk assessment tool to reverse centuries of racial injustice or gender inequality. That bar is far too high. But, one can hope to do better.
(Berk, Heidari, Jabbari, Kearns, and Roth, 2017, 35)

"

We recommend the following guidelines to promote fairness in the use of post-conviction risk and needs assessment instruments:

5. Examine the results of the post-conviction risk and needs assessment instrument for predictive bias and disparate impact across groups.

6. Apply post-conviction risk and needs assessment instrument results to individual cases in keeping with the Risk-Need-Responsivity (RNR) principles.

7. Adopt agencywide strategies to minimize the potential that local implementation of a post-conviction risk and needs assessment instrument could promote disparities.

Guideline 5: Examine the results of the post-conviction risk and needs assessment instrument for predictive bias and disparate impact across groups.

Overview

A post-conviction risk and needs assessment instrument is not necessarily biased or unfair simply because one group of people is rated higher or lower, on average, compared to another group of people.[31] Instead, consider (a) how assessment results relate to recidivism across groups and (b) how assessment results are used to inform decisions across groups. These two considerations speak to predictive bias and disparate impact, respectively. Predictive bias is present when assessment results demonstrate different levels of predictive validity across groups, whereas disparate impact is present when the assessment results are applied inequitably across groups.

These two concepts are related but are not dependent upon each other. Predictive bias relates to assessment accuracy across groups but does not necessarily lead to disparate impact. Assessment results can show some differences in predictive accuracy between groups but still demonstrate equitable, positive impacts on case decisions and outcomes.[32] For example, assessment results might demonstrate slightly better predictive accuracy for White than Black people but still contribute to less restrictive placements for both White and Black people. Further, disparate impact does not require the presence of predictive bias. Even if assessment results have similar levels of predictive accuracy across groups, they still may be used in different ways to inform case decisions and outcomes for different groups. For example, assessment results may demonstrate similar predictive accuracy for Black and White people, but judges and other decisionmakers may be more likely to deviate from assessment results in an upward direction (i.e., increase estimated risk) and impose more restrictive conditions for Black people than for White people.[33]

The Standards for Educational and Psychological Testing[34] specify that test bias exists when scores function differently for different groups of people, which implies an adverse impact on one group compared to another.[35] For these reasons, we recommend that agencies focus on whether there is evidence of disparate impact in considerations of fairness.

---

31. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84; Reynolds and Suzuki, "Bias in Psychological Assessment," 82–113.
32. See, for example, Lowder, Grommon, and Ray, *Improving the Accuracy and Fairness of Pretrial Release Decisions*; Lowenkamp, DeMichele, and Klein Warren, "Replication and Extension of the Lucas County PSA Project."
33. Orton, Hogan, and Wormith, "An Examination of the Professional Override," 0093854820942270; Marlowe et al., "Employing Standardized Risk Assessment in Pretrial Release Decisions," 361–76.
34. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, *Standards for Educational and Psychological Testing*.
35. Meade and Fetzer, "Test Bias, Differential Prediction, and a Revised Approach," 738–61.

Action Items

Establish whether the likelihood of recidivism increases in similar ways across risk levels for members of groups defined by race, ethnicity, and gender.

We recommend asking the developers of post-conviction risk and needs assessment instruments for evidence that the likelihood of recidivism increases in comparable ways across risk levels from group to group. If this information cannot be provided or the instrument was developed locally, agencies can examine the performance indicators described in Guideline 2 across groups.

Two key questions should be answered. First, do risk levels relate to the rates of recidivism as expected within groups defined by race, ethnicity, and gender (i.e., do higher observed rates of criminal behavior correspond to higher risk levels rather than lower risk levels)? It is possible—even likely—that recidivism rates will differ within a given risk level from one group to the next. What matters is whether the recidivism rates increase across risk levels *within* groups in the anticipated way. Second, do the performance indicators meet the minimum thresholds described in Guideline 2? Again, there may be some differences among groups, but what matters is that the performance indicators still meet statistical standards for predictive accuracy from group to group.[36]

Test whether assessment results identify individual risk levels and needs and predict recidivism in the same way from group to group.

A critical step in evaluating the fairness of a post-conviction risk and needs assessment instrument is determining whether the assessment results predict recidivism in the same way, regardless of group membership. Said another way, we need to test statistically whether the strength (and direction) of the relationship between assessment results and recidivism differs systematically as a function of race, ethnicity, or gender.

Agencies can use various statistical methods to find out whether the average risk rating relates to the average recidivism rate in the same way for each group. We recommend using the methods that represent the state of the art and have been applied in peer-reviewed publications that test for racial, ethnic, and gender biases in risk and needs assessment.[37] Because these methods are complex, we recommend consulting with a researcher or evaluator with specific expertise in regression analysis or other statistical methods if that expertise is not available in house.

---

36. While there has been considerable emphasis on false positives and false negatives as metrics for understanding fairness, we do not recommend their use for both pragmatic and statistical reasons, two of which we highlight here. First, calculating false positives and false negatives requires assessment results to be used to categorize people into two groups based on whether they will or will not recidivate. However, post-conviction risk and needs assessment instruments do not produce such binary classifications; instead, they place people within risk levels or categories. As a result, a single threshold or cut-off must be chosen, above which someone is designated as testing "positive" for recidivism and below which they are designated as testing "negative." This is not typically how instruments are used in practice, limiting the external validity—or practical relevance—of such metrics. Second, the false positive and false negative rates will differ dramatically as a function of the threshold selected, as well as the rate of recidivism during follow-up. Consequently, the generalizability of the results across jurisdictions and even within jurisdictions over time is very limited. For further discussion regarding the limitations of false positives and false negatives as metrics of fairness in the context of risk assessment, see Helmus and Babchishin, "Primer on Risk Assessment," 8–25; Freeman, Hu, and Jannetta, *Racial Equity.*

37. We recommend testing a moderation model, which involves conducting multiple regression analysis in which the assessment results, grouping variable (e.g., gender or race), and their interaction term are entered as predictors of recidivism. Only if the interaction term is a statistically significant predictor of recidivism is there evidence of predictive bias. See, for example, Skeem and Lowenkamp, "Risk, Race, and Recidivism," 680–712; Lowder et al., "Racial Bias and LSI-R Assessments," 210–33; Cohen and Lowenkamp, "Revalidation of the Federal PTRA," 234–60.

Compare how assessment results relate to case decisions, resource allocation, and service provision across groups.

At the core of concerns regarding the use of post-conviction risk and needs assessment is whether assessment results are applied in different ways for different groups and more specifically, whether the use of assessment results leads to more punitive and restrictive responses for marginalized groups.[38] The question that agencies must answer is whether there is evidence that the way assessment results are used to inform case decisions, resource allocation, and service provision contributes to greater racial, ethnic, or gender disparity than the strategies through which these processes are conducted otherwise.

To answer this question, we recommend, at a minimum, examining case decisions, resource allocation, and service provision across groups as part of a CQI strategy—for example, as part of routine data monitoring or case reviews every 6 months. Specifically, every 6 months, agencies should have a plan to examine the following metrics within groups defined by race, ethnicity, and gender:

- Percentage of each type of case decision.
- Assigned levels of classification, supervision, or condition.
- Average number of services provided overall.
- Percentage of each type of service.

However, to fully answer this question would require an evaluation design that allows for a systematic comparison of (1) case decisions, resource allocation, and service provision made using the results of a post-conviction risk and needs assessment instrument to (2) case decisions, resource allocation, and service provision made without the assessment results. A randomized controlled trial is the most rigorous design but challenging to do in the context of real-world practice. Alternative evaluation designs that may be more feasible include a quasi-experimental, between-groups design, or a pre-post test design. While these evaluation designs are limited in the degree to which findings speak to disparate impact that can be attributed to the assessment results (as opposed to other factors), they can still help agencies identify where there are systematic differences in case outcomes to address.

---

38. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84.

Guideline 6: Apply post-conviction risk and needs assessment instrument results to individual cases in keeping with the RNR principles.

Overview

Applying the RNR model can promote fairness by providing a structure for guiding and, specifically, limiting the scope of the use of post-conviction risk and needs assessment instruments. The RNR model[39] is widely recognized as an evidence-based framework for promoting positive public safety and case outcomes through the practical application of the results from post-conviction risk and needs assessment instruments. Briefly:

- The Risk principle entails matching the level of supervision, resources, and services with the individual's assessed level of risk of recidivism.

- The Need principle specifies that interventions should target dynamic factors and needs that increase the risk of recidivism for that individual.

- The Responsivity principle involves tailoring risk management strategies and services to a person's specific abilities, motivations, and strengths as part of the case planning process.

Together, these three principles emphasize an individualized approach that is informed by assessment results and limited in scope.

Action Items
Use assessment results to inform the appropriate level of intervention needed to manage the assessed risk of recidivism.

Consistent with the Risk principle, assessment results should inform the least restrictive level of intervention needed to manage a person's risk of recidivism. The greater the estimated level of risk, the greater the supervision, resources, and services that should be allocated and vice versa. The objective is to use the post-conviction risk and needs assessment results to help identify the minimum level of intervention, if any, that is necessary to manage a person's potential risk to public safety. Assessment results should not be used to justify a higher level of intervention than appropriate for the offense(s) of conviction.

Some post-conviction risk and needs assessment instruments provide case management recommendations regarding the type and number of hours of supervision and services required at a given level of recidivism risk. There also have been efforts to develop recommendations regarding the frequency and intensity of intervention that are not instrument specific such as the **five-level risk and needs system** ⧉. Briefly, the five-level system seeks to provide a common language to communicate information about risk and needs, and it recommends the appropriate intensity and type of risk management and intervention strategies indicated by a given risk and needs level. Other criteria for specific domains of intervention and treatment also may be relevant, such as The ASAM Criteria for the level and intensity of treatment services for people with addictions and co-occurring conditions.

39. Bonta and Andrews, *Risk-Need-Responsivity Model.*

In practice, we recommend that agencies develop local guidelines regarding the frequency and intensity of supervision and services vis-à-vis the assessment results prior to using the post-conviction risk and needs assessment instrument. For agencies that have already implemented a post-conviction risk and needs assessment instrument, development of local guidelines should be prioritized and completed over a short, but feasible, timeframe. Agencies may need to revisit these guidelines and adapt them over time as the population, availability of services, or other local resources change.

Identify the dynamic factors and needs to be addressed through intervention.

Consistent with the Need principle, interventions should target the dynamic (i.e., changeable) factors and needs that contribute to risk of recidivism for that individual. The reasons and motivations that lead to criminal behavior can differ dramatically from person to person, even among those who have the same factors present in their social environment.[40] Post-conviction risk and needs assessment instruments that include at least some dynamic risk and needs factors provide critical, person-specific information regarding behaviors, beliefs, or other factors to be targeted in intervention. Instruments that predominantly—or entirely—comprise static risk factors are more limited in their utility with respect to guiding the tailored interventions we recommend here.

