

Federal PROBATION

*a journal of correctional
philosophy and practice*

[Home](#)

The Construction and Validation of the Federal Post Conviction Risk Assessment (PCRA)

*James L. Johnson**

*Christopher T. Lowenkamp***

Scott W. VanBenschoten†

Charles R. Robinson

Office of Probation & Pretrial Services

Administrative Office of the U.S. Courts

[Principles of Effective Risk Classification](#)

[Brief History of Risk Assessment Tools](#)

[Purpose of a Risk Assessment Tool](#)

[Evolution of Risk Assessment Instruments](#)

[First Generation](#)

[Second Generation](#)

[Third Generation](#)

[Fourth Generation](#)

[Post Conviction Risk Assessment Tool](#)

[How the PCRA Came into Existence](#)

[Construction and Validation of the PCRA](#)

[Methods](#)

[Data Elements](#)

[Sample](#)

[Analysis](#)

[Findings](#)

[Discussion](#)

[Limitations and Future Research](#)

[Policy Implications](#)

[Editor's Note: The following article was accepted after under going a double-blind peer review process, in recognition of the critical importance of the Post Conviction Risk Assessment tool to the federal probation system's adoption of evidence-based practices.]

THE UNITED STATES PROBATION system was created in 1925 by the Federal Probation Act. This Act gave the U.S. Courts the power to appoint federal probation officers and the authority to sentence defendants to probation instead of a prison term. One of the primary functions of federal probation is to supervise convicted offenders who are sentenced to a term of probation or a term of supervised release following a period of imprisonment, and offenders released early from prison on parole or mandatory release by the U.S. Parole Commission or military authorities.

The federal probation and pretrial services system is organized into 94 districts within 11 regional circuits and operates under a decentralized management structure. As a result of being decentralized, each district operates with a great deal of autonomy; however, despite this autonomy, the system maintains cohesion through the Administrative Office of the U.S. Courts (AO). The AO serves as the administrative headquarters for this decentralized system and develops national policies that help districts in their efforts to protect the community and reduce recidivism.

During the past two decades, advancements in social science research, the need to use resources more efficiently and effectively, and increased expectations to reduce recidivism have sparked a major philosophical shift in the field of probation. Although probation officers are still required to monitor offender behavior and report noncompliance to the court, the general focus has shifted to reducing future criminal behavior (Alexander & VanBenschoten, 2008). Arguably, the best chance for reducing recidivism occurs when officers not only have a reliable way of distinguishing high-risk offenders from low-risk offenders but also can intervene in the criminogenic (crime supporting) needs of high-risk offenders (Andrews et al., 1990; Lowenkamp & Latessa, 2004; Bonta & Andrews, 2007; Campbell, French & Gendreau, 2007). For federal probation, this has meant looking for more effective ways to manage offenders by predicting their potential to reoffend and/or their potential dangerousness to the community (Walklate, 1999).

This article explains the process the AO used to develop a risk assessment instrument for use with its post-conviction supervision population. We provide a brief overview of the principles of effective classification and a summary of the evolution of risk assessments and then explain why the AO chose to create its own risk assessment instrument rather than use an existing instrument. However, the primary purpose of the article is twofold: (1) To present the methodology and results produced in the development of the Post Conviction Risk Assessment (PCRA) tool, and (2) to discuss limitations of the PCRA as well as future developments.

[back to top](#)

Principles of Effective Risk Classification

In general terms, the principles of effective risk classification refer to the prediction or identification of offenders most likely to violate the law or conditions of supervision during a period of criminal justice supervision, the identification of factors that can be influenced to change the likelihood of recidivism, and the acknowledgement of factors that might influence the benefits of a particular service (Van Voorhis & Brown, 1996). Risk of recidivism, criminogenic need, and general responsivity are three of the primary principles of effective classification (Andrews et al. 1990). The fourth principle, professional discretion, targets the professional's ability to look beyond the application of the first three principles when circumstances indicate a need to do so (Gottfredson, 1987).

