

Federal Probation  
September, 2012

Special Focus on: The 30<sup>th</sup> Anniversary of the Pretrial Services Act of 1982

**\*3 THE RE-VALIDATION OF THE FEDERAL PRETRIAL SERVICES RISK ASSESSMENT (PTRA)**

Timothy P. Cadigan, James L. Johnson

Christopher T. Lowenkamp

Office of Probation and Pretrial Services

Administrative Office of the U.S. Courts

University of Missouri - Kansas City

Copyright © 2012 by Timothy P. Cadigan, James L. Johnson, Christopher T. Lowenkamp

**THE UNITED STATES** Pretrial Services system was created in 10 demonstration districts by Title II of the Speedy Trial Act of 1974. The Act authorized the Director of the Administrative Office of the U.S. Courts (AO) to establish in 10 judicial districts demonstration pretrial services agencies to help reduce crime by defendants released to the community pending trial and to reduce unnecessary pretrial detention. Five of the pretrial services agencies were to be administered by the Probation Division (now the Office of Probation and Pretrial Services) and five by boards of trustees appointed by the chief judges of the district courts. Title II also instructed the Director to compile a report on the effectiveness of pretrial services in these demonstration districts.

The fourth and final report on the *of Title II of the Speedy Trial Act of 1974* was published on June 29, 1979. That report concluded that pretrial services should be expanded in the federal system. The report effectively made pretrial services the first implemented evidence-based practice in the federal probation and pretrial services system. The passage of the Pretrial Services Act of 1982 began a process of establishing pretrial services in the remaining 83 federal districts. Pretrial services cases in the District of Columbia are not classified as federal pretrial services cases by the Pretrial Services Act of 1982; thus there are only 93 pretrial services offices.

The federal pretrial services system, like all judiciary units, is highly decentralized. Each district has a great deal of autonomy, with the Administrative Office of the U.S. Courts working through a system of Judicial Conference committees to develop national policies and implement new processes and procedures like a risk assessment tool. This article explains the process used to develop the Pretrial Services Risk Assessment tool (PTRA), beginning with an overview of the literature for pretrial services risk assessments, moving to an explanation of the choice to create a federal risk assessment instrument rather than use an existing one, and concluding with the methodology and results produced in the re-validation of the PTRA.

## Literature Review

One area in which pretrial services originally led criminal justice research was actuarial risk assessment, with devices utilized in several of the larger cities, including Washington, D.C. and New York, long before post-conviction assessment devices were utilized in those cities. Unfortunately, use of such tools, while continuing in those cities, did not spread to other agencies as rapidly as they did in post-conviction assessment. Risk assessment is an area with enough significant differences between post-conviction and pretrial services to prevent much sharing between them. For example, pretrial services focuses significantly on failure to appear, which is not a focus of post-conviction; in contrast, post-conviction focuses on long-term recidivism, something which historically does not concern pretrial services. Therefore, at least theoretically, there is little crossover between the two disciplines in the area of risk assessment. While not a lot of work is being done in the literature on risk assessment in pretrial services when compared to post-conviction risk assessment literature, it is clearly the pretrial services area that has received the greatest research attention, and there are some studies of excellent quality (e.g., Toborg, Yezer, Tseng & Carpenter, 1984; Goldkamp & Gottfredson, 1988; Levin, 2006; VanNostrand, 2007; Gold-kamp & Vilcica, 2009; Lowenkamp & Whet-zel, 2009).

Toborg, Yezer, Tseng, and Carpenter provide an excellent place to begin the discussion to clearly identify the two types of selectivity bias inherent in the process. First, there is a group of arrested defendants who are detained; because of this detention, their propensity for pretrial arrest and failure-to-appear cannot be observed. This first form of bias is fairly common and is discussed in most research on pretrial services risk assessment initiatives. However, rarely seen is a discussion of the *second* form of selectivity bias, which involves defendants who are released under different scenarios: some are released without any restriction; others are released on various bond types or with various conditions that are based on individual characteristics (Toborg, Yezer, Tseng, & Carpenter, 1984:102). It is important to recognize possible errors so they can be reduced.

