DIFFERENCES IN THE PREDICTIVE ACCURACY OF IRAS-PAT ASSESSMENTS AS A FUNCTION OF AGE, SEX, AND RACE

AN UPDATED REPORT

Evan M. Lowder

Spencer G. Lawson

Chelsea Foudray

Hrishikesh Vibhute

Ashley Rodriguez

July 2022

This report was prepared for the Indiana Office of Court Services (IOCS). The views and opinions expressed herein are those of the authors and do not necessarily reflect the views and opinions of IOCS.

EXECUTIVE SUMMARY

In 2015, the Indiana Office of Court Services entered into an agreement with 11 Indiana counties to pilot the Indiana Risk Assessment System – Pretrial Assessment Tool (IRAS-PAT) in local pretrial decision-making processes. To date, evaluation of this initiative suggests that the tool was successfully implemented into local decision-making (Grommon et al., 2017) and that assessments produced by the IRAS-PAT predict key pretrial misconduct outcomes with good-to-excellent accuracy (Lowder, Lawson, Grommon, et al., 2020). In a preliminary investigation, we examined the predictive accuracy of IRAS-PAT assessments by age, sex, and race across five Indiana counties (Lowder, Lawson, & Foudray, 2020). Results of this investigation showed little evidence that IRAS-PAT assessments differentially predicted pretrial outcomes as a function of age or sex. However, there were notable differences in predictive accuracy between White and Black defendants, with IRAS-PAT assessments producing weaker predictive validity estimates for Black defendants.

The purpose of the present investigation was to replicate our earlier investigation with a 12-county validation sample. Counties include all 11 counties participating in the Indiana Pretrial Pilot Project as well as one additional county (Vigo). We first examined overall predictive accuracy by outcome. We then conducted an examination of predictive validity by age (i.e., adults aged 33 and older versus those under 33), sex (i.e., male and female defendants), and race (i.e., Black and White defendants). Where we found consistent evidence of predictive bias, we conducted an in-depth investigation to examine item-level drivers of disparate predictive accuracy. The 12-county sample included 6,919 Indiana pretrial defendants.

Summary of Findings

Overall, predictive validity findings showed IRAS-PAT assessments produced good levels of predictive accuracy for all three pretrial misconduct outcomes (any FTA, any new arrest, and any arrest). Estimates of predictive validity were slightly stronger for any new arrest and any arrest outcomes relative to any FTA. Predictive validity estimates were slightly lower than a previous 5-county validation sample (Lowder, Lawson, Grommon, et al., 2020), but comparable with predictive validity estimates from other pretrial risk assessments (Desmarais et al., 2020).

We found no evidence of predictive bias for age and limited evidence for sex. Evidence of predictive bias for sex was limited to the prediction of any FTA, where IRAS-PAT assessments predicted any FTA with greater accuracy for female defendants relative to male defendants. However, evidence of predictive bias for sex was not consistent across all analyses.

We saw evidence of predictive bias in IRAS-PAT assessments as a function of race. However, differences in predictive accuracy were smaller in magnitude relative to a previous report (Lowder, Lawson, & Foudray, 2020), driven by stronger predictive accuracy for any FTA among Black defendants and lower predictive accuracy overall for White defendants. Differences in predictive accuracy were most noticeable for arrest outcomes, including any arrest and any new arrest. There was no evidence of predictive bias for any FTA. Overall, IRAS-PAT assessments produced at least a “fair” level of predictive validity for all outcomes and by racial group.

When we examined item-level differences in predictive accuracy, there were noticeable differences in the items that contributed uniquely to the prediction of pretrial misconduct outcomes between Black and White defendants. However, only Item 4 (Unemployment) showed evidence of a different strength in predictive accuracy between Black and White defendants, particularly for arrest outcomes and relative to being fully employed.

Conclusions

IRAS-PAT assessments continued to show good evidence of predictive validity for pretrial misconduct outcomes. The tool produces equally valid assessments across age and sex, with limited exceptions. For race, IRAS-PAT assessments show weaker evidence of predictive validity for arrest outcomes among Black defendants relative to White defendants. These differences are fairly robust to the legal and extralegal characteristics of these groups.

INTRODUCTION

Determining whether to release a newly arrested defendant into the community is one of the most critical decisions during the pretrial period. Pretrial decision-making involves multiple justice system professionals making timely choices throughout the process. Front-end system decision-making has implications for subsequent outcomes. Defendants incarcerated pending trial are more likely to plead guilty, receive lengthier sentences, and subsequently recidivate more often in relation to defendants released prior to court disposition (Stevenson & Mayson, 2017). Incarceration can also disrupt housing, employment, family relationships, and ties to the community (Stevenson & Mayson, 2017). The use of actuarial risk assessments during the pretrial period has emerged as one strategy to reduce pretrial detention rates, achieve equitable non-monetary conditions of release, and minimize racial and socioeconomic disparities in release and detention decisions. The integration of these assessment tools comes at a time when communities across the United States have recognized the importance of advancing pretrial practice and policy, resulting in a national movement for pretrial and bail reform efforts.

Despite the potential of these tools to predict future pretrial misconduct (Bechtel et al., 2017; Desmarais et al., 2020; Lowder et al., 2020), pretrial risk assessment tools are not without controversy. There remain serious concerns about the potential for risk assessments to exacerbate racial and ethnic disparities in criminal justice processing (Pretrial Justice Institute, 2020). Specifically, some scholars have argued that items included in risk assessments reflect the relative disadvantage of minority defendants and, as a result, bias those defendants toward higher risk classifications (Harcourt, 2015; Starr, 2014). There have been a limited but growing number of investigations into racial and ethnic disparities in the use of risk assessments during pretrial decision-making. Risk assessment validation studies have shown evidence of lower predictive accuracy of assessments for racial and ethnic minorities (Fass et al., 2008), attributable to racial disparities in socioeconomic characteristics and criminal history, content domains that are frequently embedded within risk assessment tools (Zettler & Morris, 2015). However, some research suggests that risk can be estimated free of bias (Baglivia et al., 2019; Flores et al., 2016), particularly when assessments measure factors that protect against future misconduct (Lowder et al., 2019). Overall, there has been a lack of consistent findings on predictive accuracy of pretrial risk assessments in criminal justice settings for racial and ethnic minorities (see, e.g., Cohen & Lowenkamp, 2019; Copp et al., 2019; DeMichele & Baumgartner, 2020).

Previously, the Indiana Supreme Court established the Committee to Study Evidence-Based Pretrial Release in 2014, which developed an evidence-based pretrial program to evaluate the implementation of the Indiana Risk Assessment System – Pretrial Assessment Tool (IRAS-PAT). Guided by the National Institute of Corrections' Evidence Based Decision Making (EBDM) Framework, 11 counties in 2015 entered into an agreement with the Indiana Office of Court Services (IOCS) to develop and implement their own pretrial pilot project aimed at maximizing public safety, court appearance, and pretrial release; IRAS-PAT assessments being at the core of these local pretrial justice reform efforts. Researchers from the Indiana University Public Policy Institute, Center for Criminal Justice Research (CCJR) conducted a process evaluation of pilot counties to understand how the IRAS-PAT was adopted by participating pilot counties and to identify barriers and facilitators of implementation (Grommon et al., 2017).

As part of the pilot implementation of the IRAS-PAT, our research team has conducted county-level validations of IRAS-PAT assessments implemented in practice. All validations for the 11 pilot counties as well as an additional county (i.e., Vigo) have been completed. These county-level validations have provided useful insight into the predictive accuracy of IRAS-PAT assessments. Our local validation approach found strong predictive accuracy of IRAS-PAT assessments, with estimates meeting or exceeding standards for the performance of risk assessment tools in a justice system context (Lowder, Lawson, Grommon, et al., 2020).

However, limited sample sizes in each validation did not allow for rigorous investigation into differences in the predictive accuracy as a function of race at the county level. The ability of pretrial risk assessment tools to predict outcomes with similar accuracy between Black and White defendants, for example, has emerged as a key concern in pretrial reform efforts. Thus, the current inquiry moves to the third stage of research on Indiana's pretrial pilot project. This phase investigates differential prediction by race of IRAS-PAT assessments. This report defines differential prediction for an ordinal risk level (e.g., low, medium, or high) or continuous score (e.g., 0 to 9) as between-group differences in the slopes of regression lines. In other words, differences in slopes demonstrate that assessments are stronger predictors of pretrial misconduct outcomes for White defendants relative to Black defendants. When there is bias, this bias is typically found at higher risk levels, where Black defendants tend to be over-classified relative to their risk of misconduct. We further assessed differential prediction of IRAS-PAT assessments on the basis of sex and age.

In a preliminary investigation based on a sample of 3,539 pretrial defendants across five Indiana jurisdictions (Lowder, Lawson, & Foudray, 2020), we found no evidence of differential prediction in IRAS-PAT assessments as a function of age or sex. In contrast, we found notable and consistent evidence of differential prediction by race in IRAS-PAT assessments. IRAS-PAT assessments produced weaker predictive validity estimates for Black defendants relative to White defendants across all pretrial misconduct outcomes, but most notably for any FTA and any arrest. However, these findings were based on a five-county sample of Indiana pretrial defendants, and it is unclear whether this sample generalizes to Indiana's entire pretrial population.

To better understand how the IRAS-PAT functions across defendant characteristics, we replicated our prior report to examine the presence of differential prediction in IRAS-PAT assessments in a 12-county sample. These counties represented all 11 counties participating in Indiana's Pretrial Pilot Project and one additional county (Vigo County). This multi-jurisdictional, pooled sample includes 6,919 pretrial defendants who received a risk assessment, had court charges filed, and spent time in the community prior to court case disposition. The 11 pilot counties began using the IRAS-PAT between 2016 and 2017. Vigo county began implementing the IRAS-PAT in 2020. In each jurisdiction, a follow-up period was constructed for each defendant, which was defined by the pretrial processing period (i.e., the date of index jail release to the date of court disposition). The objective of this updated report is to test for differential prediction as an indication of bias in IRAS-PAT assessments across 12 Indiana counties, with a focus on the extent to which the IRAS-PAT provided accurate predictions on the basis of race, sex, and age. We first report on the overall predictive accuracy of IRAS-PAT assessments across pretrial misconduct outcomes.

METHODS

Data Sources

Data for this investigation were drawn from validation data from the 11 pilot counties (Allen, Bartholomew, Grant, Hamilton, Hendricks, Jefferson, Monroe, Porter, Starke, St. Joseph, and Tipton) as well as Vigo County. For each validation, data sources included jail, court, and risk assessment records. First, for each county, we received county level jail data on all admissions, associated release dates, and booking charge(s). For 11 of 12 counties, court data were drawn from Indiana's statewide court case management system, Odyssey. For one county, court data were drawn from a local records management system (i.e., Jefferson). For all sources, court data contained information on all criminal cases and case-related information (e.g., hearings, case disposition, warrants, and FTAs) processed in each county. Particularly for FTAs, we manually consulted case notes in Indiana's MyCase for all defendants in several counties when information could not be located in administrative files. Finally, we received risk assessment records from the INcite system, which included assessment date, total score, and item-level data. For two counties (i.e., Grant and Porter), we also relied on internal assessment records conducted during an initial episode of incarceration to collect pretrial data.

Sample

The sample included 6,919 pretrial defendants who were primarily White (n = 5,511, 79.7%) versus Black (n = 1,408, 20.4%) and mostly male (n = 4,963, 71.7%). To ensure consistent sample sizes across comparisons, defendants identifying with other racial groups were removed from analysis. A large proportion of defendants were ages under 33 (n = 3,734, 54.0%), with an average age of 33.7 years old (SD = 11.52, Range: 16 to 82).

Variables

IRAS-PAT. The IRAS-PAT is an actuarial assessment designed to predict risk of arrest and FTA during the pretrial period. The IRAS-PAT is a 7-item instrument measuring 1) age at first arrest, 2) number of FTA warrants in the past 24 months, 3) three or more prior jail incarcerations, 4) employment at the time of arrest, 5) residential stability, 6) illegal drug use in the past six months, and 7) a severe drug use problem. Items 1, 3, 5, 6, and 7 are scored dichotomously (i.e., 0 or 1) and items 2 and 4 are scored on a 0-2 point scale, producing a maximum total score of 9. Total scores classify defendants into three risk levels: Low (0-2), Moderate (3-5), and High (6+). Our investigation used IRAS-PAT *total scores*, *risk levels*, and *items*. Note, 699 individuals did not have item-level data and were excluded from item-level analyses.

Pretrial Misconduct Outcomes. Outcomes included *any FTA* (yes; no), *any new arrest* (yes; no), and *any arrest* (yes; no) occurring during the pretrial processing period (i.e., following initial release from jail but prior to court case disposition). Any FTA measured failure to appear at a court appearance that resulted in issuance of a warrant. The process for measuring this variable differed slightly across jurisdictions but was informed by conversations with local stakeholders. In some jurisdictions, few FTAs were recorded with accompanying event dates in court records. For these counties, we captured FTAs using triangulated jail booking and court warrant records. Specifically, we matched booking records for an FTA charge to service dates for a warrant record in court records. This process allowed us to establish an issued date for the

FTA warrant and link it to a specific court case. In other jurisdictions, we coded FTA warrants through a manual review of court case notes on Indiana's MyCase (Office of Judicial Administration, 2021). Any new arrest measured a new booking occurring during the pretrial period in which a detainee was booked on any new offense charge. Any arrest measured any booking occurring during the pretrial period.

Demographic Characteristics. Demographic variables included *race* (Black; White), *sex* (female; male), and *age* (under 33; 33 and older).

Covariates. Covariates included *county* (dummy coded, with County 1 as the reference group) and *time at risk*, which measured the total number of days from the date of pretrial release to the date of court case disposition, minus any time incarcerated in the local jail. Note, multivariable models controlled for county, but, for ease of presentation, these estimates are not shown.

Analytic Strategy

We first report on predictive validity estimates overall and then by subgroup for each of the following demographic groups: age, sex, and race. Within each set of analyses, we first conducted descriptive statistics on all study variables overall and by demographic characteristics. Second, we conducted bivariable statistics to test hypotheses of mean and proportional differences between IRAS-PAT total scores, risk levels, and pretrial misconduct outcomes across each group. We report the associated effect size estimates in text (i.e., Cramer's V, Cohen's d). Cramer's V estimates of 0.10, 0.30, and 0.50 represent small, medium, and large effect sizes, respectively (Cohen, 1988). In terms of d, Cohen (1988) suggested corresponding estimates of 0.20, 0.50, and 0.80 indicate small, medium, and large effect sizes, respectively.