Further, targeting dynamic risk and needs factors for intervention will de-emphasize historical factors that cannot be changed such as age at first arrest. In doing so, we can move away from factors that are known sources of bias and act as proxies for race, ethnicity, or gender.

Although applying the Need principle can be challenging, it is not "all or nothing." As adherence to the Need principle increases—or with better "treatment match"—the likelihood of positive case outcomes, including public safety, increases.[41] Because dynamic factors and needs can change over time, the re-administration of the post-conviction risk and needs assessment instrument at routine intervals to inform any needed amendments to case plans will promote the likelihood of success. (See Guideline 4 for more on dynamic factors and re-administration).

Maximize reductions in recidivism by tailoring the interventions to individual motivations, strengths, and abilities.

Consistent with the Responsivity principle, reductions in recidivism will be maximized by tailoring interventions to case-specific barriers or facilitators to successful habilitation, including individual motivations, abilities, and strengths. In the development and amendment of case plans, assessors may consider two types of responsivity: general responsivity and specific responsivity.

---

40. For example, one person may be heavily influenced by antisocial peers and have few prosocial contacts to buffer against these influences. Another person may also have antisocial peers, but their risk of recidivism is driven by problems related to substance use rather than the antisocial influence of these peers. An intervention focused on positive peer support, then, may mitigate risk in the former example, while a substance use intervention may have greater effectiveness in the latter. Similarly, if there is no indication of substance use as a factor, then a substance use-focused intervention may do more harm than good.

41. Singh et al., "From Risk Assessment to Risk Management," 1–9; Vieira, Skilling, and Peterson-Badali, "Matching Court-Ordered Services with Treatment Needs," 385–401; Nelson and Vincent, "Matching Services to Criminogenic Needs," 1136–53.

General responsivity refers to the use of interventions that have demonstrated effectiveness in addressing criminogenic risk factors and needs, particularly approaches that use social learning or cognitive behavioral methods.[42] The most effective interventions may use diverse, evidence-based strategies such as prosocial modeling, positive reinforcement, or problem-solving skill development that meet an individual where they are. General responsivity also emphasizes the importance of establishing a warm, respectful, trusting, and collaborative working alliance to promote positive treatment outcomes. Strategies that reflect cultural humility and a multicultural orientation, for example, may help facilitate strong working alliances and foster more just and equitable practices.[43]

Specific responsivity refers to the tailoring of services to address individual and environmental factors that may affect treatment outcomes. This may include the use of specialized interventions such as culturally tailored interventions, trauma-informed approaches, or gender-specific services. Specific responsivity also should include consideration of environmental factors such as the institutional culture, staff skills or attitudes, and barriers to service access and use. Specific responsivity represents an opportunity not only to promote positive treatment outcomes in an individual case but also to address factors that may be contributing to racial, ethnic, and gender disparities more broadly.

Guideline 7: Adopt agencywide strategies to minimize the potential that local implementation of a post-conviction risk and needs assessment instrument could promote disparities.

Overview

Ultimately, it is how a post-conviction risk and needs assessment instrument is used in practice that will determine whether it contributes to the unfair treatment of people across groups defined by race, ethnicity, and gender. Instruments differ in their contents, methods, and purposes.[44] The information used to complete the assessments—a potential source of systemic bias—also differs as a function of local policies and practices as well as record keeping.[45] As a result, the performance, meaning, and application of assessment results may differ from setting to setting and population to population.

It is unlikely that any one strategy, including the use of post-conviction risk and needs assessment instruments, will *eliminate* racial, ethnic, or gender inequities in the criminal justice system. However, strategies employed in the system should not exacerbate these inequities either. Consequently, it is imperative that agencies take the steps necessary to minimize the potential that the use of post-conviction risk and needs assessment instruments promotes disparities in their local setting, context, or jurisdiction.

---

42. Dowden and Andrews, "The Importance of Staff Practice in Delivering Effective Correctional Treatment," 203–14.
43. Mosher et al., "Cultural Humility," 221–33.
44. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84.
45. Mayson, "Bias in, Bias Out," 2122–473.

Action Items

Select and implement post-conviction risk and needs assessment instruments based on their performance, content, and context.

There is no one instrument that is "fairest." Instead, the following information should be considered to support the selection and implementation of any post-conviction risk and needs assessment instrument to ensure that it does not perpetuate inequities:

- Predictive accuracy metrics across groups, as described above.
- The implications of factors that are known sources of bias or may act as proxies for race, ethnicity, and gender.
- The context(s) in which assessment results will be used.

In the process of selecting a post-conviction risk and needs assessment instrument, agencies should consider the degree to which the instrument includes factors that are known sources of bias or may serve as proxies for race, ethnicity, or gender (e.g., criminal history, gang affiliation, employment, education level, debt, or housing stability). The information captured in these items may reflect bias or marginalization resulting from systemic and structural inequities, and, consequently, their inclusion may contribute to disparities.[46] For example, information on criminal history may reflect biases in local policing, prosecutorial, and judicial practices. There are many post-conviction risk and needs assessment instruments that include a wide range of static and dynamic factors.[47] As such, we recommend that agencies avoid instruments that emphasize factors that are known sources of bias or may serve as proxies for race, ethnicity, or gender.[48]

That said, the inclusion of such factors does not necessarily mean that the instrument will produce biased assessment results, nor does the exclusion of such factors mean that an instrument will be free from bias.[49] For these reasons, consideration of evidence regarding the performance of assessment results across groups, as described in Guideline 5, will provide information that is essential to the selection process. For example, if two or more post-conviction risk and needs assessment instruments are roughly comparable, the instrument that minimizes differences in predictive accuracy among groups should be selected.[50]

Finally, it is important to consider the context in which assessment results will be used. Certain instruments may be more appropriate for some decisions or applications than others. For instance, if the task at hand is one of classification, then a well-validated instrument that comprises primarily static factors may be acceptable (assuming there is limited evidence of group differences in assessment results). If the context also requires the development of case plans, an instrument that additionally includes dynamic risk and needs factors would be more appropriate.

---

46. Starr, "Evidence-Based Sentencing," 803–72; Harcourt, "Risk as a Proxy for Race," 237–43; Eckhouse et al., "A Unified Approach," 185–209.
47. Desmarais, Johnson, and Singh, "Performance of Recidivism Risk Assessment Instruments," 206–22.
48. Hart, "Culture and Violence Risk Assessment," 76–96.
49. Mayson, "Bias in, Bias Out," 2122–473; Skeem and Lowenkamp, "Using Algorithms," 259–78.
50. Skeem and Lowenkamp, "Using Algorithms," 259–78; Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84; Shepherd and Lewis-Fernandez, "Forensic Risk Assessment and Cultural Diversity," 427–38.

Develop and implement strategies to support equitable and safe case decisions, resource allocation, and service provision.

Agencies can—and should—develop and institute strategies to support equitable and safe case decisions, resource allocation, and service provision. The use of assessment results should be clearly articulated in local administration protocols and policies governing the use of post-conviction risk and needs assessment instruments, as described in Guidelines 2 and 9, respectively. Clear guidance on when, for whom, and how post-conviction risk and needs assessment instruments will be completed and applied to inform decisionmaking will reduce the potential that the use of assessment instruments is biased. When case decisions, resource allocation, or service provision deviate from assessment results, people managing these decisions should provide justification to explain why such deviations are appropriate. Ultimately, assessment results are just one source of information that agencies should consider during the case planning process.[51] There may be case-related issues (e.g., specific offenses for which there are blanket policies such as sex offenses) or other considerations that inform individual decisions (e.g., current caseload size, availability of placements or programming, and/or limited staff resources).

Ongoing CQI, as described in Guideline 3, will provide the opportunity to monitor the implementation and use of post-conviction risk and needs assessment results to ensure that their application supports equitable and safe case decisions, resource allocation, and service provision. If there is evidence of predictive bias or disparate impact across groups, various strategies can be implemented to increase the fairness of the process. Such strategies range from changes to the prediction model[52] to clear and direct policies for usage. We strongly advise against professionals relying on their intuition rather than the results of post-conviction risk and needs assessment instruments due to abundant evidence showing that unaided human judgments are less accurate and more biased overall.[53] The clearly documented evidence of systemic bias in the criminal justice system, in particular, requires checks and balances on personal judgment in decisionmaking. A better assessment instrument option might be one that relies less on factors that are known sources of bias or may act as proxies for race, ethnicity, and gender and instead focuses on a person's current behavior and functioning.

---

51. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84.
52. See, for example, Berk and Kuchibhotla, "Improving Fairness."
53. Grove et al., "Clinical Versus Mechanical Prediction," 19–30; Jung et al., "Simple Rules," 771–800; Lin et al., "The Limits of Human Predictions," eaaz0652.

Principle III: Transparency

The third set of guidelines speak to strategies that agencies can utilize to promote transparency in the use of post-conviction risk and needs assessment instruments. Transparency refers to how and what information about the content, structure, and application of these instruments is disseminated to stakeholders. Transparency is relevant in both the development and implementation of risk and needs assessment instruments and requires a proactive communication strategy.

"

Transparency is a necessary step to accountability. (Eaglin, 2017, 111)

"

We recommend the following guidelines to promote transparency of post-conviction risk and needs assessment instruments:

8. Provide system stakeholders with relevant information on the development, intended use, and validation of the post-conviction risk and needs assessment instrument.
9. Develop a written policy that guides the local use of the post-conviction risk and needs assessment instrument.
10. Communicate the strengths and limitations of the post-conviction risk and needs assessment instrument to the general public.

Guideline 8: Provide system stakeholders with relevant information on the development, intended use, and validation of the post-conviction risk and needs assessment instrument.

Overview

All system stakeholders should have the information they need to understand the assessment process and be able to use this information to determine for themselves whether the process is fair and the results are accurate. This means that the information must be both available and understandable. Yet it is neither realistic nor necessary for the entirety or specifics of the process to be understood by everyone. For example, a defense attorney, defendant, or community member does not necessarily need to know the specifics of a technology such as the mechanism of machine learning algorithms;[54] however, they should have enough information to be able to question the assessment content and results, and how they are being used.[55] By informing system stakeholders about the development, intended use, and validation of post-conviction risk and needs assessment instruments, we can achieve greater transparency (and accuracy) than is possible through assessments of risk and needs based on human judgments alone.

Ideally, instrument developers and researchers will make the information described below available from the outset. Indeed, the availability and accessibility of this information to the public should be key considerations when selecting post-conviction risk and needs assessment instruments.

Action Items
Articulate the purpose for which the post-conviction risk and needs assessment instrument was developed, including the intended settings, populations, and outcomes.

