Fall
2019

Validation of the Virginia Pretrial Risk Assessment Instrument–Revised in Palm Beach County, Florida

Prepared for the Palm Beach County Criminal Justice Commission

Center for Criminology and Public Policy Research
College of Criminology and Criminal Justice
Florida State University
Tallahassee, Florida

Validation of the Virginia Pretrial Risk Assessment Instrument – Revised in Palm Beach County, Florida

ESTIMATES OF OVERALL PREDICTIVE VALIDITY AND ASSESSMENT OF PREDICTIVE BIAS

JENNIFER E. COPP, THOMAS G. BLOMBERG, WILLIAM CASEY, AND GEORGE PESTA

TABLE OF CONTENTS

Introduction

Project Overview

In 2017, Palm Beach County was awarded $2 million from the MacArthur Foundation's Safety and Justice Challenge, and was tasked with implementing a range of strategies to scale back their jail population. One of five key strategies developed by the Palm Beach County Criminal Justice Commission (CJC) to safely reduce the jail population was to implement a pretrial risk assessment instrument. Together with local stakeholders, the County selected and adopted the Virginia Pretrial Risk Assessment Instrument – Revised (VPRAI-R). Florida State University was contracted to provide a local validation of the VPRAI-R, with a focus on the extent to which the tool provided race and gender neutral predictions.

The MacArthur Foundation's investment in criminal justice reform, and the use of jails in particular, stems from the realization that America's reliance on mass incarceration begins in local jails. Nationally, roughly three-quarters of a million people are in jail—nearly two-thirds of whom are awaiting trial, and thus have yet to be convicted of a crime (Zeng, 2018). Although relatively little is known about the effects of jail incarceration, scholars have linked pretrial detention to defendants' case outcomes. Based on a handful of recent studies, it appears that pretrial detention increases defendants' likelihood of conviction, primarily through an increase in guilty pleas (Dobbie, Golden, & Yang, 2018; Stevenson, 2018). The evidence also suggests that pretrial detainees are more likely to receive a custodial sentence, and to receive a sentence of greater length, than their counterparts who were released pending the adjudication of their case(s) (e.g., Gupta, Hansman, & Frenchman, 2016; Heaton, Mayson, & Stevenson, 2017; Leslie & Pope, 2017; Oleson, Lowenkamp, Wooldredge, VanNostrand, & Cadigan, 2017; Phillips, 2012; Sacks & Ackerman, 2014).

Regarding pretrial release/detention decisions, the Supreme Court ruled that "(i)n our society liberty is the norm, and detention prior to trial or without trial is the carefully limited exception" (United States v. Salerno, 481 U.S. 739, 755 (1987)). Accordingly, judges are tasked with making critical decisions that provide due process to defendants accused of violating the law, while also ensuring that the accused appear in court, and that both victims and the broader community are kept safe from potential threats/harm. Historically, judges' pretrial decisions have been largely subjective, and based on the unstructured judgement of decision-makers. More recently, validated pretrial risk assessment tools have been presented as a strategy to provide more objective information to inform judicial decision-making.

In this report, we present the findings from a validation of the VPRAI-R using local data from Palm Beach County over a nearly two-year period. Our analyses proceeded in several stages. We began by describing the data, including bivariate associations between the different risk factors included in the VPRAI-R and pretrial failure (i.e., failure to appear, new arrest, and technical violations). Next, we assessed the overall predictive validity of the VPRAI-R to determine how reliably the tool classifies defendants on the basis of their likelihood of pretrial success/failure. Finally, we considered whether the estimates provided by the VPRAI-R are race/ethnic and gender neutral. We conclude with a summary of our findings, and recommendations for future validation/examination efforts and the continued use of the VPRAI-R in Palm Beach County.

Review of Prior Research

Risk Assessments and Pretrial Decision-Making

There are currently more than two dozen different risk assessments in use across the United States, and according to recent estimates, as many as one-quarter of U.S. residents live in a jurisdiction where a validated pretrial risk assessment was in use (Pretrial Justice Institute, 2017). Existing tools share a number of similarities; however, they also differ in key ways including the factors used to predict risk, how "failure" is measured, and how defendant risk is captured. For example, many jurisdictions use defendant interviews, often relying on their pretrial services agencies, while others draw on tools developed using administrative data. Although many of the available tools produce a risk score reflecting a combined estimate of "any" pretrial failure, others produce separate estimates of the risk of failure to appear in court and rearrest. The focus on failure to appear and rearrest are based on constitutional standards for pretrial detention, which require that the use of detention be reserved for those who pose a substantial risk to public safety or are a flight risk. Yet some tools capture risks that fall beyond these constitutional standards (e.g., technical violations).

Most assessment tools include a combination of individual, social, and criminal history characteristics, factors that have been demonstrated to predict the likelihood of court appearance and rearrest (Bechtel, Lowenkamp, & Holsinger, 2011). Typical risk factors include age, education, family/peer relationships, community ties, employment, criminal history, active criminal justice status, and current charge. These factors are weighted according to their relative impact on failure, and the total score is taken as the sum across the items. Defendant scores are often translated into risk categories to facilitate recommendations (e.g., low-, medium-, and high-risk categories). Importantly, the scores produced by risk assessment tools are based on the outcomes of similarly scoring individuals who were studied during the development/validation process.

Many of the tools that are currently available were developed for use within a particular jurisdiction (e.g., Lee County, FL; Coconino County, AZ; Hennepin County, MN). There have also been a number of tools developed for broader use. Such multi-jurisdictional tools include the VPRAI, which was developed for statewide use in Virginia. Similar tools have been developed for use in other states, including Ohio, Florida, Colorado, and Indiana, in addition to the federal court system. Arnold Ventures also developed a multi-jurisdictional tool for broader use, the Public Safety Assessment (PSA), which was created using 1.5 million cases from roughly 300 jurisdictions across the United States.

Prior Research on the Virginia Pretrial Risk Assessment Instrument – Revised (VPRAI-R)

Since their development, pretrial risk assessment tools have routinely been examined to determine their predictive accuracy, impact on pretrial populations, and ability to guide recommendations regarding conditions of release. Based on this research, there is good evidence that pretrial risk assessment tools are reliable predictors of pretrial risk, including failure to appear and rearrest. This conclusion is based on a number of recent validation studies, including empirical assessments of some of the more commonly used tools, such as the VPRAI, PSA, and the Pretrial Service Risk Assessment Tool (PTRA). In general, findings have demonstrated that higher risk scores are associated with higher rates of failure to appear and rearrest. Estimates of predictive accuracy have tended to range from 0.60 to 0.75. To put these estimates into context, a tool that perfectly predicted pretrial failure would produce a value of 1, and a prediction method that provided no information to distinguish between "failures" and successes" would have an AUC of 0.5.

The VPRAI-R was designed to predict success or failure during the pretrial release period based on defendants' failure to appear, new arrest, and/or technical violation using a single score. Since its development, the tool has undergone extensive testing, in addition to numerous local validation efforts. The VPRAI was originally developed for use in Virginia; however, it has since been adopted in jurisdictions across the country. Findings from a recent validation study commissioned by the Virginia Department of Criminal Justice Services suggest that the VPRAI reliably predicts pretrial failure (AUC = .666) (Danner, VanNostrand, & Spruance, 2015). It has also been identified as a race and gender neutral tool, as it has been demonstrated to reliably predict across race/ethnic and gender categories (Danner, VanNostrand, & Spruance, 2016).

In this report, we provide findings from a validation of the VPRAI-R using data from Palm Beach County to determine whether the tool is a reliable predictor of pretrial failure in the Palm Beach County community. We discuss the implications of our findings for pretrial decision-making in Palm Beach County and make a number of recommendations for its continued use.

Data and Methods

Sample Description

The data used in this report comes from three sources, including the Booking Information Retrieval System (BIRS) of the Palm Beach County Sheriff's Office, the Palm Beach County Clerk of Courts, and Palm Beach County Pretrial Services Program. Data from the BIRS includes current arrest characteristics, demographic characteristics, and information regarding jail bookings and releases. These data were provided at the charge level and collapsed by individual identifier and booking date, which allowed us to focus on each individual booking event. When an individual faced additional charges related to that booking incident at a later point in time (whether that be hours or days later), these charges were considered part of the same booking event. This strategy enabled us to retain the most serious charge as well as whether they faced multiple felony and misdemeanor charges.