To examine differential prediction of IRAS-PAT assessments by demographic characteristics, we used a multi-pronged approach. First, we examined the Area Under the Curve (AUC) of the Receiving Operating Characteristic (ROC) statistics across each group. AUC values are commonly used to evaluate the predictive accuracy of risk assessment total scores. AUC values range from .50 to 1, with .50 indicating chance levels of classification and 1 suggesting perfect classification. AUC values below .54 are typically considered poor, .55 to .63 fair, .64 to .70 good, and .71 and above excellent. These conventions have been documented in reports adopted by the Bureau of Justice Assistance, National Institute of Justice, and National Institute of Corrections and represent benchmarks for predictive accuracy in the field of risk assessment (Desmarais & Singh, 2013).

Second, we conducted a series of logistic regression analyses to examine differential prediction as an indication of bias in IRAS-PAT assessments for each pretrial misconduct outcome, controlling for county and time at risk. For reference, odds ratios of 1.50, 3.00, and 5.00 indicate small, medium, and large effect sizes, respectively (Chen et al., 2010). Significant effects of interest are shaded in grey. To test for the presence of differential prediction in IRAS-PAT assessments by demographic characteristics, we tested for evidence of interaction effect(s) between the demographic characteristic and IRAS-PAT total scores, risk levels, and items. We employed hierarchical logistic regression models to test for improvement in model fit between a main-effects only model (e.g., race and IRAS-PAT scores as independent predictors) (Block 1) and a second model with added interaction terms (e.g., race by IRAS-PAT score) (Block 2).

Hierarchical models are useful when researchers are interested in testing how addition of a model term, such as an interaction effect between two variables, improves the overall ability of the model to predict an outcome. A significant interaction term demonstrates that a risk score produces stronger predictions of pretrial misconduct risk for one group over another. In unweighted models, we used change in -2 log likelihood statistics to assess for improvement in model fit between Block 1 and Block 2. We also present decomposed interactions (i.e., predicted probabilities and associated 95% confidence intervals).

In the race-specific analyses, we conducted two additional sets of analyses to address specific criticisms that item-level results would be due to 1) baseline differences between groups or 2) unequal sample sizes between groups, given the small number of Black defendants in the sample. First, to adjust for baseline differences between groups, we conducted propensity score matching using MatchIt in R and specifying a full matching procedure (Ho et al., 2011; Stuart & Green, 2008). Propensity scores measure the probability that a given individual will belong to a specific group (e.g., Black or White defendant) given known characteristics of that individual. For the purposes of this analysis, we matched White defendants to Black defendants based on county, age, gender, IRAS-PAT total score, time in the community, highest charge level, and charge types. Charge types were selected based on prevalence in the overall sample (i.e., $\geq$ 10% of defendants). Weights generated from the propensity score matching procedure were then used in multivariable models (i.e., weighted models).

Second, to address criticisms that item-level findings would reflect unequal sample sizes between White and Black defendants, we developed a stratified sample of White defendants. This process involved collecting a random sample of White defendants from each of the individual county samples based on the number of Black defendants from each original county sample. The stratified sample shows whether the item-level results are independent of sample size differences between Black and White defendants.

FINDINGS

Overall Predictive Accuracy of IRAS-PAT Assessments

Bivariable Comparisons. Crosstabulations of risk levels and pretrial misconduct outcomes are presented in Table 1. As shown, high risk defendants experienced higher rates of all misconduct outcomes relative to moderate and low risk defendants. Cramer's V estimates ranged from .20 to .26, corresponding to small-to-moderate effect sizes.

Table 1. Crosstabulations of Risk Level and Pretrial Misconduct Outcomes

Case Outcomes	Risk Level						Comparison	
	Low		Moderate		High			
	n	%	*n*	%	*n*	%	X^2 (df)	Cramer's V
Any FTA	148	6.0	506	15.6	333	27.2	308.35*** (2)	.21
Any new arrest	219	8.9	665	20.5	378	30.9	284.95*** (2)	.20
Any arrest	367	15.0	1,089	33.6	593	48.5	485.27*** (2)	.26

Note. ***p < .001, N = 6,919

AUC of the ROC. AUC estimates for total scores were 0.66 (SE = 0.01, 95% CI [0.65, 0.68]) for any new arrest, 0.68 (SE = 0.01, 95% CI [0.67, 0.69]) for any arrest, and 0.69 for any FTA (SE = 0.01, 95% CI [0.67, 0.70]). These estimates correspond to good levels of predictive accuracy for all outcomes. Identical AUC estimates were observed for IRAS-PAT risk levels.

Logistic Regression Models. Results of multivariable logistic regression models adjusting for time at risk in the community and county are presented in Table 2. In Model 1, each 1-point increase in IRAS-PAT total score was associated with a 1.42, 1.35, and 1.43 times increase in the rate of any FTA, any new arrest, and any arrest, respectively, *p*s < .001. In Model 2, IRAS-PAT assessments showed similar ability to discriminate between participants assessed at Low and Moderate risk levels (OR range: 2.72 to 3.03, *p*s <.001). Larger effect sizes were found for the prediction of any FTA (OR = 7.18), any new arrest (OR = 4.84), and any arrest (OR = 6.05), *p*s < .001, between defendants classified as High and Low risk.

Table 2. Logistic Regression Models of IRAS-PAT Total Scores and Risk Levels Predicting Pretrial Misconduct Outcomes

Predictor	Outcomes														
	Any FTA					Any new arrest					Any arrest				
	Estimate	*SE*	z	OR	95% CI	Estimate	*SE*	z	OR	95% CI	Estimate	*SE*	z	OR	95% CI
Model 1															
Total score	0.35	0.02	17.04***	1.42	[1.36, 1.48]	0.30	0.02	16.73***	1.35	[1.31, 1.40]	0.36	0.02	21.97***	1.43	[1.38, 1.47]
Time at risk	<0.01	<0.01	17.23***	1.00	[1.00, 1.00]	<0.01	<0.01	12.36***	1.00	[1.00, 1.00]	<0.01	<0.01	17.36***	1.00	[1.00, 1.00]
Model 2															
Risk level (Low)															
Moderate	1.04	0.10	10.05***	3.03	[2.31, 3.46]	0.99	0.08	11.46***	2.72	[2.29, 3.22]	1.10	0.07	15.21***	2.99	[2.60, 3.44]
High	1.80	0.12	15.43***	7.18	[4.80, 7.57]	1.58	0.10	15.36***	4.84	[3.96, 5.93]	1.80	0.09	20.13***	6.05	[5.08, 7.21]
Time at Risk	<0.01	<0.01	17.11***	1.00	[1.00, 1.01]	<0.01	<0.01	12.20***	1.00	[1.00, 1.00]	<0.01	<0.01	17.14***	1.00	[1.00, 1.00]

Note. For categorical variables, reference group indicated in parentheses. Models controlled for county, but estimates are not shown. CI = confidence interval for odds ratio.
***p < .001. *N* = 6,919.

Table 3 presents results of multivariable logistic regression models examining the unique predictive accuracy of IRAS-PAT items on pretrial misconduct outcomes. Notably, except for part-time employment relative to full time employment (Item 4) and only for any FTA, all items uniquely contributed to the prediction of all three pretrial misconduct outcomes. More than two FTAs in the past 24 months (Item 2; OR range: 1.61 to 2.17) and unemployment relative to full-time employment status (Item 4; OR range: 1.66 to 2.00) were consistently strong predictors of all three pretrial misconduct outcomes.

For any FTA, specifically, more than two FTAs in the past 24 months (Item 2; OR = 2.17) was the strongest unique predictor, $p <$ .001. Part-time employment (Item 4) was not significantly associated with any FTA, $p = .919$. While all items uniquely predicted any new arrest, IRAS-PAT items were weaker predictors of any new arrest overall. More than two FTAs in the past 24 months (Item 2; OR = 1.61) and unemployment relative to full-time employment status (Item 4; OR = 1.66) were the strongest item-level predictors of any new arrest, *p*s < .001. Similarly, all items uniquely predicted any arrest, although three or more prior incarcerations (Item 3; OR = 1.90) and unemployment relative to full-time employment status (Item 4; OR = 1.93) were the strongest unique predictors, *p*s < .001. Illegal drug use in the past six months (Item 6; OR = 1.28) was the weakest unique predictor, $p = .001$.

Table 3. Logistic Regression Models of IRAS-PAT Items Predicting Pretrial Misconduct Outcomes

Predictor	Outcomes														
	Any FTA					Any new arrest					Any arrest				
	Estimate	*SE*	*z*	OR	95% CI	Estimate	*SE*	*z*	OR	95% CI	Estimate	*SE*	*z*	OR	95% CI
1. Age at first arrest[a]	0.50	0.18	2.77**	1.65	[1.16, 2.35]	0.34	0.15	2.34*	1.41	[1.06, 1.89]	0.42	0.12	3.48**	1.52	[1.20 1.92]
2. Number of FTAs[b]															
1	0.69	0.11	6.25***	1.99	[1.60, 2.47]	0.28	0.10	2.68**	1.32	[1.08, 1.62]	0.51	0.09	5.56***	1.67	[1.39, 2.00]
2 or more	0.78	0.14	5.55***	2.17	[1.65, 2.86]	0.47	0.13	3.70***	1.61	[1.25, 2.06]	0.56	0.12	4.71**	1.76	[1.39, 2.22]
3. 3+ prior incarcerations[c]	0.48	0.08	5.74***	1.61	[1.37, 1.90]	0.44	0.07	5.98***	1.55	[1.34, 1.80]	0.64	0.07	9.78***	1.90	[1.67, 2.15]
4. Employed[d]															
Part-time	-0.01	0.14	-0.10	0.99	[0.75, 1.30]	0.25	0.11	2.22*	1.29	[1.03, 1.60]	0.32	0.10	3.38**	1.38	[1.15, 1.67]
Not employed	0.69	0.09	7.71***	2.00	[1.68, 2.39]	0.51	0.08	6.49***	1.66	[1.43, 1.94]	0.66	0.07	9.53***	1.93	[1.69, 2.21]
5. Residential instability[e]	0.24	0.08	2.89**	1.27	[1.08, 1.49]	0.32	0.07	4.50***	1.38	[1.20, 1.59]	0.32	0.06	4.97***	1.38	[1.21, 1.56]
6. Illegal drug use 6 mo[f]	0.29	0.09	3.10**	1.34	[1.11, 1.61]	0.35	0.08	4.23***	1.41	[1.20, 1.66]	0.25	0.07	3.48**	1.28	[1.11, 1.47]
7. Severe drug use[g]	0.39	0.10	3.82***	1.48	[1.21, 1.81]	0.38	0.09	4.18***	1.46	[1.22, 1.74]	0.45	0.08	5.51***	1.56	[1.33, 1.83]
Time at risk	<0.01	<0.01	15.27***	1.00	[1.00, 1.00]	<0.01	<0.01	11.68***	1.00	[1.00, 1.00]	<0.01	<0.01	15.62***	1.00	[1.00, 1.00]

Note. Models controlled for county, but estimates are not shown. CI = confidence interval for odds ratio. *p < .05. **p < .01. ***p < .001. N = 6,220.

[a]Item 1 reference: 33 or older

[b]Item 2 reference: No FTA warrants past 24 months

[c]Item 3 reference: Two or less prior jail incarcerations

[d]Item 4 reference: Yes, Full-time employment at time of arrest

[e]Item 5 reference: Lived at current residence past 6 months

[f]Item 6 reference: No illegal drug use during past 6 months

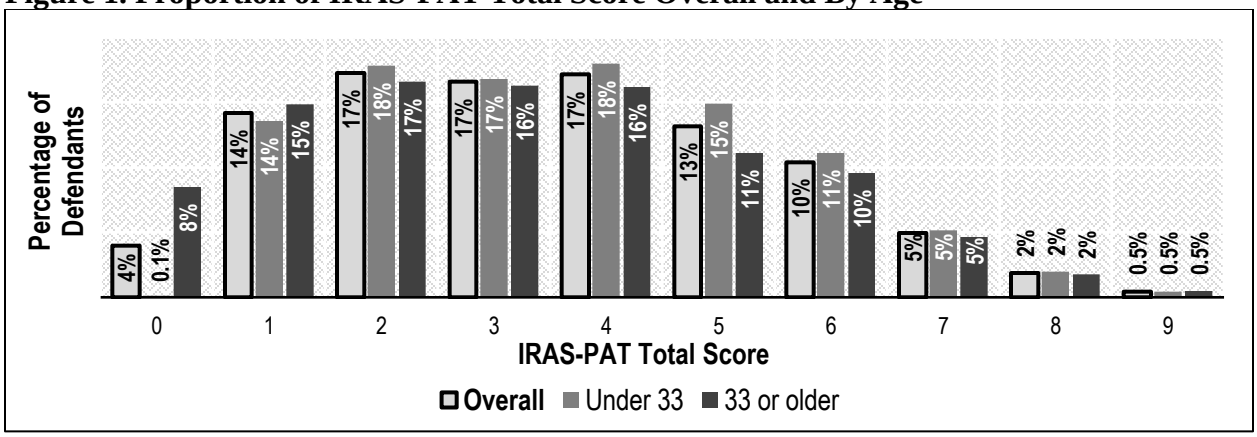
[g]Item 7 reference: No severe drug use problem

Differential Prediction of IRAS-PAT Assessments by Age

Descriptives

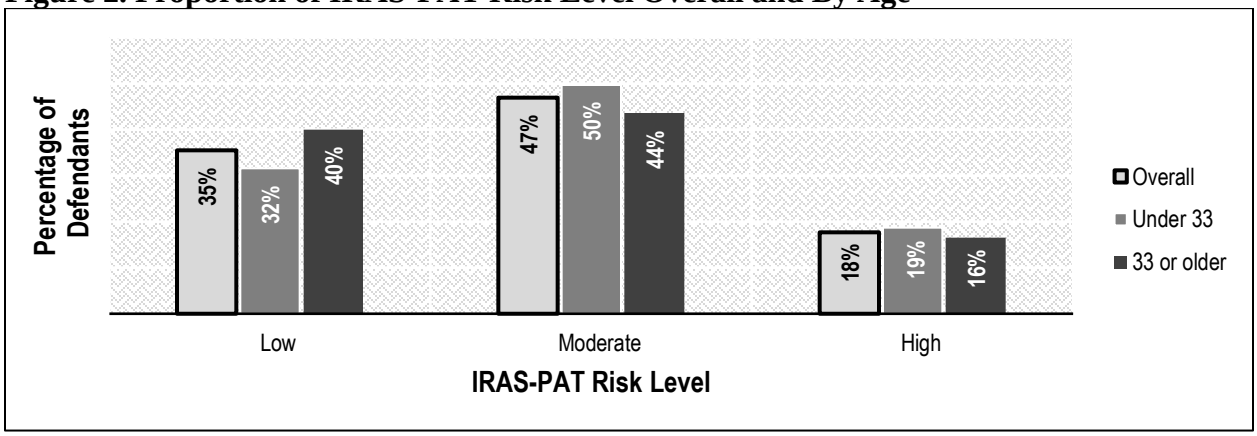
IRAS-PAT. IRAS-PAT scores averaged 3.69 (*SD* = 1.88, Range: 0 to 9) for defendants ages under 33 and 3.26 (*SD* = 2.07, Range: 0 to 9) for defendants 33 or older. Defendants ages under 33 had significantly higher IRAS-PAT scores relative to defendants who were 33 or older (t[6,491.77] = -9.12, $p < .001$, Cohen's d = -0.22). The frequency distribution of IRAS-PAT scores by age is presented in Figure 1.

