

Validation and Predictive Bias Testing of the Public Safety Assessment for Fulton County, Georgia

Matthew DeMichele, RTI International
Stephen Tueller, RTI International
Christopher Inkpen, RTI International
Debbie Dawes, RTI International
Pamela K. Lattimore, RTI International

Table of Contents

Executive Summary	4
Introduction	5
Data Sources and Sample Description	6
Charges for Those Admitted in Fulton County Jail	8
The Public Safety Assessment Validation	9
Do Detained Individuals Score Higher than Released Individuals?	11
The PSA Validation	12
Assessing Predictive Bias	14
Understanding Base Rates	14
Subgroup Differences in Base Rates	14
Implications of Base Rate Differences	15
Calibration: Equal Rates Across Groups	16
PSA Probabilities of Pretrial Outcomes for Race and Sex Subgroups	18
Discussion	25
Conclusion	26
Appendix—A	27

Exhibits

Figure 1: Fulton County Draft Historical Validation File	6
Figure 2: PSA Scores and Outcomes by Race	16
Figure 3: PSA Scores and Outcomes by Sex	17
Figure 4: Predicted PSA Outcomes by Score and Race	20
Figure 5: Predicted PSA Outcomes by Score and Sex	24
Table 1: Charges for Individuals Admitted into the Fulton County Jail (2017-2018)	8
Table 2: Risk Factor Distributions by Detained and Released Status	9
Table 3: Average PSA Scores by Release Status	11
Table 4: Frequencies of Individuals by PSA Score and Release Status	11
Table 5: Distribution of Pretrial Sample by PSA Scale Scores	12
Table 6: Outcomes Across the Scale Scores Distributions	13
Table 7: PSA Predictive Validity (AUC) Results (N=20,209 total)	13
Table 8: Pretrial Outcome Base Rates by Race and Sex	15
Table 9: Logistic Regression Results Testing Racial Bias	19
Table 10: Logistic Regression Results Testing Sex Bias	23

Acknowledgments

The authors would like to thank several individuals in Fulton County for their willingness to share data, their time, and expertise to support RTI's research. We especially thank Marlon Greathouse, Valerie Jordan, Veronica Gadson, Terrence Harris, Felicia Pack, and Zelia Lebeau. We thank Arnold Ventures for their support of this research and extend special appreciation to Kristin Bechtel and Virginia Bersch for their insights, feedback, and direction. This project involves a large team effort that has benefitted from support by Megan Comfort, Monica Shepherd, and other members of RTI's Racial Community and Justice Committee. Megan Nyce deserves special thanks for her contributions editing and formatting several drafts of this manuscript. Any errors and all points of view are the sole responsibility of the authors

Executive Summary

In this report, we present findings from a validation and predictive bias test of the Public Safety Assessment (PSA) in Fulton County, Georgia. Fulton County is committed to criminal legal system improvements that limit the use of jail incarceration, especially the unnecessary use of pretrial detention. The jurisdiction is a member of the Advancing Pretrial Policy and Research (APPR) initiative, through which they have received Training and Technical Assistance (TTA) from a team of experts led by the Center for Effective Public Policy (CEPP) and research support from RTI. As such, Fulton County officials are engaged in efforts across several branches of their local criminal legal system to develop improvements to their pretrial system. These improvements are intended to move toward risk-based pretrial decision making, reductions in pretrial detention, and conducting ongoing research on racial disparities. A tool that holds promise for facilitating such advancements is the PSA – an assessment tool that has been associated with reductions in missed court appearances and new crimes, and fewer admissions.¹ Central to Fulton County’s recent efforts are to determine whether the PSA is a tool that fits their jurisdiction (i.e., is it valid in the local context?).

The PSA is intended to provide court actors with information when making pretrial release decisions by estimating the likelihood (i.e., predicted probability) of missing a court appearance, being arrested for a new crime during pretrial release, and being arrested for a new violent crime during pretrial release. The PSA consists of 9 factors used across three scales to measure likelihood that someone will miss court with a failure to appear (FTA), and two new arrest outcomes including any new criminal arrest (NCA) and a new violent criminal arrest (NVCA). Scoring the PSA produces three scales: one per outcome (FTA, NCA, NVCA) that range from 1- 6, with 1 indicating the lowest likelihood and 6 the highest likelihood for each of the outcomes. The current report focuses on a validation and predictive bias testing of the PSA using data from adults booked into the Fulton County jail on a new charge from January 1, 2017 through December 31, 2018. On average, individuals in Fulton County score about 3 on the FTA and NCA and 2 on the NVCA (table 3), with nearly 60% of the pretrial sample scoring between 1-2 on the FTA and NCA scale and nearly 80% scoring 1-2 on the NVCA scale.

One primary concern with pretrial release assessment instruments is that people of color and females will be scored too high relative to their actual likelihood of missing a court date or committing a new crime. (i.e., that if conditions of release are predicated on risk scores, people of color or females may be disadvantaged). The results show that in Fulton County, the PSA is associated with pretrial outcomes, and we did not find evidence that the PSA exacerbates predictive bias related to race and sex. We do not find evidence that people of color are being scored higher than their actual outcome rates (i.e., no overprediction). Proper use of assessments requires ongoing research to ensure that the PSA remains valid for Fulton County.