Post-conviction risk and needs assessment instruments differ in their intended purpose, setting, population, and outcome. Some instruments were designed with a primary focus on estimating the likelihood of recidivism, while others were designed to also inform case planning, including supervision and intervention. Some instruments were designed for specific settings (e.g., jail, prison, reentry, community-based supervision) or populations (e.g., people in detention, on parole or probation, etc.), while others were designed for more general application. Many were designed to estimate general risk of recidivism, including committing a new crime or violating conditions of probation or parole. Some are focused specifically on assessing risk of violence, and others estimate risk of specific forms of violence such as sexual violence and domestic violence. Moreover, the timeframe over which instruments estimate risk may differ from days to weeks to months to years. Given this wide variation across instruments, it is important to clearly state the purpose, setting, population, and outcomes for which the post-conviction risk and needs assessment instrument was developed to ensure that it is applied as intended.

---

54. Chiao, "Fairness, Accountability and Transparency," 126–39.
55. Carlson, "The Need for Transparency," 303–29.

Explain the set of factors the assessment considers—including their definitions, scoring, and weighting—in a manner that can be understood by different audiences, particularly those who will be using the results and those who will be affected by them.

Agencies should describe the assessment's administration approach and data sources in sufficient detail so stakeholders can understand the process and any issues that may arise regarding the veracity of the information gathered (e.g., misrepresentation of circumstances or events, incomplete data, data entry errors). Although there is considerable overlap, not all post-conviction risk and needs assessment instruments include the same risk and needs factors, nor are these factors defined, scored, and weighted similarly across instruments. For example, depending on the instrument, "criminal history" may include age at first arrest, number of prior arrests, or number of convictions in the past 10 years, among many other definitions. One instrument might define "substance use" as any current drug use, while the other might define it as any lifetime alcohol or drug addiction. Because definitions, measurement, and weighting of these factors will affect risk estimates and have different implications for different groups, the general definitions, rating guidelines, and weighting must be described in plain language. This language should appear not just in technical manuals but also in other easily accessible outlets such as agency websites.

Additionally, the manner through which information is gathered to inform item ratings and the sources of this information also differ among post-conviction risk and needs assessment instruments. Some post-conviction risk and needs assessment instruments exclusively use information from official records, whereas others incorporate self-reporting. Others, still, require structured interviews with the individual being assessed or people with whom they interact (e.g., family members) and observations of behavior and functioning. Some post-conviction risk and needs assessment instruments are computerized and automated, while others are completed on paper. Again, sufficient description of the methods for information gathering should be provided for system stakeholders to consider the sources and potential issues with the information needed to complete the assessments.

Describe how risk levels are assigned.

The manner in which item ratings are combined to produce risk levels representing an estimated likelihood of recidivism differs across instruments. Some instruments use a simple checklist approach that involves adding item ratings to arrive at a total score.[56] Other instruments use an algorithmic approach that combines and weights item ratings using more advanced statistical models. These total scores are cross-referenced (by hand or via a computer program) with actuarial tables that describe probabilities or rates of recidivism seen in development, norming, or validation samples. Other instruments use a structured professional judgment approach in which assessors rate the items for their presence, severity, and relevance and use them to estimate the risk level based on their

---

56. Burgess, "Factors Determining Success or Failure on Parole," 221–34.

professional judgment (rather than computed scores). Finally, other instruments may use a checklist or statistical approach to produce an initial risk level that can be adjusted for individual case circumstances or considerations; in other words, an **adjusted actuarial method**.[57]

Agencies should identify and clearly describe the method of assigning risk levels for both lay and technical audiences. For the lay audience, a simple description of the general approach for how item ratings and risk scores relate to risk levels and how these risk levels, in turn, relate to recidivism may suffice. This information is typically included in instrument manuals, but agencies should also provide it in other easily accessible outlets (e.g., websites, information repositories, printed documents). For the technical audience, links or contacts for further detailed information on the mathematical models and training data should be provided.

Outline the training requirements for people administering the instrument, including CQI elements described previously.

As described in **Guideline 3**, all assessors must complete *all* required training before they complete a post-conviction risk and needs assessment in practice, including:

- Training on the strategies to gather and interpret information, procedures for rating items, how to interpret results, and how to apply results to inform practices.
- Completion of four to five practice cases.

All assessors should complete booster trainings, at least annually, after initial training. While instrument manuals typically contain general training requirements, agencies should also make these requirements available to local stakeholders who will be involved in, or affected by, the use of the post-conviction risk and needs assessment instrument as part of the implementation process. A summary of the local plans for implementing the initial training, case reviews, and booster training should also be outlined in a written policy and made available by the agency for review and input from stakeholders. See **Guideline 10** for more on specific strategies to support stakeholder involvement.

Publish the findings of validation studies examining the post-conviction risk and needs assessment instrument in a manner that is accessible to a variety of audiences.

The traditional approach of publishing the findings of validation studies examining post-conviction risk and needs assessment instruments in scholarly journal articles or agency reports is not sufficient. The findings must be made available to system **stakeholders** in forms that are readily accessible, understandable, and useful to them. This can be achieved in many different ways and formats; for example, in a high level, short overview of the study findings or a more detailed research brief summarizing the study purpose, methods, and findings. If such summaries or research briefs do not already exist, they should be developed through collaboration between researchers and

---

57. Picard et al., *Beyond the Algorithm*.

system stakeholders to ensure the accuracy and comprehensibility of the content. It may be helpful to consult with experts in science communication to ensure that the study findings are written in a manner that is not just accessible but also understandable to a variety of audiences. When these products are complete, agencies should make them available through posting/linking on websites, social media, information repositories, or other outlets that can be accessed by the public and are not behind a paywall.

Guideline 9: Develop a written policy that guides the local use of the post-conviction risk and needs assessment instrument.

Overview

Developing a written policy, ideally before using the instrument, will not only guide local practices but also help system stakeholders understand how the use of the post-conviction risk and needs assessment instrument may affect people in the criminal justice system. The written policy should describe the sources of information used to complete the assessments, including potential pitfalls that may exist in these sources, and the contexts in which and how the assessment results will be used. Doing so will promote greater transparency—and accountability—in the use of post-conviction risk and needs assessment instruments to inform case decisions, resource allocation, and service provision in the agency. For these reasons, developing a written policy and amending it as necessary is essential to the successful implementation of a post-conviction risk and needs assessment instrument.

Action Items

Describe the source(s) of information that will be used to complete the post-conviction risk and needs assessments locally and identify potential pitfalls, such as data quality or biases, that may exist in these sources.

The written policy should include the protocols that were developed to guide the local use of the post-conviction risk and needs assessment instrument, as discussed in Guideline 3. The protocols should describe the sources of information that will be used to complete the ratings, how and by whom that information will be gathered (e.g., record review, interviews, self-report questionnaires, etc.), and what potential concerns there may be with the data quality or potential biases that may exist in the data. For instance, there may be known issues as they relate to local record keeping for certain types of information, or there may be concerns that stem from the nature of the information source more generally. That said, it is important to balance concern regarding potential biases with the actual veracity of the information.[58] It may be valuable to conduct informal reviews of sources for accuracy of

---

58. For example, while there are reasons to question the veracity of self-reported information, we often find that self-report of criminal behavior is more—not less—accurate than official records. Similarly, collateral informants, such as family members, are often used to corroborate information; however, there may be cases in which there has been limited contact between family members and the individual being assessed and, consequently, family members may not provide accurate information on current behaviors, functioning, and circumstances.

information at both the individual and system levels. As previously noted, it is not necessary to detail all possible pitfalls in the data but, rather, to sufficiently describe the information sources for stakeholders to consider and question the veracity of data used to complete the assessments.

Define the contexts in which and how the results of the post-conviction risk and needs assessment instrument will be used to inform case decisions, resource allocation, and service provision.

Because many post-conviction risk and needs assessments have been developed for various settings and populations, it is important to define how and in which specific context(s) the results of the post-conviction risk and needs assessment instruments will be used locally. As noted above, these issues will have been addressed in the development of the local administration protocols (see Guideline 3). The task is now to ensure that these protocols are adequately described in the broader written policy. Specifically, it should include a clear description of which individuals should be assessed, when initial and repeat assessments should be completed (e.g., within 2 weeks of intake and every 6 months after), and how results will inform decisions and processes.

Create the opportunity for input on the written policy from stakeholders.

Seeking input from stakeholders contributes to transparency by creating an opportunity for individuals, various groups, and members of the public to understand and influence decisions that may affect them—directly or indirectly. To do so in a meaningful way, agencies should seek input at various points in the policy development process and on the specific issues where the input has a real potential to help inform the policy. Sometimes the opportunity for shaping the policy will be limited; at other times, there may be greater flexibility and opportunity for influencing the policy. Inviting input from stakeholders does not mean that agencies must necessarily change policy in response to the feedback gathered. Rather, it provides a forum for considering and responding to a wide range of views and concerns, as possible and appropriate. It is also an opportunity to foster trust, gain buy-in, and improve interagency and agency-community relations.

Different strategies can be used to gather input on the written policy from stakeholders. These could include individual interviews, focus groups, community cafés, study circles, written response requests (via email or other format), mail or online surveys, electronic polling, or public meetings, hearings, or workshops. In selecting strategies, agencies should consider what information stakeholders may need to make informed contributions and whether stakeholders may benefit from hearing from each other. Agencies should also determine if there are specific groups that may need additional outreach to ensure that their opinions are heard, whether there is a need to have comments on public record, and the timeframe for review and input. Regardless of the strategy, agencies should gather input from diverse stakeholders to ensure a wide range of views and concerns are considered and to promote meaningful involvement and inclusion with respect to race, ethnicity, gender, mental illness, and other

characteristics. Once agencies have gathered the input, it is their responsibility to balance and interpret it, decide whether to change the policy to address concerns or views that were shared, and report to stakeholders how their input was considered and used.

Creating the opportunity for input from stakeholders can be challenging, but the benefits are significant. It can help support better outcomes by facilitating implementation of policy that is better understood by stakeholders and reflects their interests and values. Further, gathering stakeholder input can develop system capacity to solve and manage issues that may stem from differing views and misunderstandings regarding post-conviction risk and needs assessment instruments.

Establish a process and timeline to review and update the written policy, as necessary.

Because it may be challenging to do on an ad hoc basis, we recommend that agencies establish a process and timeline to review and update the written policy, ideally during the planning period prior to using the instrument. As described in Guideline 1 in relation to revalidation efforts, we recommend that agencies identify and allocate resources and staffing to support the policy review and update during this planning process. Specifically, we recommend that agencies conduct the policy review and update following the instrument revalidation at least every 5 years. Doing so will ensure that the policy review and update can account for the findings of the revalidation in addition to other changes in the agency (e.g., staffing and resources) and local criminal justice practices, policies, and populations, as relevant. Note that agencies may also need to review and update the policy between revalidations to account for major circumstantial changes.