Data from Palm Beach County Pretrial Services includes the defendant's calculated risk score, the associated risk level, and the individual risk factors. The data were also used to determine whether the defendant violated any release conditions of the court-ordered Supervised Own Recognizance (SOR) Program operated by Pretrial Services during the pretrial period. These data were provided at the first appearance hearing level and were collapsed by individual identifier and booking date. In instances where there were multiple first appearances for the same identifier and booking date, the most recent first appearance was retained. Data from the Palm Beach County Clerk of Courts were primarily used to determine if the defendant failed to appear at their final case disposition. In addition, the court data were used to determine whether the defendant was released prior to the final disposition of their case.

Finally, we made several adjustments to our sample by dropping cases where the defendant was held for another jurisdiction or had an invalid booking, release, or disposition date (i.e., a release date or disposition date that occurred prior to the booking date). We limited our sample to defendants identified as non-Latino white, non-Latino black, and Latino, as small cell sizes precluded analyses of other racial and ethnic groups. Our final analytic sample (n=11,269) was restricted to defendants who were released prior to the final disposition of their case and who had valid data on the eight risk factors included in the VPRAI-R.

Dependent Variables

The specific aim of this validation was to determine whether the VPRAI-R accurately predicts pretrial failure among individuals released from the Palm Beach County Jail. To achieve this aim, we gathered information on four measures of pretrial failure. The first, *failure to appear (FTA)*, was a dichotomous variable indicating whether the defendant failed to appear in court. Information to construct this variable was taken from the Clerk of Court's warrant data, which identified cases where a capias warrant was issued for an FTA. Cases received a value of "1" to indicate the issuance of a capias warrant, and "0" otherwise. Second, we identified individuals who were arrested for a new crime while on pretrial release. This variable, *new arrest,* was based on data from the BIRS, and indicates whether individuals were rebooked into the Palm Beach County Jail following their initial release and prior to the final disposition of their case (1 = yes). A third indicator of pretrial failure is *technical violation,* which is taken from the Palm Beach County Pretrial Services data and indicates whether a defendant violated his or her conditions of supervised pretrial release (1 = yes). A final indicator, *any failure*, captures whether a defendant exhibited any of the three sources of failure (i.e., failure to appear, new arrest, technical violation) (1 = yes).

Independent Variables

To assess the predictive validity of the individual risk factors, we included measures for each of the eight VPRAI-R risk factors. The risk factors indicate whether the defendant was on *active community supervision*, the *charge type* of the current arrest, if the defendant had any *pending charges* at the time of their first appearance, if they had a *criminal record*, *two or more FTAs*, *two or more violent convictions,* were *unemployed at the time of arrest,* or if they had a *history of drug abuse*. Information on the eight VPRAI-R risk factors comes from the Pretrial Services data. Also contained in that data is information regarding the *pretrial risk score* and the *pretrial risk level*. The pretrial risk score (0-14) provides information regarding the total score given to each defendant based on the individual risk factors. The pretrial risk score is then transformed into the pretrial risk level (1-6) which provides information to the judge regarding the likelihood of pretrial failure, including failure to appear, new arrest, and technical violations. To assess potential race and ethnic differences in the predictive validity of the VPRAI-R we used three variables to denote race and ethnicity, including *non-Latino white*, *non-Latino black*, and *Latino*. Finally, to assess potential gender differences in the predictive validity of the VPRAI-R we used a measure indicating the gender of the defendant (male = 1).

Analytic Strategy

We began by presenting descriptive statistics for the full sample. Next, we examined bivariate associations between the eight VPRAI-R risk factors and the *any failure* outcome. Following that, we described the results of a series of logistic regression models using the individual risk factors to predict *any failure*, *failure to appear*, *new arrest*, and *technical violations*. All logistic regression models in this report provide information regarding the Area Under the Curve for the Receiver Operator Characteristics (AUC-ROC), a widely used measure of risk assessment performance (Danner et al., 2016). Next, we describe the results of a logistic regression model examining the association of the four failure outcomes and the pretrial risk score, followed by a discussion of a logistic model that uses the pretrial risk level to predict the failure outcomes. Finally, we described the results of the same logistic regression models controlling for race/ethnicity and gender to determine whether the VPRAI-R is race and gender neutral in its prediction of pretrial failure.

Sample Description

The descriptive statistics for the full sample are presented in Table 1. The average age of the sample was roughly 34. Nearly three-quarters (71.42%) of sample members were male, and the remaining quarter (28.58%) were female. The majority (83.57%) were unmarried at the time of arrest/booking. Approximately two in five (43.24%) sample members were non-Latino White, and a roughly even share were non-Latino Black (41.52%). The remaining 15.24% of sample members were of Latino origin. The vast majority (84.05%) were citizens of the United States, and a relatively small share (15.95%) were non-citizens.

In our analyses, we focused on the most serious charge. In our sample, this was typically either a non-violent felony (40.59%) or misdemeanor (44.82%) offense. Fewer sample members were charged with a violent felony (14.57%). Relatedly, few respondents fell into the high-risk categories (5 and 6) based on their overall risk assessment scores (6.95%). In contrast, more than half of sample members were categorized as low risk (1 and 2) based on the results of their risk assessment (60.39%). The remainder (32.66%) fell into the medium risk categories (3 and 4).

Overall, just over one in ten (13.51%) sample members "failed" during the pretrial period. The majority of failures came from technical violations (7.17%), followed by failure to appear (4.57%) and new arrests (1.77%).[1]

Validation Results

Validity and Practical Utility of the VPRAI-R

Bivariate results. Results of bivariate analyses are presented in Table 2. These findings are based on associations between the eight VPRAI-R risk factors and the dichotomous indicator

[1] The rate of new arrests may underestimate the true rate of new criminal activity among sample members, as it is limited to new arrests that occurred in Palm Beach County.

Table 1. Sample Descriptive Statistics (n=11,269)

	Mean/Percentage	SD	Range
Demographic Characteristics			
Age	34.34	12.52	16-90
Gender			
Male	71.42%		
Female	28.58%		
Marital Status			
Married	16.43%		
Unmarried	83.57%		
Race/Ethnicity			
Non-Latino White	43.24%		
Non-Latino Black	41.52%		
Latino	15.24%		
Citizenship Status			
Citizen	84.05%		
Non-citizen	15.95%		
Current Arrest Characteristics			
Violent Felony	14.57%		
Non-Violent Felony	40.59%		
Misdemeanor	44.82%		
Risk Level			
1	37.11%		
2	23.28%		
3	18.49%		
4	14.17%		
5	5.84%		
6	1.11%		
Any Failure	13.51%		
Failure to Appear	4.57%		
New Arrest	1.77%		
Technical Violation	7.17%		

Table 2. Descriptive and Bivariate Statistics for Eight VPRAI Risk Factors (Any Failure Outcome)

		Total		Any Failure		Chi-Square	*P*
		N	%	N	%		
Active community supervision	Yes	310	2.75	46	3.02	0.485	0.486
	No	10959	97.25	1476	96.98		
Charge Type	Felony	3474	30.83	561	36.86	30.019	0.000
	Other	7795	69.17	961	63.14		
Pending Charge	Yes	1992	17.68	446	29.30	163.463	0.000
	No	9277	82.32	1076	70.70		
Criminal History	Yes	6317	56.06	1024	67.28	89.983	0.000
	No	4952	43.94	498	32.72		
Two or more FTA	Yes	2383	21.15	453	29.76	78.357	0.000
	No	8886	78.85	1069	70.24		
Two or more violent convictions	Yes	902	8.00	137	9.00	2.376	0.123
	No	10367	92.00	1385	91.00		
Unemployed at time of arrest	Yes	3419	30.34	597	39.22	65.725	0.000
	No	7850	69.66	925	60.78		
History of drug abuse	Yes	4207	37.33	738	48.49	93.614	0.000
	No	7062	62.67	784	51.51		

of “any” pretrial failure. Significant associations indicate that the presence of these various risk factors and pretrial failure are associated with each other. In these data, six of the eight risk factors are associated with pretrial failure at the bivariate level. However, active community supervision and two or more violent convictions are not associated with pretrial failure. This suggests that individuals on community supervision are no more (or less) likely to fail than their counterparts, and moreover, that the likelihood of pretrial failure is similar between individuals with two or more violent convictions and those with one or no prior violent convictions.