¹ Lowenkamp, C., DeMichele, M., and Klein Warren, L. (2020). Replication and extension of the Lucas County PSA project. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3727443

Introduction

Pretrial assessments are tools that identify the statistical likelihood that individuals will appear in court or engage in new criminal activity if they are not held in jail while awaiting trial following a criminal arrest. Judges routinely make assumptions about the likelihood of someone returning to court or committing a new crime when making release decisions. Without a pretrial assessment, judges make ‘on the spot’ risk calculations by weighing the importance of prior criminal history, community ties, the charge, and other factors they value based on their judicial experience. Pretrial assessments add transparency by specifying the factors used for pretrial decisions, and they standardize the importance (i.e., weight) given to each factor for making pretrial decisions. Although assessments come with several advantages, they should not be implemented or used without proper research and monitoring.

When using pretrial assessments, jurisdictions should engage in ongoing research to ensure the assessment is appropriate, or “valid,” for their jurisdiction. To that end, RTI is working with Fulton County, Georgia to conduct an historical validation and bias assessment with respect to race and sex of the Public Safety Assessment (PSA). The PSA uses nine factors to produce three scores that indicate the likelihood that an individual with specific characteristics will fail to appear to court (FTA), experience a new criminal arrest (NCA), and experience a new violent criminal arrest (NVCA).² A central reason for conducting local validation and predictive bias studies is because of the uniqueness of each local jurisdiction. Although the PSA is proving to work well across jurisdictions, it is good practice to validate the PSA’s performance in each jurisdiction prior to implementation.

The purpose of this validation is to provide Fulton County with information about the predictive validity and bias with the PSA as they decide whether to use the PSA to inform pretrial release decisions. All jurisdictions are unique, so it is important to ensure that the PSA has appropriate predictive validity and does not exacerbate bias for the Fulton County pretrial population. This means we need to ensure that the set of factors used to generate PSA scores are strongly associated with each of the outcomes, and we find equal probabilities of outcomes across race and sex.

In this report, we provide a summary of our validation and tests for predictive bias of the PSA with respect to race and sex in Fulton County, Georgia. Prior PSA research in multiple jurisdictions demonstrates that the PSA meets standards of appropriate levels of predictive validity for FTA, NCA, and NVCA and does not identify any evidence of serious predictive bias.³ The current study addresses whether the PSA is valid for use in Fulton County, and whether there is evidence of predictive bias in PSA scoring and predictions. We begin by describing how we created the validation dataset and reviewing some general characteristics of individuals admitted into Fulton County’s jail.

2 The PSA factors include whether someone has a current or prior violent charge or conviction. Additionally, the PSA is unique from other pretrial risk assessments because it provides an estimate of one’s likelihood to be arrested for a violent crime during pretrial release. The list of offenses that constituted the violent crime list was developed by Fulton County justice system professionals who identified the charge codes that are considered “violent” in Fulton County. This list is included in Appendix Table 1. Arnold Ventures also has a set of 49 offenses that have identified as the “violent crime list” through prior research with other jurisdictions by Arnold Ventures (i.e., the AV list). The Fulton County violent list includes slightly fewer offenses defined as violent and thus reports a lower amount of new violent arrests compared to the AV list. RTI conducted analyses using both the Fulton list and the AV list; results did not vary substantially between the two sets of analyses and, thus, only the analyses using the Fulton County list are reported here. The results from the other analyses are available upon request.

3 Desmarais, S., Monahan, J., and Austin, J. (2021). The empirical case for pretrial risk assessment instruments. *Criminal Justice and Behavior*, 49(6): 807-816. Lowenkamp, C., DeMichele, M., and Klein Warren, L. (2020). Replication and extension of the Lucas County PSA project. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3727443

Data Sources and Sample Description

The historical validation used cases drawn from all admissions into Fulton County Jail between January 1, 2017, and December 31, 2018. The admission data were merged with court and criminal history records to develop the validation sample dataset. Outcomes were tracked in the court and criminal history data through December 31, 2019, to provide time for (most) cases to reach disposition.

Figure 1 shows how the pretrial validation sample was identified and developed for our analyses. Fulton County is considering using the PSA for all adults admitted into jail on a new criminal arrest. Thus, to create the validation dataset, we removed the data for all those who did not meet this description.