When a risk assessment tool was used, more defendants were released, on less restrictive\*4 conditions, and with no increase in failure-to-appear or rearrest rates, compared to similar defendants released without use of a risk assessment tool (Toborg, Yezer, Tseng, & Carpenter, 1984:105). The risk prediction tool Toborg, Yezer, Tseng, and Carpenter developed increased release rates by 12 percent, again with no appreciable increase in failure-to-appear or rearrest rates (Toborg, Yezer, Tseng, & Carpenter, 1984:58). Finally, their research concludes that the tool was more accurate for appearance than for safety (Toborg, Yezer, Tseng, & Carpenter, 1984:73). Risk tools, while tremendously useful in improving agency decision-making and ultimately release recommendations, have limitations. For instance, they are good at identifying groups of defendants who present various risks, but they cannot be totally accurate at the individual level (Toborg, Yezer, Tseng, & Carpenter, 1984:111). Low risk is not no risk, and that can be a difficult concept for decision-makers to support, so pretrial tools must do everything possible to limit errors. For example, when implementing a risk assessment tool, agencies need to convey to line staff the important limitation that the tool should not be followed blindly; therefore, permitting an officer to override the tool after staffing with the supervisor or some similar override methodology should be the standard.

Goldkamp and Gottfredson studied three urban jurisdictions and concluded that successful implementation of a risk assessment device requires strong judicial leadership (Goldkamp & Gottfredson, 1988:129). Gold-kamp and Gottfredson identified some ways to maximize success when strong judicial leadership was absent, through ongoing training, assessment of the officer's use of the tool, and annual or bi-annual certification of the officer's skills in using the tool. As the experience of the federal system, which lacks judicial involvement in the implementation of the risk assessment, will ultimately demonstrate, failure to involve judges makes acceptance more difficult. In addition, the Goldkamp and Gottfredson study confirmed the major findings of Toborg, Yezer, Tseng, and Carpenter's earlier research.

One of the great strengths of the Gold-kamp and Vilcica research is that it squarely takes on some of the most enduring "urban legends" of pretrial services risk assessment research. Most pretrial services agencies, including the federal

system, continue to capture data on and analyze the variable of community ties. While some of the fascination with community ties stems from its identification as an important variable in the granddaddy of all pretrial services research, the original Vera project, this variable likely endures because of its tremendous “face validity.” Its inclusion in the small number of long-standing important pretrial services variables is certainly not warranted by the research results of the last 20 years. However, most researchers merely ignore the variable of community ties, since the analysis does not bear out its value (e.g., Administrative Office of the United States Courts, 1979; VanNostrand, 2003; VanNos-trand & Keebler, 2009; Winterfield, Coggeshall, & Harrell, 2003). Goldkamp and Vilcica take on the lack of value of community ties for pretrial risk assessment in an effort to remove this variable from its lofty perch.

Goldkamp's analysis of factors influencing judicial decisions at the pretrial release decision, however, found that contrary to the intended effect of Vera-type information-based reform procedures community ties items did not play a significant role in shaping judges' actual pretrial custody decisions-and were not helpful predictors of defendant risk (Goldkamp & Vilcica, 2009: p. 124).

The seemingly “obvious” importance of including judicial officers in the development, implementation, and ongoing use of a risk assessment device is not found in virtually any other research on the topic of pretrial risk assessment. Only Goldkamp and Vilcica's findings discuss the issue of judicial involvement, not to mention endorsing the strong role it played in the Philadelphia research: “As a judicially developed and adopted policy, it stands alone in the nation in the first years of the 21st century-one might argue, in isolation-as an empirically informed approach to the problem of judicial discretion at the bail stage” (Goldkamp & Vilcica, 2009:129-30). This is an important finding for the federal system, as PTRAs were implemented without judicial involvement, which has clearly impacted the acceptance and use of the tool in the federal system.

Given Goldkamp and Vilcica's vision of pretrial justice and their desire to improve the pretrial release process and reduce judicial discretion, it is almost shocking that they missed the importance of pretrial detention and made the tool detention neutral (Gold-kamp & Vilcica, 2009:134). This is especially true since Philadelphia has operated pretrial services under federal court supervision due to jail overcrowding at various times during the 20-plus years of the guideline project in Philadelphia. Reducing unnecessary pretrial detention needs to be a core principle for pretrial services and judicial officers, given the negative consequences of pretrial detention at subsequent phases of the criminal justice system. The negative impacts on defendants have previously been documented in state, county, and local systems and will be established for the federal system in upcoming research by Oleson, Lowenkamp, and Cadigan.