Multivariate results, risk items. The results of a series of multivariate logistic regression models are presented in Table 3. The first model provides the odds ratios and p-values for the logistic regression of the individual VPRAI-R risk factors and any pretrial failure. The model chi-square, presented at the bottom of the table, suggests that as a whole, the VPRAI-R risk factors are significant predictors of pretrial failure ($\chi^2 = 291.77$, $p < .001$). Furthermore, six of the eight risk factors are significantly associated with the odds of pretrial failure in this model. Consistent with the bivariate findings described above, however, active community supervision and two or more violent convictions are not associated with the odds of any pretrial failure, controlling for the other risk factors included in the VPRAI-R assessment. The estimate for the Area under the Curve for the Receiver Operator Characteristic, a commonly used metric of risk assessment performance, is also provided for this model. The AUC-ROC value of 0.645 is comparable to other estimates of predictive validity based on the VPRAI, and is considered a “good” score according to experts in the field of risk assessment (Demarais, Johnson, & Singh, 2016). Substantively, this value represents the percent of cases in which we can expect a randomly selected defendant who “failed” to have a higher score on the assessment than a randomly selected defendant who did not fail. A value of .5 would suggest that the scores do not

Table 3. Predicting Failure Outcomes with VPRAI Risk Factors (n=11,269)

	Any Failure		Failure to Appear		New Arrest		Technical Violation	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.937	0.693	1.087	0.751	0.781	0.591	0.897	0.618
Charge Type	1.264	0.000	0.520	0.000	1.542	0.004	1.834	0.000
Pending Charge	1.940	0.000	4.361	0.000	1.510	0.014	0.902	0.289
Criminal History	1.288	0.000	0.945	0.628	1.109	0.576	1.592	0.000
Two or more FTA	1.232	0.003	1.370	0.007	1.094	0.624	1.139	0.155
Two or more violent convictions	0.831	0.072	0.655	0.023	1.472	0.090	0.834	0.170
Unemployed at time of arrest	1.365	0.000	1.376	0.001	1.178	0.279	1.333	0.000
History of drug abuse	1.253	0.001	1.051	0.651	1.159	0.391	1.356	0.000
Constant	0.085	0.000	0.031	0.000	0.011	0.000	0.036	0.000
Model Chi-Square	291.77	0.000	318.67	0.000	28.81	0.000	200.62	0.000
Nagelkerke R Square	0.033	-	0.076	-	0.014	-	0.035	-
AUC-ROC	0.645	-	0.699	-	0.601	-	0.644	-

distinguish on the basis of failure (i.e., predictions are no more accurate than a coin toss), whereas a score of 1 indicates perfect prediction.

The second model presents the findings from the logistic regression model predicting failure to appear in court. Similar to the previous model, the overall model chi-square suggests that the items included in the VPRAI-R are significant predictors of FTA ($\chi^2 = 318.67$, $p < .001$). Of the eight risk factors included in the assessment, five are significantly associated with the odds of FTA, including charge type, pending charge, two or more FTA, two or more violent convictions, and unemployed at the time of arrest. Notably, although charge type and two or more violent convictions are significantly associated with the odds of failure to appear, these effects do not operate in the expected direction. That is, the odds ratios for these factors suggest that those with a felony conviction and/or two or more violent convictions are less likely to fail to appear. Active community supervision, criminal history, and history of drug abuse are not significantly associated with the odds of pretrial failure. Despite fewer factors being associated with the odds of failure in this model relative to the model predicting “any” failure, there is a slight improvement in model fit as reflected in the large increase in the chi-square value. Furthermore, the AUC estimate for the model predicting failure to appear is a slight improvement from the any failure model. Specifically, in the logistic regression model assessing associations between the eight VPRAI-R risk factors and failure to appear, the AUC is 0.699.

The third model presents the odds ratios from the logistic regression of the individual VPRAI-R risk factors on the odds of new arrest. The overall model chi-square suggests that the VPRAI-R risk factors are significant predictors of new arrest ($\chi^2 = 28.81$, $p < .001$). However, there is a notable decline in model fit from the prior models, including those estimating the odds of any failure and failure to appear, respectively. Of the individual VPRAI-R risk factors included in this model, only two are significantly associated with the odds of new arrest, including charge type and pending charge. The remaining factors are unrelated to the odds of new arrest. The AUC estimate for this model is 0.601. Although this value still falls within an acceptable range based on the standards established in the field of risk assessment, this value represents a significant decline in overall predictability from the other models.

The fourth and final model estimates the odds of technical violation using the individual risk factors included in the VPRAI-R. The chi-square value indicates that, as a whole, the risk factors are significant predictors of technical violations ($\chi^2 = 200.62$, $p < .001$). Of the individual items, four are significant predictors of technical violations in this full model. In particular, charge type, criminal history, unemployed at time of arrest, and history of drug abuse are all significantly associated with heightened odds of technical violations. Active community supervision, pending charge, two or more FTA, and two or more violent convictions are unrelated to the odds of technical violation in this full model. The estimate of the AUC is 0.644, which is nearly identical to the model predicting any failure, and suggests that the items included in the VPRAI-R provide reliable estimates of technical violations.

Multivariate results, total risk score. The eight risk factors included in the VPRAI-R are weighted and scored. The weighting of the individual items is as follows: 1) active community supervision (yes = 2), 2) charge is felony drug, theft, or fraud (yes = 3), 3) pending charge (yes = 2), 4) criminal history (yes = 2), 5) two or more failures to appear (yes = 1), 6) two or more violent convictions (yes = 1), 7) unemployed at time of arrest (yes = 1), and 8) history of drug abuse (yes = 2). The total score is taken as the sum across these weighted individual items, and ranges from 0 to 14. Table 4 presents the odds ratios for a series of logistic regression

Table 4. Predicting Failure Outcomes with Total Risk Score (n=11,269)

	Any Failure		Failure to Appear		New Arrest		Technical Violation	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Pretrial Risk Score	1.145	0.000	1.082	0.000	1.123	0.000	1.166	0.000
Constant	0.088	0.000	0.035	0.000	0.011	0.000	0.039	0.000
Model Chi-Square	217.62	0.000	27.73	0.000	24.19	0.000	161.38	0.000
Nagelkerke R Square	0.024	-	0.007	-	0.012	-	0.028	-
AUC-ROC	0.621	-	0.580	-	0.598	-	0.635	-

models predicting pretrial failure, including any failure, failure to appear, new arrest, and technical violation, using the total risk score. The first model presents the odds ratio for the logistic regression of pretrial risk score on any failure. The fit statistics suggest that the total risk score is a significant predictor of any pretrial failure ($\chi^2 = 217.62$, $p < .001$). Specifically, a one unit change in the risk score is associated with a 15% increase in the odds of pretrial failure. This association is significant ($p < .001$). The AUC estimate for this model is 0.621, suggesting that the total risk score is a reliable predictor of pretrial failure, albeit less predictive than the model which includes the individual risk items.

The second model predicts failure to appear using the total risk score. Similar to the findings described above, the total risk score is a significant predictor of failure to appear ($\chi^2 = 27.73$, $p < .001$). The odds ratio suggests that a one-unit increase in the total risk score increases the odds of failure to appear by roughly 8%. This increase is significant at $p < .001$ level. The findings are substantively similar across the other failure outcomes examined, including new arrest and technical violation. In particular, the odds of new arrest increase approximately 12% for each one-unit increase in the risk score ($p < .001$), and a single unit increase in the risk score is associated with a nearly 17% increase in the odds of a technical violation ($p < .001$). The AUCs across failure to appear, new arrest, and technical violation in these models based on the total risk score are 0.580, 0.598, and 0.635, respectively. Taken together, these findings suggest that the overall risk score produces reliable estimates of any failure and technical violation, and marginally less reliable estimates of failure to appear and new arrest.

Multivariate results, risk level. The total score produced by the weighting of the individual risk factors (described above) is collapsed into six different risk levels. The categorization of cases into the risk levels based on the total risk score is as follows: 1) level 1 =

0-2, 2) level 2 = 3-4, 3) level 3 = 5-6, 4) level 4 = 7-8, 5) level 5 = 9-10, 6) level 6 = 11-14. The risk levels correspond to the likelihood of pretrial failure, including any pretrial failure, failure to appear, new arrest, and technical violation, such that categories 1 and 2 are low risk, 3 and 4 are medium risk, and 5 and 6 are high risk. The risk levels are used in the Pretrial Risk Management Matrix (PRMM), in conjunction with the offense type, to provide a range of release recommendations. In Table 5 we present the sample distribution across the different risk levels by pretrial failure.