Guideline 10: Communicate the strengths and limitations of the post-conviction risk and needs assessment instrument to the general public.

Overview

Agencies may use a variety of strategies to communicate to the general public the strengths and limitations of the post-conviction risk and needs assessment instrument selected for local implementation. Some strategies such as public meetings, briefings, or telephone contact require person-to-person communication, whereas others can be accomplished remotely such as through printed information (e.g., fact sheets, newsletters, bulletins), websites, information repositories, press, and social media. There is no one best or most appropriate approach for communicating this information to the general public. Instead, agencies should consider a range of factors, including the current level of knowledge and understanding of criminal justice processes, public preferences for receiving information, and forms of communication that may be more or less effective across groups as a function of accessibility, language, literacy, and trust. The key is to ensure that the information is available, understandable, and accessible across groups. At a minimum, we recommend that the information be included on agency websites.

Ideally, this would occur prior to implementation. For agencies already using a post-conviction risk and needs assessment instrument, however, this should be completed as soon as possible.

It can reasonably be expected that, without such efforts, community members will have limited knowledge and understanding of post-conviction risk and needs assessments. In addition to promoting the principle of transparency, there are benefits to ensuring that community members have the information necessary to evaluate and develop an informed opinion on the post-conviction risk and needs assessment instrument. Indeed, public opinion can have a substantial impact on the success or failure of policy implementation as it relates to post-conviction risk and needs assessment instruments or otherwise.[59]

Action Items
Make information on the instrument's purpose, content, and validation available for easy access by the general public.

Building upon the actions necessary to meet the requirements of Guideline 8, agencies should describe in lay terms the purpose, content, and validation of the post-conviction risk and needs assessment instrument selected for use. Specifically, agencies should make the following information available for easy access by the general public:

- What the post-conviction risk and needs assessment instrument is and is not designed to do.
- How item selection and weighting minimize racial, ethnic, and gender disparities in assessment results while promoting accuracy.
- How the post-conviction risk and needs assessment instrument has been evaluated, including studies of predictive bias and disparate impact, as well as any limitations or gaps in research that remain to be addressed.

Whether this requires additional communication strategies beyond those implemented under Guideline 8 will need to be determined on a case-by-case basis. To that end, agencies should evaluate whether the general public is aware of and has access to the materials, outlets, websites, etc. through which this information is currently made available. They should also assess whether the information presented is likely to be understood by a wide audience. Additionally, agencies should implement a plan for how to raise community awareness of where this information is located (e.g., through press release, social media).

Describe the process through which the post-conviction risk and needs assessment instrument was selected for implementation.

A critical aspect of transparency is outlining the process through which the post-conviction risk and needs assessment instrument was selected for implementation. Agencies should describe the issues and evidence that were considered as well as the stakeholders—individuals or groups, as appropriate—who participated in the selection process. In addition to the considerations outlined in these guidelines, there are various resources available to support agencies in instrument selection such as the Public Safety Risk Assessment Clearinghouse's

---

59. Burstein, "The Impact of Public Opinion on Public Policy," 29–40.

Tool Selector ⧉ or the guidance provided in the report on **Risk Assessment Instruments Validated and Implemented in Correctional Settings in the United States** ⧉. Whatever the process, agencies should ensure that they are documenting each step, including decisions made along the way. The goal is to describe the process in sufficient detail so that community members will understand how and why a certain instrument was selected from the many instruments available for use. Documenting this process also may benefit agencies themselves by establishing institutional knowledge that may be lost over time as a result of staffing changes or turnover.

Clearly state how the post-conviction risk and needs assessment instrument will be used locally.

Again, the requirement is not to duplicate efforts described in Guideline 9 but to make the policy available to the general public in a clear and concise manner once it is finalized and whenever it is revised. Community members need not know the details of the training and administration protocols but, rather, more generally, how and when assessments will be completed, for whom, and to inform what types of decisions. This communication may naturally flow from efforts under Guideline 9 to create opportunities for input on the written policy from stakeholders. However, if community members are not included in that process, agencies must identify the strategy(ies) that will be used to ensure that a clear statement of use is released publicly before using the instrument.

Explain how the accuracy and impact of the post-conviction risk and needs assessment instrument on case outcomes will be monitored overall and across groups.

Drawing from the plan derived to meet the requirements of Guidelines 2 and 5, prepare a short description of the methods that will be implemented to examine the performance and consequences of assessment results as they are locally used. Agencies should write this description in lay language and provide information on the general approach, key indicators of performance and impact, and, importantly, efforts that will be implemented should these efforts highlight any issues of concern. This should include a brief summary of actions that may be taken at the individual, group, or system levels.

Principle IV: Effective Communication and Use

The final three guidelines speak to strategies that agencies can utilize to promote effective communication and the use of post-conviction risk and needs assessment instrument results. The manner in which assessors communicate individual assessment results can greatly affect their impact on decisionmaking and, consequently, their effectiveness. It is only through effective communication of assessment results that they can appropriately inform case decisions, resource allocation, and service provision.[60] Improper communication of individual assessment results can undermine efforts to promote accuracy, fairness, and transparency in the use of post-conviction risk and needs assessment instruments. Communication, then, must be a central consideration in planning, training, and implementation.

> **"**
>
> *Improper risk communication can render a risk assessment that was otherwise well-conducted completely useless—or even worse than useless, if it gives consumers the wrong impression.* (Heilbrun, Dvoskin, Hart, & McNiel, 1999, 94)
>
> **"**

We recommend the following guidelines to promote the effective communication and use of post-conviction risk and needs assessment instruments:

11. Anchor communication of post-conviction risk and needs assessment results in the RNR principles.
12. Contextualize the results of the post-conviction risk and needs assessment instruments.
13. Develop a template for communicating the individual results of the post-conviction risk and needs assessment instrument to all relevant stakeholders, including the person being assessed.

60. Heilbrun et al., "Violence Risk Communication," 91–105.

Guideline 11: Anchor communication of post-conviction risk and needs assessment results in the RNR principles.

Overview

In addition to supporting fairness, as described in Guideline 6, the application of the RNR model to the post-conviction risk and needs assessment can promote effective communication and use of the results. Specifically, RNR provides a framework for helping assessors identify what information should be communicated about the assessment results and the recommended intervention to different stakeholder groups.[61] As discussed further in Guideline 13, effective communication does not mean sharing all information derived during the assessment process but, rather, focusing on what information is necessary to support decisionmaking. Indeed, when presented with too much information, decisionmakers will rely on prior experiences and personal biases, including stereotypes, to discern the relative importance and weight of the various pieces of information.[62]

Action Items

Describe assessment results as placing an individual in a particular risk level that informs the minimum level of intervention needed to mitigate their risk of recidivism rather than assigning a specific probability or likelihood of recidivism to the individual.

The results of most post-conviction risk and needs assessment instruments provide information on how the individual was rated, scored, and ranked in terms of their risk of recidivism in relation to a group of people who were assessed using the post-conviction risk and needs instrument.[63] It would be a mistake—and potentially misleading—to assign specific probability or likelihood of recidivism to the individual. Instead, assessment results should be described as placing an individual in a particular risk level that informs the minimum level of intervention needed to mitigate their assessed risk of recidivism. In keeping with the Risk principle, intervention, including supervision and services, should be commensurate with the assessed level of risk. That is, individuals at the lowest risk level should receive the least intensive intervention and those at the highest risk level, the most intensive intervention. However, the most intensive intervention should still represent the least restrictive conditions within which the risk can be managed. As such, the most intensive intervention could still be community placement and services.[64]

Judges and other decisionmakers often desire a combination of categorical and numerical information on risk.[65] Consequently, the rate of recidivism observed among those who were placed in that risk level in the norming or validation samples can be shared but with the clear specification that this is not to be understood as the individual's absolute probability or likelihood of recidivism.

---

61. Heilbrun, Newsham, and Pietruszka, "Risk Communication." 2016
62. Tversky and Kahneman, "Judgment under Uncertainty," 1124.
63. Harris, Lowenkamp, and Hilton, "Evidence for Risk Estimate Precision," 111–27.
64. Desmarais and Lowder, "Principles and Practices of Risk Assessment," 593–603.
65. Kwartner, Lyons, and Boccaccini, "Judges' Risk Communication Preferences," 185–94; Evans and Salekin, "Violence Risk Communication," 143–64.

Identify the presence of risk and protective factors that contribute to the assessment results, emphasizing the dynamic factors and needs that should be addressed through intervention.

In keeping with the Need principle, communication of assessment results should emphasize the dynamic factors and needs that contribute to recidivism risk for that individual and that, consequently, should be addressed in intervention. This communication should not just name the risk and protective factor in the abstract. Instead, it should provide a very brief operational definition of the factors as specified by the post-conviction risk and needs assessment instrument as well as a description of the specific behavior, attitude, or circumstance as it presents in the person who was assessed. These definitions and descriptions should be very short—just a few words will do in many cases—but sufficient to convey the issue that needs to be addressed.[66]

While static risk factors may be relevant to individual risk, they are not modifiable. As such, they should be communicated briefly, if at all, in relation to the initial intensity of intervention recommended, unless subsequent behavior (e.g., supervision failure, new crime) results in a higher static risk score. Unless static risk factors can be translated into some modifiable form, they should not be integrated into case planning. Further, we recommend that ratings for items that were not deemed to be relevant to individual risk (whether present or not) are excluded from communication.[67] Including these ratings may inadvertently—and mistakenly—convey that the items should be addressed through intervention.

Explain case-specific barriers or facilitators to successful habilitation, above and beyond those described previously.

Facilitate the application of the Responsivity principle by articulating any case-specific issues that may undermine or otherwise detract from the effectiveness of intervention. As discussed in Guideline 6, this may include communicating case-specific barriers or facilitators to successful habilitation that should be considered in the development and tailoring of case plans (i.e., specific responsivity), including those that relate to the individual and their environment. This may also include identifying and recommending interventions with demonstrated effectiveness in addressing the dynamic factors and needs identified during the assessment process (i.e., general responsivity).[68]

66. Heilbrun et al., "Assessing Normative Approaches to Communicating Violence Risk," 187–96.
67. Storey, Watt, and Hart, "An Examination of Violence Risk Communication," 39–55.
68. Dowden and Andrews, "The Importance of Staff Practice in Delivering Effective Correctional Treatment," 203–14.

Guideline 12: Contextualize the results of the post-conviction risk and needs assessment instrument.

Overview

The ways in which assessment results are communicated to stakeholders will determine how they are used.[69] Consequently, communicating information about the context surrounding the assessment process and its results is necessary for balancing concerns of public safety with the promotion of individual rights and habilitation in subsequent decisionmaking and intervention.[70] Risk of recidivism is not an individual trait. Rather, it will depend upon the complex interaction of a person's characteristics with their social and physical environments. To that end, it is critical that the recipient of information about assessment results understands the circumstances surrounding the assessment and its results, the situations in which risk of recidivism would be elevated, and what can be done to prevent it.