Given that risk of failure to appear is only relevant in pretrial, we can't rely on post-conviction risk assessment research to establish it. Levin merged data from the Bureau of Justice Statistics State Court Processing Statistics (SCPS) program, which compiles criminal justice data (including pretrial) from the 75 largest counties in the nation, with Bureau of Justice Assistance survey data from 200 of the nations pretrial programs. The merged datasets enabled him to study over 1,500 defendants on conditional release in 28 counties during 2000 and 2002. That research revealed that a defendants odds of failing to appear in a county that uses a quantitative risk assessment are .40 times lower than the odds faced by a defendant appearing in a county that uses qualitative risk assessment (Levin, 2006:10). In addition, if the county uses some mix of quantitative and qualitative measures, defendants are still less likely to fail to appear (Levin, 2006:10). This result is particularly relevant to the federal system, because it is the approach now employed. Finally, if the county uses some mix of quantitative and qualitative measures, defendants are also less likely to be rearrested (Levin, 2006:11).

The literature on pretrial services risk assessment clearly establishes several important premises: “objective risk assessment produces more non-cash release recommendations” (Cooprider, 2009:15); “Notwithstanding a broader definition of ‘pretrial failure’ and cutting field contacts in half, violation rates declined or remained stable since the implementation of objective risk assessment” (Cooprider, 2009:15); and predictive items identified in pretrial services risk assess-

ment research change over time and therefore must be re-validated on an ongoing basis to ensure their integrity and effectiveness (e.g., VanNostrand, 2003; VanNostrand & Keebler, 2009; Siddiqi, 2002).

One example of an established risk assessment finding likely to change is a relatively consistent finding in risk prediction research in the city of New York for the past 20 years: the predictive value of having a telephone in the residence of the defendant. Given the changes in telecommunications in the past decade, \*5 from the dominance of landline technology to increasing reliance on cell phone technology, it seems unlikely that future research will continue to find great predictive value for a landline phone in the defendants residence (Siddiqi, 2002:2). Fortunately for citizens in New York City, the agency providing pretrial services has an excellent research operation that re-validates their risk prediction tool every three to five years as warranted. Ongoing re-validation is an essential step for all pretrial risk assessments and is the motivation for this research.

### **Pretrial Services Risk Assessment Tool**

Actuarial risk assessments are new to the federal pretrial services system; in fact, this is the first tool developed and implemented in the federal pretrial services system since its inception in the early 1980s. One tool was previously developed for use in the federal pretrial services system by Dr. John Gold-kamp and Dr. Barbara Meierhoefer. The tool was effective at identifying cases appropriate for release, tested effectively in 12 districts, and was submitted to the Judicial Conference Committee on Criminal Law for national implementation (Meierhoefer, 1994). Unfortunately, because it was named "Recommendation Guidelines" and was presented to the judges within two years of the implementation of Sentencing Guidelines, the tool was rebuked as too limiting to judicial discretion in the pretrial release decision. It took almost 18 years to overcome issues generated by the name of this tool.

The Administrative Office of the U.S. Courts works closely with the Office of Federal Detention Trustee, a Justice Department agency charged with administering and controlling the costs of pretrial detention in the federal system. That relationship led to a significant piece of research funded by the Office of Federal Detention Trustee using United States Court data and expertise to assist the researcher. The report on that research is titled *Pretrial Risk Assessment in the Federal Court* and has already led to the most significant improvement in the federal pretrial services system since its inception: the development and implementation of an actuarial risk assessment tool.

In addition to recommending a risk assessment tool, the Office of Federal Detention Trustee Report contains a number of interesting findings relevant to the operation of the federal pretrial services system. One of the primary goals of the system, reduction of unnecessary detention, is not being promoted by the staff, as they recommend detention more often than judicial officers actually detain defendants. Similarly, recommendations of detention by pretrial services officers rose each year, from 56 percent in 2001 to 64 percent in 2007. The report also observes that the risk posed by the defendants released increased slightly, from 2.85 in 2001 to 3.1 in 2007, as measured by the Risk Prediction Index (RPI). The Risk Prediction Index is a post-conviction measure of risk that was developed by the Federal Judicial Center and was implemented in federal pretrial services in 2004. However, it was only applied to or required to be completed on defendants who were released and subject to a condition of pretrial services supervision. For cases prior to 2004, the researcher abstracted the Risk Prediction Index score from the post-conviction record.