Table 5. "Any" Pretrial Failure, by VPRAI-R Risk Level (n = 11,269)

	Full Sample		Any Failure	
Risk Levels (Defendant's Score)	N	%	N	%
1 (0 - 2)	4,182	37.11%	336	8.03%
2 (3 - 4)	2,623	23.28%	352	13.42%
3 (5 - 6)	2,084	18.49%	369	17.71%
4 (7 - 8)	1,597	14.17%	313	19.60%
5 (9 - 10)	658	5.84%	128	19.45%
6 (11 - 14)	125	1.11%	24	19.20%
Base Rate				13.51%

The majority of defendants fall into the lowest risk level (1 and 2), with relatively few in the high-risk categories (5 and 6). More specifically, three-fifths of defendants (60.39%) were categorized as low risk, whereas fewer than one in ten (6.95%) defendants were categorized as high risk. The remaining one-third of defendants (32.66%) were in the medium risk category. The last column examines the distribution of pretrial failure across the different risk levels. The general trend suggests that the risk of pretrial failure increases across the different risk categories. For example, roughly 8% of defendants in the lowest risk category "fail" as compared to more nearly 20% in the higher risk categories. It is noteworthy, however, that the upward

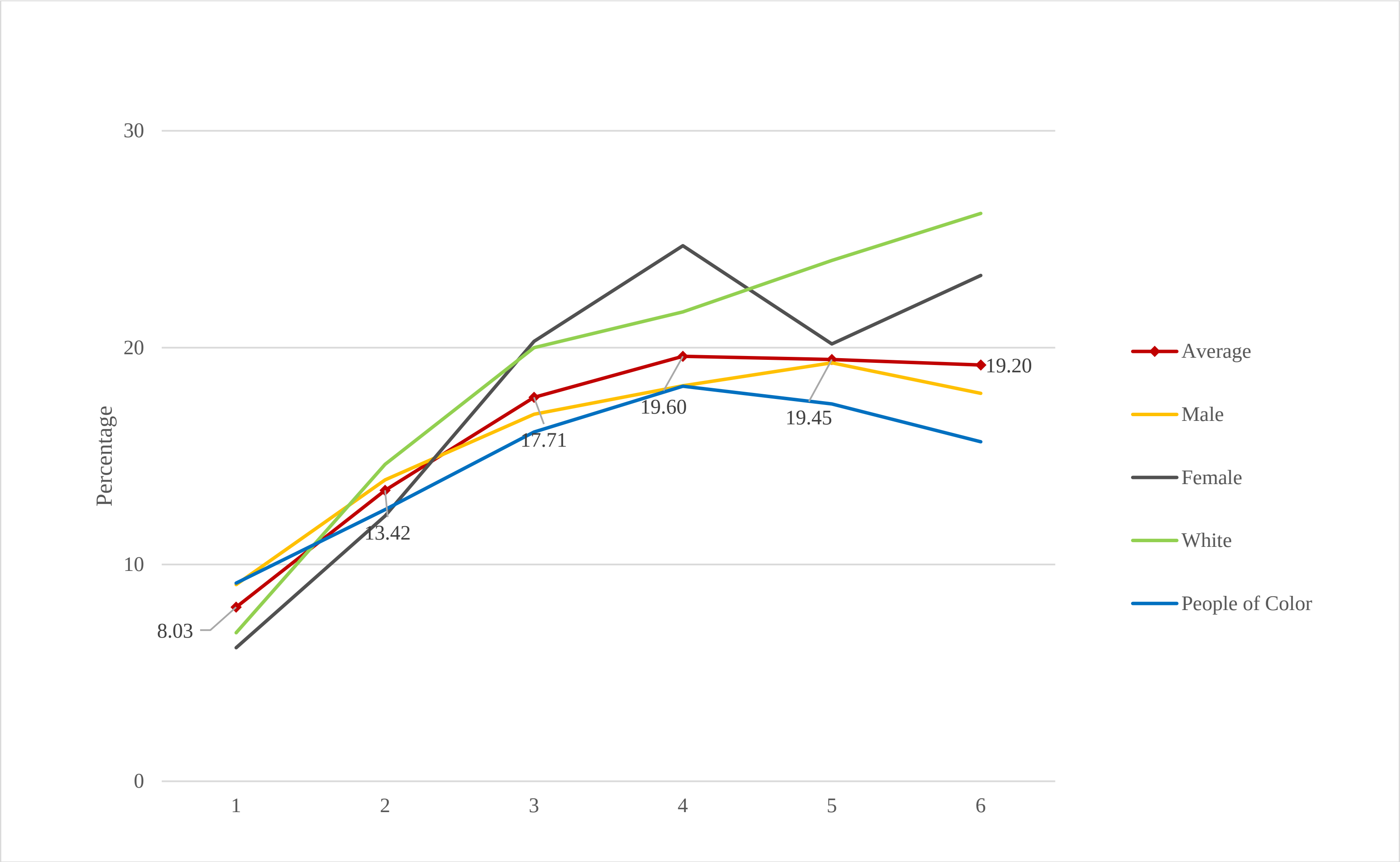

Figure 1. The Risk of "Any" Pretrial Failure Across Risk Categories, by Race/Ethnicity and Gender (n = 11,269)

trend in pretrial failure plateaus after risk category 4. This suggests that the VPRAI-R may be over-predicting risk for defendants in the higher-risk categories, as they appear to be at relatively low risk in terms of pretrial failure relative to their risk categorization.

To determine whether there was any systematic patterning to the over-classification of risk on the basis of race/ethnicity and gender, we plotted the failure rates across the risk levels (see Figure 1). The trend for white defendants resembles the expected trajectory, as the percent who "failed" pretrial increased monotonically across the different risk categories. The trend for people of color sharply contrasts that expected trajectory, as the percent who failed peaks among members of risk category 4, and then declines. It is also noteworthy that more than 1 in 4 white defendants in the highest risk category (6) failed, as compared to less than 16% of people of color in that same category. These descriptive findings suggest some potential differences in terms of predictive validity on the basis of race, an issue we return to later in this report.

In Table 6, we present the findings from a series of logistic regression models that use the pretrial risk level to estimate the odds of pretrial failure. The first model presents the odds ratio for the logistic regression of pretrial risk level on any failure. According to the model, pretrial risk level is a significant predictor of any pretrial failure ($p < .001$). Moreover, a one-unit increase in risk level (i.e., a move from one level to the next) is associated with a 31% increase in pretrial failure. Similar findings hold for the other outcomes examined in these analyses, including failure to appear, new arrest, and technical violation. In particular, a one-unit increase in risk level is associated with a 14% increase in failure to appear, a 28% increase in new arrest, and a 38% increase in technical violations. The AUC estimates suggest that the risk levels provide reliable estimates of any failure and technical violations, while the risk level is a slightly less reliable predictor of failure to appear and new arrest.

Table 6. Predicting Failure Outcomes with Pretrial Risk Level (n=11,269)

	Any Failure		Failure to Appear		New Arrest		Technical Violation	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Pretrial Risk Level	1.311	0.000	1.139	0.000	1.276	0.000	1.375	0.000
Constant	0.079	0.000	0.035	0.000	0.010	0.000	0.034	0.000
Model Chi-Square	181.05	0.000	15.03	0.000	22.52	0.000	147.01	0.000
Nagelkerke R Square	0.020	-	0.004	-	0.011	-	0.025	-
AUC-ROC	0.611	-	0.560	-	0.594	-	0.630	-

Taken together, the results of the bivariate and multivariate analyses suggest that the individual items included in the VPRAI-R, the total risk score, and the different risk levels used in the PRMM are significant predictors of pretrial failure, including any failure, failure to appear, new arrest, and technical violation. The AUC estimates suggest that the tool is a generally reliable predictor of pretrial failure; however, there is some variability in the reliability of estimates across the different outcomes and model specifications. The most reliable predictions were produced based on models using the individual risk items. Yet the most meaningful estimates of reliability are those based on the pretrial risk level, as this reflects the categories used in the PRMM to provide recommendations. We discuss potential sources of this variability and implications in the summary and conclusions section of this report.