Action Items
State the likelihood and, when possible, the type(s) of criminal behavior anticipated in the absence of interventions over the timeframe(s) specified by the instrument.

Communication of assessment results should include a clear statement on the likelihood of recidivism anticipated in the absence of intervention, as estimated using the post-conviction risk and needs assessment instrument. This may include a simple statement of the assessment results reporting placing an individual in a particular risk level, as described in Guideline 11. To the extent possible, however, the type(s) of criminal behavior anticipated in the absence of interventions, over the timeframe(s) specified by the instrument, should be clearly described. Recidivism is not one type of behavior. Instead, the behavior that would constitute "recidivism" varies in nature, frequency, and severity. For this reason, it may not be sufficient to make a general statement regarding risk level. Instead, the nature (e.g., nonviolent, violent, sexually violent) and severity of the anticipated behavior(s) as well as the potential harm to victim(s) should be specified to the extent possible. Further, communication should specify over what timeframe(s) and in what setting(s) the assessment results are intended to estimate risk of recidivism, if specified by the instrument. Because risk of recidivism is time and context dependent, a statement regarding how the level of risk might change over time or across settings can help inform case decisions and prioritize resource allocation and intervention. For example, such a statement could help identify outcomes that need more immediate intervention for victims and public safety compared to those that may help support successful habilitation in the long term.

---

69. Heilbrun et al., "Violence Risk Communication," 91–105.
70. Ignelzi et al., "Best Practices," 452–54.

The degree and specificity to which such information is known to assessors may differ. Some post-conviction risk and needs assessment instruments, for example, may produce different risk estimates for different types of recidivism (e.g., any criminal behavior, violent behavior, technical violations/infractions) and over different timeframes (e.g., 1 month, 6 months, 2 years, 5 years), while others may only speak to risk of recidivism in a general or aggregate way.

Define the parameters of the assessment results.

In addition to addressing the context of the assessed risk, the context of the assessment itself should be communicated. This information will help support evaluations of the accuracy and fairness of the assessment results, promote transparency, and, ultimately, provide for the appropriate application of assessment results. To that end, the purpose of the assessment should be clearly stated—not only the conditions that prompted the assessment (e.g., intake to a new facility) but also the decision(s) and processes the assessment results are intended to inform (e.g., custody level, case planning, program placement, etc.). There also should be a clear but brief description of the administration protocols, including the sources of information used and any concerns or limitations regarding that information (see Guideline 3). As previously noted, data sources can reflect bias and affect the accuracy and fairness of the assessment results. As such, a cautionary statement regarding confidence in the accuracy of the current assessment results may be warranted and, if so, factors that affected confidence such as mixed or inconsistent information in the data sources that could not be resolved. Finally, conditions that would prompt re-assessment should be specified (see Guideline 4 for more on re-assessment).

Identify the type and approximate intensity of interventions that are likely to reduce the anticipated risk of recidivism and support successful case outcomes.

The estimated likelihood of recidivism reflects the absence of intervention. Consequently, it is necessary to specify the minimum level of intervention needed to manage that risk of recidivism (Risk principle), the interventions that are likely to be successful for that individual (Need principle), and any case-specific considerations for promoting the effectiveness of the intervention (Responsivity principle; see Guideline 6). It is also important to clearly communicate how assessment results inform those interventions. Again, assessment results will only improve outcomes if they are used to inform decisionmaking, resource allocation, and service provision in meaningful ways. For these reasons, identifying the type and intensity of interventions that are likely to be effective is a critical step in effective risk communication.

Few post-conviction risk and needs assessment instruments will produce specific recommendations for risk management and intervention. Instead, and as discussed in Guideline 6, agencies should develop local guidelines describing the frequency, type, and intensity of supervision and services that can be quickly referenced to inform this communication before using the instrument.

Guideline 13: Develop a template for communicating the individual results of the post-conviction risk and needs assessment instrument to all relevant stakeholders, including the person being assessed.

Overview

A template for communication helps outline and structure what information assessors will share with different groups of stakeholders about the assessment process and results. The development and implementation of a standard template for the written communication of individual assessment results will improve comprehension and use of the results and can reduce assessor effort and time. Specifically, having a standard template for written communication may help overcome barriers to effective communication and use of assessment results by reducing chances for factual error, misrepresentation of assessment information, or presentation of misleading or irrelevant information. Using a standard template for written communication can also streamline the presentation of assessment results for ease of understanding and increase the predictability of information that will be communicated.

While the focus here is on the development of a standard template for written communication, we suggest that this template also be used as the foundation for oral communication of the findings such as in a courtroom or in meetings.

Action Items

Provide a structure and format for presenting the assessment results in a manner that is clear, concise, predictable, and consistent across assessors and cases.

The exact structure and format of the written communication template may vary from agency to agency. However, the following information should be included:

- A brief statement on the instrument that was used and the sources of information used to complete the assessment.
- The estimated risk levels.
- The identified risk and protective factors, needs, and case-specific barriers or facilitators to successful community reintegration.
- The recommendations for intervention (if appropriate).

Also, we recommend against selectively reporting individual item ratings. Doing so may overburden the audience and unintentionally emphasize individual factors in ways that are inconsistent with their contributions to the overall risk estimates.

Importantly, assessors should receive training on the communication template and how to use it as part of the pre-service training process (see Guideline 3) to maximize its use and effectiveness. Other stakeholders also should receive a brief training on the template to increase the predictability of information communicated about the assessment and to support their comprehension of assessment results. If possible, integrate the template into existing electronic reporting tools or as a fillable form to promote implementation of the template with fidelity.

Use communication strategies that promote comprehension and reduce the impact of potentially problematic information.

Assessment information can be complex and difficult to understand. Agencies must make efforts to promote comprehension through the use of evidence-based communication strategies. In particular, people tend to comprehend more and make more informed decisions when the important information is easy to evaluate and understand. To that end, assessment results should be presented in accordance with cognitive expectations (e.g., higher numbers mean greater risk).[71] Further, less is indeed more when it comes to the communication of assessment results.[72] Consequently, only the most relevant information about risk and needs should be communicated such as the risk factors and needs that were present and relevant rather than those that were not. Efforts also should be made to avoid technical jargon and information that requires inferences or interpretation;[73] for example, it is better to provide the estimated likelihood of recidivism as a percentage rather than as a number out of 100 (or some other denominator).[74]

Additionally, the typical presentation of historical information first—followed by information about the current case and present functioning—may unintentionally emphasize and anchor decisionmaking in what has occurred in the past (i.e., criminal history) rather than the present circumstances and current functioning of the individual.[75] For this reason, we recommend structuring the template to follow the Situation, Background, Assessment, and Recommendation (SBAR) communication strategy[76] such that information on the current case and circumstances is presented first, followed by the background or historical case information, then the assessment results (i.e., estimated risk level), and last, recommendations for supervision and services. Finally, communication of assessment information should avoid language that may be biased and inadvertently perpetuate prejudicial beliefs. Guidelines for bias-free language should be consulted in the development of the communication template.[77]

71.   Peters et al., "Numeracy Skill," 741–48.
72.   Peters et al., "Less Is More," 169–90.
73.   Heilbrun et al., "Violence Risk Communication," 91–105.
74.   Monahan et al., "Communicating Violence Risk," 121–26.
75.   Dror, "A Hierarchy of Expert Performance," 121–27.
76.   Thomas, Bertram, and Johnson, "The SBAR Communication Technique," 176–80.
77.   See, for example, "Bias-Free Language," APA Style, accessed March 25, 2021, https://apastyle.apa.org/style-grammar-guidelines/bias-free-language/.

Tailor communication to the target audience, with the potential for different templates for different stakeholders, but avoid sharing assessment results beyond relevant stakeholders.

For communication to be effective, it must be tailored to the target audience. This may mean developing different templates or modifying the standard template for different stakeholders. For example, how information on the assessment process and results is shared with a judge may differ from how this information is shared with the individual who was assessed or a service provider to whom the individual may be referred. Some audiences may need detailed information, while others may only need a high level summary. Tailoring communication requires knowing how the audience prefers to receive information, what information is relevant to them and their decisionmaking, and their level of knowledge about post-conviction risk and needs assessment, generally, and the instrument used, specifically.[78] It also is important to consider factors that may affect communication accessibility such as literacy, preferred language(s), or abilities.

There may be gaps in knowledge regarding these various communication preferences, needs, and barriers. Agencies can most easily gather the information needed to help tailor communication to the target audience by involving diverse stakeholders in the template development process.

Share the template with stakeholders for review and feedback prior to finalizing it.

Sharing the template with stakeholders prior to using the instrument will afford them the opportunity for input that can be used to promote the appropriateness, acceptability, and effectiveness of the template across audiences. For agencies that have already implemented a post-conviction risk and needs assessment instrument, developing and sharing a communication template with stakeholders for review and feedback should be prioritized and completed over a short, but feasible, timeframe. As discussed in relation to the written policy (Guideline 9), seeking input from stakeholders can also promote transparency in the use of post-conviction risk and needs assessment instruments, foster trust, and improve relations. Again, inviting input from stakeholders does not mean that agencies must necessarily change the template in response to the feedback, but it does provide the opportunity for ensuring that the assessment information is being communicated as intended; the organization of information makes sense to the audience; the desired content is included; and the language and format are appropriate. With such input, the template may be better received and given greater consideration by stakeholders in their decisionmaking.

78. Heilbrun, "Prediction Versus Management Models Relevant to Risk Assessment," 347; Schopp, "Communicating Risk Assessments," 939.

Glossary

Acceptable levels of accuracy: The accuracy with which the results of post-conviction risk and needs assessment instruments predict the outcome they were intended to predict (e.g., recidivism) indicated by area under the curve (AUC) values of .64–.71. (See good validity.)

Accuracy: The degree to which results of post-conviction risk and needs assessment instruments predict the recidivism outcomes they were designed to predict.

Adjusted actuarial method: An actuarial approach to post-conviction risk and needs assessment in which the statistically derived risk estimate can be adjusted for individual case circumstances or considerations through the use of professional judgment (i.e., professional or clinical override) to increase or decrease the risk estimate.

Algorithmic (or actuarial) approach: An approach to post-conviction risk and needs assessment that combines and weights item ratings using statistical models that produce risk levels representing an estimated likelihood of recidivism. The total scores are cross-referenced (by hand or via computer program) with actuarial tables that describe probabilities or rates of recidivism seen in development, norming, or validation samples.

Area under the curve (AUC): In this context, a predictive validity performance indicator measuring the probability that a randomly selected person who recidivated during follow-up would have received a higher risk score or level using a given risk assessment approach than a randomly selected person who did not recidivate during follow-up.