The study commissioned by the Office of Federal Detention Trustee tested for effectiveness the conditions of release known as alternatives to detention (substance abuse testing and treatment, third-party custody, halfway house placement, location (electronic) monitoring, and mental health treatment); the report contains a number of findings based on that analysis. First, low-risk defendants placed on location monitoring had an increased risk of failure compared to similar defendants who were not placed on location monitoring (Van-Nostrand & Keebler, 2009:32). In addition, location monitor-

ing was greatly overused on low-risk defendants. The only alternative to detention to positively impact defendants at all levels of risk, provided there was a demonstrated need, was mental health treatment (VanNostrand & Keebler, 2009:32). All four other alternatives to detention negatively impacted low-risk defendants (VanNostrand & Keebler, 2009:31-33).

What impact does over-supervising or over-treating low-risk federal defendants have on their outcomes? For the most part we have operated under the assumption that “it can’t hurt” to have conditions in place. Unfortunately the research demonstrates that unnecessary alternatives to detention placed on low-risk federal defendants can and do hurt defendant outcomes by increasing their failure rates.

First, the lower risk defendants, risk levels 1 and 2, are the most likely to succeed if released pending trial and in most cases release should be recommended. An alternative to detention, with the exception of mental health treatment when appropriate, generally decreases the likelihood of success for this population and should be recommended sparingly (VanNostrand & Keebler, 2009:10).

In some areas, for example location monitoring, level one defendants (the best risks) on location monitoring were 112 percent more likely to fail than if they were not on this type of monitoring (VanNostrand & Keebler, 2009:32). The quick refrain from most pretrial services professionals is: Of course there are more violations, due to the technical violations being counted as failures. However, this analysis did not include technical violations; it included only failure-to-appear and rearrest violations. In addition, the finding is not limited to location monitoring; substance abuse testing and treatment defendants are 41 percent more likely to fail. There are similar results for third-party custodians and halfway house placements. On average defendants released to the alternatives to detention program who were lower risk, risk levels 1 and 2, were less likely to be successful pending trial, while defendants in the moderate to higher risk levels (risk levels 3, 4, & 5) were more likely to be successful if released to the alternatives to detention program (Van-Nostrand & Keebler, 2009:31). VanNostrand and Keebler establish, apparently for the first time with hard national pretrial services data, the risk principle in federal pretrial services, which states “that the intensity of the program should be modified to match the risk level of the defendant” (Dowden & Andrews, 2004:1).

### **Federal Risk Assessment**

One of the major recommendations of the Office of Federal Detention Trustee research is that the pretrial services system should develop and implement an actuarial risk assessment tool. The Office of Probation and Pretrial Services hired a staff person proficient in the development of actuarial devices and ultimately developed the tool internally. The developed tool was piloted in several districts and the formal implementation of the tool began in January 2009. Currently there are 89 districts “live” using the tool on a majority of cases, 93 districts trained, and 93 with personnel certified in using the Pretrial Services Risk Assessment tool. National implementation was completed in all 93 districts by September 2011. Early results from the implementation show that the tool increases officer recommendations in favor of release, which is the desired goal of the implementation. There has as yet been no identified impact from the tool on release rates.

The Pretrial Services Risk Assessment \*6 tool was constructed using the same archival data employed in the Office of Federal Detention Trustee research. The PTR tool is an objective, actuarial instrument that provides a consistent and valid method of predicting risk of failure-to-appear, new criminal arrest, and technical violations that lead to revocation while on pretrial release. The instrument contains 11 scored and 9 unscored items. The unscored items are for future revisions to the instrument, and this research addresses the issues raised by the unscored items. The unscored items are rated as either A or B and do not contribute to the current overall risk score. The scored items are given a number of points (0, 1, or 2). The points from the items are then added up to give an overall score. When administered correctly, the Pretrial

Services Risk Assessment provides a score that allows for classification into a risk category. Those risk categories are then associated with rates of failure-to-appear, new criminal arrest, and technical violations leading to revocation.