Figure 1. Proportion of IRAS-PAT Total Score Overall and By Age



The larger proportion of defendants ages under 33 classified at higher risk relative to defendants ages 33 or older is also depicted in Figure 2. As shown, slightly higher proportions of defendants ages under 33 were classified at Moderate and High risk (n = 1,857, 49.7% and n = 699, 18.7%, respectively) relative to defendants ages 33 or older (n = 1,388, 43.6% and n = 524, 16.4%, respectively). Conversely, more defendants ages 33 or older were classified at Low risk (n = 1,273, 40.0%) relative to defendants ages under 33 (n = 1,178, 31.6%).

Figure 2. Proportion of IRAS-PAT Risk Level Overall and By Age

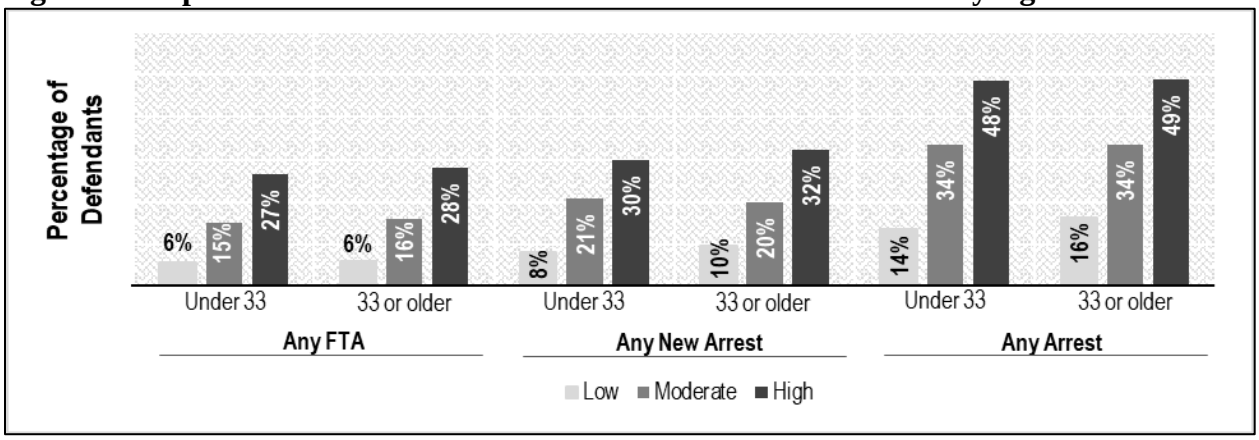


Pretrial Misconduct Outcomes. Defendants ages under 33 and defendants ages 33 or older did not diverge significantly from one another on pretrial misconduct outcomes. Following jail release, but prior to case disposition, 14.3% of defendants ages under 33 had any failure to appear for any court hearing (*n* = 535), and 18.5% had at least one new arrest (*n* = 691). About one-third of defendants ages under 33 had any arrest prior to case disposition (30.0%, *n* = 1,121). Similar rates of any FTA (*n* = 452, 14.2%), any new arrest (*n* = 571, 17.9%), and any arrest (*n* = 928, 29.1%) were observed among defendants ages 33 or older.

Frequency Distributions of Risk Level and Pretrial Misconduct Outcomes by Age

Figure 3 presents the frequency distribution of IRAS-PAT risk level and pretrial outcomes stratified by age. As shown, defendants ages under 33 and 33 or older had, on average, similar rates of pretrial misconduct at each risk level. See Appendix Table 1A for crosstabulations of risk level and pretrial misconduct outcomes by age.

Figure 3. Proportion of Risk Level and Pretrial Misconduct Outcome by Age



Differential Prediction Analyses

AUC of the ROC. In Table 4, we present AUC values and their associated conventions separately by age and outcome. As shown, IRAS-PAT assessments produced similar predictive accuracy estimates across both groups for all pretrial misconduct outcomes.

Table 4. AUC Values by Pretrial Misconduct Outcome and Age

Pretrial Outcomes	Under 33 *n* = 3,734			33 or older *n* = 3,185		
	AUC (SE)	95% CI	Convention	AUC (SE)	95% CI	Convention
Any FTA	0.68 (0.01)	[0.65, 0.70]	Good	0.70 (0.01)	[0.67, 0.72]	Good
Any New Arrest	0.66 (0.01)	[0.64, 0.68]	Good	0.67 (0.01)	[0.64, 0.69]	Good
Any Arrest	0.69 (0.01)	[0.67, 0.70]	Good	0.68 (0.01)	[0.66, 0.70]	Good

Logistic Regression Models. Table 5 presents results of a series of unweighted logistic regression analyses of IRAS-PAT total scores and age modeling pretrial misconduct outcomes while controlling for county and time at risk. After conducting a main-effects only model, we examined whether age moderated the effect of total score on pretrial misconduct outcomes in Block 2. Together, the addition of these interactions did not contribute to a significant improvement in model fit over Block 1, *p*s ≥ .120. The age by total score interaction effects were not statistically significant in any of the pretrial misconduct models, *p*s ≥ .120. The findings suggest no evidence of differential prediction by age in IRAS-PAT assessments.

Table 5. Unweighted Logistic Regression Models of IRAS-PAT Total Scores and Age Predicting Pretrial Misconduct

Predictor	Unweighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Total Score	1.42	[1.36, 1.47]	<.001	1.35	[1.31, 1.40]	<.001	1.43	[1.39, 1.48]	<.001
Under 33 (33 or older)	1.01	[0.88, 1.18]	.847	1.01	[0.88, 1.15]	.921	0.93	[0.83, 1.05]	.236
Time at Risk	1.00	[1.00, 1.01]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Under 33 X Total Score	0.99	[0.92, 1.07]	.897	1.02	[0.95, 1.09]	.588	1.05	[0.99, 1.11]	.120
Δ-2LL		0.02 (1)			0.29 (1)			2.42 (1)	

Note. *N* = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. Δ-2LL reflects improvement in model fit upon addition of the interaction term(s) in Block 2. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio.

As shown in Table 6, we examined whether age moderated the effect of risk level on pretrial misconduct outcomes. Together, the addition of these interactions in Block 2 did not contribute to a significant improvement in model fit over the main-effects only model (i.e., Block 1), *p*s ≥ .318. The age by risk level interaction effects were not statistically significant in any of the pretrial misconduct models, *p*s ≥ .154. Similarly, the findings suggest no evidence of differential prediction by age in IRAS-PAT assessments.

Table 6. Unweighted Logistic Regression Models of IRAS-PAT Risk Levels and Age Predicting Pretrial Misconduct

Predictor	Unweighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Risk Level (Low)									
Moderate	2.81	[2.30, 3.45]	<.001	2.71	[2.28, 3.21]	<.001	3.00	[2.61, 3.46]	<.001
High	6.01	[4.78, 7.55]	<.001	4.83	[3.95, 5.91]	<.001	6.08	[5.10, 7.24]	<.001
Under 33 (33 or older)	1.04	[0.90, 1.21]	.561	1.03	[0.90, 1.17]	.663	0.97	[0.86, 1.08]	.554
Time at Risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Under 33 X Moderate	1.09	[0.73, 1.62]	.682	1.28	[0.91, 1.79]	.154	1.18	[0.89, 1.56]	.254
Under 33 X High	1.07	[0.69, 1.66]	.775	1.11	[0.75, 1.62]	.609	1.21	[0.87, 1.69]	.260
Δ-2LL		0.17 (2)			2.29 (2)			1.67 (2)	

Note. *N* = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. Δ-2LL reflects improvement in model fit upon addition of the interaction term(s) in Block 2. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio.

Summary of Age Findings

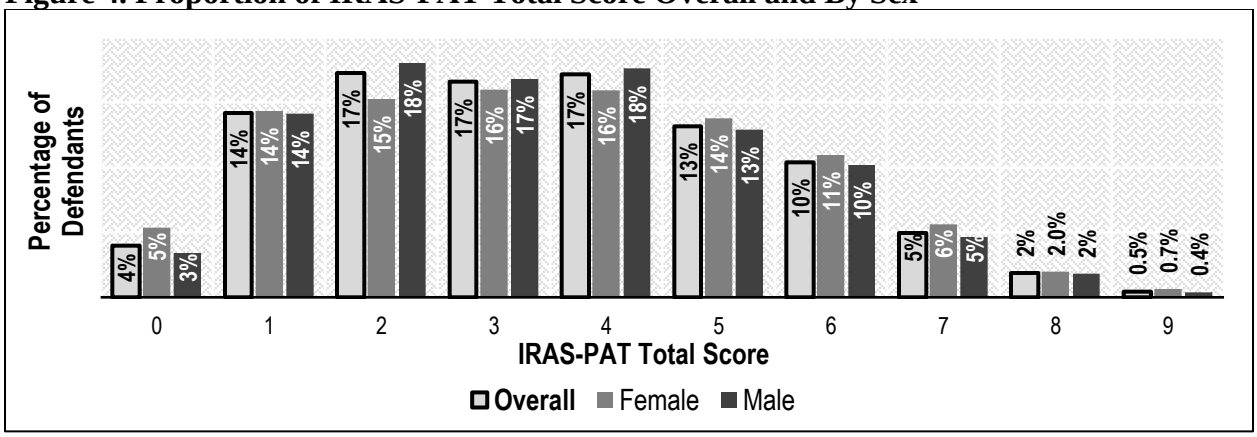
Together these results provide evidence that IRAS-PAT assessments produce similar predictive accuracy for defendants ages under 33 and 33 or older. There were no substantive differences in misconduct rates between both groups assessed at each risk level.

Differential Prediction of IRAS-PAT Assessments by Sex

Descriptives

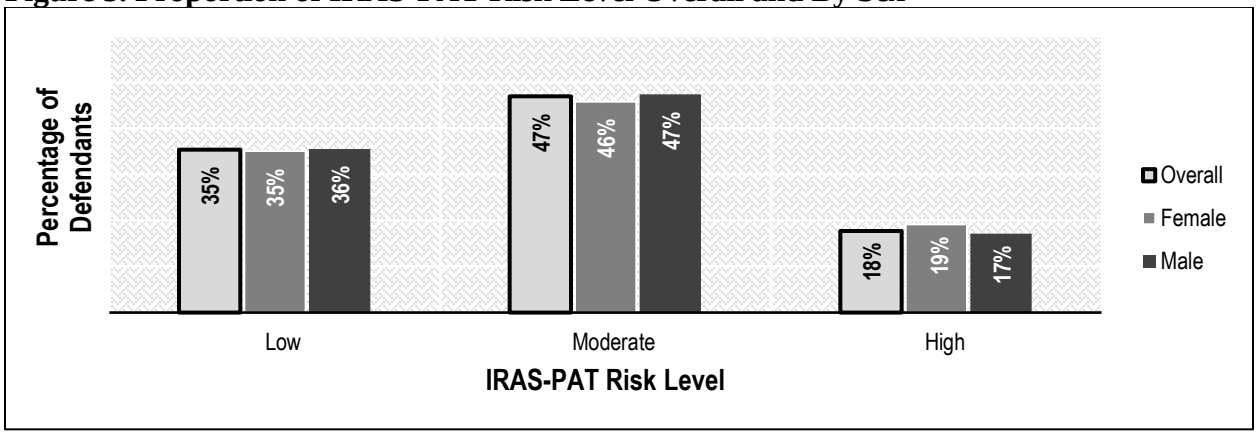
IRAS-PAT. IRAS-PAT scores averaged 3.53 (*SD* = 2.07, Range: 0 to 9) for female defendants and 3.48 (*SD* = 1.94, Range: 0 to 9) for male defendants. There was no significant difference between female and male defendants' IRAS-PAT total scores (*t*[3393.78] = -0.94, *p* = .348, Cohen's *d* = -0.03). The frequency distribution of IRAS-PAT scores by sex is presented in Figure 4.

Figure 4. Proportion of IRAS-PAT Total Score Overall and By Sex



As shown in Figure 5, female and male defendants had on average similar rates of being classified at IRAS-PAT risk levels.

Figure 5. Proportion of IRAS-PAT Risk Level Overall and By Sex

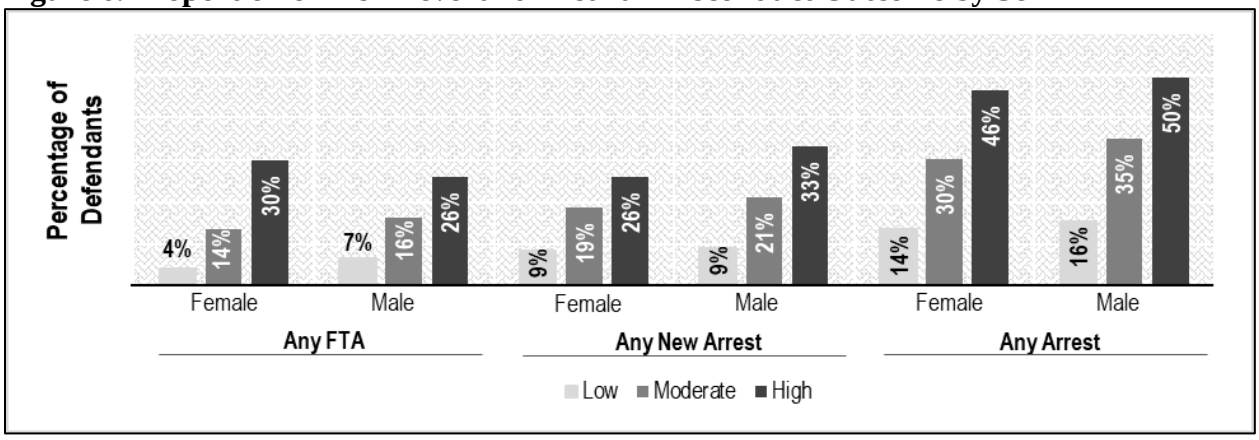


Pretrial Misconduct Outcomes. Female and male defendants diverged significantly from one another on arrest outcomes. Female defendants had lower rates of any arrest (*n* = 537, 27.5%) compared to male defendants (*n* = 1,512, 30.5%), $\chi^2(1) = 6.10$, *p* = .013, Cramer's *V* = -0.03. Similarly, female defendants were slightly less likely to have any new arrest (*n* = 325, 16.6%) relative to male defendants (*n* = 937, 18.9%). However, there was no significant differences between female and male defendants in rates of any FTA (*n* = 264, 13.5% and *n* = 723, 14.6%, respectively).