The principles of effective risk classification suggest that agencies should use actuarial assessment tools to identify dynamic risk factors, especially in high-risk offenders, while also identifying potential barriers to treatment (Bonta & Andrews, 2007; Latessa et al., 2010). Actuarial risk assessments rest on three factors: (1) certain individual characteristics and behaviors are statistically predictive of future involvement in criminal behavior; (2) the more risk factors an offender has, the greater the likelihood of future criminal behavior; and (3) when properly validated and administered, actuarial risk predictions are more accurate than clinical predictions (Meehl, 1954; Sawyer, 1966; Gottfredson, 1987; Andrews and Bonta, 1994). Andrews and Bonta (1998) argue that it is the combined assessment of risk and need that improves the ability to predict who is likely to offend and outlines what interventions should take place to reduce risk and subsequently recidivism.

[back to top](#)

Brief History of Risk Assessment Tools

Purpose of a Risk Assessment Tool

The assessment of offenders has long been acknowledged as a necessary component for criminal justice practitioners who are responsible for assessing and managing offenders. In the field of probation, the primary purpose for using a risk assessment tool is to help keep communities safe from offenders who are most likely to reoffend. Although security was the primary reason for the development of risk assessment instruments, the ability to classify offenders at the appropriate risk level is also beneficial. Consequently, risk assessment tools help probation officers identify which offenders need intensive interventions and what needs should be targeted by the interventions.

Evolution of Risk Assessment Instruments

The evolution of risk assessment is described as following a generational path that started with the most basic form of assessment and has progressed to a more complex form of risk assessment (Bonta & Wormith, 2007). Each generation utilized the best available methods to predict the risk of recidivism and then applied the results of the assessment to supervision strategies. This tradition continues today, with researchers continually refining their understanding of criminal behavior and the associated enhancements to risk/needs prediction tools (VanBenschoten, 2008).

First generation

For most of the 20th century, professional judgment or intuition was the most common method used to predict criminal behavior. This form of assessment involved an unstructured interview with the offender and a review of official documentation (Bonta, 1996; Van Voorhis & Brown, 1996; Andrews & Bonta, 2006; Connolly, 2003). Guided by their own professional training and experience, probation officers and clinical professionals made judgments about who required enhanced supervision or correctional programming (Bonta & Andrews, 2007). One of the inherent weaknesses of such an unstructured process is the lack of a quantitative way to determine how decisions are reached, which leads to a lack of consistency and agreement resulting in low inter-rater reliability (O'Rourke, 2008). In other words, the same interview conducted by different interviewers could net dramatically different results; therefore, the conclusions and recommendations regarding the offender could vary depending on the interviewer (Wardlaw & Millier, 1978; Monahan, 1981; Van Voorhis & Brown, 1996).

Second generation

Although second-generation risk tools have been available since the late 1920s, it was not until the 1970s that the assessment of risk began to depend more upon actuarial, evidence-based science and less on professional judgment and intuition. Second generation risk assessments are often referred to as actuarial methods (O'Rourke, 2008). Actuarial risk assessments consider individual items (e.g., history of substance abuse) that have been demonstrated to increase the risk of reoffending and assign these items quantitative scores (Bonta & Andrews, 2007). Burgess (1928) established the first of these models. In the Burgess method, each variable in the model can be scored as a "point," and the prediction is based on the aggregate number of points assigned to an offender (Connolly, 2003). For example, the presence of a risk factor may receive a score of one and its absence a score of zero. The scores on the items can then be summed—the higher the score, the higher the risk that the offender will reoffend (Bonta & Andrews, 2007). This technique gives equal weight to all predictors, even though there may be unequal effects. There is little research, if any, indicating that more complex (i.e., weighted) scoring methods produce better prediction than simple (i.e., unweighted) methods (Gottfredson 1987).

Third generation

Recognizing the limitations of second-generation risk assessment, research began to develop in the late 1970s and early 1980s on assessment instruments that included dynamic risk factors (Bonta & Wormith, 2007). The third generation of assessment is commonly referred to as risk-need assessments (Andrews & Bonta, 1995; Bonta & Andrews, 2007). These instruments combined the static predictor variables of the second-generation instruments with dynamic criminogenic need items (e.g., present employment, criminal friends, and family relationships) that were sensitive to

changes in an offender's circumstances (Connolly, 2003; Bonta & Andrews, 2007). Third-generation risk assessment tools exceed statistical risk prediction by adding the element of need identification. As previous instruments assisted in decision-making regarding supervision conditions, third-generation assessments help identify areas that require intervention to mitigate recidivism risk while under supervision (Van Voorhis & Brown, 1996).