Race and Gender Neutrality of the VPRAI-R

The next section of the report focuses on whether the estimates produced by the VPRAI-R are race/ethnic and gender neutral, beginning with a focus on race/ethnicity. Determinations of race and gender neutrality were based on comparisons of predictive validity across race/ethnicity and gender, as well as whether the VPRAI-R systematically produced higher risk scores for defendants from a particular race/ethnic or gender category than expected based on their "true risk." Using the example of race, a racially biased tool would provide a similar risk score for white and black defendants despite black defendants' lower likelihood of pretrial failure. Recall that the VPRAI has demonstrated predictive parity across race and gender in prior analyses (see Danner et al., 2016). The goal of the current analyses was to assess whether similar parity in predictions exist using data from Palm Beach County.

Multivariate results, risk items (race/ethnicity). Table 7 presents the odds ratios from a series of logistic regression models predicting pretrial failure, including any failure, failure to appear, new arrest, and technical violations. In addition to the individual risk factors included in the VPRAI-R, these models included indicators of race/ethnicity. These include dichotomous variables indicating whether the defendant is non-Latino black or Latino (non-Latino white is the reference category). The models shed light on whether the overall predictability of the model changes following the inclusion of race/ethnicity, and moreover, whether the observed associations between the risk factors and pretrial failure change as a result of including race/ethnicity in the model. In the first model predicting any failure, the associations between the risk factors and any failure are largely similar to those presented above (see Table 3). Furthermore, the model estimate of predictive validity is identical to the model estimated without race (AUC = 0.645), suggesting that the addition of race to the model does not alter the overall reliability. However, race is significantly associated with the odds of any failure. That is, non-Latino black defendants, relative to their non-Latino white counterparts, are about 14% less likely to “fail” pretrial. The subsequent models predict the separate pretrial failure outcomes, including failure to appear, new arrest, and technical violation. Similar to the any failure outcome, the odds ratios for the different risk factor items are similarly predictive of failure to appear in models with and without the race indicators. Furthermore, the indicators of race/ethnicity were not significantly associated with the odds of failure to appear, net of the individual risk items. A similar pattern was observed for new arrest; that is, the associations between charge type, pending charge, and new arrest remained significant and positive in the model predicting new arrest following the inclusion of race/ethnicity. Consistent with models that included only the risk factors, the remaining items were not significant. Controlling for the

Table 7. Predicting Failure Outcomes with VPRAI Risk Factors and Race and Ethnicity (n=11,269)

	Any Failure		Failure to Appear		New Arrest		Technical Violation	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.939	0.703	1.078	0.775	0.775	0.579	0.904	0.645
Charge Type	1.267	0.000	0.526	0.000	1.518	0.005	1.838	0.000
Pending Charge	1.958	0.000	4.428	0.000	1.478	0.021	0.913	0.348
Criminal History	1.296	0.000	0.964	0.751	1.082	0.672	1.602	0.000
Two or more FTA	1.264	0.001	1.414	0.003	1.032	0.866	1.184	0.068
Two or more violent convictions	0.848	0.110	0.675	0.036	1.406	0.137	0.859	0.254
Unemployed at time of arrest	1.366	0.000	1.382	0.001	1.167	0.308	1.330	0.000
History of drug abuse	1.244	0.001	1.052	0.649	1.167	0.371	1.334	0.001
Black	0.861	0.016	0.848	0.115	1.348	0.058	0.789	0.004
Latino	0.916	0.310	1.245	0.092	0.799	0.378	0.754	0.018
Constant	0.090	0.000	0.031	0.000	0.010	0.000	0.041	0.000
Model Chi-Square	297.61	0.000	327.00	0.000	35.24	0.000	211.51	0.000
Nagelkerke R Square	0.033	-	0.078	-	0.018	-	0.036	-
AUC-ROC	0.645	-	0.704	-	0.611	-	0.648	-

individual risk factors, race/ethnicity is not significantly associated with the odds of new arrest. The final model assesses associations between the individual risk factors and technical violation, controlling for race/ethnicity. Net of race/ethnicity, the associations between the individual risk factors and technical violation remain largely unchanged. Yet race/ethnicity emerged as a significant predictor of technical violation in this model. Specifically, the odds ratios suggest that, controlling for the full roster of risk factors, non-Latino black and Latino defendants are less likely to violate the conditions of SOR than their non-Latino white peers. That race emerged as a significant predictor of technical violation is reflected in the slight increase in predictive validity produced by this model (AUC = 0.648).

Given that race/ethnicity did emerge as a significant predictor of pretrial failure, controlling for the other risk factors included in the VPRAI-R, we further assessed the issue of variable predictions on the basis of race/ethnicity in a series of supplemental models (see Appendix Tables A1-A4). These findings indicate where some of the associations between the risk factors and pretrial failure differ across race/ethnic groups. For example, criminal history is positively associated with the odds of any pretrial failure among non-Latino white and Latino defendants, yet unrelated to the odds of pretrial failure among non-Latino black defendants. Similarly, a history of drug use is positively associated with any pretrial failure among non-Latino whites, but does not emerge as a significant predictor among non-Latino blacks. Overall, the estimates of predictive validity suggest that the VPRAI-R is a better predictor of pretrial failure among non-Latino white defendants (AUC = 0.689) relative to defendants of color (AUCs = 0.610 and 0.625 for non-Latino black and Latino defendants, respectively). Nevertheless, the AUC estimates for all race/ethnic groups fall within an acceptable range.

Multivariate results, risk level (race/ethnicity). Table 8 explores the potential for race/ethnic bias in the prediction of pretrial failure based on the risk levels used in the PRMM. Included in the table are separate models predicting any failure, failure to appear, new arrest, and technical violation by race/ethnicity. The first set of models present odds ratios for logistic regressions of pretrial risk level on pretrial failure for non-Latino white defendants, followed by models for non-Latino black and Latino defendants, respectively. Across the majority of models, pretrial risk level emerged as a significant predictor of pretrial failure. However, there were a few exceptions. In particular, pretrial risk level was unrelated to the odds of new arrest among white defendants. Furthermore, pretrial risk level was not associated with the odds of pretrial failure for non-Latino black or Latino defendants. Consistent with the models presented above, pretrial risk level appears to be a better predictor of pretrial failure among non-Latino white defendants relative to defendants of color, with one exception. That is, pretrial risk level emerged as a stronger predictor of new arrest among non-Latino black and Latino defendants as compared to whites. However, this finding should be interpreted with caution given the relatively small numbers of defendants who were arrested for a new crime during the pretrial period after disaggregating the sample on the basis of race/ethnicity (e.g., 20 Latino defendants, or 0.18% of the sample, were arrested for a new crime during the pretrial period).

Multivariate results, risk items (gender). The next set of models examine the potential for differential prediction on the basis of gender (Table 9). To examine this possibility, we estimated a series of logistic regression models, which included the individual risk items included in the VPRAI-R, as well as an indicator for gender (male). The findings presented here are very similar to the estimates presented in Table 3, suggesting that the addition of gender to the model did little to modify any of the associations between the individual risk factors and

Table 8. Predicting Failure Outcomes with Pretrial Risk Level by Race and Ethnicity

	Any Failure		Failure to Appear		New Arrest		Technical Violation	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
White (n=4,873)								
Pretrial Risk Level	1.453	0.000	1.267	0.000	1.080	0.390	1.559	0.000
Constant	0.065	0.000	0.026	0.000	0.013	0.000	0.028	0.000
Model Chi-Square	144.11	0.000	20.56	0.000	0.72	0.396	129.62	0.000
Nagelkerke R Square	0.037	-	0.012	-	0.001	-	0.048	-
AUC-ROC	0.647	-	0.602	-	0.536	-	0.675	-
Black (n=4,679)								
Pretrial Risk Level	1.207	0.000	1.045	0.394	1.340	0.000	1.232	0.000
Constant	0.096	0.000	0.041	0.000	0.010	0.000	0.043	0.000
Model Chi-Square	37.68	0.000	0.72	0.396	17.57	0.000	26.06	0.000
Nagelkerke R Square	0.010	-	0.000	-	0.018	-	0.011	-
AUC-ROC	0.577	-	0.526	-	0.613	-	0.584	-
Latino (n=1,717)								
Pretrial Risk Level	1.284	0.000	1.158	0.075	1.387	0.042	1.323	0.000
Constant	0.081	0.000	0.043	0.000	0.006	0.000	0.032	0.000
Model Chi-Square	18.60	0.000	3.04	0.081	3.80	0.051	12.00	0.000
Nagelkerke R Square	0.015	-	0.004	-	0.017	-	0.016	-
AUC-ROC	0.599	-	0.552	-	0.645	-	0.618	-

