Determining Racial Equity in Pretrial Risk Assessment

Howard Henderson
Texas Southern University
Jack Sevil
Texas Southern University
Danielle Lessard
West Virginia State College
David Rembert
Prairie View A&M University, TX

ON THE HEELS of the George Floyd Movement, advocacy groups and activist journalists called for the dismantling of pretrial risk assessments. They argued that Black and Brown defendants were being unfairly classified as high-risk threats to public safety. The negative attention directed at pretrial risk assessments swept pretrial justice by storm and overshadowed empirically supported benefits to pretrial assessments in favor of the previously established judicial determinations of risk.

A few of the more urban counties responded by returning to judicial determinations of pretrial release and supervision. In fact, some have gone so far as to remove pretrial risk assessment requirements, placing the determination of risk and needs squarely in the hands of county judges, a practice that was more common over four decades ago (Sanchez & Strenio, 2022; Desmarais & Lowder, 2019; Rachlinski et al., 2008; Guthrie et al., 2007). Unfortunately, the statutory removal of pretrial risk assessments came without adequate interest or support from academic researchers. In fact, there have yet to be any determinations of the extent to which judicial determinations of risk and need differ from or improve upon risk assessment instruments (Desmarais & Lowder, 2019; Gottfredson, 1999).

At this moment, findings of racial bias in pretrial risk assessment are ambiguous, at best, as scholars continue to debate the nexus of bias in the instruments (Desmarais et al., 2021). One camp holds that Black and Hispanic persons score higher on these assessments than White persons (Desmarais et al., 2021). Others note that racial bias in pretrial assessments is inevitable because racial minority groups have a much higher likelihood of being arrested, thus ensuring that risk assessment instruments with criminal history items will inevitably score them at higher risk than others (Eckhouse et al., 2019; Mayson, 2019). The third group holds that minority groups are more likely to be over-classified (i.e., false positives) and White persons are at greater risk of being under-classified (i.e., false negatives), rates of error that often go unexamined in risk assessment validations (Rembert et al., 2014; Singh & Fazel, 2010; Whiteacre, 2006). Lowder et al. (2021) have suggested further research to understand the nature and extent of racial bias in pretrial risk assessment, as these assessments have consequences for individuals, communities, and the overall legitimacy of the criminal justice system.

The ability of pretrial risk assessments to equitably predict outcomes has garnered limited attention in the academic literature. Despite the limited research, the results remain mixed, and examinations of predictive error are far fewer (Desmarais et al., 2021; DeMichele, 2020; Bechtel, 2017; 2011). Given the demographics of the pretrial system, these gaps are all the more troubling as the system seeks to maintain public safety and reduce racial/ethnic disparities.

In general, researchers have found pretrial risk assessments to be valid predictors of pretrial success, court appearances, rearrest, and violent crime, despite a few findings of racial/ethnic predictive inequities (Desmarais et al., 2021; DeMichele et al., 2020). Most of the racial bias pretrial risk assessment research finds instrument validity, though not as good for racial/ethnic groups, is in the fair to good category for these groups.

Despite the limited focus of pretrial risk assessment research on racial group validations, most prior analyses have hinged on group classification proportionality and regression analysis, leaving bias, as measured by error, mostly unexamined (Rembert et al., 2013; Singh & Fazel, 2010; Whiteacre, 2006). As a result, very little is understood about the degree of bias expressed in pretrial risk assessment. Understanding the impact of bias in pretrial risk assessment is ever more pertinent

TABLE 1.
Sample Characteristics

Variable	n	(%)
Defendant Race		
Black	69	(19.66)
White	242	(68.95)
Unknown	40	(11.40)
Defendant Sex		
Female	86	(24.50)
Male	226	(64.39)
Unknown	39	(11.11)
Court		
District	208	(59.26)
Circuit	93	(26.50)
District and Circuit	10	(2.85)
Unknown	40	(11.40)

TABLE 2.
Sample Risk Classification and Highest Charge

Variable	n	(%)
Risk classification		
Low	25	(7.12)
Moderate	220	(62.68)
Moderate/High	2	(0.57)
High	100	(28.49)
N/A	3	(0.85)
Unknown	1	(0.28)
Defendant Highest Charge		
Animal Cruelty	2	(0.57)
Assault	71	(20.23)
Murder	2	(0.57)
Burglary and Theft	21	(5.98)
Drug Possession	94	(26.78)
Child Abuse	4	(1.14)
Disorderly Conduct	3	(0.85)
Driving Offenses	15	(4.27)
Obstruction of Justice	42	(11.97)
Firearms	8	(2.28)
Forgery	9	(2.56)
Harassment	2	(0.57)
Intoxicated Endangerment	2	(0.57)
Property Destruction	2	(0.57)
Sexual Offenses	7	(1.99)
Robbery	2	(0.57)
Order Violation	60	(17.09)
Unknown	5	(1.42)

when considering the deleterious effect bias has on pretrial detention and sentencing decisions (Jackson et al., 2013; Zinger, 2004). Fair predictions also serve as the crux of rehabilitative efforts and appropriate supervision levels.

There remains ambiguity regarding predictive racial disparities in pretrial risk assessments. As these debates continue, it is pertinent to keep in mind that all these instruments should be validated on their local population, and that jurisdictions should never adhere to blind adoption. As such, we examine racial differences in predictive accuracy of a pretrial release risk assessment instrument. To do so, we use a convenience sampling of 351 defendants who had been administered a pretrial release risk assessment within an East Coast county. The goal of the current research is to examine the ability of a pretrial risk assessment instrument to predict supervision outcomes and to understand the extent to which error is equitably distributed.