Fourth generation

The last few years has seen the introduction of fourth-generation risk assessment instruments. These new risk assessment instruments go beyond the third-generation risk-need assessments. Not only do fourth-generation instruments include risk-need assessments, they also assess a broader range of risk factors along with responsivity factors important to treatment for integration into the case management plan (Bonta & Andrews, 2007; Bonta & Wormith, 2007). Some examples of responsivity factors include reading and cognitive abilities, race, gender, motivation to change, as well as external factors such as treatment setting and counselor characteristics (Andrews et al., 1990; Bonta & Wormith, 2007). One other aspect of fourth-generation risk assessments is the attempt to explicitly link identified needs with supervision and treatment services (Bonta & Wormith, 2007).

[back to top](#)

Post Conviction Risk Assessment Tool

Actuarial risk assessments are not new to the federal probation system; in fact, they have been part of the supervision process since the early 1980s. To better assist probation officers in identifying high-risk offenders and intervening in their criminogenic needs, the AO chose to develop a risk assessment instrument tailored specifically to its population of offenders. The Post Conviction Risk Assessment (PCRA) is an actuarial risk and needs assessment tool developed from data collected on federal offenders who started a term of supervision between October 1, 2005 and August 13, 2009. This tool is designed to target treatment interventions prioritized by risk, need, and responsivity.

How the PCRA Came into Existence

In the *Strategic Assessment of the Federal Probation and Pretrial Services System* (hereafter cited as IBM, 2004), the authors identified shortcomings with the AO's use of the Risk Prediction Index (RPI).¹ One of the concerns expressed by the authors was the RPI's static nature, which causes a disconnection between the risk score and case management (IBM, 2004). Put another way, if an offender's risk to recidivate changes during the course of supervision, the RPI does not reflect this change; therefore, officers are not able to consistently and effectively interpret those changes and provide the proper supervision response.

To address the RPI's shortcomings, the Strategic Assessment recommended that the AO research other data-driven supervision tools (IBM, 2004). The desire to meet this recommendation, coupled with emerging criminal justice literature about more advanced risk assessment tools, influenced the AO to develop its own Research to Results (R2R) effort. During the R2R effort, 16 of the 94 federal probation districts were awarded funding to implement evidence-based practices² into their districts. Of those 16 districts, five districts chose to use a commercially available risk and needs tool to conduct risk assessments. In addition, AO staff members met with developers of three commonly used off-the-shelf risk/needs tools (LS/CMI, COMPAS, RMS)³ to better understand the advantages and disadvantages of each tool.

Since the federal criminal justice system represents a distinctive population and since specific trailer assessments for special needs populations (such as sex offenders) are also required, it became obvious that more flexibility would be needed. At the conclusion of the experimentation and information gathering stage, the AO assembled a panel of experts to examine the options of

purchasing a commercially available tool or building a new tool. After much discussion, the consensus of the group was to build a new tool with data specific to federal probation.

[back to top](#)

Construction and Validation of the PCRA

Methods

Data used to construct and validate the PCRA came from federal presentence reports (PSR), existing risk assessments, criminal history record checks, and PACTS.⁴ Criminal history records or rap sheets were used to identify any new arrest after the start of supervision. The five R2R districts that were using a commercially available risk assessment tool were asked by the AO to provide data to assist in the development of the PCRA.⁵ Each district provided a list of offenders who had received an assessment using an off-the-shelf risk prediction instrument and who also had a completed PSR. In total, the five districts submitted a list of 4,746 offenders, from which 479 cases were randomly selected.⁶ Districts were then asked to provide rap sheets on the randomly selected cases. PACTS was the main source of data for scored elements on the PCRA; it included data on roughly 100,000 offenders.