Figure 1: Fulton County Draft Historical Validation File

Filtering criteria

The following bookings are omitted from PSA factor eligibility:

Post-trial bookings:	5,055
Probation/parole only:	1,456
Process-related charges ⁱ :	3,141
Juvenile defendants:	865

Sub-total of filtered bookings: 10,517

Bookings are omitted due to the following considerations:

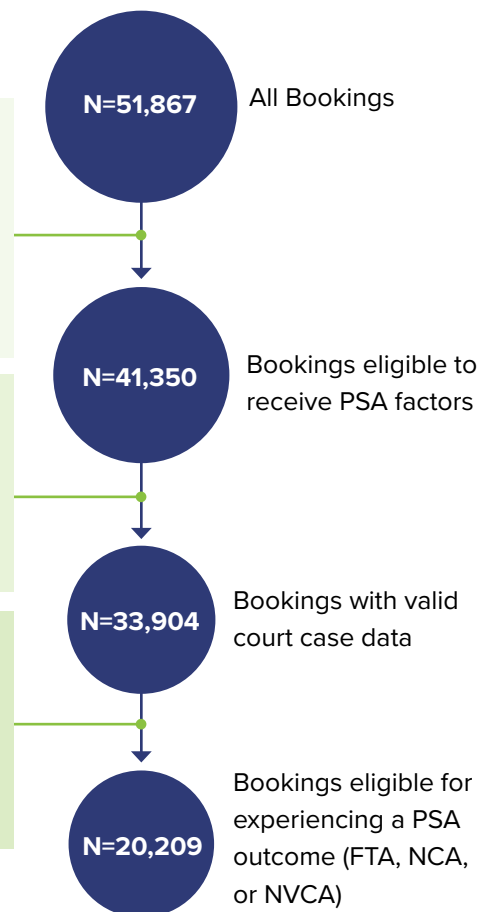
Lack of corresponding court case ⁱⁱ :	5,568
Subsequent booking for court case ⁱⁱⁱ :	1,067
Insufficient data on charge dispositions ^{iv} :	811

Sub-total of Filtered bookings: 7,446

Bookings omitted from outcome analysis:

Detained pretrial:	11,164
No court data on disposition ^v :	10
Pretrial release reason indicates detention, or other status ^{vi} :	2,521

Sub-total of Filtered bookings: 13,695



Date range for cohort: All bookings from 2017-2018

Additional documentation

ⁱ Bookings that only have process-related charges associated with the booking refer to foreign county warrants, correctional transfers, ICE detainees, bookings related to a criminal appeal, or transfers to programs (e.g. substance abuse treatment programs). These are omitted from analyses as they do not have new criminal charges that can be linked to a court case.

ⁱⁱ These bookings are omitted for outcome analysis due to a lack of having an associated court case with any of the charges linked to the booking. Without information on the outcomes for these charges, it is not feasible to assert that the bookings and subsequent releases were pretrial in nature nor is it possible to include them for outcome analyses given that there is no date range for the charges being dropped.

ⁱⁱⁱ These bookings are omitted from the outcome analysis as they were subsequent (second or higher order) bookings associated with a court case during the date range for the outcome study. The first bookings associated with the court case are already included in the cohort.

^{iv} These bookings have a booking status derived from information about prosecutor decisions on relevant charges. However, these 811 bookings lack dispositions for charges associated with the booking v These bookings have court cases associated with the charges but do not have a disposition or hearing date that can be used to create a date range within which an outcome (e.g. new criminal activity or failure to appear) can occur.

^v The release reasons associated with these bookings indicate that despite being released, the person in question may have been transferred to detention elsewhere or returned to a different law enforcement agency. For this reason, these bookings are omitted from the outcome analysis as they may not be at risk to commit new criminal activity or fail to appear.

There were 51,867 admissions into the jail in 2017-2018. Of these, 10,517 cases were removed because they were post-trial admissions (5,055), admissions for juvenile defendants (865), admissions for probation or parole violations only (1,456) or were for other process-related charges (3,141). An additional 7,446 cases were removed due to data quality issues. All data processing decisions were coordinated with Fulton County officials to ensure that our decisions aligned with local policies and practice. After removing these cases, the remaining pretrial sample totaled 33,904 individuals. The final step in producing the validation file was to identify people who were released pretrial and, thus, could experience one or more of the PSA outcomes (FTA, NCA, and NVCA). The final validation data file included 20,209 individuals who were admitted into and released from the Fulton County jail during the between January 1, 2017, and December 31, 2018.

Charges for Those Admitted in Fulton County Jail

Table 1 provides information on the charges for the people admitted into the Fulton County jail between January 1, 2017, and December 31, 2018. The information is for all PSA-eligible individuals—the released and detained pretrial sample (N = 33,904). About 60% of the individuals admitted into Fulton County jail were released pretrial. Most of the people admitted into the jail were people of color⁴ (N = 28,850, 85%), about 35 years old on average, and male (N = 26,458, 78%).

People in Fulton County were admitted for a variety of different charge categories. In order of prevalence, the most serious charge associated with an admission was classified as a violent (33%), property (31%), drug (19%), public order (10%), other arrest-related (8%), or other charge.⁵ Although the average length of stay in the Fulton County jail was 34 days, people who were detained pretrial spent an average 69 days in jail. By contrast, people who were released pretrial spent an average of 10 days in jail, and nearly 50% of them were released within 4 days.