TABLE 3.
Sample Violation Year and Pretrial Status

Variable	n	(%)
Defendant Violation Year		
2021	47	(13.39)
2022	55	(15.67)
No Violation	249	(70.94)
Defendant Pretrial Status		
Active	155	(44.16)
Completed	112	(31.91)
Removed	84	(23.93)

TABLE 4.
Distribution of Risk Classification Groups

	Total		Removed	
Risk classification	n	(%)	n	(%)
Low	25	(7.12)	2	(2.38)
Moderate	220	(62.68)	46	(54.76)
Moderate / High	2	(0.57)	2	(2.38)
High	100	(28.49)	34	(40.48)
N/A	3	(0.85)	0	(0.00)
Unknown	1	(0.28)	0	(0.00)
Total	351	(99.99)*	84	(100.00)

Note. * Due to rounding, the percentage does not equal 100%

TABLE 5.
Sample Classification Errors

	False Positives		False Negatives	
	%	(n)	%	(n)
Defendants	92.59	(25)	7.41	(2)

Methods

Participants

In this study, we used a convenience sampling (N) of 351 pretrial defendants. Table 1 demonstrates characteristics of subsamples (n) within the sample (N = 351). The majority of the pretrial defendants were White (68.95 percent), male (64.39 percent), and awaiting district court trial (59.26 percent).

Table 2 highlights the risk classification determined using the Pre-Trial Release Risk Assessment (PTRA) tool and *defendant's highest charge*. Slightly more than 1 percent of the sample (1.14 percent) had no identified risk classification. Still, most defendants had been classified as Moderate Risk. Moreover, most defendants' highest charge was reported as a drug possession (26.78 percent).

Table 3 demonstrates defendants' violation year and pretrial status. Most defendants had unreported violation years (70.94 percent) due to having no reported violations. Moreover, almost half (44.16 percent) of defendants were currently under pretrial supervision, with another quarter (23.93 percent) already removed.

Additionally, Table 4 demonstrates sample characteristics of persons awaiting trial by risk classification and defendant pre-trial status. This tabulation demonstrates that most defendants were assessed as having Moderate Risk (62.86 percent), and the majority of removed defendants also had Moderate Risk (54.76 percent).

Table 5 highlights classification errors—or incorrect predictions using the PTRA—within the sample as either false positives or false negatives. Specifically, false positives were instances where defendants were classified as having high risk but successfully completed their pretrial diversion term, false negatives were instances where pretrial defendants were classified as having low risk but were removed from supervision. False positives were more common within the sample than false negatives. That is, only 2.38 percent (n = 2) of persons removed were classified as having low risk, but 7.71 percent (n = 27) of high-risk persons were incorrectly predicted.

Last, Table 6 (next page) demonstrates classification errors as a cross tabulation of defendant race. The majority (75 percent) of false positives were White defendants, and all false negatives were White defendants (n = 2). Moreover, 8.7 percent (n = 6) and 0% (n = 0) of Black defendants were either false positives or false negatives, respectively. Regarding White defendants, 7.44 percent (n = 18) and .83 percent (n = 2) of defendants were false positives or false negatives, respectively.

Materials

The current study used a convenience sampling of 351 pretrial criminal defendants supervised by the county detention center. Data were received within a Microsoft Excel workbook. The file contained defendants' demographic information, as well as responses to the PTRA. The dataset comprised 14 categories of information with both non-numerical and numerical data, including gender, race, court, case, risk level, risk score, highest charge, notes, violation, next court date, and whether the pretrial term was completed. The risk level of each defendant was calculated using the PTRA, a risk assessment tool designed to score defendants based on risk of unsuccessfully completing their pretrial supervision term. The PTRA was administered to defendants independently from the current study, wherein defendants were scored based on six categories: 1) their most serious current offense, with a maximum of nine points; 2) additional considerations, with a maximum of two points; 3) their current legal status, with a maximum of six points; 4) the severity of their prior convictions, with a maximum of nine points; 5) supervision, failures to appear, or probation violations within the past 10 years, with a maximum of eight points; and 6) mitigating factors, with a maximum of four points able to be subtracted from a defendant's score. Based on these scores, defendants were classified as either: 1) high risk, with 14 points or more; 2) moderate risk, with between 6 and 13 points; and 3) low risk, with 5 points or fewer.

Procedures

Upon request, the county detention center provided data. We screened, cleaned, and coded the data and derived 31 coded risk assessment variables. Zero represented the absence of a phenomenon, and 1 represented the observation of that phenomenon. Coded variable groups included 1) defendant's pretrial status, comprising a) active, b) completed, or c) removed; 2) risk classification, comprising a) Low Risk, b) Moderate Risk, c) Moderate / High Risk, d) High Risk, and e) unknown; 3) defendant sex, comprising a) female, b) male, and c) unknown; 4) defendant race, comprising a) Black, b) White, and c) unknown; 5) court, comprising a) circuit, b) district, and c) unknown; 6) defendant highest charge, comprising a) animal cruelty, b) assault, c) murder, d) burglary and theft, e) drug possession, f) child abuse, g) disorderly conduct, h) driving offense, i) obstruction of justice, j) firearms, k) forgery, l) harassment, m) intoxicated endangerment, n) property destruction, o) sexual offenses, p) robbery, q) order violation, and r) unknown; and 7) defendant violation year, comprising a) 2021, b) 2022, and c) no violation.

Defendant violation year was derived from the violation variable where exact dates and times of defendants' violations were reported. Violation year was created to limit categories of violation date.