Data Elements

There are two sets of items included on the PCRA: scored and not scored. The first set of items are rated and scored and thus contribute to an offender's risk score. Rated and scored items used to develop the PCRA were based on prior research in the area of predicting criminal behavior (for example Gendreau, Little, & Goggin, 1996; Simourd & Andrews, 1994; Hubbard & Pratt, 2002; Andrews & Bonta, 2006) that were also available in PACTS. Based on a review of extant research, data elements related to criminal history, peer associations, family, employment, substance abuse, and attitudes were selected from PACTS. As a result of bivariate analyses, some interval and ratio variables (e.g., age, prior arrests, education, and drug and alcohol problems) were collapsed into ordinal measures. Multivariate models and completeness of data were used to identify the most predictive and practical data elements to be included on the instrument. Variables included on the PCRA had a significance level of .10 or below (see [Table 1](#)).

The second set of data elements are rated but not scored and do not contribute to an offender's risk score. These items were identified as potentially predictive in a smaller sample of offenders from five of the R2R districts. With the exception of peer relationships, which came from the COMPAS and RMS, data elements came from the PSR. A total of 104 elements were collected from the PSR; however, four of those elements were personal identifiers (i.e., first name, last name, middle initial, and PACTS number). Additional rated but not scored items were added based on probation officers' input on what data they need to supervise a case (see [Appendix 1](#)). A total of 29 factors were identified as potential predictors and included on the assessment. These potential predictors were included as "test items" and future analysis will determine whether these items will become rated and scored PCRA items.⁷

Sample

In order to construct and validate the PCRA, the researchers devised three sample groups. A construction group was created for the construction of the instrument, and two validation groups⁸ were created for the validation of the instrument. These groups were created using an existing analysis file from PACTS data that contained 185,297 offenders on probation or supervised release.⁹ The construction group was created from data obtained from the initial case plan.¹⁰ Using a near 50/50 split, data from the first case plan was divided into two sample groups; one became the construction sample and the other became the first validation group. One validation group (Validation) was taken from the initial case plan the offender receives during his or her term of supervision and the second validation group was taken from subsequent case plans (hence the

name Subsequent Case Plan). Both the construction (N=51,428) and validation (N=51,643) groups comprised offenders who started a term of supervised release or probation on or after October 1, 2005. The subsequent case plan group comprised 193,586 case plan periods.

Analysis

A fairly straightforward and traditional approach was used in the development of the PCRA. Multivariate logistic regression models¹¹ were used to determine which items were superfluous. As a result, the total number of items included in the multivariate model was reduced to ensure that statistical significance and direction of the relationship were maintained. Once the multivariate model was finalized, bivariate cross tabulations were used to assign appropriate weights. This method was chosen due to its transparency and, to date, there is little research indicating the superiority of complex weighting structures over dichotomous coding risk factors (see Gottfredson & Gottfredson, 1979; Silver, Smith, and Banks, 2000; Gottfredson & Snyder, 2005; Harcourt, 2007).¹² The bivariate cross tabulations are presented in [Appendices 2–4](#).

Once the final scoring algorithm was determined, a score was calculated with a cutoff score developed by visual inspection of the data. Although the data cutoffs were fairly evident in the data, alternate cutoffs were tested with confirmation of best fit as determined through the use of chi-square statistics. A final set of analyses was conducted to determine how changes or stability in risk category from the beginning to the end of supervision was correlated with change in the probability of a new arrest.

[back to top](#)

Findings

[Table 1](#) displays the results of a multivariate model predicting arrest during the initial case plan period using a split sample from the construction sample. As [Table 1](#) shows, many of the variables included in the multivariate model were statistically significant at the .001 level. Odds ratios in the model also appear to be consistent with existing research that support well-accepted beliefs that alcohol and drug problems, unemployment, poor attitude (not motivated to change), criminal history, and lack of social support increase an offender's chances of getting re-arrested. Females appear to have a decreasing effect on the likelihood of re-arrest, which is also consistent with much of the existing research on gender and crime (Gendreau et al., 1996).

From the multivariate analysis, variables were selected for inclusion on the risk assessment instrument (see [Appendix 5](#)). To gain a better understanding of the bivariate relationships between the significant predictors in the multivariate model, we conducted a series of cross tabulations. Those results are reported in [Appendices 2–4](#). In general, the bivariate cross-tabulations allowed us to assign 1 or 2 points to each of the factors. Although this approach may seem counter to prevailing wisdom on the development of weights for risk assessment, there is evidence that suggests that this approach produces an instrument that still outperforms clinical approaches to prediction (Dawes, 1979) and is more robust across time and sample variations (Gottfredson and Snyder, 2005; and McEwan, Mullen, & MacKenzie, 2009).