Table 1: Charges for Individuals Admitted into the Fulton County Jail (2017-2018)

Most serious charge overall category	Total admitted N (percentage)	Released N (percentage)
Arrested – Other	2,545 (7.5%)	1873 (9.3%)
Drug Offenses	6,324 (18.7%)	3,979 (19.7%)
Other Offenses	29 (0.1%)	20 (0.1%)
Property Offenses	10,502 (31.0%)	5,760 (28.5%)
Public Order Offenses	3,331 (9.8%)	1,816 (9%)
Violent Offenses	11,173 (33.0%)	6,761 (33.5%)
Total	33,904 (100%)	20,209 (100%)

Most released individuals were successful during pretrial as about 84%, 76%, and 93% did not have an FTA, an NCA, or an NVCA respectively.

⁴ Our categorization of people's racial and ethnic identities relies on the categories used in the Fulton County criminal legal system data. In this report, we categorize people as "White" if that is only racial/ethnic identifier available for them. People who are identified as belonging to all other racial/ethnic categories are categorized as "people of color." For more on the complexities of these issues, see [Using Data to Explore Racial and Ethnic Disparities in the Criminal Legal System](#).

The restriction of gender categories to male and female in most criminal legal system data results in the erasure of people whose gender identity does not correspond to these biology-based definitions. We acknowledge this limitation of our analyses and highlight the need for more inclusive data collection (see Greene 2018 for a discussion of how such "categorical exclusions" reproduce hardship through the denial of gender-segregated resources).

⁵ Charge categories are derived from the National Corrections Reporting Program broad charge categories using the most Serious offense charged for everyone.

PSA Validation

The Public Safety Assessment Validation

The PSA was created to inform pretrial decisions about the likelihood of FTAs, NCAs, and NVCA, so we focus on these three outcomes in this study.⁶ We understand that pretrial agencies and courts are interested in other important outcomes besides missed court appearances and new arrests, but a key step to using a pretrial assessment is to identify the specific purpose for the instrument.

The PSA consists of 9 factors which can be found in administrative databases, does not require an interview, and excludes items about socioeconomic factors to reduce disparities. This means that RTI can score the PSA by extracting and linking data from administrative records from the jail, courts, and the criminal history repository.

Table 2 shows the number and proportion of individuals that have each of the PSA risk factors. Some of the 9 factors are used for each of the three PSA scales—as indicated in the three righthand columns in Table 2—to produce one score per outcome— an FTA scale and score, an NCA scale and score, and an NVCA scale and score. Each scale ranges from 1 to 6 indicating lowest to highest probability of experiencing one of the outcomes. In this report, when we refer to risk, we are only referring to the probability of someone experiencing one of these outcomes.

Table 2: Risk Factor Distributions by Detained and Released Status

PSA Factor	Factor Labels	Overall N(%)	Detained N(%)	Released N(%)	FTA	NCA	NVCA
1. Age at current arrest	<23	4,918(14.5)	1,697(12.4)	3,221(15.9)		X	
	>=23	28,986(85.5)	11,998(87.6)	16,988(84.1)			
2. Current violent Offense ⁷	No	22,743(67.1)	9,149(66.8)	13,594(67.3)			X
	Yes	11,161(32.9)	4,546(33.2)	6,615(32.7)			
2a. Current violent offense and <= 20 years old	No	32,896(97.0)	13,270(96.9)	19,626(97.1)			X
	Yes	1,008(3.0)	425(3.1)	583(2.9)			
3. Pending charge at the time of the arrest	No	26,942(79.5)	10,054(73.4)	16,888(83.6)	X	X	X
	Yes	6,962(20.5)	3,641(26.6)	3,321(16.4)			
4. Prior misdemeanor conviction	No	14,883(43.9)	4,115(30.0)	10,768(53.3)		X	
	Yes	19,021(56.1)	9,580(70.0)	9,441(46.7)			

⁶ Coding of violent charges were selected by members of the Fulton County staff participating in PSA validation efforts

⁷ Factor 2 coding for current violent offense at booking uses a list of statutes developed by Fulton County to designate a charge as “violent” or “non-violent” (see Table 1 in Appendix). Differences between this list of statutes and the charges that are designated as a Violent Offense in the National Corrections Reporting Program explain the discrepancy between the count of admissions with associated violent charges listed in tables 2 and 3.