Results

To determine the instrument's predictive accuracy, we calculated a tetrachoric correlation (*rtet*) analysis. We found that *rtet* was favorable over a Pearson product-moment correlation, as the risk classification and defendant pretrial status measures were categorical as opposed to continuous. Risk classification and defendant pretrial status being dichotomous, we could not test the normality and linearity assumptions necessary to examine a Pearson correlation coefficient. Table 7 demonstrates our *rtet* results. Most relationships were significant using Alpha (α) = .05, 2-tailed. The only non-significant finding was the association between Moderate / High Risk classification and completed defendant pretrial status (*rtet* = -.01, probability p = .8). Additionally, we removed Case 314 from our analysis because no information regarding the defendant's risk classification was reported.

To estimate the predictive accuracy of the risk classification model with N = 351, we constructed scatterplots using false positive rates and true positive rates statistics. Visual analysis of this plot demonstrated that PTRA accurately predicted pretrial program removal and completion. Moreover, we calculated receiver operating characteristics (ROC) and summed them to produce an area under the curve (AUC) statistic. Table 8 demonstrates these findings, including the p, standard error (*SE*), and margin of error (*ME*). We used Hanley and McNeil's (1982) formula for calculating AUC *SE*. Even more sparse are formulas for calculating AUC *ME*. As such, we used a common formula to calculate *ME* for regression models. Specifically, *ME* can be calculated by multiplying the t-crit by the *SE* of β. To derive the AUC *ME*, we multiplied z-critical values derived from the Mann-Whitney U statistics by the AUC *SE*. Table 8 shows the PTRA as a good predictor of pretrial outcomes. This is because an AUC statistic of 1 indicates perfect predictability of the analyzed tool. An AUC of 0.5 suggests no discrimination, 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered excellent, and more than 0.9 is considered outstanding. Moreover, we found each AUC statistic to be significant using a Mann-Whitney U-test. This means there is leastways a 99.99 percent probability that analyzed samples were similarly distributed.

Due to Moderate / High Risk not being a risk classification prescribed by the PTRA, we

TABLE 6.
Sample Classification Errors

	False Positives		False Negatives	
Defendant Race	%	(n)	%	(n)
Black	25	(6)	0	(0)
White	75	(18)	100	(2)
Total	100	(24)	100	(2)

Note. Total of false positives is less than n in Table 7 because one defendant's race was unreported.

TABLE 7.
Correlations Between Risk Classification and Outcome

	Low Risk	Moderate Risk	Moderate / High Risk	High Risk	Completed	Removed
Low Risk	1	-	-	-	-	-
Moderate Risk	.4*	1	-	-	-	-
Moderate / High Risk	-.52*	-.72*	1	-	-	-
High Risk	-.97*	-.8*	-.93*	1	-	-
Completed	-.34*	.89*	-.01	.99*	1	
Removed	-.93*	-.95*	-.96*	-1*	.93*	1

Note. * Indicated significance using α = .05, 2-tailed.

calculated alternate AUC statistics by delegating $n = 2$ defendants classified as Moderate / High Risk to the appropriate classification(High Risk). Table 9 demonstrates AUC statistics between the instrument's predictive accuracy for Black and White pretrial defendants without the moderate/high risk classification. The AUC statistics were all higher than those that included Moderate / High Risk. Only the total sample AUC remained significant using the Mann-Whitney U test, but this only suggests that samples used to derive the AUC may be differently distributed and should not be interpreted as a sole determinant of accuracy.

TABLE 8.
Area Under the Curve Statistics

Sample	AUC	*p*	*SE*	*ME*
Black	.86	<.01	.26	.52
White	.85	<.01	.16	.31
Total	.85	<.01	.14	.27

TABLE 9.
Area Under the Curve for Black and White Pretrial Defendants

Sample	AUC	*p*	*SE*	*ME*
Black	.92	>.99	.17	.34
White	.86	>.99	.15	.29
Total	.87	<.01	.14	.27

TABLE 10.
Pre-Trial Release Risk Assessment Multiple Generalized Regression Output

	Estimate	*SE*	*df*	*t*	*t*-crit	*ME*
Intercept	4.45*	0.08	276	55.58	1.97	0.16
Defendant Race						
Black	2.63*	0.14	276	18.78	1.97	0.28
White	1.82*	0.10	276	18.17	1.97	0.16
Defendant Sex						
Female	1.77*	0.14	276	12.63	1.97	0.20
Male	2.68*	0.08	276	33.48	1.97	0.28
Court						
District	0.36	0.24	276	1.5	1.97	0.47
Circuit	1.72*	0.29	276	5.94	1.97	0.57
District Circuit	2.36*	1.01	276	2.34	1.97	1.99
Defendant Highest Charge						
Animal Cruelty	2.55	8.40	276	0.3	1.97	16.54
Assault	0.37	0.59	276	0.63	1.97	1.16
Murder	6.79	8.53	276	0.8	1.97	16.79
Burglary and Theft	0.93	1.25	276	0.74	1.97	2.46
Drug Possession	1.47*	0.53	276	2.77	1.97	1.04
Child Abuse	-0.42	4.36	276	-0.1	1.97	8.58
Disorderly Conduct	0.66	5.72	276	0.12	1.97	11.26
Driving Offenses	1.32	1.54	276	0.86	1.97	3.03
Obstruction of Justice	-1.92*	0.76	276	-2.52	1.97	1.50
Firearms	1.24	2.40	276	0.51	1.97	4.72
Forgery	1.93	2.67	276	0.72	1.97	5.26
Harassment	2.24	8.39	276	0.27	1.97	16.52
Intoxicated Endangerment	-6.50	16.47	276	-0.39	1.97	32.42
Property Destruction	-7.31	16.70	276	-0.44	1.97	32.88
Sexual Offenses	-3.08	3.13	276	-0.98	1.97	6.16
Robbery	2.42	8.44	276	0.29	1.97	16.61
Order Violation	1.76*	0.65	276	2.71	1.97	1.28
Pre-Trial Status						
Active	1.19*	0.13	276	9.18	1.97	0.26
Removed	2.52*	0.16	276	15.75	1.97	0.31
Completed	0.73*	0.14	276	5.23	1.97	0.28

Note. * indicates significance using $\alpha = .05$, 2-tailed.