[Table 2](#) presents the descriptive statistics on the risk assessment score, which can theoretically range from 0 to 19. There are 15 scored items. The scoring for each of the 15 items is displayed in detail in [Appendix 5](#). [Table 2](#) presents the number of cases in each sample, minimum and maximum values, mean, and standard deviation of the linear risk score. There are no significant differences in the length of the prediction period or average risk score for the construction sample and first validation sample (6.46 and 6.43, respectively). However, there are differences in the mean risk score between the subsequent case plan sample and construction sample and subsequent case plan sample and first validation sample. The difference in prediction periods is a matter of policy, as the first case plan period is approximately 6 months while the third case plan is completed 12 months after the second case plan or 18 months after the beginning of supervision.

The lower mean risk score might simply be a function of lower-risk offenders surviving supervision to the third and subsequent case plan periods. At any rate, there could be some debate that the difference in risk scores is not practically significant, and this argument might be valid since all three mean scores fall into the low-risk category.

Table 3 presents the distribution of risk categories by the type of sample used. In all three samples, low and low-moderate risk offenders accounted for at least 85 percent of the cases, whereas high-risk offenders accounted for only 1 percent. There was no statistically significant difference between the construction sample and the validation sample at an alpha level of .01. However, there was a significant difference between the second validation sample (subsequent case plan) and the construction sample as well as between the second validation and the first validation sample. This is likely due to higher-risk offenders having a greater likelihood of revocation and thereby failing to survive to the second and subsequent case plan periods. This finding, like that of the linear risk score, might be more an issue of sample size rather than holding practical significance. The change in the percentage of low-risk cases seems to be what drives the overall significant chi-square test.

The next set of analyses focused on assessing the PCRA's predictive ability. AUC-ROC (Area of the Curve-Receiver Operating Characteristics)¹³ was chosen as the measure to assess prediction in large part because it is not impacted by base rates. Another convenient property of the AUC-ROC over a correlation coefficient is that AUC-ROC is a singular measure and does not have differing calculations depending on level of measurement of the variables being evaluated (Rice & Harris, 2005). Table 4 displays the AUC-ROC between risk scores and re-arrests. A fourth sample (long-term follow-up) that includes initial case plan data on all offenders placed on supervision between September 30, 2005 and September 30, 2006 is introduced in Table 4. The data therefore allow for a follow-up period between three and four years. As Table 4 shows, the AUC for each of the four sample groups is close to or exceeds the AUC-ROC value associated with large effect sizes (Rice & Harris, 2005). The AUC for the second validation sample rose to .73, while the AUC for the long-term follow-up sample rose even higher to .78. Based on these results, the PCRA appears to have very good predictive validity in terms of accurately classifying offenders' risk level.

To put the AUC values into practical terms,¹⁴ we calculated the failure¹⁵ rates by each category of risk for each sample. These results are presented in Table 5. With the exception of the long-term follow-up sample, the failure rates were relatively unchanged for a risk category across samples. For example, low-moderate risk offenders failed at a rate of 13 percent in both the construction and initial validation samples, and at 12 percent in the subsequent case plan sample. However, in the long-term follow-up sample, the low-moderate risk group's failure rate increased significantly to 42 percent. Overall, the failure rate for the long-term follow-up group was 44 percent, but the failure rate was significantly higher for high-risk offenders in this same group. Moderate-risk offenders failed at a rate of 71 percent and high-risk offenders had an 83 percent failure rate. The uniform increase in failure rates across categories of risk and across the various samples continues to support the validity of the PCRA.