When a defendant or material witness is arrested or summoned to appear before the court for an initial appearance, the magistrate judge typically requires a pretrial services report based on the investigation conducted by the pretrial services officer. The officer interviews the defendant to gather information for the report, the length of which varies somewhat, due to time constraints. The pretrial services report contains defendant case information, including residence, family ties, employment history, financial resources, health (including mental health and substance abuse histories), and criminal history. Based on this information, the officer will provide the court with an assessment of whether or not the defendant is likely to appear for court proceedings in the future or presents a danger to the community. Finally, the last section of the report provides the officer's recommendation to the court for the release or detention of the defendant. The recommendation should be based on the Pretrial Services Risk Assessment, although the officer can depart from the tool's recommendation after staffing the results with his or her supervisor.

The implementation of the tool has generated great debate over the finding, represented in the scores of "0" for defendants charged with violent offenses, that violent defendants in fact performed better than most other defendants in terms of rearrest, failure-to-appear, and technical violations leading to revocation of pretrial release in the construction research. The results found in the federal study are consistent with other similar findings: "defendants charged with more serious offenses do not pose a high risk of rearrest pending trial" (Austin, Krisberg, & Litsky, 1984:30; VanNostrand & Keebler, 2009:21; Toborg, Yezer, Tseng, & Carpenter, 1984:56). However, this validation research further refines that initial finding, showing violent defendants failing at higher rates than other defendant offense categories.

To better assist pretrial services officers in identifying high-risk defendants, the AO chose to develop a risk assessment instrument tailored specifically to its population of defendants. In doing so the AO looked at two existing tools: one operational in the state of Virginia and one used in the District of Columbia. After reviewing them, the AO concluded that its population of defendants differed enough from that of other pretrial services populations (for example, only federal courts address immigration charges) to warrant development of a tool using federal data. The Pretrial Services Risk Assessment (PTRA) is an actuarial risk and needs assessment tool developed from data collected on federal defendants who started a term of supervision between October 1, 2000 and September 30, 2007. This tool is designed to identify and categorize cases by risk of failure-to-appear, rearrest, and technical violations leading to revocation (FTA/NCA/Revocation).

## **Construction and Validation of the PTRA**

### **Data**

The archival data used to construct and validate the PTRA came from the Probation and Pretrial Services Automated Case Tracking System (PACTS). [FN1] Criminal history records or rap sheets were used to identify any new arrest after the defendant's release. PACTS was the main source of data for scored elements on the PTRA; it included data on 565,178 defendants. The data was extracted from PACTS in June 2008 and consists of all persons charged with criminal offenses in the federal courts between October 1, 2001 and September 20, 2007 (FY 2001- FY 2007) who were processed by the federal pretrial services system. The prospective data for the re-validation was extracted from PACTS in June 2012 and consists of all persons charged with criminal offenses in the federal courts between October 1, 2010 and September 30, 2011 (FY 2011) who were processed by the federal pretrial services system and from the Electronic Reporting System (ERS), which officers use to complete the PTRA.

### Data Elements

There are two sets of items included on the PTRAs: scored and not scored. The first set of items are rated and scored and thus contribute to a defendant's risk score. Rated and scored items used to develop the PTRAs were based on prior research by VanNostrand and the original construction research (Lowenkamp & Whetzel, 2009), and were available in PACTS. Using the extant research as a guide, available data elements models were constructed; the most predictive elements were ultimately included based solely on the data. Those elements are felony conviction (most predictive of available criminal history measures), pending felonies or misdemeanors, prior failures to appear, current charge, seriousness of current charge, employment, substance abuse, age, citizenship, education level, and home ownership. As a result of bivariate analyses, some interval and ratio variables were collapsed into ordinal measures. In the prior construction research, multivariate models and completeness of data were used to identify the most predictive and practical data elements to be included on the instrument.