To estimate the probability of predicting PTRA score by race, we controlled for four other independent variable (IV) groups. We calculated a 27-predictor multiple generalized regression (MGR) analysis and included 1) defendant race, 2) defendant sex, 3) court, 4) defendant highest charge, and 5) pretrial status. An MGR analysis calculates scalar directional relationships between a non-dichotomous dependent variable (DV) and multiple IVs, with each relationship accounting for others within the model. Prior to constructing the MGR, we scanned the DV—PTRA score—for missing values. Subsequently, $n = 47$ cases within the dataset were identified as missing PTRA scores. We removed these $n = 41$ cases from analyses as a necessity for calculating beta coefficients (β).

β were calculated using the Moore-Penrose generalized method. The Moore-Penrose generalized method is a type of pseudo-inversion that assumes linearity between residuals and z-scores, as well as homoscedasticity of residuals and predicted values ($\hat{y}$). We constructed a normal probability plot with residuals and z-values, which, upon visual inspection, indicated linearity. That is, a linear relationship was observable between the error terms for predicted values. Additionally, we constructed a scatterplot with residuals and $\hat{y}$, which, upon visual inspection, indicated homoscedasticity. That is, $\hat{y}$ and residual error terms did not linearly relate. We used Moore-Penrose inversion because it provides the same output as generalized inversion when determinant > 0 but remains interpretable for determinant = 0. As such, data were appropriate for regression modelling.

An analysis of variance (ANOVA) indicated the 27-predictor MGR accounted for a non-significant 18.04 percent of the variance in PTRA score, calculated as $R2 = .18$, $F(278,277) = .22$, $p = > .99$, $\alpha = .05$, 2-tailed. This means there is leastways a 99.99 percent probability that the 18.04 percent of the variance in PTRA scores accounted for by the IVs may be due to sampling error.

Table 10 provides the output of the MGR, including β, *SE*, *p*, and *ME* for all predictors within the model. Both Black ($\beta = 2.63$, $p = < .01$, 2-tailed) and White ($\beta = 1.82$, $p = < .01$, 2-tailed) defendant race were significant positive predictors of PTRA score using α

= .05, 2-tailed. That is, on average, being a Black defendant within pretrial supervision predicted greater PTRA scores than being a White defendant. Being significant, these results indicate there is leastways a 95 percent probability that the observed relationships are not related to sampling error. Additionally, having a removed pretrial status predicted, on average, more PTRA points than an active or completed status.

To estimate the probability of predicting removal from pretrial supervision using five IV groups, comprising 28 IVs, we conducted a multiple probability of outcome generalized regression (MPOGR) analysis. The IV groups included 1) risk classification, 2) defendant sex, 3) defendant race, 4) court, and 5) defendant highest charge. An MPOGR calculates the non-scalar directional relationship between a binary DV and multiple IVs, with each relationship accounting for others within the model. Prior to constructing the MPOGR, the DV—removal—was scanned for missing values. Subsequently, n = 41 cases within the dataset were identified as missing values for removal. These n = 41 cases were removed from analyses as a necessity for β calculation. Further, we calculated β for n = 310 to calculate $\hat{y}$, odds (eL), and probability of outcome (p(X)) necessary for MPOGR. p(X) replaced the binary DV values within the model, creating a MPOGR.

We calculated β using the Moore-Penrose generalized method. Due to the model being a probability of outcome regression, no assumption testing was required, and data were deemed appropriate for regression modelling.

An analysis of variance (ANOVA) indicated that the 29-predictor MPOGR accounted for a significant 2.86 percent of the variance in removals, calculated as predicted $R2$ ($\hat{y}R2$) = .03, $F(281,280) = 34.04$, p = < .01, α = .05, 2-tailed. This means there is leastways a 95 percent probability that the 2.86 percent of the variance in removals accounted for by the IV was not related to sampling error.

Table 11 provides the output of the MPOGR, including β, SE, p, and ME for all predictors within the model. Only three predictors within the current model were non-significant, using α = .05, 2-tailed: 1) animal cruelty (β = 0, p = 1), 2) child abuse (β = 0, p = 1), and 3) robbery (β = -.01, p = .16). Regarding risk classification, defendants being classified as High Risk was the strongest predictor of pretrial supervision failure. Specifically, being classified as High Risk predicted a significant 9 percent probability of unsuccessfully completing pretrial supervision, whereas being classified as having Low Risk predicted a significant 2 percent probability. Moreover, Black race and female sex were both the greatest predictors of removal, predicting a significant 11 percent probability of unsuccessfully completing pretrial supervision. Being significant, these results indicate there is leastways a 95 percent probability the observed relationships are not related to sampling error.

Additionally, we calculated an MPOGR analysis to estimate the probability of predicting PTRA false positives by defendant race using n = 310. The false positive MPOGR included defendant race alongside three other IV groups to control for possible relationships. That is, the false positive MPOGR comprised 24 IV groups: 1) defendant race, 2) defendant sex, 3) court, and 4) highest charge. The IV

TABLE 11.
Pretrial Removal Multiple Probability of Outcome Generalized Regression Output