Survival analysis was conducted for each risk category and the survival curves associated with those analyses are displayed in Figure 1. All possible data points, regardless of follow-up time, were used in the analysis.¹⁶ The follow-up period ranged from 0 to 60 months. Survival rates for each risk category are displayed at 6 months, 12 months, 36 months, and 60 months. As Figure 1 shows, high-risk offenders have a very steep decrease in survival, as only 69 percent survived the first 6 months of supervision. As time passes, survival rates continue to drop rapidly for high-risk offenders, as only 46 percent survived at 12 months and only 17 percent at 36 months. After 60 months of supervision, a mere 6 percent of the high-risk offenders remain. In contrast to high-risk offenders, low-risk offenders have a significantly different experience on supervision. For example, while the survival rate for high-risk offenders was only 17 percent at 36 months, 90 percent of the low-risk offenders survived at this time period. Moreover, the survival rate for low-risk offenders decreased only 5 percentage points through 60 months to 85 percent.

Low-moderate risk offenders have a survival curve that is almost precisely between the survival curves of the low- and moderate-risk cases. Interestingly, the survival curve for the moderate-risk offenders seems to follow a form that is closer to the high-risk offenders than to the lower-risk offenders. Note that the survival rates continue to grow throughout the follow-up period for each group, and each curve (with the exception of low-risk offenders) shows little sign of leveling off.

One of the major benefits of third- and fourth-generation risk assessment is the ability to measure change in risk over time. While many of the risk factors on the PCRA would be considered stable, some would also be considered acute (for a full discussion see Serin, Lloyd, & Hanby, 2010; Serin, Mailloux, & Wilson, 2010). Therefore, analyses were conducted that compared actual failure rates based on changes in initial and subsequent PCRA assessments. [Table 6](#) outlines changes in failure rates based on first and last case plan assessment categories. The failure rates are based on the risk category for the last case plan period of the offender's supervision term; therefore, to be included in this table, the offender had to have at least two case plan periods that allowed for the scoring of the PCRA. According to the results presented in [Table 6](#), not surprisingly, offenders in the higher risk categories (moderate and high) failed at a higher rate than offenders in the lower risk categories (low and low-moderate). However, offenders whose risk rating increased while under supervision appear to fail at a higher rate than offenders who maintained their initial rating through to their last assessment. For example, low-moderate risk offenders whose risk category increased to moderate had a failure rate of 41 percent, whereas low-moderate risk offenders who remained low-moderate risk or were reassessed as low risk had a failure rate between 16 and 18 percent. Similarly, moderate-risk offenders who continued to be moderate risk had a 38 percent failure rate, while those who were reassessed as low-moderate had an 18 percent failure rate and moderate-risk offenders reassessed as high risk had a 61 percent failure rate.

[back to top](#)

Discussion

As previously stated, the purpose of this article is twofold: (1) To present the methodology and results produced in the development of the PCRA and (2) to discuss limitations of the PCRA as well as future developments. This article has provided details on the methods, measures, and sample used in the development of the PCRA. A fairly traditional model was followed in the development of the PCRA. Our efforts were supported by a relatively large dataset and fairly complete data. The sample was fairly representative of the population served and allowed for a construction and two validation samples. The overall results have demonstrated that the PCRA provides adequate predictive validity both in the short term (6–12 months) as well as in longer follow-up periods (up to 48 months).

Multivariate analysis (see [Table 1](#)) of proposed predictors revealed that 15 factors were significantly related to the outcome of interest (new arrest). Seven additional factors tested were determined to be unrelated to a prediction of new arrest once the effects of the other factors were controlled. One additional measure, being female, was found to be significantly related to a new arrest. Subsequent models, not reported here, indicated that the addition of gender to the models yielded no increase in the predictive validity of the model. In addition, non-significant differences were noted in the AUCs between males and females for each sample (i.e., construction, validation, subsequent case plan, and long-term follow-up). Therefore, we concluded that the instrument performs equally well for males and females, even though the failure rates for males might be slightly higher than for females with similar risk scores.

The creation of the risk score and categories allowed for the identification of four risk categories: low, low-moderate, moderate, and high. Approximately 80 percent of each sample was made up of low and low-moderate risk offenders. Much smaller percentages were identified in each sample as moderate and high risk (approximately 12 percent and 1 percent, respectively). Due to the distribution of risk categories being heavily skewed toward lower risk, the validity of the instrument may be brought into question. However, it should be noted that a current validated risk

prediction instrument used in the federal system (RPI) yields a similarly skewed distribution. Analysis of failure rates by risk score and category using the PCRA yielded AUC-ROC values over the traditionally accepted value of .70 and an AUC value for the long-term follow-up over .78. All of the AUC-ROC values were close to or exceeding the value associated with large effect sizes. Practically speaking, the instrument provided categorizations that are associated with the group failure rates that are differentiated and meaningful for meeting the risk principle (see [Tables 4 and 5](#)).