PSA Factor	Factor Labels	Overall N(%)	Detained N(%)	Released N(%)	FTA	NCA	NVCA
5. Prior felony conviction	No	20,956(61.8)	6,479(47.3)	14,477(71.6)		X	
	Yes	12,948(38.2)	7,216(52.7)	5,732(28.4)			
5a. Prior conviction (misdemeanor or felony)	No	13,414(39.6)	3,479(25.4)	9,935(49.2)	X		X
	Yes	20,490(60.4)	10,216(74.6)	10,274(50.8)			
6. Prior violent conviction	No	23,237(68.5)	7,655(55.9)	15,582(77.1)		X	X
	Yes, 1 or 2	4,436(13.1)	2,235(16.3)	2,201(10.9)			
	Yes, 3 or more	6,231(18.4)	3,805(27.8)	2,426(12.0)			
7. Prior FTA in the past 2 years	No	24,865(73.3)	9,231(67.4)	15,634(77.4)	X	X	
	Yes, just 1	4,087(12.1)	2,044(14.9)	2,043(10.1)			
	Yes, 2 or more	4,952(14.6)	2,420(17.7)	2,532(12.5)			
8. Prior FTA older than 2 years	No	22,798(67.2)	7,793(56.9)	15,005(74.2)	X		
	Yes	11,106(32.8)	5,902(43.1)	5,204(25.8)			
9. Prior sentence to incarceration >= 14 days	No	23,323(68.8)	7,503(54.8)	15,820(78.3)		X	
	Yes	10,581(31.2)	6,192(45.2)	4,389(21.7)			

Note. An “X” in the last three columns indicates that the indicated PSA factor is included in the PSA scale for the indicated pretrial outcome: FTA (failure to appear), NCA (new criminal arrest), or NVCA (new violent criminal arrest).

Table 2 includes important information to understand how people score on each of the factors, which is the foundation for the rest of the study. Pretrial assessments rely on the accumulation of the presence of specific factors such that more factors (generally) equate to higher scores and higher scores (generally) equate to increased likelihood of missed court or arrest for a new crime.

About 33% of the sample (N = 11,161) have a current violent charge and 21% have a pending charge at the time of this booking. About 60% (N = 20,490) have a conviction for a prior crime, including 56% who have a prior misdemeanor conviction (N = 19,021) and 38% who have a prior felony conviction (N = 12,948). About 31% have a prior violent conviction (N = 10,667). There are two factors included in the PSA scales related to prior FTAs. This sample included 27% (N = 9,039) who had an FTA in the past 2 years and 33% (N = 11,106) who had an FTA earlier than the past 2 years.

Using these data, we calculated the three PSA scores for everyone in the pretrial sample. Everyone received a score of between 1 and 6 for each of the three PSA outcomes—FTA, NCA, and NVCA. Table 2 shows more extensive criminal histories for the detained compared to the released.

Do Detained Individuals Score Higher than Released Individuals?

The first question to address once the scores are calculated is whether there are differences in scores for people who were released pretrial and people who were not released. Recall these are historical data, and the PSA was not used in Fulton during this period, so the PSA did not inform the release decisions for the individuals in the sample. Table 3 shows that detained individuals had higher average PSA scores than those who were released. This suggests that even without an assessment tool release decisions in Fulton County, on average, did reflect the likelihood of FTA, NCA, and NVCA. The Table includes results for a statistical test known as Cohen's d that shows moderate differences in the FTA and NCA average scores and small differences in the NVCA scores between the released and detained groups.

Table 3: Average PSA Scores by Release Status

Release Status	Average PSA Score (scores range 1-6)		
	FTA	NCA	NVCA
Detained	3.06	3.30	2.28
Released	2.41	2.47	1.87
Cohen's d	0.46	0.57	0.37

Average scores are important, as is the fact that decisionmakers in Fulton County are more likely to release individuals who would have lower PSA scores. However, the differences in average scores are not large, suggesting that many individuals with low scores were detained and many individuals with high scores were released. Table 3 provides the FTA, NCA, and NVCA scale scores for the detained and released samples. What is apparent is how many lower scoring individuals were detained. This is especially important because we defined release as anyone released at any point prior to their final case disposition, which means that the detained individuals are people who were held in jail throughout their entire pretrial period (60 days on average for individuals who were detained pretrial).

Table 4: Frequencies of Individuals by PSA Score and Release Status

Release Status	Score	FTA ¹		NCA		NVCA	
		Detained	Released	Detained	Released	Detained	Released
Fulton ²	1	2,299	7,437	1,819	5,698	4,476	9,564
	2	2,909	4,362	2,535	6,220	4,422	6,026
	3	3,360	3,706	2,767	3,462	2,106	2,779
	4	2,533	2,528	3,760	3,150	1,889	1,441
	5	1,997	1,804	1,954	1,213	778	392
	6	597	372	860	466	*	*

1. The frequencies for FTA do not change based on the violent list because none of the violent PSA factors are used to compute the FTA score.

2. The Fulton violent list, used in the creation of factors 2 and 6 along with the NVCA measure, was used (See Exhibit A4 in the Appendix for list of charges included as violent)

* Cell sizes less than 16 are not displayed as statistics calculated from small cells are unreliable.

People appear to have been released across the risk continuum as there is not a clear risk threshold. Instead, some people scoring low were detained and some people scoring high were released.⁸ For example, 2,299 people with a score of 1 on the FTA scale were detained for their entire pretrial period, while there were more than 8,500 people detained until disposition who scored between 1 and 3 on the FTA scale. As there are 13,695 people who were detained pretrial, nearly two-thirds of these individuals scored a 3 or less on the FTA scale. There is a similar pattern across the other scales, with many individuals who scored less than 3 on the NCA and NVCA scales having been detained until case disposition in Fulton County. These results suggest that there are opportunities for a pretrial assessment to assist with improving release decisions in Fulton County.