Frequency Distributions of Risk Level and Pretrial Misconduct Outcomes by Sex

Figure 6 presents the frequency distribution of IRAS-PAT risk level and pretrial outcomes stratified by sex. As shown, female and male defendants had, on average, similar rates of pretrial misconduct at each risk level, with some slight differences at High risk across the misconduct outcomes. See Appendix Table 2A for crosstabulations of risk level and pretrial misconduct outcomes by sex.

Figure 6. Proportion of Risk Level and Pretrial Misconduct Outcome by Sex



Differential Prediction Analyses

AUC of the ROC. In Table 7, we present AUC values and their associated conventions separately by sex and outcome. As shown, IRAS-PAT assessments produced similar predictive accuracy estimates across arrest outcomes for female and male defendants, with a significant difference in any FTA, $p = .014$. For any FTA, IRAS-PAT total scores produced good levels of predictive validity for male defendants and excellent levels for female defendants.

Table 7. AUC Values by Pretrial Misconduct Outcome and Sex

Pretrial Outcomes	Female *n* = 1,956			Male *n* = 4,963		
	AUC (SE)	95% CI	Convention	AUC (SE)	95% CI	Convention
Any FTA	0.72 (0.02)	[0.69, 0.75]	Excellent	0.67 (0.01)	[0.65, 0.69]	Good
Any New Arrest	0.66 (0.02)	[0.62, 0.69]	Good	0.67 (0.01)	[0.65, 0.68]	Good
Any Arrest	0.69 (0.01)	[0.68, 0.72]	Good	0.68 (0.01)	[0.66, 0.69]	Good

Logistic Regression Models. Table 8 presents results of a series of unweighted logistic regression analyses of IRAS-PAT total scores and sex modeling pretrial misconduct outcomes while controlling for county and time at risk. After conducting a main-effects only model, we examined whether sex moderated the effect of total score on pretrial misconduct outcomes in Block 2. Together, the addition of these interactions did not contribute to a significant improvement in model fit over Block 1 for the arrest outcomes, *p*s ≥ .122. Conversely, Block 2 had significant improvement in model fit over Block 1 for the any FTA model, $p = .049$. However, the sex by total score interaction effects were not statistically significant in any of the pretrial misconduct models, *p*s ≥ .051. The findings suggest no evidence of differential prediction by sex in IRAS-PAT assessments.

Table 8. Unweighted Logistic Regression Models of IRAS-PAT Total Scores and Sex Predicting Pretrial Misconduct

Predictor	Unweighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Total Score	1.42	[1.36, 1.48]	<.001	1.36	[1.31, 1.41]	<.001	1.43	[1.39, 1.48]	<.001
Female (Male)	0.87	[0.73, 1.02]	.093	0.78	[0.67, 0.90]	.001	0.76	[0.67, 0.86]	<.001
Time at Risk	1.00	[1.00, 1.01]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Female X Total Score	1.09	[1.00, 1.19]	.051	0.94	[0.88, 1.02]	.121	0.98	[0.92, 1.05]	.562
∆-2LL		3.87* (1)			2.39 (1)			0.34 (1)	

Note. *N* = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. ∆-2LL reflects improvement in model fit upon addition of the interaction term(s) in Block 2. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio. *p <.05.

As shown in Table 9, we examined whether sex moderated the effect of risk level on pretrial misconduct outcomes. In the FTA model, Block 2 had significant improvement in model fit over Block 1, $p = .042$. The sex by IRAS-PAT risk level interaction effect on any FTA was statistically significant, $p = .023$. Specifically, female defendants assessed at High risk had higher rates of predicted FTA (28%) relative to male defendants assessed at High risk (25%). In contrast, female defendants assessed at Low risk had slightly lower rates of predicted FTA (5%) compared to male defendants assessed at Low risk (7%). For the arrest outcomes, the addition of these interactions in Block 2 did not contribute to a significant improvement in model fit over the main-effects only model (i.e., Block 1), ps $\geq .218$. The sex by risk level interaction effects were not statistically significant in any of the arrest models, ps $\geq .109$. Overall, the findings suggest limited evidence of differential prediction by sex in IRAS-PAT assessments.

Table 9. Unweighted Logistic Regression Models of IRAS-PAT Risk Levels and Sex Predicting Pretrial Misconduct

Predictor	Unweighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Risk Level (Low)									
Moderate	2.83	[2.31, 3.46]	<.001	2.72	[2.29, 3.23]	<.001	3.00	[2.60, 3.46]	<.001
High	6.07	[4.83, 7.62]	<.001	4.90	[4.00, 5.99]	<.001	6.14	[5.15, 7.31]	<.001
Female (Male)	0.87	[0.74, 1.03]	.109	0.79	[0.68, 0.91]	.001	0.77	[0.68, 0.87]	<.001
Time at Risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Female X Moderate	1.26	[0.78, 2.05]	.349	0.91	[0.62, 1.33]	.624	0.88	[0.64, 1.20]	.414
Female X High	1.82	[1.09, 3.05]	.023	0.70	[0.46, 1.08]	.109	0.88	[0.61, 1.27]	.499
∆-2LL		6.33* (2)			3.05 (2)			0.72 (2)	

Note. N = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. ∆-2LL reflects improvement in model fit upon addition of the interaction term(s) in Block 2. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio. *p <.05.

Summary of Sex Findings

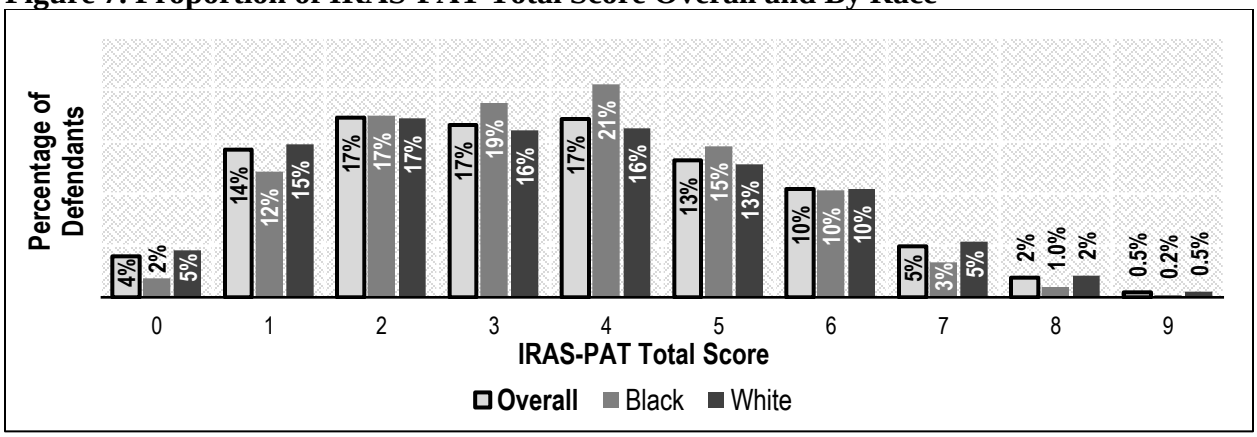
In conclusion, the findings suggest IRAS-PAT assessments produce similar predictive accuracy for female and male defendants. The only exception to this trend was in the prediction of any FTA, where IRAS-PAT assessments produced stronger predictive accuracy for female defendants relative to male defendants. This trend was driven primarily by female defendants classified at High risk, who had much higher rates of FTA relative to male defendants and relative to defendants classified at Low risk. Both any arrest and new arrest outcomes showed similar predictive accuracy for male and female defendants.

Differential Prediction of IRAS-PAT Assessments by Race

Descriptives

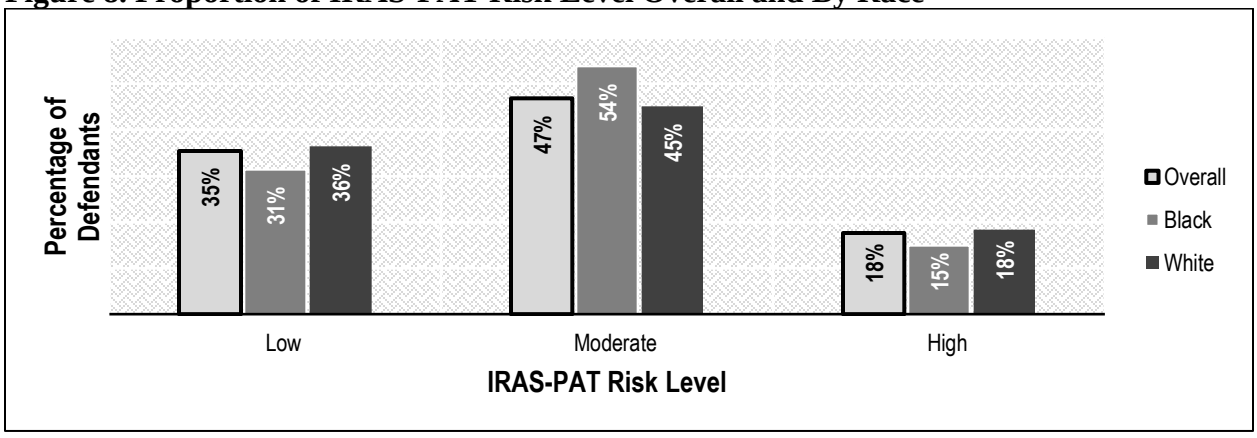
IRAS-PAT. IRAS-PAT scores averaged 3.52 (*SD* = 1.77, Range: 0 to 9) for Black defendants and 3.48 (*SD* = 2.03, Range: 0 to 9) for White defendants. There was no significant difference between Black and White defendants' IRAS-PAT total scores (t[2,445.74] = -0.906, p = .365, Cohen's d = -0.02). The frequency distribution of IRAS-PAT scores by race is presented in Figure 7.

Figure 7. Proportion of IRAS-PAT Total Score Overall and By Race



As shown in Figure 8, slightly higher proportions of White defendants were classified at Low and High risk (n = 2,009, 36.4% and n = 1,014, 18.4%, respectively) relative to Black defendants (n = 442, 31.4% and n = 209, 14.8%, respectively). Conversely, more Black defendants were classified at Moderate risk (n = 757, 53.8%) relative to White defendants (n = 2,488, 45.2%).

Figure 8. Proportion of IRAS-PAT Risk Level Overall and By Race

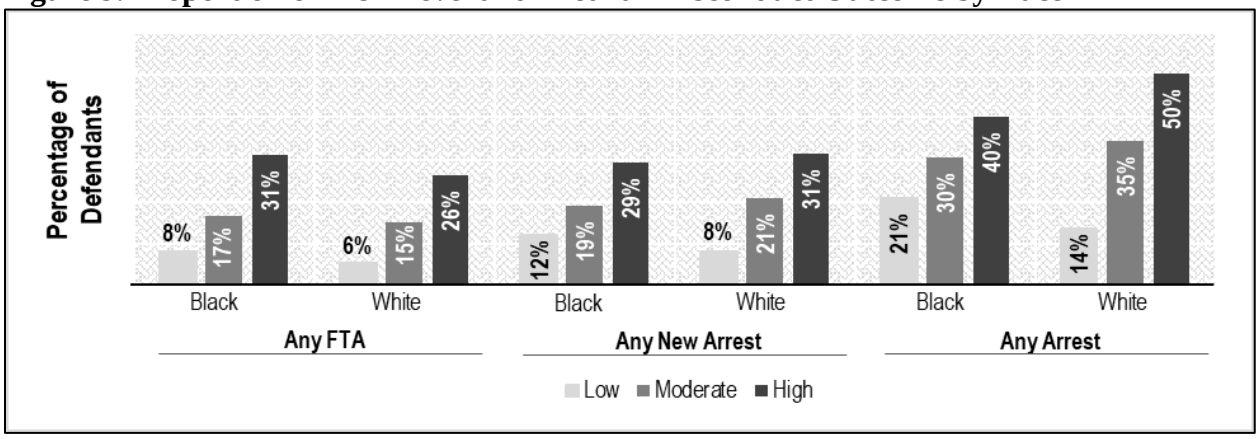


Pretrial Misconduct Outcomes. Black defendants had higher rates of any FTA (n = 227, 16.1%) relative to White defendants (n = 760, 13.8%), $\chi^2(1) = 4.98$, $p = .026$, Cramer's $V = 0.03$. There were no significant differences between Black and White defendants in rates of any arrest (n = 406, 28.8% and n = 1,643, 29.8%, respectively) and any new arrest (n = 260, 18.5% and n = 1,002, 18.2%, respectively).

Frequency Distributions of Risk Level and Pretrial Misconduct Outcomes by Race

Figure 9 presents the distribution of IRAS-PAT risk level and pretrial outcomes by race. Rates of pretrial misconduct were higher for Black defendants assessed at Low risk relative to White defendants, particularly for any arrest. At Moderate risk, Black defendants had slightly higher rates of any FTA compared to White defendants. Conversely, White defendants had higher rates of arrest outcomes relative to Black defendants. At High risk, Black defendants had higher rates of any FTA relative to White defendants. However, Black defendants had lower rates of misconduct across the arrest outcomes compared to White defendants. See Appendix Table 3A for crosstabulations of risk level and pretrial misconduct outcomes by race.

Figure 9. Proportion of Risk Level and Pretrial Misconduct Outcome by Race



Differential Prediction Analyses

AUC of the ROC. In Table 10, we present AUC values and their associated conventions separately by race and outcome. As shown, IRAS-PAT assessments produced weaker predictive accuracy estimates for Black defendants relative to White defendants, particularly for the arrest outcomes, ps $\leq .002$. There were no significant differences in AUC estimates for any FTA, p = .219.