Table 9. Predicting Failure Outcomes with VPRAI Risk Factors and Gender (n=11,269)

	Any Failure		Failure to Appear		New Arrest		Technical Violation	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.937	0.695	1.086	0.752	0.780	0.589	0.897	0.617
Charge Type	1.261	0.000	0.520	0.000	1.526	0.005	1.830	0.000
Pending Charge	1.936	0.000	4.364	0.000	1.497	0.017	0.901	0.282
Criminal History	1.281	0.001	0.946	0.636	1.088	0.651	1.585	0.000
Two or more FTA	1.233	0.003	1.370	0.007	1.095	0.618	1.139	0.154
Two or more violent convictions	0.823	0.059	0.656	0.024	1.418	0.127	0.826	0.152
Unemployed at time of arrest	1.379	0.000	1.373	0.001	1.222	0.190	1.345	0.000
History of drug abuse	1.245	0.001	1.053	0.642	1.132	0.472	1.348	0.001
Male	1.075	0.268	0.984	0.878	1.338	0.103	1.069	0.442
Constant	0.081	0.000	0.031	0.000	0.009	0.000	0.035	0.000
Model Chi-Square	293.00	0.000	318.70	0.000	31.59	0.000	201.22	0.000
Nagelkerke R Square	0.033	-	0.076	-	0.016	-	0.035	-
AUC-ROC	0.647	-	0.698	-	0.607	-	0.646	-

pretrial failure. Furthermore, gender was unrelated to the pretrial outcomes examined in these analyses, including any failure, failure to appear, new arrest, and technical violation.

Supplemental models (Appendix Tables A5-A8) examined the associations between the individual risk factors included in the VPRAI-R and pretrial failure in separate gender models. The results of these models identified predictors that operated somewhat differently for male and female defendants. For example, criminal history was associated with heightened odds of pretrial failure among female defendants, yet was only a marginally significant predictor of pretrial failure among male defendants. Furthermore, two or more FTAs was associated with a significant increase in the odds of failure to appear among female defendants; however, two or more FTAs was unrelated to the odds of failure to appear among male defendants. A few other minor differences can be observed in the appendix tables. Overall, the results of the separate gender models suggest that VPRAI-R risk factors are more reliable predictors of pretrial failure among female defendants relative to their male counterparts. In particular, the AUC estimates for the logistic regression model predicting any pretrial failure among females was 0.707, as compared to an AUC estimate of 0.622 among males. This pattern of slightly more reliable prediction among female defendants held across the different measures of pretrial failure, with one exception. Specifically, the AUC estimates for male and female defendants were similar in models predicting new arrest (AUC = 0.607 and 0.606, respectively).

Multivariate results, risk level (gender). In order to further assess whether the VPRAI-R produces gender neutral predictions, we examined associations between the pretrial risk level and pretrial failure in separate gender models. These findings are presented in Table 10. In models estimating the odds of any pretrial failure, the pretrial risk level emerges as a significant predictor among male and female defendants. However, there is some variability in the

Table 10. Predicting Failure Outcomes with Pretrial Risk Level by Gender

	Any Failure		Failure to Appear		New Arrest		Technical Violation	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Male (n=8,048)								
Pretrial Risk Level	1.242	0.000	1.052	0.199	1.283	0.000	1.309	0.000
Constant	0.093	0.000	0.042	0.000	0.010	0.000	0.040	0.000
Model Chi-Square	84.88	0.000	1.63	0.201	18.44	0.000	77.40	0.000
Nagelkerke R Square	0.013	-	0.001	-	0.012	-	0.018	-
AUC-ROC	0.588	-	0.528	-	0.595	-	0.609	-
Female (n=3,221)								
Pretrial Risk Level	1.524	0.000	1.398	0.000	1.187	0.131	1.572	0.000
Constant	0.053	0.000	0.022	0.000	0.009	0.000	0.023	0.000
Model Chi-Square	111.05	0.000	29.63	0.000	2.15	0.143	74.96	0.000
Nagelkerke R Square	0.046	-	0.025	-	0.005	-	0.049	-
AUC-ROC	0.666	-	0.641	-	0.567	-	0.683	-

predictive validity of the pretrial risk level by gender across the different pretrial outcomes. For example, the pretrial risk level is significantly associated with the odds of failure to appear for female, but not male, defendants. Furthermore, pretrial risk level is a significant predictor of new arrest for male, but not female, defendants. Overall, the AUC estimates suggest that the pretrial risk level is a better predictor of pretrial failure among female, relative to male, defendants. Given the low rate of new arrest in the sample, and among female defendants in particular, the finding of the separate gender models estimating the risk of new arrest should be interpreted with caution (e.g., 43 women, or 0.38% of the sample, were arrested for a new crime during the pretrial period).

Summary and Conclusions

Our objective was to provide a local validation of the VPRAI-R using data from Palm Beach County, Florida, to determine whether the VPRAI-R provides reliable estimates of pretrial failure (i.e., failure to appear, new arrest, technical violation) in the Palm Beach County context. In addition to providing estimates of the overall predictive validity, a secondary goal was to consider whether there was evidence of predictive bias on the basis of race and/or gender. In the following paragraphs, we discuss the key findings of the validation study. We conclude with some recommendations for future validation efforts, in addition to considerations for the continued use of the VPRAI-R as a pretrial decision-making tool in Palm Beach County.

We analyzed data covering a roughly two-year period from September 2017 through June 2019, and focused on the subset of resolved cases where the defendant was released at some point prior to the final disposition of his or her case (n = 11,269). Although the sample included defendants charged with more serious offenses, the most serious charge in the majority of cases was either a non-violent felony (40.59%) or a misdemeanor (44.82%). Accordingly, the majority of cases fell into the low risk categories (1 and 2) based on their VPRAI-R assessment scores, suggesting that they were unlikely to “fail” during the pretrial period. Indeed, the overall rate of pretrial failure was low, as 13.51% of defendants either failed to appear in court, were arrested for a new crime, or violated a condition of SOR during the pretrial period. Conversely, nearly 9 in ten defendants (86.49%) “succeeded” in the period leading up to the adjudication of their cases.

Our models estimated the predictive validity of the VPRAI-R in a few different ways. We began by estimating a model with each of the eight risk factors included in the assessment, followed by a measure of the total risk score. A final set of analyses included an indicator of the defendants’ risk level. Across these different specifications, our findings indicated that the VPRAI-R provides reliable estimates of the likelihood of “any” pretrial failure. The estimates of predictive validity were highest in the models based on the individual risk items, followed by the total score and risk level. Although these scores fall within the same range, the estimate for the risk level is the most relevant, as these values guide the recommendations made using the PRMM. The estimate of the overall predictive validity of the VPRAI-R based on the risk levels was 0.611. This suggests that the level of a randomly selected defendant who failed would be higher than a randomly selected defendant who did not fail approximately 60% of the time. Based on the criteria for practical significance used in the field of risk assessment, the ability of the VPRAI-R risk levels to distinguish between cases on the basis of risk in the Palm Beach context is fair.

In addition to providing estimates of predictive validity for any failure, we provided separate estimates of the reliability of risk estimates across the different measures of pretrial failure, including failure to appear, new arrest, and technical violation. In the models including the full range of risk factors, the most reliable estimates of pretrial failure came from the model predicting failure to appear. Yet in the models assessing the impact of pretrial risk level on pretrial failure, risk level appeared to more accurately predict technical violations. Thus, given that pretrial risk level is used to inform release/detention recommendations using the PRMM, our findings suggest that the VPRAI-R most reliably predicts technical violations (AUC = 0.630), relative to failure to appear and new arrest (AUCs = 0.560 and 0.594, respectively), in Palm Beach County.