Between-groups design: An evaluation design in which one compares outcomes between two or more groups that receive different interventions to measure the effectiveness of an intervention; for example, comparing placement decisions of one group of people who were assessed using a post-conviction risk and needs assessment instrument (i.e., intervention group) with another group that was not (i.e., comparison group).

Bias-free language: Language that demonstrates inclusive treatment of people and sensitivity with respect to race, ethnicity, gender, age, and other categories or identities. It involves avoiding terminology that may be hurtful, offensive, or perpetuate prejudicial beliefs.

Case review: Part of the continuous quality improvement (CQI) process, case reviews examine fidelity to the rating and scoring guidelines, adherence to the implementation protocols, and concordance between assessment results and case decisions, resource allocation, and service provision.

Checklist approach: An approach to post-conviction risk and needs assessment that involves simply adding item ratings to arrive at a total score of the number of items endorsed as present, where lower scores reflect lower risk of recidivism and higher scores reflect higher risk of recidivism.

Continuous quality improvement (CQI): A structured process that expands upon basic quality assurance methods, examining aggregate data on processes, practices, and outcomes to identify areas for improvement at the organizational or system level and to implement needed improvements.

Disparate impact: When results of post-conviction risk and needs assessment instruments are applied inequitably across groups, leading to adverse agency or system-level responses to one group of people, such as a group defined by race, ethnicity, or gender, as compared to another group.

Dynamic risk factors: Factors that contribute to risk but can change over time (e.g., social networks, thinking patterns, housing, substance use, finances, etc.), also called criminogenic needs. Dynamic risk factors not only add to the predictive ability of an assessment instrument, they represent those areas that can be changed through programming and interventions.

Effective communication and use: When the results of the post-conviction risk and needs assessment instruments are shared, discussed, and applied with strategies that promote understanding, accuracy, transparency, and positive case outcomes.

Evaluation: The systematic investigation of the results of a post-conviction risk and needs assessment instrument to determine its performance and effect on case decisions, resource allocation, and service provision.

Excellent agreement: Concordance among assessors who administer a post-conviction risk and needs assessment instrument indicated by (1) observed agreement of 90 percent or greater, (2) Kappa of .75–1.00, or (3) intra-class coefficient (ICC) of .75–1.00.

Excellent validity: The accuracy with which the results of post-conviction risk and needs assessment instruments predict recidivism indicated by area under the curve (AUC) values of .71–1.00.

Fairness: The equitable use of results from post-conviction risk and needs assessment instruments to inform case decisions, resource allocation, and service provision overall. This principle considers the degree to which assessment results have similar meanings and applications across groups, as it relates to racial, ethnic, and gender disparities in post-conviction processes.

Fidelity: The degree to which a post-conviction risk and needs assessment instrument is used as intended, including adherence to scoring guidelines, administration protocols, and local policies for use in practice.

General responsivity: Subprinciple of the Responsivity principle positing that the use of cognitive social learning methods will be most effective at reducing recidivism.

Good agreement: Concordance among assessors who administer a post-conviction risk and needs assessment instrument indicated by (1) observed agreement of 80 percent or greater, (2) Kappa of .60–.74, or (3) intra-class coefficient (ICC) of .60–.74.

Good validity: Accuracy with which the results of post-conviction risk and needs assessment instruments predict recidivism indicated by area under the curve (AUC) values of .64–.71. (See acceptable levels of accuracy.)

"Gold standard": An assessment completed by an instrument developer or other expert that serves as the criterion against which to compare the accuracy of ratings completed by an assessor in the context of training or use in practice.

Group level: Characteristics, knowledge, attitudes, skills, behaviors, or other attributes of multiple people together.

Individual level: Characteristics, knowledge, attitudes, skills, behaviors, or other attributes of a single person.

Inter-rater reliability: The degree to which assessors who administer a post-conviction risk and needs assessment instrument achieve the same results when assessing the same person. This is a property of the assessment results rather than of the post-conviction risk and needs assessment instrument itself.

Intra-class coefficient (ICC): The measure of inter-rater reliability representing the strength of agreement among multiple assessors on continuous variables (e.g., total scores), statistically corrected for chance.

Item: Component of a post-conviction risk and needs assessment instrument that is used to document the presence and/or severity of a risk or needs factor.

Kappa: Measure of inter-rater reliability representing the percentage of categorizations (e.g., low, moderate, or high risk) upon which multiple assessors agreed, statistically corrected for chance.

Minimum level of intervention: The lowest amount and intensity of supervision, resources, and services that is necessary to manage an identified level of recidivism risk.

Need principle: The principle of the Risk-Need-Responsivity model positing that treatment and case management should target the identified dynamic risk factors and criminogenic needs that can be positively impacted through services, supervision, and supports to reduce recidivism. The greater the number of dynamic risk factors and criminogenic needs are addressed through interventions, the greater positive impact those interventions will have on reducing recidivism.

Norming: In the development of an actuarial post-conviction risk and needs assessment instrument, the process through which population-based recidivism rates for each risk level or category are established. Individual assessment results are then compared against these risk levels or categories.

Observed agreement: The measure of inter-rater reliability representing the percentage of categorizations (e.g., low, moderate, or high risk) upon which multiple assessors agreed.

Observed rates of criminal behavior: The proportion of people within each risk level who went on to recidivate divided by the total number of people who were rated at that risk level.

Outcomes: In the context of post-conviction risk and needs assessment validation, the specific form(s) of recidivism that is being forecasted (e.g., general offending, violent crime, sexual violence).

Performance thresholds: Well-established scientific standards for measuring the strength or degree of agreement among assessors (i.e., inter-rater reliability) or between the assessment results and recidivism (i.e., predictive validity).

Population: The specific group(s) of people in the criminal justice system (e.g., people in detention, on parole or probation, etc.) for which a risk and needs assessment instrument is intended and validated for use.

Predictive bias: When the results of a post-conviction risk and needs assessment instrument consistently demonstrate different levels of predictive validity across groups (e.g., race, ethnicity, gender).

Predictive validity: The accuracy with which results of the post-conviction risk and needs assessment instrument forecast the outcomes they were intended to predict (e.g., recidivism). This is a property of the assessment results rather than of the assessment instrument itself.

Pre-post test design: An evaluation design in which the outcome of interest is assessed at least two times (i.e., pre-test and post-test) in order to measure the effectiveness of a new treatment or intervention; for example, recidivism rates or detention rates are examined before and after the implementation of a post-conviction risk and needs assessment instrument.

Protective factors: Characteristics of a person (e.g., attitudes, substance use), their environment (e.g., neighborhood, family, peers), or situation (e.g., housing, employment) that is associated with a decrease in the likelihood of recidivism.

Protocols for administration: Written documentation that describes for whom post-conviction risk and needs assessments will be completed and by whom, the sources of information that should be used to complete the assessments, what decisions and processes they inform, and when re-assessments should be conducted.

Purpose: The primary goal of implementing a risk and needs assessment instrument (e.g., predicting the likelihood of recidivism, informing case planning, etc.).

Quasi-experimental design: A type of between-groups evaluation design in which one compares outcomes between two or more groups to measure the effectiveness of a given intervention. In this evaluation design, there is no random assignment; rather, participants are assigned to groups based on other criteria. For example,

such an evaluation might involve comparing placement decisions in a jurisdiction where a post-conviction risk and needs assessment instrument has been implemented (i.e., the intervention group) to placement decisions in another jurisdiction that has not implemented a post-conviction risk and needs assessment instrument (i.e., the comparison group).

Randomized controlled trial (RCT): A type of between-groups evaluation design in which one compares outcomes between two or more groups of participants who are randomly assigned to receive different interventions to measure the effectiveness of an intervention. For example, such an evaluation might involve comparing placement decisions for participants who were randomly assigned to be assessed using a post-conviction risk and needs assessment instrument (i.e., the intervention group) to placement decisions for participants who were randomly assigned not to be assessed using a post-conviction risk and needs assessment instrument (i.e., the control group) within one jurisdiction.

Responsivity principle: The principle of the Risk-Need-Responsivity model positing that individual and system-level efforts to provide cognitive behavioral treatment and reduce barriers to positive learning outcomes (e.g., tailoring to reading ability, motivation, strengths) will promote the effectiveness of interventions in reducing recidivism.

Risk and needs assessment: The process of estimating the likelihood of future criminal behavior and identifying the dynamic risk and needs factors that may serve as treatment targets in the development of risk management and treatment plans.

Risk and needs assessment instrument: An instrument—composed of empirically or theoretically based risk (and in some tools also protective) factors—used to estimate the likelihood of future criminal behavior and to inform decisionmaking following convictions.

Risk-Need-Responsivity (RNR) principles: The RNR principles are a set of research-based guiding principles that, when implemented correctly, can help reduce reoffending and violations of conditions of probation and parole and help policymakers, administrators, and practitioners determine how to allocate resources, deliver services, and provide the right people with the right supports and services to have the greatest impact on recidivism and public safety.

Risk principle: The principle of the Risk-Need-Responsivity model dictating that the level and intensity of supervision, treatment, and other services should be proportionate to a person's assessed level of risk of recidivism.

Risk screening instrument: A short, easily administered set of items to quickly identify (1) individuals who are at potentially heightened risk of recidivism and who should, therefore, receive a more in-depth, comprehensive risk and needs assessment (i.e., screened "in") versus (2) individuals who pose limited risk of recidivism and, thus, do not need to be evaluated further (i.e., screened "out").

Setting: The specific location or stage of criminal justice processing (e.g., jail, prison, reentry, community-based supervision, etc.) in which a risk and needs assessment instrument is intended and validated for use.

Specific responsivity: The subprinciple of the Responsivity principle emphasizing the importance of considering and addressing individual and environmental characteristics that may act as barriers to intervention effectiveness; for example, building relevant staff skills, addressing prejudicial beliefs among staff, or "fine-tuning" services or interventions such as modifying cognitive behavioral treatment to account for a cognitive impairment associated with mental illness.

Stakeholders: An individual or group with a vested interest in a criminal justice agency's work, including professionals who work within or with the criminal justice system, such as judges, attorneys, service providers, and probation/parole officers, as well as people in the criminal justice system and their families.

Static risk factors: Factors that are unchanging or that cannot be changed through deliberate intervention (e.g., age, prior offenses). Static risk factors contrast with dynamic risk factors (or criminogenic needs), which can be used to inform the targets of supervision and human service interventions.

Structured professional judgment: An approach to post-conviction risk and needs assessment in which assessors estimate risk by considering a set number of factors that are empirically and theoretically associated with the outcome of interest. Total scores are not used to make the final judgments of risk; instead, assessors consider the relevance of each item to the person being assessed as well as whether there are any case-specific factors not explicitly included in the list.