Variable	Estimate	*SE*	*df*	*t*	*t*-crit	*ME*
Intercept	0.20*	0.00	281	62.53	1.97	0.01
Risk classification						
Low	0.02*	0.01	281	3.30	1.97	0.01
Moderate	0.05*	0.01	281	8.53	1.97	0.01
Moderate / High	0.20*	0.01	281	24.60	1.97	0.01
High	0.09*	0.01	281	15.23	1.97	0.01
Sex						
Female	0.11*	0.00	281	65.11	1.97	0.00
Male	0.09*	0.00	281	53.50	1.97	0.00
Race						
Black	0.11*	0.00	281	63.39	1.97	0.00
White	0.09*	0.00	281	54.39	1.97	0.00
Court						
Circuit	0.06*	0.00	281	44.27	1.97	0.00
District	0.06*	0.00	281	43.63	1.97	0.00
District Circuit	0.09*	0.00	281	41.25	1.97	0.00
Highest Charge						
Animal Cruelty	0.00	0.01	281	0.00	1.97	0.00
Assault	0.07*	0.00	281	15.11	1.97	0.01
Murder	-0.02*	0.01	281	-2.79	1.97	0.01
Burglary and Theft	0.10*	0.00	281	20.58	1.97	0.01
Drug Possession	0.03*	0.00	281	6.54	1.97	0.01
Child Abuse	0.00	0.01	281	0.00	1.97	0.01
Disorderly Conduct	0.02*	0.01	281	3.12	1.97	0.01
Driving Offenses	0.10*	0.00	281	20.10	1.97	0.01
Obstruction of Justice	0.07*	0.00	281	14.94	1.97	0.01
Firearms	0.15*	0.01	281	28.29	1.97	0.01
Forgery	0.09*	0.01	281	16.66	1.97	0.01
Harassment	0.02*	0.01	281	2.80	1.97	0.01
Intoxicated Endangerment	0.05*	0.01	281	5.52	1.97	0.02
Property Destruction	0.03*	0.01	281	3.30	1.97	0.02
Sexual Offenses	0.02*	0.01	281	3.66	1.97	0.01
Robbery	-0.01	0.01	281	-1.40	1.97	0.01
Order Violation	0.09*	0.00	281	19.37	1.97	0.01

Note. All SE are > 0 but are presented as 0 when below .01.
* indicates significance using α = .05, 2-tailed

risk classification group was not included within the false positive MPOGR because risk classification was used to calculate the false positive DV. We again used the Moore-Penrose generalized method to calculate β.

The ANOVA indicated the 24-predictor MPOGR accounted for a significant 4.6 percent of the variance in false positives ($\hat{y}R2$ = .05, $F(285,284) = 20.81$, $p = < .01$, $\alpha = .05$, 2-tailed). This means there is leastways a 95 percent probability that 4.6 percent of the variance in false positives accounted for by the IVs was not related to sampling error.

Table 12 provides the output of the false positive MPOGR, including the intercept, β, *SE*, *p*, and *ME* for all predictors. All predictors and the intercept were significant. Both Black and White defendant race predicted the same probability of false positive on the PTRA (12%). This means we observed no difference in the probability of being classified as high risk and completing pretrial supervision between Black and White defendants.

Last, we calculated the same MPOGR analysis to estimate the probability of predicting false negatives by defendant race using n = 310. The ANOVA indicated the 24-predictor MPOGR accounted for a non-significant 73.36 percent of the variance in false negatives ($\hat{y}R2$ = .73, $F(285,284) = .36$, $p = > .99$, α = .05, 2-tailed). This means there is more than a 99.99 percent probability that variance in false positives accounted for by the false negative MPOGR was somehow related to sampling error.

TABLE 12.
False Positive Multiple Probability of Outcome Generalized Regression Output

Variable	Estimate	*SE*	*df*	*t*	*t*-crit	*ME*
Intercept	0.25*	0.00	285	209.61	1.97	0.00
Defendant Race						
Black	0.12*	0.00	285	171.60	1.97	0.00
White	0.12*	0.00	285	178.90	1.97	0.00
Defendant Sex						
Female	0.12*	0.00	285	172.00	1.97	0.00
Male	0.12*	0.00	285	182.09	1.97	0.00
Court						
Circuit	0.08*	0.00	285	115.45	1.97	0.00
District	0.09*	0.00	285	138.98	1.97	0.00
District / Circuit	0.07*	0.00	285	58.77	1.97	0.00
Defendant Highest Charge						
Animal Cruelty	-0.08*	0.00	285	-18.50	1.97	0.01
Assault	-0.07*	0.00	285	-25.02	1.97	0.01
Murder	0.15*	0.00	285	34.68	1.97	0.01
Burglary and Theft	-0.07*	0.00	285	-23.79	1.97	0.01
Drug Possession	-0.05*	0.00	285	-18.02	1.97	0.01
Child Abuse	-0.08*	0.00	285	-22.17	1.97	0.01
Disorderly Conduct	-0.08*	0.00	285	-20.66	1.97	0.01
Driving Offenses	-0.06*	0.00	285	-19.91	1.97	0.01
Obstruction of Justice	-0.07*	0.00	285	-24.69	1.97	0.01
Firearms	-0.05*	0.00	285	-15.61	1.97	0.01
Forgery	-0.05*	0.00	285	-15.33	1.97	0.01
Harassment	-0.08*	0.00	285	-18.49	1.97	0.01
Intoxicated Endangerment	-0.08*	0.01	285	-14.64	1.97	0.01
Property Destruction	-0.09*	0.01	285	-16.44	1.97	0.01
Sexual Offenses	-0.08*	0.00	285	-24.42	1.97	0.01
Robbery	-0.08*	0.00	285	-18.53	1.97	0.01
Order Violation	-0.07*	0.00	285	-24.96	1.97	0.01

Note. All SE are >0, but are presented as 0 when below .01.
* Indicates significance using α = .05, 2-tailed

Table 13 (next page) provides the output of the false negative MPOGR, including the intercept, β, *SE*, *p*, and *ME* for all predictors. Both Black and White defendant race were positive predictors of false negatives; however, both were equal as in the false positive MPOGR (11%). This means we observed no difference in the probability of Black and White defendants being classified as low risk and being removed from pretrial supervision.