The second set of data elements are rated but not scored and do not contribute to a defendant's risk score. These items were identified as potentially predictive by the Pretrial Services Work Group (PSWG). One additional rated but not scored item was added based on pretrial services officers' input on what data they felt strongly needed to be added: alcohol abuse. A total of 9 factors were identified as potential predictors and included on the assessment. These potential predictors were included as "test items" and the analysis determined that these items, for the most part, do not warrant becoming rated and scored PTRAs items. [FN2]

### Sample

That re-validation file contained 32,455 defendants for whom PTRAs have been completed in 2011, the first full year of operations. The total number of cases with PTRAs completed is 32,475, and the number of \*7 PTRAs cases opened and disposed of is 5,077. The cases were opened between October 1, 2010, and September 30, 2011. Given that PTRAs were validated using archival data and officers have now completed assessments prospectively, it is important to ensure that the tool is still valid.

### Findings

Table 1 displays the results of the test questions in relation to new criminal activity (NCA) and failure to appear (FTA), while Table 2 displays the results of the test questions in relation to NCA/FTA/Revocation. Adding current alcohol abuse and the various measures of foreign ties to the risk score produced no increase in the predictive ability of the TRA. Therefore, the authors recommend to the decision-making body that the nine unscored items not be added to the PTRAs and the collection of those items be discontinued.

**TABLE 1**

*Test Questions In Relation to FTA/NCA*

Question	Failure Rate	
	Yes	No

	N	%	N	%
<b>Current Alcohol</b>	33/625	5.3	200/4450	4.5
<b>Foreign Ties</b>	51/965	5.3	182/4110	4.4
<b>Foreign Ties Who Contact with Ties</b>	43/744	5.8	169/3719	4.5
<b>Foreign Citizen</b>	27/426	6.3	196/4381	4.5
<b>Passport</b>	60/1547	3.9	170/3431	5.0
<b>Foreign Financial Interests</b>	7/123	5.7	155/4808	4.6
<b>Travelled Outside US</b>	79/1928	4.1	148/3029	4.9
<b>Foreign Travel For Business &amp; Pleasure</b>	13/183	7.0	220/4941	4.5

**TABLE 2.**

*Test Questions in Relation to FTA/NCA /Revocation*

Question	Failure Rate			
	Yes		No	
	N	%	N	%
<b>Current Alcohol</b>	132/625	21.1	597/4450	13.4
<b>Foreign Ties</b>	92/965	9.5	636/4110	15.5
<b>Foreign Ties Who Contact with Ties</b>	73/744	9.8	557/3719	15.0



<b>Foreign Citizen</b>	38/428	8.9	650/4381	14.8
<b>Passport</b>	127/1547	8.2	581/3431	16.9
<b>Foreign Financial Interests</b>	12/123	9.8	691/4808	14.4
<b>Travelled Outside US</b>	189/1928	9.8	514/3029	17.0
<b>Foreign Travel For Business &amp; Pleasure</b>	20/183	10.9	713/4941	14.4

Table 3 presents descriptive statistics and total scores for the two instrument scales contained in the tool: Criminal History and Other.

**TABLE 3.**

*Descriptive Statistics for Criminal History Sub-score, Other Factors Sub-score, and Total PTRA Score*

	N	Mean	SD	Min	Max
Criminal History Score	5077	3.32	1.54	0	9
Other Factor Score	5077	2.82	1.32	0	6
PTRA Score	5077	6.17	2.46	0	14

Table 4 presents descriptive statistics and total scores for both outcomes contained in the tool: FTA/NCA and FTA/NCA/Revocation. As the table shows, the majority of defendants released in the federal system are successful.

**TABLE 4.**

*Descriptive Statistics for Outcomes*

	N	Percent Failing	Percent Not Failing
--	---	-----------------	---------------------

FTA/NCA	5077	4.5	95.5
FTA/NCA/Revocation	5077	143	857

The next set of analyses focused on assessing the PTRA's predictive ability. AUC-ROC (Area under the Curve-Receiver Operating Characteristics) [FN3] 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 5 displays the AUC-ROC between risk scores and FTA/NCA/Violation revocation. As Table 5 shows, the AUC for the FTA/NCA outcomes only is .69. The AUC for the validation of all three outcome measures rose to .71. Based on these results, the PTRA appears to have very good predictive validity in terms of accurately classifying defendants' risk level.