The PSA Validation

The previous sections described the entire pretrial sample, but the validation study focuses only on those individuals who were released pretrial and who could have experienced an FTA, NCA, or NVCA while awaiting disposition of their case. To develop the validation dataset, we removed the data for the 13,695 individuals detained for their entire pretrial period, resulting in a validation sample of 20,209.

The first consideration for the validation is to examine the risk distribution of cases across the validation sample. In other words, what are the number and proportion of cases that had each of the scale scores. Table 5 shows that most people who were released pretrial have low PSA scores, with more than half of the individuals scoring at or below a 2 across each of the scales. Specifically, about 58% had scores of 1 or 2 on the FTA scale, 59% had scores of 1 or 2 on the NCA scale, and 77% had scores of 1 or 2 on the NVCA scale. Relatedly, few people who were released had high scores of 5 or 6—only about 11% for the FTA scale, 8% for the NCA scale, and 2% for the NVCA scale.

Table 5:
Distribution of
Pretrial Sample
by PSA Scale
Scores

Scale Score	FTA		NCA		NVCA	
	N	%	N	%	N	%
1	7,437	36.80	5,698	28.20	9,564	47.30
2	4,362	21.60	6,220	30.80	6,026	29.80
3	3,706	18.30	3,462	17.10	2,779	13.80
4	2,528	12.50	3,150	15.60	1,441	7.10
5	1,804	8.90	1,213	6.00	392	1.90
6	372	1.80	466	2.30	7	0.03
Total	20,209	100%	20,209	100%	20,209	100%

The next consideration is to assure that the outcome rates increase as the scores increase—as higher scores are expected to be associated with higher risk levels. In other words, there should be higher success rates for lower scores and lower success rates for higher scores. Table 6 shows that for the most part the success rates decrease as the score increases (e.g., the percentage of cases without an FTA goes down as the score goes up—91% of those with an FTA score of 1 had no FTA compared to 61% of those with an FTA score of 6. There is, however, some clustering of FTA outcome rates in the mid-range scores of 3 to 5. The NCA and NVCA scales show stronger relationships between the scores and outcomes. Across

⁸ Kleinberg

the NCA scores, the outcome rate (i.e., no NCA) ranges from about 13 (score = 1) to 52% (score = 6). Across the NVCA scores, the outcome rate ranges from 4% to 29%. Overall, we see that the PSA scores work as intended—with higher scores associated with more risk or, equivalently, less success.

Table 6:
Outcomes Across
the Scale Scores
Distributions

Scale Score	FTA	% FTA	NCA	% NCA	NVCA	% NVCA
1	687	9.20	748	13.10	403	4.20
2	717	16.40	1261	20.80	454	7.50
3	777	21.00	1012	29.20	312	11.20
4	540	21.30	1141	36.20	246	17.10
5	438	24.30	543	44.80	84	21.40
6	146	39.30	241	51.70	2	28.60
Total	3305	16.35	4946	24.47	1501	7.43

We now turn to statistics that are used for instrument validation studies. Validation studies usually include estimation and references to the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC). The AUC is one way to evaluate how well scales perform. This evaluation metric shows how well an assessment works by plotting the correct high-risk predictions or true positives against the incorrect high-risk predictions or false positives at each of the scale scores. The higher the AUC value the better the assessment performs at distinguishing between individuals with and without an outcome. The AUC ranges between 0 and 1, with 0 meaning that the assessment predicts all successful cases as failures and all cases that fail as successes (i.e., it is always wrong). Conversely, an AUC of 1 means the assessment can correctly distinguish all successes and failures—providing a perfect prediction of outcomes. Finally, an AUC of 0.5 means the assessment does not distinguish between successful and unsuccessful cases and the assessment is no better than chance. Thus, for an assessment to be useful, the AUC value should be larger than 0.5 and the closer to 1 the better the assessment is.

Table 7 shows the AUCs are 0.62 or 0.65 for all three assessment scales⁹. All the AUC values using the Fulton data are similar to the national sample (see the last row) and are similar to other pretrial risk assessments which find AUC scores in the range of 0.6 to 0.7. The AUCs for the PSA are in the fair range and within acceptable predictive ranges.

Table 7: PSA
Predictive Validity
(AUC) Results
(N=20,209 total)

	FTA	NCA	NVCA
N with the outcome (%)	3,305 (16.4%)	4,946 (24.5%)	1,501 (7.4%)
AUC	0.62	0.65	0.65

⁹ An AUC of 0.65 can also be understood to mean that if you select two random individuals from the dataset and one had an outcome, and the other did not that there is a 65 % chance the person with the outcome had a higher score.