System level: Organizations, policies, laws, practices, and structures that comprise a system such as the criminal justice system.

Systemic bias: Disparities in criminal justice system responses to one group of people with a protected characteristic (e.g., race, ethnicity, and gender) compared to another group, stemming from both current and historical discriminatory policies and practices. An example of systemic bias is higher rates of conviction among Black people compared to White people despite similar rates of criminal behavior. Although we have chosen to use the term "systemic bias" here, it is often interchangeable with "structural bias."

Transparency: The degree to which information about the content, structure, and application of post-conviction risk and needs assessment instruments is disseminated to stakeholders in an understandable manner.

Validation: An empirical evaluation used to determine the predictive validity of the results of a post-conviction risk and needs assessment instrument. (See predictive validity.)

Written communication template: A template that outlines and structures what information assessors will share with stakeholders about the assessment process and results in written communications (e.g., reports).

Bibliography

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014.

Barnes, Ashlee R., Nordia A. Campbell, Valerie R. Anderson, Christina A. Campbell, Eyitayo Onifade, and William S. Davidson. "Validity of Initial, Exit, and Dynamic Juvenile Risk Assessment: An Examination across Gender and Race/Ethnicity." *Journal of Offender Rehabilitation* 55, no. 1 (2016): 21–38. https://doi.org/10.1080/10509674.2015.1107004.

Berk, Richard, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. "Fairness in Criminal Justice Risk Assessments: The State of the Art." *arXiv* (2017). arXiv:1703.09207.

Berk, Richard A., and Arun Kumar Kuchibhotla. "Improving Fairness in Criminal Justice Algorithmic Risk Assessments Using Conformal Prediction Sets." *arXiv: Applications* (2020). https://arxiv.org/abs/2008.11664v1.

Bonta, James, and Donald A. Andrews. *Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation*. Ottawa, Canada: Public Safety Canada, 2007. https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/rsk-nd-rspnsvty/rsk-nd-rspnsvty-eng.pdf.

Breaux, Ariel, Yenys Castillo, Lara Guzmán-Hosta, Antoinette Kavanaugh, Rebecca Rivas, Danielle Rynczak, and Mark Worthen. "Race in Forensic Evaluations: The AP-LS Practice Committee Offers Practical Considerations." *American Psychology-Law Society Newsletter* (May 2021). https://ap-ls.wildapricot.org/resources/EmailTemplates/2021_05%20May%20AP-LS%20Newsletter/ClinicalForensicPractice5_21.pdf.

Buchanan, Alec, Renee Binder, Michael Norko, and Marvin Swartz. "Psychiatric Violence Risk Assessment." *American Journal of Psychiatry* 169, no. 3 (2012): 340. https://doi.org/10.1176/appi.ajp.2012.169.3.340.

Burgess, Ernest W., "Factors Determining Success or Failure on Parole." In *The Workings of the Indeterminate Sentence Law and the Parole System in Illinois*, edited by A. W. Bruce, E.W. Burgess, J. Landesco, and A.J. Harno: 221–34. Springfield: Illinois Board of Parole, 1928.

Burstein, Paul. "The Impact of Public Opinion on Public Policy: A Review and an Agenda." *Political Research Quarterly* 56, no. 1 (2003): 29–40. https://doi.org/10.1177%2F106591290305600103.

Carlson, Alyssa M. "The Need for Transparency in the Age of Predictive Sentencing Algorithms." *Iowa Law Review* 103 (2017): 303–29.

Chiao, Vincent. "Fairness, Accountability and Transparency: Notes on Algorithmic Decision-Making in Criminal Justice." [In English], *International Journal of Law in Context* 15, no. 2 (2019): 126–39. https://doi.org/10.1017/S1744552319000077.

Cicchetti, Domenic V. "The Precision of Reliability and Validity Estimates Re-Visited: Distinguishing between Clinical and Statistical Significance of Sample Size Requirements." *Journal of Clinical and Experimental Neuropsychology* 23, no. 5 (2001): 695–700. https://doi.org/10.1076/jcen.23.5.695.1249.

Cohen, Thomas H., and Christopher T. Lowenkamp. "Revalidation of the Federal PTRA: Testing the PTRA for Predictive Biases." *Criminal Justice and Behavior* 46, no. 2 (2019): 234–60. https://psycnet.apa.org/doi/10.1177/0093854818810315.

de Vogel, Vivienne, and Michiel de Vries Robbé. "Adapting Risk Assessment Tools to New Jurisdictions." In *International Perspectives on Violence Risk Assessment*, edited by Jay P. Singh, Stål Bjørkly, and Seena Fazel. American Psychology-Law Society Series, 26–39. New York: Oxford University Press, 2016.

Desmarais, Sarah L., and Evan M. Lowder. "Principles and Practices of Risk Assessment in Mental Health Jail Diversion Programs." *CNS Spectrums* 25, no. 5 (Oct 2020): 593–603. https://doi.org/10.1017/S1092852919001652.

Desmarais, Sarah L., Kiersten L. Johnson, and Jay P. Singh. "Performance of Recidivism Risk Assessment Instruments in U.S. Correctional Settings." *Psychological Services* 13, no. 3 (Aug 2016): 206–22. https://doi.org/10.1037/ser0000075.

Douglas, Kevin S., Randy K. Otto, Sarah L. Desmarais, and Randy Borum. "Clinical Forensic Psychology." *Handbook of Psychology, Second Edition* 2 (2012). https://doi.org/10.1002/9781118133880.hop202008.

Douglas, Kevin S., Jennifer L. Skeem, and Elizabeth Nicholson. "Research Methods in Violence Risk Assessment." In *Research Methods in Forensic Psychology*, edited by Barry Rosenfeld and Stephen D. Penrod, 325–46. John Wiley & Sons, Inc., 2011.

Dowden, Carig, and D. A. Andrews. "The Importance of Staff Practice in Delivering Effective Correctional Treatment: A Meta-Analytic Review of Core Correctional Practice." [In English], *International Journal of Offender Therapy and Comparative Criminology* 48, no. 2 (April 2004): 203–14. https://doi.org/10.1177/0306624x03257765.

Dror, Itiel E. "A Hierarchy of Expert Performance." *Journal of Applied Research in Memory and Cognition* 5, no. 2 (2016): 121–27. https://doi.org/10.1016/j.jarmac.2016.03.001.

Eckhouse, Laurel, Kristian Lum, Cynthia Conti-Cook, and Julie Ciccolini. "Layers of Bias: A Unified Approach for Understanding Problems with Risk Assessment." *Criminal Justice and Behavior* 46, no. 2 (2018): 185–209. https://doi.org/10.1177/0093854818811379.

Evans, Stephanie A., and Karen L. Salekin. "Violence Risk Communication: What Do Judges and Forensic Clinicians Prefer and Understand?" *Journal of Threat Assessment and Management* 3, no. 3-4 (2016): 143–64. https://doi.org/10.1037/tam0000062.

Feinstein, Alvan R., and Domenic V. Cicchetti. "High Agreement but Low Kappa: I. The Problems of Two Paradoxes." *Journal of Clinical Epidemiology* 43, no. 6 (1990): 543–49.

Freeman, Kelly Roberts, Cathy Hu, and Jesse Jannetta. *Racial Equity and Criminal Justice Risk Assessment.* Washington, DC: Urban Institute, 2021. https://www.urban.org/research/publication/racial-equity-and-criminal-justice-risk-assessment.

Gottfredson, Stephen D., and Laura J. Moriarty. "Statistical Risk Assessment: Old Problems and New Applications." *Crime & Delinquency* 52, no. 1 (2006): 178–200. https://journals.sagepub.com/doi/10.1177/0011128705281748.

Grove, William M., David H. Zald, Boyd S. Lebow, Beth E. Snitz, and Chad Nelson. "Clinical Versus Mechanical Prediction: A Meta-Analysis." *Psychological Assessment* 12, no. 1 (2000): 19–30.

Hamilton, Melissa. *Risk Assessment Tools in the Criminal Legal System – Theory and Practice: A Resource Guide.* Washington, DC: National Association of Criminal Defense Lawyers, 2020. https://www.nacdl.org/Document/RiskAssessmentReport.

Hanson, R. Karl, Guy Bourgon, Robert J. McGrath, Daryl G. Kroner, David A. D'Amora, Shenique S. Thomas, and Lahiz P. Tavarez. *A Five-Level Risk and Needs System: Maximizing Assessment Results in Corrections through the Development of a Common Language.* New York: The Council of State Governments Justice Center, 2017.

Harcourt, Bernard E., "Risk as a Proxy for Race: The Dangers of Risk Assessment." *Federal Sentencing Reporter* 27, no. 4 (2015): 237–43.

Harris, Grant T., Christopher T. Lowenkamp, and N. Zoe Hilton. "Evidence for Risk Estimate Precision: Implications for Individual Risk Communication." *Behavioral Sciences & the Law* 33, no. 1 (February 2015): 111–27. https://doi.org/10.1002/bsl.2158.

Hart, Stephen D., "Culture and Violence Risk Assessment: The Case of Ewert V. Canada." *Journal of Threat Assessment and Management* 3, no. 2 (2016): 76–96. https://doi.org/10.1037/tam0000068.

Heilbrun, Kirk. "Prediction Versus Management Models Relevant to Risk Assessment: The Importance of Legal Decision-Making Context." *Law and Human Behavior* 21, no. 4 (1997): 347–59. https://doi.org/10.1023/A:1024851017947.

Heilbrun, Kirk, Joel Dvoskin, Stephen Hart, and Dale McNiel. "Violence Risk Communication: Implications for Research, Policy, and Practice." *Health, Risk & Society* 1, no. 1 (1999): 91–105. https://doi.org/10.1080/13698579908407009.

Heilbrun, Kirk, Rebecca Newsham, and Victoria Pietruszka. "Risk Communication: An International Update." In *International Perspectives on Violence Risk Assessment.* New York: Oxford University Press, 2016.

Heilbrun, Kirk, Melanie L. O'Neill, Tomika N. Stevens, Lisa K. Strohman, Quinten Bowman, and Yi-Wen Lo. "Assessing Normative Approaches to Communicating Violence Risk: A National Survey of Psychologists." *Behavioral Sciences & the Law* 22, no. 2 (March 2004): 187–96. https://doi.org/10.1002/bsl.570.

Helmus, L. Maaike, and Kelly M. Babchishin. "Primer on Risk Assessment and the Statistics Used to Evaluate Its Accuracy." *Criminal Justice and Behavior* 44, no. 1 (2017): 8–25. https://doi.org/10.1177/0093854816678898.

Ignelzi, James, Bob Stinson, James Raia, Thomas Osinowo, Larry Ostrowski, and Jennifer Schwirian. "Best Practices: Utilizing Risk-of-Violence Findings for Continuity of Care." *Psychiatric Services* 58, no. 4 (2007): 452–54. https://doi.org/10.1176/ps.2007.58.4.452.