**TABLE 5.**

*Failure rates by risk category and AUC-ROC values*

Risk Category	N	%	% FTA/NCA	% FTA/NCA/Revoke
Category I	1372	27.0	1.3	3.4
Category II	1406	27.7	3.4	8.5
Category III	1401	27.6	6.7	20.5
Category IV	698	13.8	12.5	29.9
Category V	200	3.9	11.6	31.5
AUC-ROC Risk Category	5077		0.68	071
AUC-ROC Total Score	5077		0.69	0.71

Table 5 presents failure rates by risk category and associated AUC-ROC values. The results for the first four categories were expected based on the construction research. To put the AUC values into practical terms, we calculated the failure rates by two sets of outcome measures: FTA/NCA, the statutory standard, and FTA/NCA/Revocation, the standard preferred by judicial officers. These results are presented in Table 5. The uniform increase in failure rates across categories of risk and across the various samples continues to support the validity of the PTRA. However, in Category V the FTA/NCA rate was twice as high in the original sample as it was in this sample. All looks good, except that Category

might not really be different from Category IV, or perhaps we are supervising Category differently now and driving their failure rates down. It is speculative now, it may hold true, as we do further analysis in the future.

In Table 6 we collapsed Category IV and Category V from Table 5 into one category and reran outcomes and AUC-ROC values. This was done for completeness, since the change in the failure rates could have resulted from a concerted effort to provide more services to the highest-risk defendants, thereby driving their failure rates down. Obviously \*8 interpretation is key here, and if the plausible is true we should not collapse Category V into Category IV. Therefore, this is a significant decision. It should be noted that the reduction to four categories did not add to AUC-ROC values produced by the existing instrument, which is why we will continue to look at this in future research.

**LTABLE 6.**

*LFailure rates by risk category and AUC-ROC values with Category IV and V collapsed*

Risk Category	N	%	% FTA/NCA	% FTA/NCA/Revoke
Category I	1372	27.0	1.3	3.4
Category II	1406	27.7	3.4	8.5
Category III	1401	27.6	67	20.5
Category IV	898	14.7	12.3	30.3
AUC-ROC	5077		0.69	071

**Discussion**

As previously stated, the purpose of this article is threefold: (1) to present the methodology and results produced in the re-validation of the PTRAs; (2) to discuss the implications of the research on the unscored items currently collected in the PTRAs; and (3) to discuss future developments. Overall, the instrument as administered by officers does as well as the construction and validation samples. Even though the foreign ties items did not improve prediction, officers and the court still might want to know about the nature of foreign ties. The sample, though small, was fairly representative of the population served and allowed for re-validation of the existing tool items. Thus the overall results have demonstrated that the PTRAs provide adequate predictive validity.

The creation of the risk score and categories allowed for the re-validation of five risk categories: 1 through 5. Practically speaking, the instrument provided categorizations that are associated with the group failure rates that are differentiated and meaningful for meeting the risk principle.

*Limitations and Future Research*

Although this study was fairly comprehensive in scope, the dataset was small and thus may not be representative of the population served. In addition, there are a number of limitations and areas for future research that deserve mention. First, 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. We did not generate analysis for individual districts due to small samples of data at the district level. Subsequent analysis could focus on assessing AUC values between risk scores and NCA/FTA/Revocation to ensure appropriateness of fit at the district level.

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 defendants 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 NCA measured by 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 a defendant by trying to predict the severity and type of NCA.

#### *Policy Implications*

Notwithstanding the limitations discussed above, two major policy implications stem from this research. First, the federal pretrial services system now has a re-validated risk assessment tool for use on defendants under its jurisdiction. The instrument can be used to identify higher-risk defendants for enhanced services (see VanNostrand & Keebler, 2009) and also to reduce services to low-risk defendants, conserving those resources for high-risk defendants. The second major policy implication is the apparent need to add dynamic factors. Data analyzed in this study focused on static factors associated with changes in NCA/FTA/Revocation rates. Therefore, the addition of dynamic factors would seem to provide officers with an essential tool to monitor and reassess risk in a standardized way to ensure that supervision and services are having intended impacts. If intended impacts are not being achieved, then officers would be able to modify supervision services to reduce the risk and refine supervision methodologies.