Assessing Predictive Bias

Once validity of the PSA is assessed (and confirmed), the next consideration is to test for predictive bias. Bias testing is a process to determine whether the PSA provides equal results for different race and sex groups. Assessing predictive bias is complex as it can be measured in multiple ways. The purpose of testing for predictive bias is to determine whether the chances of someone having an outcome are similar across subgroups (i.e., for the PSA, for sex and race subgroups). This means, for example, that a Black person with a score of 2 on the FTA scale will have the same chance of an FTA as a White person with a score of 2. More generally, for any score on each of the three risk scales, the probability of the outcome should be equal regardless of race or sex. If the probabilities of success within risk score differ among different race or sex groups using the PSA could create ethical and practical challenges related to decisions about detention and supervision conditions.¹⁰

Understanding Base Rates

In this report, base rates are the prevalence rate of each of the three pretrial outcomes in Fulton County. The base rate converts the number or count of each outcome to a proportion of the released population. Or, more specifically, each base rate is the percentage of people in the validation sample who had an FTA, NCA, or NVCA during the pretrial period. Table 9 shows that the overall base rates are 16.4%, 24.5%, and 7.4% for FTA, NCA, and NVCA, respectively. Although the overall base rates provide information about the pretrial outcomes in Fulton County, the overall base rates do not reveal anything about predictive bias.

Subgroup Differences in Base Rates

For predictive bias testing we need to understand if and how base rates differ between the subgroups of interest. For example, are there differences in FTA rates between White individuals and people of color or males and females? Base rates are unlikely to be identical between groups, so we use statistical tests to determine whether differences in the FTA, NCA, and NVCA base rates between White individuals and people of color and males and females are “significant.” Specifically, statistical significance is used to provide confidence that observed differences between groups are “real” or true differences and not merely a difference that could be expected to be observed due to chance.

To determine significance, we apply a threshold known as a p-value (see Table 8, column 7). For the Fulton analyses, we use a p-value of ≤ 0.001 that provides strong confidence that any identified differences are true differences between the groups and not artifacts of the samples.¹¹ With this standard, Table 8 shows that there are, in fact, differences in the base rates for the three pretrial outcomes for our subgroups. People of color and males have significantly higher base rates for FTAs, NCAs, and NVCA than White individuals and females.

¹⁰ To provide a hypothetical example, suppose that males with a score of 4 on the NCA had a 30% chance of experiencing a new arrest during the pretrial period and females had a 10% chance. If conditions of release are based on the score of 4, females posing much lower risk to public safety would be exposed to more substantial conditions than males.

¹¹ The 0.001 threshold means that there is less than a 0.1% chance that a difference between groups is by chance.

Table 8: Pretrial Outcome Base Rates by Race and Sex

Scale	Outcome	People of Color (N=17,040)	White (N=3,169)	All Released (N=20,209)	Group Difference	p- value
FTA	Yes FTA	2855 (16.8%)	450 (14.2%)	3305 (16.4%)	2.6%	<0.001
NCA	Yes NCA	4299 (25.2%)	647 (20.4%)	4946 (24.5%)	4.8%	<0.001
NVCA	Yes NVCA	1349 (7.9%)	152 (4.8%)	1501 (7.4%)	3.1%	<0.001
		Male (N=14,831)	Female (N=5,375)	All Released (N=20,209)	Group Difference	p- value
FTA	Yes FTA	2589 (17.5%)	715 (13.3%)	3305 (16.4%)	4.2%	<0.001
NCA	Yes NCA	3957 (26.7%)	988 (18.4%)	4946 (24.5%)	8.3%	<0.001
NVCA	Yes NVCA	1192 (8.0%)	309 (5.7%)	1501 (7.4%)	2.3%	<0.001

Note that we are unable to say why there are differences in base rates among these groups and whether these differences fully reflect actual differences in the outcomes that are of concern to the PSA. Criminal arrests are only proxies for actual criminalized behavior and reflect not only individual behavior (or suspicion of behavior) but the actions and responses of law enforcement and criminal legal system agents. Prior research shows that neighborhoods that are predominately populated by people of color are policed more heavily, have higher prevalence of violent crimes, and people of color are more likely to be convicted than their White counterparts. Additionally, research shows that males tend to be more likely to engage in criminalized acts, display more risky behavior, and are more prone to violence. Further, missing court appearance does not necessarily imply willful absences, but instead may reflect differences in resources such as transportation or childcare that may make it more difficult for individuals of some groups to successfully meet court appearances.¹²

Implications of Base Rate Differences

Base rate differences in PSA outcomes in the Fulton County data do not indicate predictive bias. However, base rate differences are important for testing predictive bias because when base rates differ there are some bias tests that are inappropriate.¹³ Technically, differences in base rates between groups result in differences in the possible classification errors between subgroups, which means that error metrics are not a good metric to consider when assessing predictive bias.

¹² Kohler-Hausmann, I. (2020). "Nudging people to court." *Science* 370(6517): 658-659.

¹³ Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2), 153–163. DeMichele, M. and Baumgartner, P. (2021). Bias testing of the Public Safety Assessment: Error rate balance between Whites and Blacks for new arrests. *Crime and Delinquency*, 1-26. Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *Quarterly Journal of Economics*, 133, 237–293.