Jung, Jongbin, Connor Concannon, Ravi Shroff, Sharad Goel, and Daniel G. Goldstein. "Simple Rules to Guide Expert Classifications." *Statistics in Society* 183, no. 3 (2020): 771–800. https://rss.onlinelibrary.wiley.com/doi/10.1111/rssa.12576.

Kwartner, Phylissa P., Phillip M. Lyons, and Marcus T. Boccaccini. "Judges' Risk Communication Preferences in Risk for Future Violence Cases." *International Journal of Forensic Mental Health* 5, no. 2 (2006): 185–94. https://doi.org/10.1080/14999013.2006.10471242.

Lin, Zhiyuan J., Jongbin Jung, Sharad Goel, and Jennifer Skeem. "The Limits of Human Predictions of Recidivism." *Science Advances* 6, no. 7 (Feb 2020): eaaz0652. https://doi.org/10.1126/sciadv.aaz0652.

Lloyd, Caleb D., R. Karl Hanson, Dylan K. Richards, and Ralph C. Serin. "Reassessment Improves Prediction of Criminal Recidivism: A Prospective Study of 3,421 Individuals in New Zealand." [In English], *Psychological Assessment* 32, no. 6 (2020): 568–81. https://doi.org/10.1037/pas0000813.

Lowder, Evan M., Megan M. Morrison, Daryl G. Kroner, and Sarah L. Desmarais. "Racial Bias and LSI-R Assessments in Probation Sentencing and Outcomes." [In English], *Criminal Justice and Behavior* 46, no. 2 (2019): 210–33. https://doi.org/10.1177/0093854818789977.

Lowder, Evan M., Eric Grommon, and Bradley R. Ray. *Improving the Accuracy and Fairness of Pretrial Release Decisions: A Multi-Site Study of Risk Assessments Implemented in Four Counties*. Washington, DC: Bureau of Justice Programs, U.S. Department of Justice, 2020.

Lowenkamp, Christopher T., Matthew DeMichele, and Lauren Klein Warren. "Replication and Extension of the Lucas County PSA Project." *Advancing Pretrial Policy and Research*. 2020.

Marlowe, Douglas B., Timothy Ho, Shannon M. Carey, and Carly D. Chadick. "Employing Standardized Risk Assessment in Pretrial Release Decisions: Association with Criminal Justice Outcomes and Racial Equity." *Law and Human Behavior* 44, no. 5 (Oct 2020): 361–76. https://doi.org/10.1037/lhb0000413.

Mayson, Sandra G., "Bias in, Bias Out." *The Yale Law Journal* 128, no. 8 (2019): 2122–473.

Meade, Adam W., and Michael Fetzer. "Test Bias, Differential Prediction, and a Revised Approach for Determining the Suitability of a Predictor in a Selection Context." *Organizational Research Methods* 12, no. 4 (2009/10/01 2009): 738–61. https://doi.org/10.1177/1094428109331487.

Mills, Jeremy F., Michael N. Jones, and Daryl G. Kroner. "An Examination of the Generalizability of the LSI-R and VRAG Probability Bins." *Criminal Justice and Behavior* 32, no. 5 (2005): 565–85. https://doi.org/10.1177/0093854805278417.

Monahan, John, Kirk Heilbrum, Eric Silver, Erik Nabors, Jonathan Bone, and Paul Slovic. "Communicating Violence Risk: Frequency Formats, Vivid Outcomes, and Forensic Settings." *The International Journal of Forensic Mental Health* 1, no. 2 (2002): 121–26. https://doi.org/10.1080/14999013.2002.10471167.

Mosher, David K., Joshua N. Hook, Laura E. Captari, Don E. Davis, Cirleen DeBlaere, and Jesse Owen. "Cultural Humility: A Therapeutic Framework for Engaging Diverse Clients." *Practice Innovations* 2, no. 4 (2017): 221–33. https://doi.org/10.1037/pri0000055.

Nelson, Rebecca J., and Gina M. Vincent. "Matching Services to Criminogenic Needs Following Comprehensive Risk Assessment Implementation in Juvenile Probation." *Criminal Justice and Behavior* 45, no. 8 (2018): 1136–53. https://doi.org/10.1177%2F0093854818780923.

Orton, Laura C., Neil R. Hogan, and J. Stephen Wormith. "An Examination of the Professional Override of the Level of Service Inventory–Ontario Revision." *Criminal Justice and Behavior* (2020): 0093854820942270. https://doi.org/10.1177/0093854820942270.

Partnership on AI. *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System* (San Francisco: Partnership on AI, 2019). https://partnershiponai.org/wp-content/uploads/2021/08/Report-on-Algorithmic-Risk-Assessment-Tools.pdf.

Peters, Ellen, Nathan Dieckmann, Anna Dixon, Judith H Hibbard, and CK Mertz. "Less Is More in Presenting Quality Information to Consumers." *Medical Care Research and Review* 64, no. 2 (2007): 169–90. https://doi.org/10.1177/1077558707064002 0301.

Peters, Ellen, Judith Hibbard, Paul Slovic, and Nathan Dieckmann. "Numeracy Skill and the Communication, Comprehension, and Use of Risk-Benefit Information." *Health Affairs* 26, no. 3 (May/June 2007): 741–48. https://doi.org/10.1377/hlthaff.26.3.741.

Picard, Sarah, Matt Watkins, Michael Rempel, and Ashmini G. Kerodal. *Beyond the Algorithm: Pretrial Reform, Risk Assessment, and Racial Fairness*. New York: Center for Court Innovation, 2019. https://www.courtinnovation.org/sites/default/files/media/document/2019/Beyond_The_Algorithm.pdf.

Reynolds, Cecil R., and Lisa A. Suzuki. "Bias in Psychological Assessment: An Empirical Review and Recommendations." In *Handbook of Psychology: Assessment Psychology*, Vol. 10, 2nd Ed., 82–113. (Hoboken, NJ: John Wiley & Sons, Inc., 2013). https://doi.org/10.1002/0471264385.wei1004.

Rice, Marnie E., and Grant T. Harris. "Comparing Effect Sizes in Follow-up Studies: Roc Area, Cohen's D, and R." *Law and Human Behavior* 29, no. 5 (2005): 615–20. https://doi.apa.org/doi/10.1007/s10979-005-6832-7.

Schopp, Robert F. "Communicating Risk Assessments: Accuracy, Efficacy, and Responsibility." *American Psychologist* 51, no. 9 (1996): 939–44. https://doi.org/10.1037/0003-066X.51.9.939.

Shepherd, Stephane M., and Roberto Lewis-Fernandez. "Forensic Risk Assessment and Cultural Diversity: Contemporary Challenges and Future Directions." *Psychology, Public Policy, and Law* 22, no. 4 (2016): 427–38. https://doi.org/10.1037/law0000102.

Singh, Jay P., "Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer." [In English], *Behavioral Sciences and the Law* 31, no. 1 (January/February 2013): 8–22. https://doi.org/10.1002/bsl.2052.

Singh, Jay P., Sarah L. Desmarais, Brian G. Sellers, Tatiana Hylton, Melissa Tirotti, and Richard A. Van Dorn. "From Risk Assessment to Risk Management: Matching Interventions to Adolescent Offenders' Strengths and Vulnerabilities." [In English], *Children and Youth Services Review* 47 (December 2014): 1–9. https://doi.org/10.1016/j.childyouth.2013.09.015.

Singh, Jay P., Suzanne Yang, Edward P. Mulvey, and Ragee Group. "Reporting Guidance for Violence Risk Assessment Predictive Validity Studies: The Ragee Statement." *Law and Human Behavior* 39, no. 1 (February 2015): 15–22. https://doi.org/10.1037/lhb0000090.

Skeem, Jennifer L., and Christopher T. Lowenkamp. "Risk, Race, and Recidivism: Predictive Bias and Disparate Impact." *Criminology: An Interdisciplinary Journal* 54, no. 4 (2016): 680–712. https://doi.org/10.1111/1745-9125.12123.

Skeem, Jennifer, and Christopher Lowenkamp. "Using Algorithms to Address Trade-Offs Inherent in Predicting Recidivism." *Behavioral Sciences & the Law* 38, no. 3 (2020): 259–78. https://onlinelibrary.wiley.com/doi/10.1002/bsl.2465.

Slobogin, Christopher. "Primer on Risk Assessment for Legal Decision-Makers." *Vanderbilt Criminal Justice Program* (2020). https://scholarship.law.vanderbilt.edu/faculty-publications/1182/.

Smith, William R. "The Effects of Base Rate and Cutoff Point Choice on Commonly Used Measures of Association and Accuracy in Recidivism Research." *Journal of Quantitative Criminology* 12, no. 1 (1996/03/01 1996): 83–111. https://doi.org/10.1007/BF02354472.

Starr, Sonja B. "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination." *Stanford Law Review* 66, no. 4 (2014): 803–72.

Storey, Jennifer E., Kelly A. Watt, and Stephen D. Hart. "An Examination of Violence Risk Communication in Practice Using a Structured Professional Judgment Framework." *Behavioral Sciences & the Law* 33, no. 1 (January 2015): 39–55. https://doi.org/10.1002/bsl.2156.

Thomas, Cynthia M., Evelyn Bertram, and Doreen Johnson. "The SBAR Communication Technique: Teaching Nursing Students Professional Communication Skills." *Nurse Educator* 34, no. 4 (2009): 176–80. https://doi.org/10.1097/NNE.0b013e3181aaba54.

Tversky, Amos, and Daniel Kahneman. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185, no. 4157 (1974): 1124. https://doi.org/10.1126/science.185.4157.1124.

Vieira, Tracey A., Tracey A. Skilling, and Michele Peterson-Badali. "Matching Court-Ordered Services with Treatment Needs: Predicting Treatment Success with Young Offenders." [In English], *Criminal Justice and Behavior* 36, no. 4 (April 2009): 385–401. https://doi.org/10.1177/0093854808331249.

Vincent, Gina M., Laura S. Guy, and Thomas Grisso. *Risk Assessment in Juvenile Justice: A Guidebook for Implementation*. Chicago: John D. and Catherine T. MacArthur Foundation, 2012.

Vincent, Gina M., and Jodi L Viljoen. "Racist Algorithms or Systemic Problems? Risk Assessments and Racial Disparities." *Criminal Justice and Behavior* 47, no. 12 (2020): 1576-84. https://doi.org/10.1177%2F0093854820954501.

Zottola, Samantha A., Sarah L. Desmarais, Evan M. Lowder, and Sarah E. Duhart Clarke. "Evaluating Fairness of Algorithmic Risk Assessment Instruments: The Problem with Forcing Dichotomies." *Criminal Justice and Behavior* (August 2021). https://doi.org/10.1177%2F00938548211040544.