#### **\*9 References**

- Administrative Office of the United States Courts. (1979, June 19). *Fourth report on the implementation of Title II of the Speedy Trial Act of 1974*. Washington, D.C.: Administrative Office of the United States Courts.
- Austin, J., Krisberg, B. & Litsky, P. (1984, June 1). *Evaluation of the field test of supervised release; final report*. Washington, D.C.: National Council on Crime and Delinquency.
- Coopriider, K. (2009). Pretrial risk assessment and case classification: A case study. *Federal Probation*, 73(1), 12-15.
- Dowden, C, & Andrews, D. A. (2004). The importance of staff practice in delivering effective correctional treatment: a metaanalytic review of core correctional practice. *International Journal of Offender Therapy and Comparative Criminology*, 48(2), 203-214.
- Goldkamp, J. S., & Gottfredson, M. R. (1988). *Development of bail/pretrial release guidelines in Maricopa County Superior Court, Dade County Circuit Court and Boston Municipal Court*. Washington, D.C.: National Institute of Justice.
- Goldkamp, J. S., & Gottfredson, M. R. (1988, June & December). *Guidelines for bail and pretrial release: Volumes I and II*. Washington, D.C.: National Institute of Justice.

Goldkamp, J. S., & Vilcica, E.R. (2009). Judicial discretion and the unfinished agenda of American bail reform: Lessons from Philadelphia's evidence-based judicial strategy. *Studies in law, politics, and society volume 47: New perspectives on crime and criminal justice* (pp. 115-157). Emerald Publishing: Bingley, UK.

Levin, D. (2006, November). *Examining the efficacy of pretrial release conditions, sanctions, and screening with the state court processing statistics dataserries*. Washington, D.C.: Pretrial Justice Institute.

Lowenkamp, C. T., & Whetzel, J. (2009). The development of an actuarial risk assessment instrument for U.S. Pretrial Services. *Federal Probation*, 73(2), 33-36.

Meierhoefer, B. (1994, April). *Meeting the goals of the bail reform act of 1984: A preliminary study*. Washington, D.C.: Federal Judicial Center.

Rice, M. E.; Harris, G. T. (2005, August). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology*, 73(4), 762.

Siddiqi, Q. (2002, June). *Prediction of pretrial failure to appear and an alternative pretrial release risk-classification scheme in New York City: A reassessment study*. New York: New York City Criminal Justice Agency.

Toborg, M. A., Yezer, A. M., Tseng, P., & Carpenter, B. L. (1984, June). *Pretrial release assessment of danger and flight: Method makes a difference*. McLean, VA.: Lazer Management Group.

VanNostrand, M. (2003). *Assessing risk among pretrial defendants in Virginia: The Virginia pretrial risk assessment instrument*. Richmond, VA: Virginia Department of Criminal Justice Services.

VanNostrand, M., & Keebler, G. (2007). Our journey to pretrial justice. *Federal Probation*, 71(2), 20-25.

VanNostrand, M., & Keebler, G. (2009, April). *Pretrial risk assessment in the federal court*. Washington, D.C.: Office of Federal Detention Trustee.

Winterfield, L., Coggeshall, M., & Harrell, A. (2003, April). *Development of an empirically-based risk assessment instrument: Final report*. Washington, D.C.: Urban Institute. *Journal of Criminology and Penology*, 6, 185-90.

[FN1]. PACTS (Probation/Pretrial Services Automated Case Tracking System) is an electronic case management tool used by probation and pretrial services officers in all 94 federal districts to track federal defendants and offenders. At the end of each month, districts submit case data into a national repository that is accessible to the Administrative Office of the U.S. Courts (AO), Office of Probation and Pretrial Services.

[FN2]. This research presents results on the unscored or test items; however, policy decisions concerning ultimate changes to the PTRS will be determined by the appropriate group or committee, not the authors.

[FN3]. The AUC measures the probability that a score drawn at random from one sample or population (e.g., defendants with a re-arrest) is higher than that drawn at random from a second sample or population (e.g., defendants with no re-arrest). The AUC can range from .0 to 1.0 with .5 representing the value associated with chance prediction. Values equal to or greater than .70 are considered good.

76-SEP Fed. Probation 3

END OF DOCUMENT