Table 10. AUC Values by Pretrial Misconduct Outcome and Race

Pretrial Outcomes	Black n = 1,408			White n = 5,511		
	AUC (SE)	95% CI	Convention	AUC (SE)	95% CI	Convention
Any FTA	0.66 (0.02)	[0.63, 0.70]	Good	0.69 (0.01)	[0.67, 0.71]	Good
Any New Arrest	0.61 (0.02)	[0.58, 0.65]	Fair	0.67 (0.01)	[0.66, 0.69]	Good
Any Arrest	0.60 (0.02)	[0.57, 0.63]	Fair	0.70 (0.01)	[0.68, 0.71]	Good

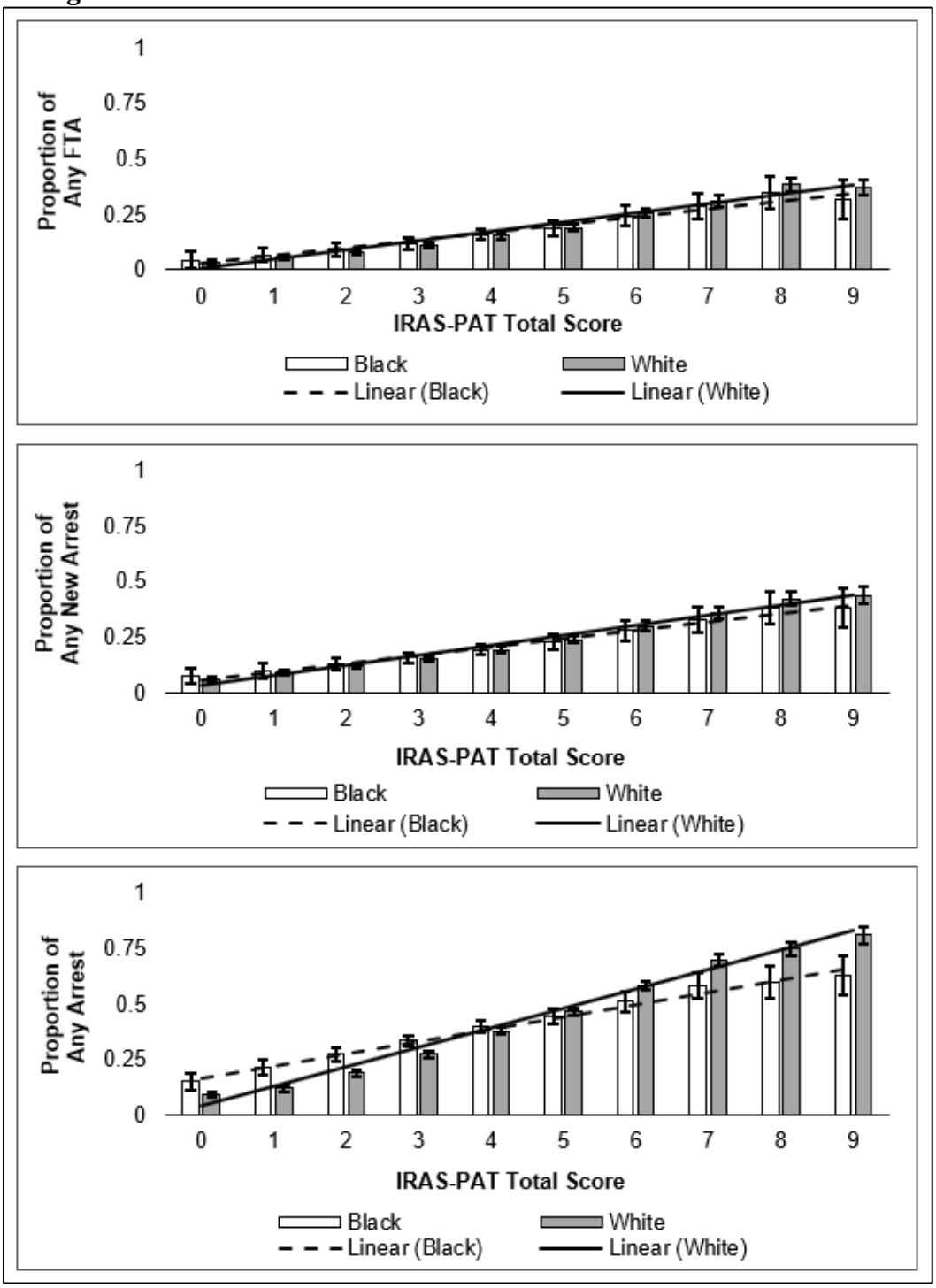
Logistic Regression Models. Table 11 presents results of a series of unweighted logistic regression analyses of IRAS-PAT total scores and race modeling pretrial misconduct outcomes while controlling for county and time at risk. The results from the main-effects only model (i.e., Block 1) showed strong predictive validity of IRAS-PAT assessments across pretrial misconduct outcomes. In Block 2, we examined whether race moderated the effect of total score on pretrial misconduct outcomes. Together, the addition of these interactions contributed to a significant improvement in model fit over Block 1 in only the any arrest model, $p = .005$. While race by total score was not a statistically significant term in the any FTA and any new arrest models, $ps \geq .216$, we observed significant race by total score interaction effects in the any arrest model, $p = .004$. This finding suggests some evidence of differential prediction by race in IRAS-PAT assessments. These results are displayed in Figure 10, graphically depicting the differences in predicted pretrial misconduct for the same risk estimate across Black and White defendants. Black defendants assessed at higher risk had lower rates of any arrest relative to White defendants. In contrast, Black defendants assessed at lower risk had higher rates of any arrest relative to White defendants. Visually, the trend lines are not parallel in the any arrest figure (compared to the parallel trend lines seen in the figures for any FTA and any new arrest), suggesting that the incremental gain in rate of any arrest from Low risk to High risk differs between Black and White defendants. In the weighted analyses, the race by total score interaction term in the any arrest model was no longer significant. See Appendix Table 4A for weighted models.

Table 11. Unweighted Logistic Regression Models of IRAS-PAT Total Scores and Race Predicting Pretrial Misconduct

Predictor	Unweighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Total Score	1.42	[1.36, 1.48]	<.001	1.35	[1.31, 1.40]	<.001	1.43	[1.38, 1.47]	<.001
Black (White)	1.05	[0.87, 1.26]	.645	1.00	[0.85, 1.19]	.981	1.12	[0.97, 1.30]	.129
Time at Risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Black X Total Score	0.94	[0.85, 1.04]	.216	0.95	[0.87, 1.04]	.265	0.89	[0.83, 0.96]	.004
∆-2LL		1.52 (1)			1.24 (1)			8.08** (1)	

Note. *N* = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. ∆-2LL reflects improvement in model fit upon addition of the interaction term(s) in Block 2. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio. **p <.01. (two-tailed)

Figure 10. Predicted Probabilities of Pretrial Misconduct by IRAS-PAT Total Scores and Race, Unweighted



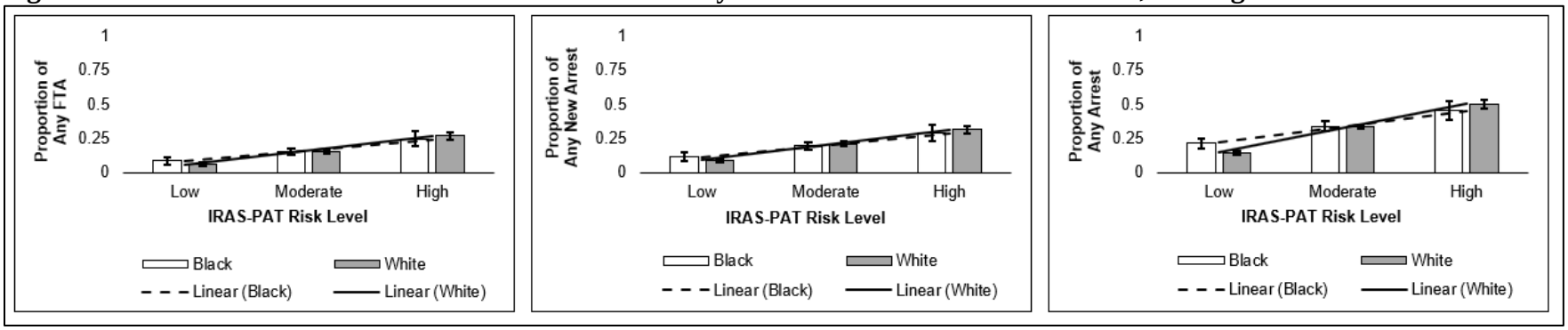
Next, we examined differential prediction of IRAS-PAT risk levels by race. The results from the main-effects only model (i.e., Block 1) showed that IRAS-PAT risk levels had a similar ability to predict pretrial misconduct outcomes (Table 12). In Block 2, we examined whether race moderated the effect of risk level on pretrial misconduct outcomes. For Moderate and High risk levels, we observed significant race by risk level interaction effects in the any new arrest model ($ps \leq .047$) and any arrest model ($ps \leq .002$). These findings suggest some evidence of differential prediction by race in IRAS-PAT assessments. However, the addition of these interactions contributed to a significant improvement in model fit over Block 1 in the any arrest model only, $p = .001$. As shown in Figure 11, Black defendants classified at High risk had lower predicted rates of any arrest relative to White defendants. In contrast, Black defendants classified at Low risk had higher predicted rates of any arrest relative to White defendants. The trend lines in the any arrest figure are not parallel, which suggests the risk scores produce weaker predictions of any arrest risk for Black over White defendants. In the weighted analyses, we found no substantive changes in the arrest models. Moreover, the race by High risk level was a statistically significant term in the any FTA model. See Appendix Table 5A for weighted models.

Table 12. Unweighted Logistic Regression Models of IRAS-PAT Risk Levels and Race Predicting Pretrial Misconduct

Predictor	Unweighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Risk Level (Low)									
Moderate	2.82	[2.30, 3.45]	<.001	2.72	[2.29, 3.23]	<.001	2.98	[2.59, 3.43]	<.001
High	6.03	[4.80, 7.57]	<.001	4.85	[3.96, 5.93]	<.001	6.05	[5.08, 7.21]	<.001
Black (White)	1.04	[0.86, 1.25]	.710	0.99	[0.84, 1.17]	.891	1.11	[0.96, 1.29]	.152
Time at Risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Black X Moderate	0.67	[0.42, 1.07]	.097	0.63	[0.42, 0.93]	.020	0.59	[0.42, 0.82]	.002
Black X High	0.60	[0.35, 1.03]	.064	0.62	[0.38, 0.99]	.047	0.48	[0.31, 0.73]	.001
∆-2LL		3.66 (2)			5.79 (2)			13.77** (2)	

Note. *N* = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. ∆-2LL reflects improvement in model fit upon addition of the interaction term(s) in Block 2. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio. **p <.01. (two-tailed)

Figure 11. Predicted Probabilities of Pretrial Misconduct by IRAS-PAT Risk Levels and Race, Unweighted



Together, these results provide evidence that IRAS-PAT assessments produce differential predictive accuracy and over-classification at High risk level for Black defendants, particularly for arrest outcomes. However, these results do not tell us why IRAS-PAT assessments may be producing different levels of predictive accuracy between Black and White defendants, including whether specific IRAS-PAT items may be driving trends. As such, we additionally examined differences in the predictive accuracy of IRAS-PAT items by race.

Item-Level Analyses

To further examine differential prediction of IRAS-PAT assessments by race, we conducted an item-level analysis of the IRAS-PAT. This involved first conducting logistic regression models of IRAS-PAT items predicting pretrial misconduct outcomes separately for Black and White defendants. These models are presented in Table 10-12. Significant odds ratios are shaded. Due to missing item-level data, the analytic sample in each model deviates from the original distribution of Black (n = 1,408) and White (n = 5,511) defendants. For White defendants, we explored item-level functioning in three separate samples. The first sample was unweighted, representing all White defendants in the original sample who had item-level data. Second, to address concerns about baseline differences between Black and White defendants, we matched White defendants to Black defendants using propensity score weighting. Third, to address potential differences as a function of sample size, which was considerably larger for White defendants in the unweighted sample, we conducted analyses using a stratified sub-sample of White defendants. These analytic decisions are discussed below.

Summary of Item-Level Analytic Considerations	
Regression Without the Use of Propensity Scores (Unweighted)	Multivariable models unweighted by propensity scores, which do not address concerns about unbalanced groups, are helpful because they provide a baseline model for comparison to other models that integrate estimation and sampling considerations.
Using Propensity Scores as Weights in a Regression	Multivariable models weighted by propensity scores statistically balance individuals on a specific set of covariates based on their likelihood of being in a group. Black defendants were assigned a propensity score of 1, whereas White defendants were assigned a score above or below 1 that was calculated in the propensity score analysis based on the observed covariates. We then weighted estimations using these propensity scores.
Regression with a Stratified Sample	To balance the racial groups based on sample size, we developed a stratified sample of White defendants. We randomly selected White defendants within each county to create a sample of White defendants that was equal to the proportion of Black defendants within the respective county. This resulted in a stratified sample of 1,408 White defendants in addition to the original 1,408 Black defendants. Due to missing item-level data across racial groups, the final stratified sample used in the analyses consisted of 1,295 White defendants and 1,287 Black defendants.

Next, we present item-level predictive validity for Black defendants and White defendants based on the three distinct samples (i.e., unweighted, weighted, and stratified). These comparisons are provided in Table 13-15 for each of the three pretrial misconduct outcomes. Non-significant items are listed as "ns." Item-level predictors that are significant predictors for each group are highlighted in gray. We additionally present the odds ratio associated with each significant item-level predictor. Importantly, odds ratios allow for within-model comparison of which items have the most unique explanatory power for pretrial misconduct outcomes; however, they are not readily comparable across models.

As shown in Table 13, two or more FTAs in the past 24 months (Item 2, OR = 2.14), three or more prior incarcerations (Item 3, OR = 1.73), and unemployment (Item 4, OR = 1.75) were the only unique predictors of any FTA for Black defendants. For White defendants in the unweighted model, most items were significant predictors of any FTA. Specifically, age at first arrest (Item 1, OR = 1.76), a history of FTAs in the past 24 months (Item 2, ORs = 2.18), three or more prior incarcerations (Item 3, OR = 1.58), unemployment (Item 4, OR = 2.08), residential instability (Item 5, OR = 1.28), illegal drug use in the past 6 months (Item 6, OR = 1.39), and severe drug use problem (Item 7, OR = 1.53) contributed uniquely to the prediction of any FTA for White defendants. Part-time employment (Item 4) was not a significant predictor of any FTA. As shown, following application of weights and stratification, most items remained significant predictors of any FTA for White defendants, except for age at first arrest (Item 1) and severe drug use problem (Item 7). Across models, number of FTAs (Item 2) and unemployment (Item 4) were the strongest predictors among White and Black defendants. See Appendix Table 6A for complete models.

Table 13. Summary of Logistic Regression Models of IRAS-PAT Items Predicting Any FTA, by Race

Predictor	Odds Ratio Status for Any FTA			
	Black	White		
	n = 1,287	Unweighted *n* = 4,928	Weighted *n* = 4,928	Stratified *n* = 1,295
Item 1 - Age at first arrest	ns	**1.76**	ns	ns
Item 2 - Number of FTAs				
1	ns	**2.18**	**2.74**	**2.39**
2 or more	**2.14**	**2.18**	**2.45**	**2.59**
Item 3 - 3+ prior incarcerations	**1.73**	**1.58**	**1.44**	**1.78**
Item 4 - Employed				
Part-time	ns	ns	ns	ns
Not employed	**1.75**	**2.08**	**2.34**	**2.96**
Item 5 - Residential instability	ns	**1.28**	**1.40**	**1.48**
Item 6 - Illegal drug use 6 months	ns	**1.39**	**1.76**	**1.89**
Item 7 - Severe drug use	ns	**1.53**	ns	ns

Table 14 presents item-level results by sample for the prediction of any new arrest. As shown in Table 14, two or more FTAs in the past 24 months (Item 2, OR = 3.33), unemployment (Item 4, OR = 1.40), and illegal drug use in the past 6 months (Item 6, OR = 1.49) uniquely predicted any new arrest for Black defendants. Age at first arrest (Item 1), one prior FTA (Item 2), three or more prior incarcerations (Item 3), part-time employment (Item 4), and residential instability (Item 5) did not contribute uniquely to the prediction of any new arrest for Black defendants. For White defendants in the unweighted model, one prior FTA (Item 2, OR = 1.29), three or more prior incarcerations (Item 3, OR = 1.71), part-time employment and unemployment (Item 4, OR Range: 1.31 to 1.76), residential instability (Item 5, OR = 1.40), illegal drug use in the past 6 months (Item 6, OR = 1.41), and severe drug use problem (Item 7, OR = 1.44) were unique predictors of any new arrest. Age at first arrest (Item 1) and two or more FTAs in the past 24 months (Item 2) were not significant predictors of any new arrest for White defendants. As shown, following application of weights and stratification, item-level predictions were much more comparable to those of Black defendants. Specifically, four items remained significant predictors of any new arrest for White defendants: history of FTAs (Item 2), three or more prior incarcerations (Item 3), unemployment (Item 4), and illegal drug use in the past 6 months (Item 6). See Appendix Table 7A for complete models.