Discussion

We aimed to determine if relationships among risk classification categories were measurable. As such, we calculated a tetrachoric correlation. From our tetrachoric correlation, we found measurable relationships amongst PTRA risk classification categories. Specifically, each PTRA risk classification category significantly correlated with every other category. Being significant, this means classification categories of the PTRA are related and do not rank individuals' risk randomly. Additionally, the tetrachoric correlation addressed the second research question to determine if relationships between PTRA risk classification categories and pretrial supervision completion and removal were measurable. Relationships between PTRA risk classification categories and pretrial supervision completion and removal were measurable; however, one relationship was non-significant. The relationship between PTRA moderate/high risk classification and pretrial supervision completion may be due to sampling error. This means, unlike low, moderate, and high-risk classification categories, Moderate / High Risk may be randomly ranking individuals' risk.

We found that being a Black pretrial defendant led to higher scale scores on the PTRA than being a White defendant. This supports previous research suggesting Black defendants score higher on assessments similar to the PTRA (Desmarais et al., 2021). Moreover, previous research has suggested such findings are likely because Black persons are more likely than White persons to be arrested (Henderson et al., 2015). Still, the current research does not demonstrate such scoring is the result of racial bias. This is because, like previous research, the current research evidences that risk assessments are leastways equitable amongst Black and White criminal defendants (Bechtel et al., 2017; Bechtel et al., 2011; DeMichele, 2020; Desmarais, 2021). This is an important finding given that racial inequities are present in total scale score, but the disparity is not solely explainable by racial bias, and it is worthy of future research considerations. Still, removal from pretrial supervision is associated with racial disparity.

The four PTRA risk classification categories significantly predicted removal from pretrial supervision. This means that PTRA's prediction of which individuals will be removed from pretrial supervision is likely not random. Further, being classified as a High-Risk individual was the best predictor of removal from pretrial supervision. This supports the risk assessment tool as effective at delineating individuals of higher risk.

Prediction for removal from pretrial supervision was more likely for Black and White pretrial defendants. Being significant, this result has a 95 percent probability of not being due to sampling error. This is important, given that Black persons were predicted at a greater probability (11 percent) of removal than White persons (9 percent). As such, when controlling for PTRA risk classification, sex, court, and highest charge, being Black was a greater probability of being removed from pretrial supervision.

Previous research suggested Black defendants were more often wrongly classified than White defendants by risk assessment tools (Rembert et al., 2014; Singh & Fazel, 2010; Whiteacre, 2006). The sixth and seventh research questions aimed to determine if this notion was supported within our sample. The sixth research question was addressed using the false positive regression analysis. The results demonstrated that race did significantly predict false positives on the PTRA, but races did not differ in prediction. As such, no racial disparities were observed in predicting false positives. Last, we aimed to determine if race predicted false negatives on the PTRA by using the false negative regression analysis. Like the false positive regression, race did significantly predict false negatives; however, racial groups did not differ in prediction. This means no racial disparities were observed in predicting false negatives. Ultimately, the current results demonstrate that Black defendants have a greater probability of being removed from pretrial supervision but not of being falsely classified. Our AUC findings demonstrating that the PTRA was more accurate among Black defendants within the sample also support this assertion.

This research contributes to a growing understanding of racial equity in pretrial risk assessment instruments. As noted within the current review of the literature, research investigating racial appropriateness of risk assessment instruments is limited and contentious (Bechtel, et al. 2017; Bechtel et al., 2011; DeMichele et al., 2020; Desmarais et al., 2021). The current research is not contentious, as the evidence is clear for this population. The PTRA risk classification items collectively form a tool that predicts success in pretrial supervision. Being imperfect, the PTRA produces both false positive and false negative predictions that can have undue effects on defendants' lives. Despite this, the PTRA does not falsely classify defendants disproportionately by race for this population. Still, Black criminal defendants do experience racial disparities in removal and PTRA total scale scoring. Consequently, there is a need to better understand how race and the PTRA intersect.

TABLE 13.
False Negative Multiple Probability of Outcome Generalized Regression Output

Variable	Estimate	*SE*	*df*	*t*	*t*-crit	*ME*
Intercept	0.21*	0	285	326.79	1.97	0
Defendant Race						
Black	0.11*	0	285	292.62	1.97	0
White	0.11*	0	285	309.79	1.97	0
Defendant Sex						
Female	0.11*	0	285	291.95	1.97	0
Male	0.11*	0	285	303.15	1.97	0
Court						
Circuit	0.07*	0	285	187.48	1.97	0
District	0.07*	0	285	200.63	1.97	0
District / Circuit	0.07*	0	285	109.08	1.97	0
Defendant Highest Charge						
Animal Cruelty	0	0	285	0	1.97	0
Assault	0	0	285	0	1.97	0
Murder	0	0	285	0	1.97	0
Burglary and Theft	0	0	285	0	1.97	0
Drug Possession	0	0	285	0	1.97	0
Child Abuse	0	0	285	0	1.97	0
Disorderly Conduct	0	0	285	0	1.97	0
Driving Offenses	0	0	285	0	1.97	0
Obstruction of Justice	0.01*	0	285	6.55	1.97	0
Firearms	0	0	285	0	1.97	0
Forgery	0	0	285	0	1.97	0
Harassment	0	0	285	0	1.97	0
Intoxicated Endangerment	0	0	285	0	1.97	0.01
Property Destruction	0	0	285	0	1.97	0.01
Sexual Offenses	0	0	285	0	1.97	0
Robbery	0	0	285	0	1.97	0
Order Violation	0	0	285	0	1.97	0

Note. All SE are >0, but are presented as 0 when below .01.
* Indicates significance using α = .05, 2-tailed