The final analysis conducted in this study related to the dynamic nature of the PCRA. Recall from [Table 6](#) that changes in actual failure rates were associated with changes in risk category from the initial assessment to the last assessment. This finding is rather important, as it provides the opportunity to track meaningful changes in risk that occur throughout the supervision process. Moreover, [Table 6](#) confirmed that the PCRA identifies and measures dynamic risk factors that, apparently, when changed through supervision, services, or some other unmeasured process (natural desistance), lead to commensurate reductions in actual failure rates. The dynamic nature of the PCRA adds to its usefulness in developing case plans throughout the life of the supervision term.

Limitations and Future Research

Although this study was fairly comprehensive in scope and the dataset used was large and representative of the population served, there are a number of limitations and areas for future research that deserve mention. First, while the dataset was large and comprehensive, we have not investigated how scoring algorithms might be adjusted for each district. As with any measure, there is a distribution of AUC values when that test is calculated for each district. Data from 17 districts generated AUC values below .70; however, only three districts had 95 percent confidence intervals that failed to cross the .70 threshold. While this finding may have been due to small samples in some districts, subsequent analysis should focus on bringing AUC values between risk scores and re-arrests up to larger values.

A second limitation is that the data used in this research came from an administrative dataset. While it proved useful for our initial task of creating and validating a risk assessment instrument, it will be important to conduct similar validation analyses once we have an ample sample of offenders that were actually assessed using the assessment protocol.

The third limitation involves the nature of the outcome measure being predicted. In this research we focused exclusively on the likelihood of a re-arrest and not the severity of the offense. We found it important to assess and determine the likelihoods of re-arrest as a first step in the assessment process. Because we do recognize that there is more than one dimension to an assessment in the criminal justice system, future analysis will focus on predicting the dangerousness of an offender.

Fourth, while the PCRA is apparently dynamic, with changes in risk associated with changes in actual failure rates, it may not be sensitive enough for use on a monthly or shorter schedule. Due to the high value associated with a dynamic risk assessment, it will be necessary to make the PCRA more sensitive to change, or supplement it with a more sensitive trailer assessment that increases its utility as a guide to service allocation.

Finally, because rated but not scored items outnumber scored items on the assessment, future analysis will review the impact of rated but not scored items. For example, the PCRA currently has only one scored item in the area of cognitions. As a result of current testing on 80 self-report items that relate to criminal thinking styles, the number of scored items in the area of cognitions will likely increase. Continued analyses on rated but not scored items will also increase the understanding of the impact of self-reported attitudes, as well as guide adjustments to algorithms based on district, gender, and race differences, if relevant.

Policy Implications

Notwithstanding the limitations discussed above, there are two major policy implications that stem from this research. First, the federal probation system now has a dynamic fourth-generation risk assessment for use on offenders under its jurisdiction. The instrument can be used to identify higher-risk offenders for enhanced services (see Andrews et al., 1990) and can also be used to identify targets for change to be addressed by external service providers. The second major policy implication is the apparent necessity for ongoing reassessment. Data analyzed in this study indicate that changes in levels of risk are associated with changes in actual failure rates. Therefore officers need to monitor risk in a standardized way to ensure that supervision and services are having intended impacts. If intended impacts are not being achieved, then officers will be able to modify supervision services to reduce the risk of recidivism.

[back to top](#)

[Endnotes](#)

[References](#)

The articles and reviews that appear in *Federal Probation* express the points of view of the persons who wrote them and not necessarily the points of view of the agencies and organizations with which these persons are affiliated. Moreover, *Federal Probation's* publication of the articles and reviews is not to be taken as an endorsement of the material by the editors, the Administrative Office of the U.S. Courts, or the Federal Probation and Pretrial Services System. Published by the Administrative Office of the United States Courts www.uscourts.gov
[Publishing Information](#)