Further examination of the factors contributing to the different estimates of predictive validity based on different measures of risk and "failure" outcomes revealed some potential explanations for these discrepancies. First, attention to the associations between the different risk factors and pretrial failure revealed that the predictors of the diverse types of pretrial failure vary. For example, criminal history is a significant predictor of technical violations, but is unrelated to the likelihood of court appearance. Defendants' employment status at the time of arrest is associated with the odds of court appearance and technical violation, but did not emerge as a significant predictor of new arrest. History of drug abuse is associated with an increased likelihood of technical violations, but does not predict failure to appear or new arrest. In sum, despite some overlap in predictors, there appear to be different mechanisms underlying the likelihoods of failure to appear, new arrest, and technical violation.

Second, that the estimates of pretrial failure varied according to the measure of pretrial risk appears to stem primarily from the inclusion of "two or more violent convictions" in the model. This factor, which was positively associated with pretrial failure in prior validations of the VPRAI, is negatively associated with pretrial risk in models accounting for the full range of risk factors based on the data from Palm Beach County. In particular, defendants with two or more violent convictions are about one-third less likely to fail to appear ($p < .05$), and roughly 17% less likely to fail in general ($p < .10$), than their peers with less than two violent convictions, net of the other risk factors included in the VPRAI-R.

That two or more violent convictions was negatively associated with the odds of pretrial failure net of the other indicators of pretrial risk, and unrelated to pretrial risk at the bivariate level, runs counter to expectations based on prior criminological research and the risk assessment literature. It was also noteworthy that the association between active community supervision and pretrial failure was null at both the bivariate and multivariate level. One potential explanation for these findings is that individuals who were on active community supervision at the time of their arrest, or have two or more violent convictions, are less likely to appear in the analytic sample, as our sample is limited to individuals who were released at some point pending the disposition of their case. This explanation seems plausible; less than 3% of defendants included in our analytic sample were on active community supervision at the time of their arrest as compared to roughly 6% of those who remained in detention. Similarly, 8% of our sample members, relative to 23% of those detained, had two or more violent convictions. These findings reflect a limitation of the data used to develop/validate risk assessments, which is that the most high risk defendants are not released pretrial, and thus do not appear in the data used to assess the factors associated with pretrial risk.

With respect to race, our findings indicate that the VPRAI-R is a more reliable predictor of pretrial failure for non-Latino white defendants relative to defendants of color. These differences appear to stem largely from differences in associations between risk factors and pretrial failure across race/ethnic groups. In particular, charge type is positively associated with pretrial failure among white defendants, but is unrelated to the odds of failure among defendants of color. In addition, history of drug abuse emerged as a significant predictor of pretrial failure among whites, yet was unrelated to the likelihood of pretrial failure among defendants of color.

In models assessing the separate failure outcomes by race, it was also revealed that two or more violent convictions was negatively associated with the odds of failure to appear for non-Latino black defendants. In contrast, two or more violent convictions was unrelated to the odds of court appearance among the other race/ethnic groups. That defendants of color, and black defendants in particular, accumulated points for charge type, history of drug abuse, and two or more violent convictions despite these factors being unrelated to their likelihood of pretrial failure (or negatively related in the case of two or more violent convictions) suggests that the VPRAI-R may be over-predicting risk for these groups. This conclusion is supported by the findings presented in Figure 1, which document that the likelihood of pretrial failure among non-Latino black and Latino defendants in the highest risk category is about 40% lower than that of their white counterparts.

We also examined the potential for predictive bias on the basis of gender. Our findings suggest that the VPRAI-R provides more reliable estimates of pretrial failure among female defendants relative to males. Although associations between risk factors and any pretrial failure were similar between males and females, there were a couple of exceptions. Active community supervision was negatively associated with pretrial failure among female defendants ($p < .10$) and unrelated to the odds of failure among male defendants. In addition, two or more violent convictions was negatively associated with odds of failure among male defendants ($p < .10$) and unrelated to pretrial failure among females. Together with the findings of our racial bias analyses, these findings suggest that the VPRAI-R may be producing overestimates of risk for male defendants—and male defendants of color in particular.

On the basis of these findings, we provide the following suggestions for the future use and empirical evaluation of the VPRAI-R in Palm Beach County. First, findings of bivariate and multivariate models indicated that some of the predictors included in the VPRAI-R are unrelated to the odds of pretrial failure. The obvious implication of these findings is that these factors be removed from the model predicting the likelihood of failure. But such a conclusion is premature, as the current analyses is based on relatively few cases. Although our sample size of 11,269 is generally sufficient to draw conclusions, the cell sizes become small for some of the risk factors included in the model. The least common risk factors are also those that appear to be unrelated to the odds of failure at the bivariate level, including active community supervision and two or more violent convictions. Thus, it will be important to reevaluate associations between the risk factors and pretrial failure at a later point in time once additional data points have been accumulated.

Second, given that the associations between the risk factors and pretrial failure varied across the different outcome variables, there may be some utility in considering separate estimates of the likelihood of failure to appear, new arrest, and technical violation. Other widely available tools (e.g., the PSA) produce separate estimates of pretrial outcomes in light of the fact that the factors contributing to court appearance and new criminal activity differ. Relatedly, whether the risk of technical violation should be included in the overall estimates of pretrial risk should be addressed. There are substantive arguments against the inclusion of technical violations; however, an empirical consideration is that not all defendants are at risk of technical violation, and thus it is illogical to compute estimates that incorporate this element of risk across the full sample.

Finally, given the potential bias with respect to race/ethnicity in pretrial decisions, an important next step is to focus on the population of defendants detained pretrial to determine whether there are any systematic biases that are identifiable in the data. That defendants of color are being assessed as higher risk despite their relatively lower risk of reoffending is in line with current critiques of pretrial risk assessments, and warrants further investigation. Attention to the specific risk factors that are most common among members of the higher risk groups by race, and consideration of whether/how pretrial decisions disproportionately result in detention for members of certain groups, are two important first steps in this regard.

In sum, based on analyses of nearly two years of data from Palm Beach County, we conclude that the VPRAI-R produces reliable estimates of pretrial failure, with some variability on the basis of pretrial failure type and defendant subpopulation. Specifically, the results indicate that the VPRAI-R is a fair predictor of pretrial failure overall, yet the VPRAI-R risk levels are not similarly indicative of pretrial risk for defendants of color and whites or for female and male defendants. Important next steps are to reassess associations between the risk factors included in the VPRAI-R and pretrial failure once additional data becomes available, and to further consider the sources of variability in the predictive validity of estimates documented in this report. Whereas we have focused on the population of released defendants in order to determine whether the VPRAI-R accurately classifies defendants on the basis of pretrial risk, attention to the population of defendants who remained detained during the pretrial period may help disentangle some of the sources of variability and/or predictive bias.

References

Bechtel, K., Lowenkamp, C. T., & Holsinger, A. (2011). Identifying the predictors of pretrial failure: A meta-analysis. *Federal Probation*, *75*(2), 78-88.

Danner, M.J.E., VanNostrand, M., & Spruance, L.M. (2015). Risk-based pretrial release recommendation and supervision guidelines. St. Petersburg, FL: Luminosity, Inc.

Danner, M.J.E., VanNostrand, M., & Spruance, L.M. (2016). Race and gender neutral pretrial risk assessment, release recommendations, and supervision: VPRAI and Praxis revised. St. Petersburg, FL: Luminosity, Inc.

Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in US correctional settings. *Psychological Services*, *13*(3), 206-222.

Dobbie, W., Goldin, J., & Yang, C. S. (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review, 108*(2), 201-240.

Gupta, A., Hansman, C., & Frenchman, E. (2016). The heavy costs of high bail: Evidence from judge randomization. *The Journal of Legal Studies, 45*(2), 471-505.

Heaton, P., Mayson, S., & Stevenson, M. (2017). The downstream consequences of misdemeanor pretrial detention. *Stan. L. Rev.*, *69*, 711.

Leslie, E., & Pope, N. G. (2017). The unintended impact of pretrial detention on case outcomes: Evidence from New York City arraignments. *The Journal of Law and Economics, 60*(3), 529-557.

Phillips, M. T. (2012). *A decade of bail research in New York City*. New York: New York Criminal Justice Agency, Inc.

Oleson, J. C., Lowenkamp, C. T., Wooldredge, J., VanNostrand, M., & Cadigan, T. P. (2017). The sentencing consequences of federal pretrial supervision. *Crime & Delinquency*, *63*(3), 313-333.

Pretrial Justice Institute. (2017). The state of pretrial justice in America. Rockville, MD: Pretrial Justice Institute.