Limitations

Despite the contributions of this research, there are a few limitations that must be noted. The extant pretrial literature has used varying outcome measures (i.e., rearrest, conviction, and pretrial failure), yet all of these have been shown to be directly impacted by racial/ethnic, gender, and class disproportionalities. As a result, it should be assumed that any pretrial outcome measure used, any criminal justice outcome, for that matter, potentially is exacerbated by these demographic criminal justice realities (Holsinger, Lowenkamp, & Latessa, 2006; Vincent, Chapman, & Cook, 2011). In our sample, though Whites comprised the majority, Blacks had the greatest likelihood of pretrial failure. It should be noted that Blacks represent only 5.3 percent of this county's population; thus they are overrepresented in the pretrial population and are failing. Therefore, it is plausible to assume the possibility of the overrepresentations affecting our results and ultimately biasing the degree to which the predictors affect the outcome measure (Warren, Chiricos, & Bales, 2012). In short, our results are dependent upon the outcome measure of choice. Therefore, we recommend that future research seeking to determine predictive equity be contextualized within the context of the various motivations for pretrial success and/or failure closure types (such as positive drug test, rearrests, etc.) and their potential intervening variables.

We must also note that our outcome measure is subject to potential treatment effects recommended by the assessment instrument and/or officer directives (Hosp, Hosp, & Dole, 2011). In effect, there are potential treatment implementations and supervision effects that could affect the likelihood of pretrial failure, but that we were not able to measure. Consequently, we recommend that further pretrial risk assessment validations examine the impact of treatment modalities. We also recognize an additional limitation here in the focus on intersectionality concerning race and other factors. In particular, intersectionality research often focuses on the particular, rather than the universal (e.g., "being Black" vs. "being human"), which may serve to reify racial differences at the same time racial problems are being isolated (Mitchell, 2013). The tendency of quantitative research is to focus on the particular and, as such, supplemental qualitative or mixed-methods research may be warranted, as these approaches may more adequately be able to consider both the particular *and* universal.

Recommendations

The PTRA appears suitable for predicting pretrial supervision completion, but it is imperfect, and as such, should serve in an advisory capacity to inform decisions. Therefore, we posit two recommendations: First, the use of a moderate/high risk classification should be discontinued. Untested subcategories are not recommended. Specifically, only two pretrial defendants were classified as Moderate / High Risk, and both were removed from pretrial supervision. The most accurate risk classification within the current study was High Risk, which is guideline-prescribed.

Secondly, PTRA items should be adjusted to predict relationships with specific offenses committed while under pretrial supervision. Specific offenses include 1) animal cruelty, 2) child abuse, and 3) robbery. While relationships were observed between most offenses and removal, those between removal and animal cruelty, child abuse, and robbery were non-significant, and could be a result of randomness. To inform PTRA item adjustments that better predict pretrial success, we recommend individual item-level data reporting. This would allow validation of each item within the PTRA tool. Thereafter, specific suggestions for adjustments can be posited. Further, such data is usable to assess reliability and validity of the PTRA alongside accuracy, which has been tested herein via AUC.

Conclusion

In conclusion, the PTRA functions as a risk classification tool for a particular county. Further research that evaluates the PTRA's validity and reliability will assist in the ability to understand how this instrument functions among various populations. To aid in this endeavor, administrators of risk assessment tools should both report and provide item-based data for future research to explore.

Our study finds that there are racial disparities in scoring and removal from pretrial supervision programs, but they may not warrant a return to judicial determinations of risk, primarily because we cannot hold that these disparities are the result of racial bias. Rather, research suggests other factors may better explain racial disparities here. Such disclosures are important because racial disparities are of greater societal and systemic concern. Future research should venture towards uncovering those factors impacting racial disparities in pretrial risk assessments and the extent to which bias in the instruments compares to judicial determinations of risk.

References

Advancing Pretrial Policy and Research. (2021). *Pretrial research summary*. Center for Effective Public Policy.

Barron, B. (2014). *Pretrial decision-making: How a model pretrial services program changed Allegheny County's criminal justice system*.

Bechtel, K., Holsinger, A., Lowenkamp, C., & Warren, M. (2017). A meta-analytic review of pretrial research: Risk assessment, bond type, and interventions. *American Journal of Criminal Justice, 42*(2), 443-467. https://doi.org/10.1007/s12103-016-9367-1

Bechtel, K., Lowenkamp, C., & Holsinger, A. (2011). Identifying the predictors of pretrial failure: A meta-analysis. *Federal Probation, 75*(2), 1-21. https://heinonline.org/HOL/LandingPage?handle=hein.journals/fedpro75&div=27&id=&page=

DeMichele, M., Baumgartner, P., Wenger, M., Barrick, K., & Comfort, M. (2020). Public safety assessment: Predictive utility and differential prediction by race in Kentucky. *Criminology and Public Policy, 19*(2), 409-431. https://doi.org/10.1111/1745-9133-12481

Desmarais, S. (2020). The role of risk assessment in the criminal justice system: Moving beyond a return to the status quo. *Harvard Data Science Review, 2*(1), 2-5. https://doi.org/10.1162/99608f92.181cd09f

Desmarais, S., Zottola, S., Clarke, S., & Lowder, E. (2021). Predictive validity of pretrial risk assessments: A systematic review of the literature. *Criminal Justice and Behavior, 48*(4), 398-420. https://doi.org/10.1177/0093854820932959

Desmarais, S., & Lowder, E. (2019). *Pretrial risk assessment tools: A primer for judges, prosecutors, and defense attorneys*. Safety and Justice Challenge.