Table 14. Summary of Logistic Regression Models of IRAS-PAT Items Predicting Any New Arrest, by Race

Predictor	Odds Ratio Status for Any New Arrest			
	Black	White		
	n = 1,284	Unweighted *n* = 4,928	Weighted *n* = 4,928	Stratified *n* = 1,295
Item 1 – Age at first arrest	ns	ns	ns	ns
Item 2 - Number of FTAs				
1	ns	**1.29**	**1.83**	ns
2 or more	**3.33**	ns	**2.16**	ns
Item 3 - 3+ prior incarcerations	ns	**1.71**	ns	**1.60**
Item 4 - Employed				
Part-time	ns	**1.31**	ns	ns
Not employed	**1.40**	**1.76**	**2.36**	**1.88**
Item 5 - Residential instability	ns	**1.40**	ns	ns
Item 6 - Illegal drug use 6 months	**1.49**	**1.41**	**1.38**	**1.69**
Item 7 - Severe drug use	ns	**1.44**	ns	ns

Table 15 presents item-level results by sample for the prediction of any arrest. For Black defendants, two or more FTAs in the past 24 months (Item 2, OR = 2.26), three or more prior incarcerations (Item 3, OR = 1.40), unemployment (Item 4, OR = 1.63), and a severe drug use problem (Item 7, OR = 1.88) were the strongest unique predictors of any arrest (Table 15). Age at first arrest (Item 1), one prior FTA (Item 2), part-time employment (Item 4), residential instability (Item 5), and illegal drug use in the past 6 months (Item 6) did not contribute uniquely to the prediction of any arrest for Black defendants. For White defendants in the unweighted model, age at first arrest (Item 1, OR = 1.51), a history of FTAs in the past 24 months (Item 2, OR Range: 1.63 to 1.72) three or more prior incarcerations (Item 3, OR = 2.06), employment status (Item 4, OR Range: 1.43 to 2.04), residential instability (Item 5, OR = 1.43), illegal drug use in the past 6 months (Item 6, OR = 1.33), and severe drug use problem (Item 7, OR = 1.49) contributed uniquely to the prediction of any arrest. As shown, following application of weights and stratification, five items remained significant predictors of any arrest for White defendants: age at first arrest (Item 1), history of FTAs (Item 2), three or more prior incarcerations (Item 3), unemployment (Item 4), and illegal drug use in the past 6 months (Item 6). See Appendix Table 8A for complete models.

Table 15. Summary of Logistic Regression Models of IRAS-PAT Items Predicting Any Arrest, by Race

Predictor	Odds Ratio Status for Any Arrest			
	Black	White		
	n = 1,287	Unweighted *n* = 4,928	Weighted *n* = 4,928	Stratified *n* = 1,295
Item 1 – Age at first arrest	ns	**1.51**	ns	**1.73**
Item 2 - Number of FTAs				
1	ns	**1.72**	**1.82**	**1.47**
2 or more	**2.26**	**1.63**	**2.33**	ns
Item 3 - 3+ prior incarcerations	**1.40**	**2.06**	**1.33**	**1.95**
Item 4 - Employed				
Part-time	ns	**1.43**	ns	ns
Not employed	**1.63**	**2.04**	**2.42**	**2.02**
Item 5 - Residential instability	ns	**1.43**	ns	ns
Item 6 - Illegal drug use 6 months	ns	**1.33**	**1.36**	**1.82**
Item 7 - Severe drug use	**1.88**	**1.49**	ns	ns

Based on the evidence that some items functioned differently by race, we examined whether race moderated the effect of IRAS-PAT items on pretrial misconduct outcomes in combined models. For these comparisons, we reported on the sample of White defendants that was weighted (i.e., matched) to the Black defendant sample based on a number of legal and extralegal characteristics. Table 16 presents results of hierarchical weighted logistic regression models of IRAS-PAT items for the sample overall, adding race as a covariate (Block 1) and item interactions with race (Block 2).

Across all pretrial misconduct outcomes, four items uniquely contributed to the prediction of all three outcomes: history of FTAs (Item 2, OR Range: 1.70 to 2.38), three or more prior incarcerations (Item 3, OR Range: 1.29 to 1.49), unemployment (Item 4, OR Range: 2.11 to 2.22), and illegal drug use in the past 6 months (Item 6, OR Range: 1.29 to 1.61).

In Block 2, we examined the interaction effects of IRAS-PAT items on pretrial misconduct outcomes. While race by IRAS-PAT items were not statistically significant interaction terms in the any FTA model, *p*s ≥ .114, we observed evidence of differential prediction by race for rearrest outcomes. For any new arrest, race by unemployment (Item 4, $p = .020$) was a significant interaction effect. Black defendants who were unemployed had lower predicted rates of any new arrest relative to White defendants (Figure 12).

Similarly, for any arrest, race by unemployment (Item 4, $p = .036$) emerged as significant interaction effect. Black defendants who were unemployed had lower predicted rates of any arrest relative to White defendants (Figure 12). Overall, results provide some evidence that items, specifically Item 4, may be contributing to differences in predictive accuracy for rearrest outcomes between Black and White defendants. See Appendix Table 9A for unweighted models.

Figure 12. Predicted Probabilities of Rearrest Outcomes by IRAS-PAT Items and Race, Weighted

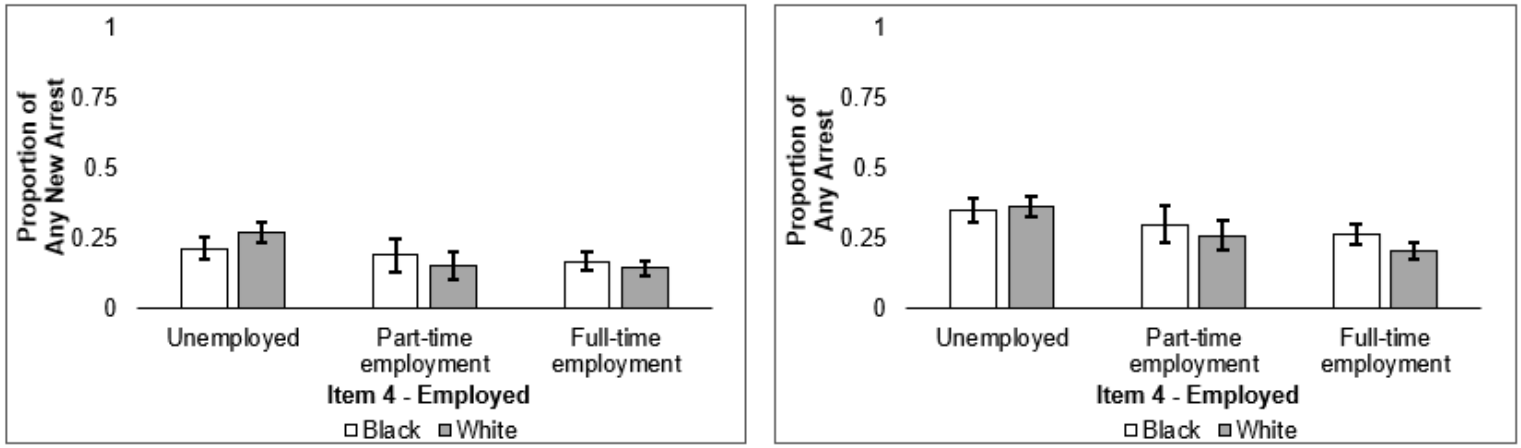


Table 16. Weighted Logistic Regression Models of Race and IRAS-PAT Items Predicting Pretrial Misconduct

Predictor	Weighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Black[a]	0.95	[0.75, 1.21]	.695	0.94	[0.76, 1.17]	.576	1.11	[0.92, 1.34]	.285
Item 1 - Age at first arrest[b]	1.31	[0.76, 2.26]	.330	1.52	[0.90, 2.56]	.116	1.40	[0.90, 2.18]	.141
Item 2 - Number of FTAs[c]									
1	2.34	[1.64, 3.35]	<.001	1.71	[1.21, 2.41]	.002	1.70	[1.26, 2.31]	.001
2 or more	2.40	[1.56, 3.68]	<.001	2.38	[1.57, 3.60]	<.001	2.35	[1.60, 3.47]	<.001
Item 3 - 3+ prior incarcerations[d]	1.49	[1.15, 1.94]	.003	1.29	[1.01, 1.65]	.040	1.35	[1.09, 1.68]	.007
Item 4 - Employed[e]									
Part-time	0.97	[0.63, 1.48]	.889	1.11	[0.76, 1.61]	.580	1.36	[1.00, 1.85]	.046
Not employed	2.22	[1.67, 2.94]	<.001	2.11	[1.63, 2.72]	<.001	2.20	[1.75, 2.77]	<.001
Item 5 - Residential instability[f]	1.36	[1.05, 1.75]	.021	1.21	[0.95, 1.53]	.115	1.25	[1.01, 1.55]	.040
Item 6 - Illegal drug use 6 mo[g]	1.61	[1.21, 2.14]	.001	1.39	[1.08, 1.78]	.011	1.29	[1.03, 1.61]	.025
Item 7 - Severe drug use[h]	1.31	[0.95, 1.82]	.104	1.28	[0.94, 1.75]	.120	1.39	[1.04, 1.86]	.026
Time at risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Black X Item 1	0.70	[0.24, 2.04]	.511	1.45	[0.51, 4.13]	.481	0.85	[0.39, 1.88]	.695
Black X Item 2 – 1 FTA	0.61	[0.33, 1.13]	.114	0.71	[0.40, 1.27]	.247	0.85	[0.51, 1.42]	.539
Black X Item 2 – 2 or more FTAs	0.95	[0.46, 1.98]	.898	1.31	[0.66, 2.61]	.441	1.02	[0.53, 1.98]	.956
Black X Item 3	1.17	[0.74, 1.86]	.495	0.88	[0.58, 1.35]	.566	1.01	[0.69, 1.48]	.955
Black X Item 4 – Part-time	0.87	[0.41, 1.88]	.732	1.11	[0.56, 2.17]	.769	0.88	[0.51, 1.51]	.634
Black X Item 4 – Not Employed	0.75	[0.46, 1.23]	.257	0.59	[0.38, 0.92]	.020	0.65	[0.44, 0.97]	.036
Black X Item 5	0.87	[0.55, 1.37]	.536	1.10	[0.72, 1.67]	.664	0.90	[0.62, 1.31]	.574
Black X Item 6	0.69	[0.43, 1.13]	.139	1.05	[0.67, 1.63]	.845	0.84	[0.57, 1.24]	.381
Black X Item 7	0.94	[0.51, 1.73]	.834	1.05	[0.60, 1.84]	.869	1.46	[0.88, 2.43]	.142

Note. *N* = 6,220. Block 1 controlled for county, but estimates are not shown. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio. [a]Reference: White. [b]Item 1 reference: 33 or older. [c]Item 2 reference: No FTA warrants past 24 months. [d]Item 3 reference: Two or less prior jail incarcerations. [e]Item 4 reference: Yes, Full-time employment at time of arrest. [f]Item 5 reference: Lived at current residence past 6 months. [g]Item 6 reference: No illegal drug use during past 6 months. [h]Item 7 reference: No severe drug use problem.

Summary of Race Findings

Overall, several findings emerged from the race-specific analyses:

- Relative to our previous 5-county investigation (Lowder, Lawson, & Foudray, 2020), we found *less* evidence that IRAS-PAT assessments predicted pretrial misconduct outcomes differentially by race. In fact, there was no evidence of differential prediction for any FTA.
- For total scores, differences in predictive accuracy were most notable for any arrest. When we matched defendants on a variety of legal and extralegal factors, this difference was no longer significant.
- Differences in predictive accuracy were most notable for risk estimates, particularly for arrest outcomes (any arrest and any new arrest). These differences persisted after matching defendants on legal and extralegal factors.
- Item-level findings showed fewer differences in predictive accuracy between Black and White defendants after matching defendants on legal and extralegal characteristics. Some items emerged as unique predictors for White defendants, but not Black defendants, including age at first arrest (Item 1), 3+ prior incarcerations (Item 3), and illegal drug use in the past 6 months (Item 6).
- However, item-level differences largely did not translate into varying levels of item-level predictive accuracy in combined models. In fact, only one item showed evidence of differential prediction by race: unemployment (Item 4) relative to full-time employment.

CONCLUSION

The purpose of this investigation was to examine the predictive accuracy of IRAS-PAT assessments by age, sex, and race in a pooled dataset of 12 Indiana Counties. Overall, our findings provided limited evidence of predictive bias in IRAS-PAT assessments as a function of age or sex. However, there were some differences in predictive accuracy as a function of race, particularly for arrest outcomes. Below we summarize and discuss these findings in greater detail.

Overall Predictive Accuracy

Age

Overall, findings showed little evidence of differences in predictive validity as a function of age (i.e., under 33 vs. 33 and older). There were minimal differences in AUC estimates (i.e., < 0.02 difference) and no evidence of statistically significant differences in predictive accuracy across all analyses. These findings are aligned with our previous 5-county results (Lowder, Lawson, & Foudray, 2020). However, we are not aware of any prior systematic investigation examining predictive bias in pretrial risk assessments as a function of age.

Sex

Similar to age, we found limited evidence suggesting predictive bias in IRAS-PAT assessments between male and female defendants. There was a slight difference in the predictive accuracy of any FTA by sex, with IRAS-PAT assessments producing stronger assessments of any FTA for female defendants relative to male defendants. This effect was driven primarily by the high proportion of female defendants classified at High risk who had an FTA during the pretrial period. Overall, our findings were largely consistent with our previous investigation and with existing research on the predictive validity of pretrial risk assessments as a function of sex (Desmarais et al., 2020).