Sacks, M., & Ackerman, A. R. (2014). Bail and sentencing: Does pretrial detention lead to harsher punishment? *Criminal Justice Policy Review, 25*(1), 59-77.

Stevenson, M. (2018). Distortion of justice: How the inability to pay bail affects case outcomes. *The Journal of Law, Economics, and Organization, 34*(4), 511-542.

Zeng, Z. (2018). *Jail inmates in 2016*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Appendix: Supplementary Tables

Appendix Table A1. Predicting Any Failure with VPRAI Risk Factors – Race and Ethnicity

	Any Failure – White (n=4,873)		Any Failure – Black (n=4,679)		Any Failure – Latino (n=1,717)	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.531	0.038	1.380	0.137	0.783	0.619
Charge Type	1.372	0.001	1.153	0.117	1.246	0.187
Pending Charge	2.503	0.000	1.619	0.000	1.874	0.001
Criminal History	1.277	0.020	1.232	0.074	1.478	0.029
Two or more FTA	1.312	0.019	1.311	0.008	1.071	0.743
Two or more violent convictions	0.919	0.648	0.834	0.177	0.891	0.722
Unemployed at time of arrest	1.546	0.000	1.256	0.010	1.155	0.389
History of drug abuse	1.451	0.000	1.080	0.447	1.137	0.481
Constant	0.074	0.000	0.095	0.000	0.088	0.000
Model Chi-Square	227.49	0.000	72.78	0.000	29.67	0.000
Nagelkerke R Square	0.058	-	0.019	-	0.023	-
AUC-ROC	0.689	-	0.610	-	0.625	-

Appendix Table A2. Predicting Failure to Appear with VPRAI Risk Factors – Race and Ethnicity

	FTA – White (n=4,873)		FTA – Black (n=4,679)		FTA – Latino (n=1,717)	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.471	0.154	1.764	0.102	1.256	0.716
Charge Type	0.519	0.000	0.376	0.000	0.969	0.901
Pending Charge	6.363	0.000	3.916	0.000	2.576	0.000
Criminal History	0.840	0.334	1.077	0.695	0.946	0.833
Two or more FTA	1.399	0.078	1.441	0.034	1.469	0.200
Two or more violent convictions	0.816	0.544	0.564	0.027	1.050	0.916
Unemployed at time of arrest	1.369	0.039	1.350	0.044	1.548	0.058
History of drug abuse	1.488	0.022	0.811	0.221	0.840	0.525
Constant	0.025	0.000	0.031	0.000	0.041	0.000
Model Chi-Square	194.53	0.000	141.53	0.000	23.53	0.003
Nagelkerke R Square	0.111	-	0.083	-	0.032	-
AUC-ROC	0.740	-	0.719	-	0.644	-

Appendix Table A3. Predicting Technical Violations with VPRAI Risk Factors – Race and Ethnicity

	TV – White (n=4,873)		TV – Black (n=4,679)		TV – Latino (n=1,717)	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.454	0.065	1.453	0.173	0.639	0.545
Charge Type	2.049	0.000	1.812	0.000	1.169	0.508
Pending Charge	1.152	0.318	0.682	0.013	1.163	0.587
Criminal History	1.669	0.000	1.341	0.066	2.117	0.004
Two or more FTA	1.310	0.055	1.172	0.243	1.002	0.994
Two or more violent convictions	1.049	0.826	0.777	0.169	0.950	0.904
Unemployed at time of arrest	1.617	0.000	1.168	0.195	0.860	0.545
History of drug abuse	1.434	0.005	1.215	0.154	1.357	0.223
Constant	0.032	0.000	0.044	0.000	0.034	0.000
Model Chi-Square	160.25	0.000	58.81	0.000	19.73	0.011
Nagelkerke R Square	0.060	-	0.025	-	0.027	-
AUC-ROC	0.691	-	0.623	-	0.636	-

Appendix Table A4. Predicting New Arrest with VPRAI Risk Factors – Race and Ethnicity

	NA – White (n=4,873)		NA – Black (n=4,679)		NA – Latino (n=1,717)	
	Odds Ratio	*P*	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	1.678	0.392	0.505	0.344	1.000	-
Charge Type	1.377	0.201	1.397	0.103	3.583	0.006
Pending Charge	1.253	0.475	1.593	0.032	1.602	0.415
Criminal History	0.990	0.970	1.157	0.615	1.604	0.368
Two or more FTA	0.698	0.375	1.292	0.268	0.203	0.132
Two or more violent convictions	0.570	0.444	1.728	0.032	1.000	-
Unemployed at time of arrest	1.208	0.457	1.221	0.327	0.948	0.919
History of drug abuse	1.037	0.899	1.209	0.422	1.421	0.516
Constant	0.013	0.000	0.011	0.000	0.006	0.000
Model Chi-Square	5.31	0.7241	25.95	0.001	13.12	0.041
Nagelkerke R Square	0.007	-	0.026	-	0.061	-
AUC-ROC	0.593	-	0.630	-	0.720	-

Appendix Table A5. Predicting Any Failure with VPRAI Risk Factors – Gender

	Any Failure – Female (n=3,221)		Any Failure – Male (n=8,048)	
	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.473	0.069	1.103	0.588
Charge Type	1.349	0.014	1.222	0.004
Pending Charge	2.518	0.000	1.769	0.000
Criminal History	1.538	0.001	1.169	0.064
Two or more FTA	1.398	0.019	1.179	0.043
Two or more violent convictions	1.189	0.521	0.813	0.065
Unemployed at time of arrest	1.449	0.001	1.332	0.000
History of drug abuse	1.361	0.019	1.226	0.008
Constant	0.061	0.000	0.098	0.000
Model Chi-Square	169.89	0.000	145.43	0.000
Nagelkerke R Square	0.071	-	0.022	-
AUC-ROC	0.707	-	0.622	-

Appendix Table A6. Predicting Failure to Appear with VPRAI Risk Factors – Gender

	FTA – Female (n=3,221)		FTA – Male (n=8,048)	
	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.960	0.934	1.103	0.750
Charge Type	0.566	0.011	0.499	0.000
Pending Charge	6.460	0.000	3.699	0.000
Criminal History	0.842	0.431	0.962	0.781
Two or more FTA	1.876	0.005	1.211	0.163
Two or more violent convictions	0.763	0.584	0.691	0.069
Unemployed at time of arrest	1.650	0.006	1.229	0.083
History of drug abuse	1.280	0.235	0.992	0.952
Constant	0.021	0.000	0.037	0.000
Model Chi-Square	163.13	0.000	175.97	0.000
Nagelkerke R Square	0.135	-	0.059	-
AUC-ROC	0.781	-	0.676	-

Appendix Table A7. Predicting Technical Violations with VPRAI Risk Factors – Gender

	TV – Female (n=3,221)		TV – Male (n=8,048)	
	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	0.281	0.080	1.118	0.628
Charge Type	2.016	0.000	1.758	0.000
Pending Charge	0.881	0.532	0.908	0.385
Criminal History	2.231	0.000	1.371	0.005
Two or more FTA	1.136	0.493	1.139	0.216
Two or more violent convictions	1.417	0.277	0.777	0.085
Unemployed at time of arrest	1.389	0.030	1.322	0.002
History of drug abuse	1.443	0.033	1.334	0.004
Constant	0.026	0.000	0.042	0.000
Model Chi-Square	102.23	0.000	112.39	0.000
Nagelkerke R Square	0.067	-	0.026	-
AUC-ROC	0.701	-	0.626	-

Appendix Table A8. Predicting New Arrest with VPRAI Risk Factors – Gender

	NA – Female (n=3,221)		NA – Male (n=8,048)	
	Odds Ratio	*P*	Odds Ratio	*P*
Active community supervision	1.000	-	0.974	0.955
Charge Type	1.652	0.130	1.489	0.017
Pending Charge	1.495	0.309	1.516	0.026
Criminal History	1.538	0.241	0.974	0.902
Two or more FTA	1.019	0.965	1.110	0.607
Two or more violent convictions	1.363	0.683	1.417	0.150
Unemployed at time of arrest	0.833	0.580	1.366	0.069
History of drug abuse	0.879	0.739	1.225	0.298
Constant	0.010	0.000	0.012	0.000
Model Chi-Square	5.80	0.563	23.93	0.002
Nagelkerke R Square	0.013	-	0.016	-
AUC-ROC	0.606	-	0.607	-