Eckhouse, L., Lum, K., Conti-Cook, C., & Ciccolini, J. (2019). Layers of bias: A unified approach for understanding problems with risk assessment. *Criminal Justice and Behavior, 46*(2), 185-209. https://doi.org/10.1177/0093854818811379

Goel, S., Shroff, R., Skeem, J., & Slobogin, C. (2021). The accuracy, equity, and jurisprudence of criminal risk assessment. In *Research handbook on big data law* (pp. 14-15). Edward Elgar Publishing, 2021.

Gottfredson, D. (1999). *Effects of judges' sentencing decisions on criminal careers*. United States Department of Justice, Office of Justice Programs, National Institute of Justice. https://www.ncjrs.gov/pdffiles1/nij/178889.pdf

Guthrie, C., Rachlinski, J., & Wistrich, A. (2007). Blinking on the bench: How judges decide cases. *Cornell Law Review, 93*(1), 29-31. https://heinonline.org/HOL/LandingPage?handle=hein.journals/clqv93&div=4&id=&page=

Hanley, J., & McNeil, B. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology, 143*(1), 29-36. https://doi.org/10.1148/radiology.143.1.7063747

Henderson, H., & Miller, H. (2013). The (twice) failure of the Wisconsin risk needs assessment instrument. *Criminal Justice Policy Review, 24*(2), 198 – 220.

Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2006). Exploring the validity of the Level of Service Inventory–Revised with Native American offenders. *Journal of Criminal Justice, 34*, 331–337.

Hosp, J., Hosp, M., & Dole, J. (2011). Potential bias in predictive validity of universal screening measures across disaggregation subgroups. *School Psychology Review, 40*, 108–131.

Jackson, J., Huq, A., Bradford, B., & Tyler, T. (2013). Monopolizing force? Police legitimacy and public attitudes toward the acceptability of violence. *Psychology, Public Policy, and Law, 19*(4), 1-31. https://doi.org/10.1037/a0033852

Lipsey, M., & Wilson, D. (1993). The efficacy of psychological, educational, and behavioral treatment. *American Psychologist, 48*(12), 1181-1209. https://doi.org/10.1037/0003-066X.48.12.1181

Lovins, B., & Lovins, L. (2016). *Validation of a pretrial risk assessment tool.* Riverside Pretrial Assistance to California Counties (PACC) Project. https://www.crj.org/assets/2017/07/6_Riverside_Validation_Final_Report_5-3-16.pdf

Lowder, E., Diaz, C., Grommon, E., & Ray, B. (2021). Effects of pretrial risk assessment on release decisions and misconduct outcomes relative to practice as usual. *Journal of Criminal Justice, 73*, 1-14. https://doi.org/10.1016/j.jcrimjus.2020.101754

Mamalian, C. (2011). *State of the science of pretrial risk assessment*. Pretrial Justice Institute. United States Department of Justice. https://biblioteca.cejamericas.org/bitstream/handle/2015/5466/%5BPJI%5DStateoftheScienceofPretrialRiskAssessment%282011%29.pdf?sequence=1&isAllowed=y

Mayson, S. (2019). Bias in, bias out. *Yale Law Journal, 128*, 2218-2300. https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1010&context=aidr_collection

Mitchell, E. (2013, September 12). I am a woman and a human: A marxist-feminist critique of intersectionality theory. Unity and Struggle. Retrieved September 17, 2013, from http://gatheringforces.org/2013/09/12/iam-a-woman-and-a-human-a-marxist-feminist-critique-of-intersectionality-theory/

Picard, S., Watkins, M., Rempel, M., & Kerodal, A. (2015). *Beyond the algorithm: Pretrial reform, risk assessment, and racial fairness*. Center for Court Innovation. https://www.courtinnovation.org/sites/default/files/media/document/2019/Beyond_The_Algorithm.pdf

Rachlinksi, J., Johnson, S., Wistrich, A., & Guthrie, C. (2008). Does unconscious racial bias affect trial judges? *Notre Dame Law Review, 84*(1195), 1208-1224. https://heinonline.org/HOL/LandingPage?handle=hein.journals/tndl84&div=31&id=&page=

Rembert, D., Henderson, H., & Pirtle, D. (2014). Differential racial/ethnic predictive validity. *Youth Violence and Juvenile Justice, 12*(2), 152-166. https://doi.org/10.1177/1541204013485606

Sanchez, S., & Strenio, S. (2022). Stakeholder perceptions and potential barriers to pretrial release reform. *Journal of Restoration, Rehabilitation, and Reentry, 2022*(1), 1-19. https://digitalcommons.wou.edu/jr3/vol2022/iss1/1/

Singh, J., & Fazel, S. (2010). Forensic risk assessment: A metareview. *Criminal Justice and Behavior, 37*(9), 965-988. https://doi.org/10/1177/0093854810374274

Smith, S. (2022). *Public detention, pretrial release, and public safety.* Arnold Ventures. http://craftmediabucket.s3.amazonaws.com/uploads/AVCJIReport_PretrialDetentionPretrialReleasePublicSafety_Smith_v3-1.pdf

Vincent, G. M., Chapman, J., & Cook, N. E. (2011). Risk-needs assessment in justice: Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior*, 38, 42–62.

Warren, P., Chiricos, T., & Bales, W. (2012). The imprisonment penalty for young Black and Hispanic males: A crime-specific analysis. *Journal of Research in Crime and Delinquency, 49*, 56–80.

Whiteacre, K. (2006). Testing the Level of Service Inventory-Revised (LSI-R) for racial/ethnic bias. *Criminal Justice Policy Review, 17*(3), 330-342. https://doi.org/10.1177/0887403405284766

Zinger, I. (2004). Actuarial risk assessment and human rights: A commentary. *Canadian Journal of Criminology of Criminal Justice, 46*(5), 607-620. https://doi.org/10.3138/cjccj.46.5.607