Race

We found that IRAS-PAT assessments produced weaker assessments of any arrest and any new arrest risk for Black defendants relative to White defendants. However, overall, we found weaker evidence of differential predictive accuracy by race compared to our previous 5-county investigation (Lowder, Lawson, & Foudray, 2020). Importantly, IRAS-PAT assessments showed no evidence of predictive bias for any FTA between White and Black defendants, in contrast to our previous findings (Lowder, Lawson, & Foudray, 2020). Overall, IRAS-PAT assessments produced fair to good levels of predictive accuracy for Black defendants and good levels of predictive accuracy for White defendants. For example, whereas total scores produced AUC estimates ranging from 0.60-0.66 for Black defendants, AUC estimates for White defendants ranged from 0.67-0.70. Results of multivariable models suggested that differences were most notable for IRAS-PAT risk estimates versus total scores, where High and Moderate risk Black defendants had lower relative rates of any arrest and any new arrest relative to White defendants and those assessed at Low risk.

To explore whether IRAS-PAT items would help explain differences in predictive accuracy for arrest outcomes between White and Black defendants, we conducted several item-level analyses. After accounting for the oversampling of White defendants and potential between-group differences, unique item-level predictors of any new arrest were largely similar (i.e., Item 2 – 2+ prior FTA; Item 4 – Unemployed; and Item 6 – Illegal drug use in past 6 months), with one exception. Item 3 (3+ prior incarcerations) remained a significant predictor of any arrest for White defendants but not Black defendants. In the prediction of any arrest, unique item-level predictors again were largely similar across groups after adjustments to the White sample, with two exceptions. Item 7 (severe drug use) was a unique predictor for Black defendants, but not White defendants. Item 1 (age at first arrest) was a unique predictor for White defendants, but not Black defendants. When we examined whether the strength of item-level predictions differed across racial groups, only one item showed evidence of a differential association with arrest outcomes. Item 4 (unemployment), although a unique predictor for both groups, predicted arrest outcomes to a weaker degree among Black defendants compared to White defendants.

Differences in item-level predictors across racial groups could reflect several possibilities. Age at first arrest and a prior incarceration history were more consistent predictors among White defendants, suggesting these items may have less predictive utility among Black defendants. From prior research, we know that Black individuals, particularly youth, are arrested at younger ages (Lau et al., 2018) and are disproportionately incarcerated relative to White individuals (Abrams et al., 2012; Bales & Piquero, 2012; Gelman et al., 2007; Kutateladze et al., 2014), even after controlling for a variety of legal and extralegal characteristics. Together, these findings suggest that a prior criminal history may be more indicative of disproportionate likelihood of being involved in the criminal-legal system rather than actual risk for repeated misconduct.

Unemployment, however, was the only factor that showed differential item-level predictive validity for both arrest outcomes. Specifically, employment was a stronger unique predictor of arrest outcomes for White defendants relative to Black defendants. Employment status is commonly measured on pretrial risk assessment tools (Desmarais et al., 2020). Yet, prior research suggests several reasons why employment outcomes may differ across racial groups. Primarily, history of justice involvement has been seen to have more detrimental effects on post-release employment outcomes for Black individuals relative to White individuals (Decker et al., 2015), decreasing the likelihood that Black individuals who are justice-involved will have full-time employment. Further, there is some evidence that the effect of employment on recidivism for Black individuals is influenced by contextual factors (e.g., community-level unemployment rates; Wang et al., 2010), providing some evidence that unemployment may be less indicative of individual risk for recidivism among Black individuals.

More broadly, it is important to note that although these findings suggest evidence of predictive bias, they do not provide evidence that IRAS-PAT assessments are having a disparate impact on Black defendants relative to White defendants. To the contrary, recent findings from Indiana suggest that the use of IRAS-PAT assessments had a similar impact on pretrial decision-making for Black and White defendants relative to decision-making as usual (Lowder, Grommon, & Ray, 2020). Overall, IRAS-PAT assessments are predicting pretrial misconduct with good accuracy for its pretrial population as a whole (Lowder, Lawson, Grommon, et al., 2020). Scholars have noted that there are inherent tradeoffs to the fairness and accuracy of risk

assessments, particularly when there is evidence of different rates of misconduct among different pretrial populations (Berk et al., 2021; Kleinberg et al., 2016). Achieving accurate assessments of risk can come at a cost to fairness, and vise-versa. Jurisdictions must decide how to weigh these considerations by prioritizing public safety at the cost of fairness or achieving fairness at the cost of compromising public safety (Corbett-Davies et al., 2017).

Limitations and Future Directions

These findings are based on a 12-county sample of Indiana pretrial defendants. The sample includes all 11 counties that participated in Indiana's Pretrial Pilot Project, with one additional county (Vigo). However, the sample is not fully representative of Indiana's pretrial population. Similar to our 5-county report (Lowder, Lawson, & Foudray, 2020), Black individuals were underrepresented in the sample. However, after applying several sampling corrections, we still observed evidence of differential prediction between Black and White defendants. Beyond IRAS-PAT items, we did not examine whether other characteristics of defendants (e.g., offense type) or their case processing (e.g., time until filing; time until disposition) might explain different levels of predictive validity. We also had limited information on dynamic risk factors (i.e., factors that are changeable) that may be more sensitive to actual changes in risk among defendants regardless of racial identification. Dynamic factors may also be less reflective of the relative and structural disadvantages experienced by Black defendants, compared to static items such as criminal history.

It remains to be seen whether incorporation of other items to discount IRAS-PAT scores could improve the predictive performance of IRAS-PAT assessments for Black defendants. As it stands, the tool measures only 7 items, most of which are fairly static in nature and place heavy emphasis on criminal history. One challenge is that such considerations would need to be implemented for all defendants and irrespective of race. Measures that could predict outcomes more accurately regardless of race or measures that could account for broader social disadvantage may provide two possible avenues for improved predictive performance of the IRAS-PAT. However, as mentioned previously, IRAS-PAT assessments predict pretrial outcomes well overall for defendants broadly. Any modifications to the tool or scoring criteria to improve fairness could have implications for predictive accuracy more broadly.

References

Abrams, D. S., Bertrand, M., & Mullainathan, S. (2012). Do judges vary in their treatment of race? *The Journal of Legal Studies*, *41*(2), 347–383. https://doi.org/10.1086/666006

Baglivia, M., Wolff, K. T., & Jackowski, K. (2019). The usefulness of a general risk assessment, the Static Risk Assessment (SRA), in predicting pretrial failure: Examining predictive ability across gender and race. *Justice Evaluation Journal*, 1–26.

Bales, W. D., & Piquero, A. R. (2012). Racial/ethnic differentials in sentencing to incarceration. *Justice Quarterly*, *29*(5), 742–773. https://doi.org/10.1080/07418825.2012.659674

Bechtel, K., Holsinger, A. M., Lowenkamp, C. T., & Warren, M. J. (2017). A meta-analytic review of pretrial research: Risk assessment, bond type, and interventions. *American Journal of Criminal Justice*, *42*(2), 443–467.

Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2021). Fairness in criminal justice risk assessments: The state of the art. *Sociological Methods & Research*, *50*(1), 3–44. https://doi.org/10.1177/0049124118782533

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studie. *Communications in Statistics - Simulation and Computation*, *39*(4), 860–864. https://doi.org/10.1080/03610911003650383

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.

Cohen, T. H., & Lowenkamp, C. T. (2019). Revalidation of the Federal PTRA: Testing the PTRA for predictive biases. *Criminal Justice and Behavior*, *46*(2), 234–260. https://doi.org/10.1177/0093854818810315

Copp, J. E., Blomberg, T. G., Casey, W., & Pesta, G. (2019). *Validation of the Virginia Pretrial Risk Assessment Instrument – Revised in Palm Beach County, Florida*. Center for Criminology and Public Policy Research. http://criminology.fsu.edu/wp-content/uploads/Final_Validation_Dissemination.pdf

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 797–806. https://doi.org/10.1145/3097983.309809

Decker, S. H., Ortiz, N., Spohn, C., & Hedberg, E. (2015). Criminal stigma, race, and ethnicity: The consequences of imprisonment for employment. *Journal of Criminal Justice*, *43*(2), 108–121. https://doi.org/10.1016/j.jcrimjus.2015.02.002

DeMichele, M., & Baumgartner, P. (2020). Bias testing of the Public Safety Assessment: Error rate balance between Whites and Blacks for new arrests. *Crime & Delinquency*, 1–26. https://doi.org/10.1177/0011128720965116

Desmarais, S. L., & Singh, J. P. (2013). *Risk assessment instruments validated and implemented in correctional settings in the United States.* Council of State Governments Justice Center. https://csgjusticecenter.org/wp-content/uploads/2020/02/Risk-Assessment-Instruments-Validated-and-Implemented-in-Correctional-Settings-in-the-United-States.pdf

Desmarais, S. L., Zottola, S. A., Duhart Clarke, S. E., & Lowder, E. M. (2020). Predictive validity of pretrial risk assessments: A systematic review of the literature. *Criminal Justice and Behavior*. https://doi.org/10.1177/0093854820932959

Fass, T. L., Heilbrun, K., DeMatteo, D., & Fretz, R. (2008). The LSI-R and the COMPAS: Validation data on two risk-needs tools. *Criminal Justice and Behavior*, *35*(9), 1095–1108.

Flores, A. W., Bechtel, K., & Lowenkamp, C. T. (2016). False positives, false negatives, and false analyses: A rejoinder to "machine bias: There's software used across the country to predict future criminals. And it's biased against blacks." *Federal Probation*, *80*(2), 38–46.

Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City Police Department's "Stop-and-Frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association*, *102*(479), 813–823. https://doi.org/10.1198/016214506000001040

Grommon, E., Ray, B., Sapp, D., & Thelin, R. (2017). *Process evaluation of the IRAS-PAT pilot program implementation*. Indiana University Public Policy Institute, Center for Criminal Justice Research. https://www.in.gov/publicdefender/files/IRAS-PAT_Report_Pr1%20(3).pdf

Harcourt, B. E. (2015). Risk as a proxy for race: The dangers of risk assessment. *Federal Sentencing Reporter*, *27*(4), 237–243.

Ho, D., Stuart, E. A., King, G., & Imai, K. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8). https://doi.org/10.18637/jss.v042.i08

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2016). Inherent trade-offs in the fair determination of risk scores. *ArXiv*. https://doi.org/1609.05807

Kutateladze, B. L., Andiloro, N. R., Johnson, B. D., & Spohn, C. C. (2014). Cumulative disadvantage: Examining racial and ethnic disparity in prosecution and sentencing. *Criminology*, *52*(3), 514–551. https://doi.org/10.1111/1745-9125.12047

Lau, K. S. L., Rosenman, M. B., Wiehe, S. E., Tu, W., & Aalsma, M. C. (2018). Race/ethnicity, and behavioral health status: First arrest and outcomes in a large sample of juvenile offenders. *The Journal of Behavioral Health Services & Research*, *45*(2), 237–251. https://doi.org/10.1007/s11414-017-9578-3

Lowder, E. M., Grommon, E., & Ray, B. (2020). *Improving the accuracy and fairness of pretrial release decisions: A multi-site study of risk assessments implemented in four counties*. George Mason University. Report prepared for the National Institute of Justice in fulfillment of 2018-R2-CX-0023

Lowder, E. M., Lawson, S. G., & Foudray, C. (2020). *Differences in the predictive accuracy of IRAS-PAT assessments as a function of age, sex, and race: Final report*. Indiana Office of Court Services. https://www.in.gov/courts/iocs/files/predictive-bias-report.pdf

Lowder, E. M., Lawson, S. G., Grommon, E., & Ray, B. R. (2020). Five-county validation of the Indiana Risk Assessment System – Pretrial Assessment Tool (IRAS-PAT) using a local validation approach. *Justice Quarterly*, *37*(7), 1241–1260. https://doi.org/10.1080/07418825.2020.1829006

Lowder, E. M., Morrison, M. M., Kroner, D. G., & Desmarais, S. L. (2019). Racial bias and LSI-R assessments in probation sentencing and outcomes. *Criminal Justice and Behavior*, *46*(2), 210–233.

Office of Judicial Administration. (2021). MyCase-Odyssey Public Access. https://public.courts.in.gov/mycase/#/vw/Search

Pretrial Justice Institute. (2020). *Updated position on pretrial risk assessment tools*. Author. https://www.pretrial.org/wp-content/uploads/Risk-Statement-PJI-2020.pdf

Starr, S. B. (2014). Evidence-based sentencing and the scientific rationalization of discrimination. *Stanford Law Review*, *66*, 803–872.

Stevenson, M. T., & Mayson, S. G. (2017). *Bail reform: New directions for pretrial detention and release* ((SSRN Scholarly Paper ID 2939273)). Social Science Research Network. https://papers.ssrn.com/abstract=2939273

Stuart, E. A., & Green, K. M. (2008). Using full matching to estimate causal effects in nonexperimental studies: Examining the relationship between adolescent marijuana use and adult outcomes. *Developmental Psychology*, *44*(2), 395–406. https://doi.org/10.1037/0012-1649.44.2.395

Wang, X., Mears, D. P., & Bales, W. D. (2010). Race-specific employment contexts and recidivism. *Criminology*, *48*(4), 1171–1211. https://doi.org/10.1111/j.1745-9125.2010.00215.x

Zettler, H. R., & Morris, R. G. (2015). An exploratory assessment of race and gender-specific predictors of failure to appear in court among defendants released via a pretrial services agency. *Criminal Justice Review, 40*(4), 417–430.

Appendix

Age

Table 1A. Crosstabulations of Risk Classifications and Pretrial Outcomes by Age

Pretrial Outcomes	Risk Level					
	Low		Moderate		High	
	Under 33 *n* = 1,178	33 or older *n* = 1,273	Under 33 *n* = 1,857	33 or older *n* = 1,388	Under 33 *n* = 699	33 or older *n* = 524
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Any FTA	69 (5.9)	79 (6.2)	281 (15.1)	225 (16.2)	185 (26.5)	148 (28.2)
Any New Arrest	96 (8.1)	123 (9.7)	387 (20.8)	278 (20.0)	208 (29.8)	170 (32.4)
Any Arrest	160 (13.6)	207 (16.3)	623 (33.5)	466 (33.6)	338 (48.4)	255 (48.7)

Sex

Table 2A. Crosstabulations of Risk Classifications and Pretrial Outcomes by Sex

Pretrial Outcomes	Risk Level					
	Low		Moderate		High	
	Female *n* = 685	Male *n* = 1,766	Female *n* = 895	Male *n* = 2,350	Female *n* = 376	Male *n* = 847
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Any FTA	30 (4.4)	118 (6.7)	122 (13.6)	384 (16.3)	112 (29.8)	221 (26.1)
Any New Arrest	58 (8.5)	161 (9.1)	169 (18.9)	496 (21.1)	98 (26.1)	280 (33.1)
Any Arrest	93 (13.6)	274 (15.5)	270 (30.2)	819 (34.9)	174 (46.3)	419 (49.5)

Race

Table 3A. Crosstabulations of Risk Classifications and Pretrial Outcomes by Race

Pretrial Outcomes	Risk Level					
	Low		Moderate		High	
	Black *n* = 442	White *n* = 2,009	Black *n* = 757	White *n* = 2,488	Black *n* = 209	White *n* = 1,014
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Any FTA	36 (8.1)	112 (5.6)	126 (16.6)	380 (15.3)	65 (31.1)	268 (26.4)
Any New Arrest	54 (12.2)	165 (8.2)	145 (19.2)	520 (20.9)	61 (29.2)	317 (31.3)
Any Arrest	92 (20.8)	275 (13.7)	203 (30.4)	859 (34.5)	84 (40.2)	509 (50.2)

Table 4A. Weighted Logistic Regression Models of IRAS-PAT Total Scores and Race Predicting Pretrial Misconduct

Predictor	Weighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Total Score	1.45	[1.36, 1.55]	<.001	1.37	[1.29, 1.46]	<.001	1.39	[1.31, 1.46]	<.001
Black (White)	0.97	[0.78, 1.20]	.769	0.92	[0.75, 1.13]	.428	1.07	[0.90, 1.28]	.432
Time at Risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Black X Total Score	0.92	[0.82, 1.03]	.160	0.93	[0.84, 1.03]	.172	0.92	[0.83, 1.01]	.068

Note. *N* = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio.

Table 5A. Weighted Logistic Regression Models of IRAS-PAT Risk Levels and Race Predicting Pretrial Misconduct

Predictor	Weighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Risk Level (Low)									
Moderate	3.00	[2.16, 4.18]	<.001	2.75	[2.07, 3.66]	<.001	2.79	[2.19, 3.55]	<.001
High	7.20	[4.97, 10.42]	<.001	5.44	[3.87, 7.63]	<.001	5.37	[3.97, 7.26]	<.001
Black (White)	0.95	[0.77, 1.18]	.651	0.91	[0.74, 1.11]	.336	1.06	[0.89, 1.26]	.540
Time at Risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Black X Moderate	0.64	[0.37, 1.14]	.128	0.62	[0.38, 1.00]	.049	0.62	[0.41, 0.93]	.022
Black X High	0.51	[0.27, 0.96]	.037	0.53	[0.30, 0.95]	.033	0.53	[0.32, 0.89]	.016

Note. *N* = 6,919. For categorical variables, reference group indicated in parentheses. Block 1 controlled for county, but estimates are not shown. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio.

Table 6A. Logistic Regression Models of IRAS-PAT Items Predicting Any FTA, by Race

Predictor	Any FTA											
	Black			White								
	n = 1,287			Unweighted *n* = 4,928			Weighted *n* = 4,928			Stratified *n* = 1,295		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Item 1 - Age at first arrest[a]	1.04	[0.41, 2.60]	.939	1.76	[1.19, 2.59]	.004	1.36	[0.74, 2.50]	.324	1.87	[0.85, 4.09]	.118
Item 2 - Number of FTAs[b]												
1	1.51	[1.00, 2.30]	.052	2.18	[1.69, 2.81]	<.001	2.74	[1.73, 4.34]	<.001	2.39	[1.41, 4.05]	.001
2 or more	2.14	[1.19, 3.84]	.011	2.18	[1.60, 2.98]	<.001	2.45	[1.47, 4.08]	.001	2.59	[1.41, 4.77]	.002
Item 3 - 3+ prior incarcerations[c]	1.73	[1.24, 2.40]	.001	1.58	[1.31, 1.91]	<.001	1.44	[1.04, 1.99]	.029	1.78	[1.21, 2.61]	.003
Item 4 - Employed[d]												
Part-time	0.84	[0.48, 1.49]	.558	1.03	[0.75, 1.41]	.859	0.99	[0.59, 1.66]	.980	0.93	[0.45, 1.93]	.855
Not employed	1.75	[1.23, 2.49]	.002	2.08	[1.70, 2.56]	<.001	2.34	[1.65, 3.23]	<.001	2.96	[1.97, 4.45]	<.001
Item 5 - Residential instability[e]	1.23	[0.88, 1.71]	.218	1.28	[1.06, 1.54]	.010	1.40	[1.02, 1.92]	.039	1.48	[1.03, 2.13]	.036
Item 6 - Illegal drug use 6 mo[f]	1.21	[0.86, 1.72]	.272	1.39	[1.12, 1.73]	.003	1.76	[1.23, 2.51]	.002	1.89	[1.23, 2.89]	.003
Item 7 - Severe drug use[g]	1.08	[0.66, 1.77]	.759	1.53	[1.22, 1.92]	<.001	1.28	[0.88, 1.88]	.198	1.52	[0.99, 2.35]	.056
Time at risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001

Note. Models controlled for county, but estimates are not shown. OR = odds ratio; CI = confidence interval for odds ratio. [a]Item 1 reference: 33 or older. [b]Item 2 reference: No FTA warrants past 24 months. [c]Item 3 reference: Two or less prior jail incarcerations. [d]Item 4 reference: Yes, Full-time employment at time of arrest. [e]Item 5 reference: Lived at current residence past 6 months. [f]Item 6 reference: No illegal drug use during past 6 months. [g]Item 7 reference: No severe drug use problem.

Table 7A. Logistic Regression Models of IRAS-PAT Items Predicting Any New Arrest, by Race

Predictor	Any New Arrest											
	Black			White								
	n = 1,284			Unweighted *n* = 4,928			Weighted *n* = 4,928			Stratified *n* = 1,295		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Item 1 - Age at first arrest[a]	1.90	[0.72, 5.01]	.191	1.33	[0.98, 1.80]	.068	1.47	[0.83, 2.61]	.183	1.39	[0.76, 2.53]	.283
Item 2 - Number of FTAs[b]												
1	1.39	[0.90, 2.13]	.134	1.29	[1.02, 1.63]	.033	1.83	[1.18, 2.82]	.007	1.39	[0.85, 2.27]	.184
2 or more	3.33	[1.88, 5.89]	<.001	1.31	[0.99, 1.74]	.060	2.16	[1.32, 3.54]	.002	1.14	[0.62, 2.11]	.675
Item 3 - 3+ prior incarcerations[c]	1.15	[0.84, 1.58]	.388	1.71	[1.45, 2.01]	<.001	1.33	[0.98, 1.80]	.066	1.60	[1.15, 2.25]	.006
Item 4 - Employed[d]												
Part-time	1.17	[0.72, 1.90]	.530	1.31	[1.02, 1.69]	.033	1.06	[0.66, 1.70]	.802	1.16	[0.68, 2.01]	.585
Not employed	1.40	[1.00, 1.95]	.048	1.76	[1.48, 2.10]	<.001	2.36	[1.72, 3.24]	<.001	1.88	[1.33, 2.66]	<.001
Item 5 - Residential instability[e]	1.36	[1.00, 1.86]	.050	1.40	[1.19, 1.64]	<.001	1.18	[0.88, 1.58]	.268	1.27	[0.92, 1.75]	.142
Item 6 - Illegal drug use 6 mo[f]	1.49	[1.08, 2.07]	.016	1.41	[1.17, 1.70]	<.001	1.38	[1.01, 1.89]	.042	1.69	[1.19, 2.42]	.004
Item 7 - Severe drug use[g]	1.43	[0.90, 2.29]	.134	1.44	[1.18, 1.76]	<.001	1.23	[0.85, 1.77]	.276	1.31	[0.89, 1.95]	.175
Time at risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001

Note. Models controlled for county, but estimates are not shown. OR = odds ratio; CI = confidence interval for odds ratio. [a]Item 1 reference: 33 or older. [b]Item 2 reference: No FTA warrants past 24 months. [c]Item 3 reference: Two or less prior jail incarcerations. [d]Item 4 reference: Yes, Full-time employment at time of arrest. [e]Item 5 reference: Lived at current residence past 6 months. [f]Item 6 reference: No illegal drug use during past 6 months. [g]Item 7 reference: No severe drug use problem.

Table 8A. Logistic Regression Models of IRAS-PAT Items Predicting Any Arrest, by Race

Predictor	Any Arrest											
	Black			White								
	n = 1,287			Unweighted *n* = 4,928			Weighted *n* = 4,928			Stratified *n* = 1,295		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Item 1 - Age at first arrest[a]	1.22	[0.64, 2.36]	.546	1.51	[1.17, 1.95]	.001	1.43	[0.86, 2.39]	.167	1.73	[1.00, 2.98]	.049
Item 2 - Number of FTAs[b]												
1	1.46	[1.00, 2.12]	.050	1.72	[1.40, 2.12]	<.001	1.82	[1.23, 2.69]	.003	1.47	[0.93, 2.31]	.009
2 or more	2.26	[1.29, 3.96]	.004	1.63	[1.26, 2.12]	<.001	2.33	[1.48, 3.69]	<.001	1.22	[0.68, 2.16]	.503
Item 3 - 3+ prior incarcerations[c]	1.40	[1.06, 1.85]	.018	2.06	[1.78, 2.38]	<.001	1.33	[1.01, 1.75]	.042	1.95	[1.43, 2.65]	<.001
Item 4 - Employed[d]												
Part-time	1.22	[0.80, 1.85]	.355	1.43	[1.16, 1.78]	.001	1.40	[0.96, 2.03]	.083	1.38	[0.87, 2.19]	.170
Not employed	1.63	[1.21, 2.19]	.001	2.04	[1.75, 2.38]	<.001	2.42	[1.81, 3.22]	<.001	2.02	[1.47, 2.78]	<.001
Item 5 - Residential instability[e]	1.19	[0.91, 1.57]	.211	1.43	[1.24, 1.65]	<.001	1.28	[0.98, 1.67]	.069	1.33	[0.99, 1.78]	.056
Item 6 - Illegal drug use 6 mo[f]	1.17	[0.88, 1.56]	.276	1.33	[1.13, 1.56]	.001	1.36	[1.02, 1.80]	.034	1.82	[1.32, 2.51]	<.001
Item 7 - Severe drug use[g]	1.88	[1.25, 2.83]	.003	1.49	[1.25, 1.78]	<.001	1.26	[0.89, 1.77]	.193	1.24	[0.86, 1.79]	.246
Time at risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.01]	<.001

Note. Models controlled for county, but estimates are not shown. OR = odds ratio; CI = confidence interval for odds ratio. [a]Item 1 reference: 33 or older. [b]Item 2 reference: No FTA warrants past 24 months. [c]Item 3 reference: Two or less prior jail incarcerations. [d]Item 4 reference: Yes, Full-time employment at time of arrest. [e]Item 5 reference: Lived at current residence past 6 months. [f]Item 6 reference: No illegal drug use during past 6 months. [g]Item 7 reference: No severe drug use problem.

Table 9A. Unweighted Logistic Regression Models of Race and IRAS-PAT Items Predicting Pretrial Misconduct

Predictor	Unweighted								
	Any FTA			Any New Arrest			Any Arrest		
	OR	95% CI	*p*	OR	95% CI	*p*	OR	95% CI	*p*
Block 1									
Black[a]	1.06	[0.87, 1.30]	.579	1.00	[0.83, 1.19]	.960	1.12	[0.96, 1.31]	.151
Item 1 - Age at first arrest[b]	1.63	[1.15, 2.34]	.006	1.41	[1.06, 1.89]	.020	1.50	[1.19, 1.90]	.001
Item 2 - Number of FTAs[c]									
1	1.98	[1.60, 2.46]	<.001	1.32	[1.08, 1.62]	.007	1.66	[1.38, 1.98]	<.001
2 or more	2.17	[1.65, 2.86]	<.001	1.61	[1.25, 2.06]	<.001	1.76	[1.39, 2.23]	<.001
Item 3 - 3+ prior incarcerations[d]	1.61	[1.37, 1.89]	<.001	1.55	[1.34, 1.80]	<.001	1.88	[1.66, 2.14]	<.001
Item 4 - Employed[e]									
Part-time	0.98	[0.75, 1.29]	.909	1.29	[1.03 1.60]	.026	1.38	[1.14, 1.66]	.001
Not employed	2.00	[1.67, 2.38]	<.001	1.66	[1.43, 1.94]	<.001	1.93	[1.69, 2.21]	<.001
Item 5 - Residential instability[f]	1.27	[1.08, 1.49]	.004	1.38	[1.20, 1.59]	<.001	1.37	[1.21, 1.56]	<.001
Item 6 - Illegal drug use 6 mo[g]	1.34	[1.11, 1.61]	.002	1.41	[1.20, 1.66]	<.001	1.28	[1.11, 1.47]	<.001
Item 7 - Severe drug use[h]	1.49	[1.22, 1.82]	<.001	1.46	[1.22, 1.74]	<.001	1.58	[1.35, 1.86]	<.001
Time at risk	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001	1.00	[1.00, 1.00]	<.001
Block 2									
Black X Item 1	0.56	[0.21, 1.48]	.240	1.55	[0.57, 4.19]	.390	0.77	[0.39, 1.53]	.461
Black X Item 2 - 1 FTA	0.74	[0.45, 1.20]	.217	0.99	[0.62, 1.57]	.970	0.87	[0.58, 1.30]	.491
Black X Item 2 - 2 or more FTAs	1.05	[0.55, 2.00]	.880	2.04	[1.14, 3.67]	.017	1.36	[0.76, 2.43]	.297
Black X Item 3	1.08	[0.74, 1.57]	.701	0.70	[0.49, 0.98]	.040	0.66	[0.49, 0.90]	.008
Black X Item 4 - Part-time	0.85	[0.44, 1.61]	.614	0.93	[0.55, 1.59]	.798	0.89	[0.56, 1.39]	.599
Black X Item 4 - Not Employed	0.84	[0.56, 1.27]	.410	0.79	[0.55, 1.14]	.204	0.78	[0.57, 1.08]	.140
Black X Item 5	0.95	[0.65, 1.39]	.809	0.92	[0.65, 1.29]	.610	0.82	[0.60, 1.11]	.194
Black X Item 6	0.86	[0.57, 1.29]	.467	1.01	[0.70, 1.45]	.978	0.87	[0.63, 1.19]	.377
Black X Item 7	0.79	[0.46, 1.33]	.372	0.93	[0.58, 1.50]	.781	1.27	[0.83, 1.94]	.269
∆-2LL		6.73 (9)			12.01 (9)			16.50 (9)	

Note. *N* = 6,220. Block 1 controlled for county, but estimates are not shown. ∆-2LL reflects improvement in model fit upon addition of the interaction term(s) in Block 2. All model terms from Block 1 were included in Block 2; however, only unique terms are shown. OR = odds ratio. CI = confidence interval for odds ratio. [a]Reference: White. [b]Item 1 reference: 33 or older. [c]Item 2 reference: No FTA warrants past 24 months. [d]Item 3 reference: Two or less prior jail incarcerations. [e]Item 4 reference: Yes, Full-time employment at time of arrest. [f]Item 5 reference: Lived at current residence past 6 months. [g]Item 6 reference: No illegal drug use during past 6 months. [h]Item 7 reference: No severe drug use problem.