Calibration: Equal Rates Across Groups

Calibration examines whether equal proportions of the subgroups (e.g., race, sex) have the outcomes of interest (i.e., FTA, NCA, NVCA). Thus, calibration is an important check for predictive bias by determining if the PSA scores and outcomes are related to one another in similar ways across, in our case, race and sex—in other words, do we see similar patterns in the PSA scores and outcomes for people of color and White individuals and males and females?

Figure 2 shows the proportions of people of color and White groups that experienced FTAs, NCAs, and NVCA for each score (1 through 6). Two issues are important when examining these plots. First, is there an increase in risk (proportion of failures) as the scores increase? Second, are there similar patterns between people of color and White individuals (and males and females in Figure 3). The shaded areas show the 99% confidence intervals around the point estimates for each score level—if the blue and red areas overlap, difference between the point estimates for each group (shown by the dots on the lines) for a score is not statistically significant.

Figure 2 suggests that the PSA achieves both criteria for the two race subgroups: (1) higher scores are associated with higher rates of the outcomes and (2) there is little difference between results for people of color and White individuals. For an FTA score of 1, between 7% and 10% of White individuals and people of color experienced an FTA and there are small increases along the FTA scale with some flattening between scores of 4 and 5. Importantly for examination of predictive bias, the scores do not differ significantly across the scale scores between groups. For the NCA scales, Figure 2 shows about a 5%-to-10%-point increase in NCA rate between each score for the NCA scale with about 10% of those with NCA scores of 1 and 50% of those with NCA scores of 6 experiencing a new arrest. Again, there are no significant differences in the relationship by race. Finally, Figure 2 shows that the NVCA scale demonstrates similar patterns to the other outcomes, albeit with much lower overall rates.

Figure 2: PSA Scores and Outcomes by Race

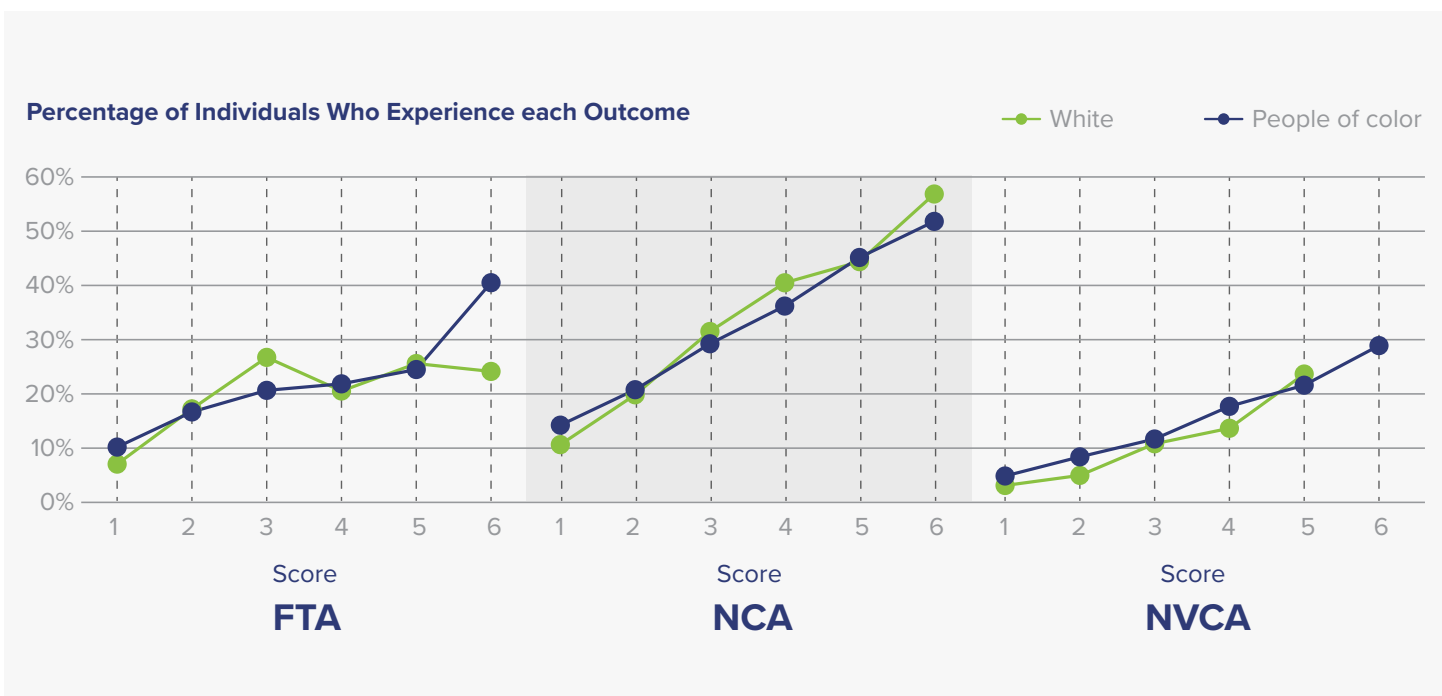
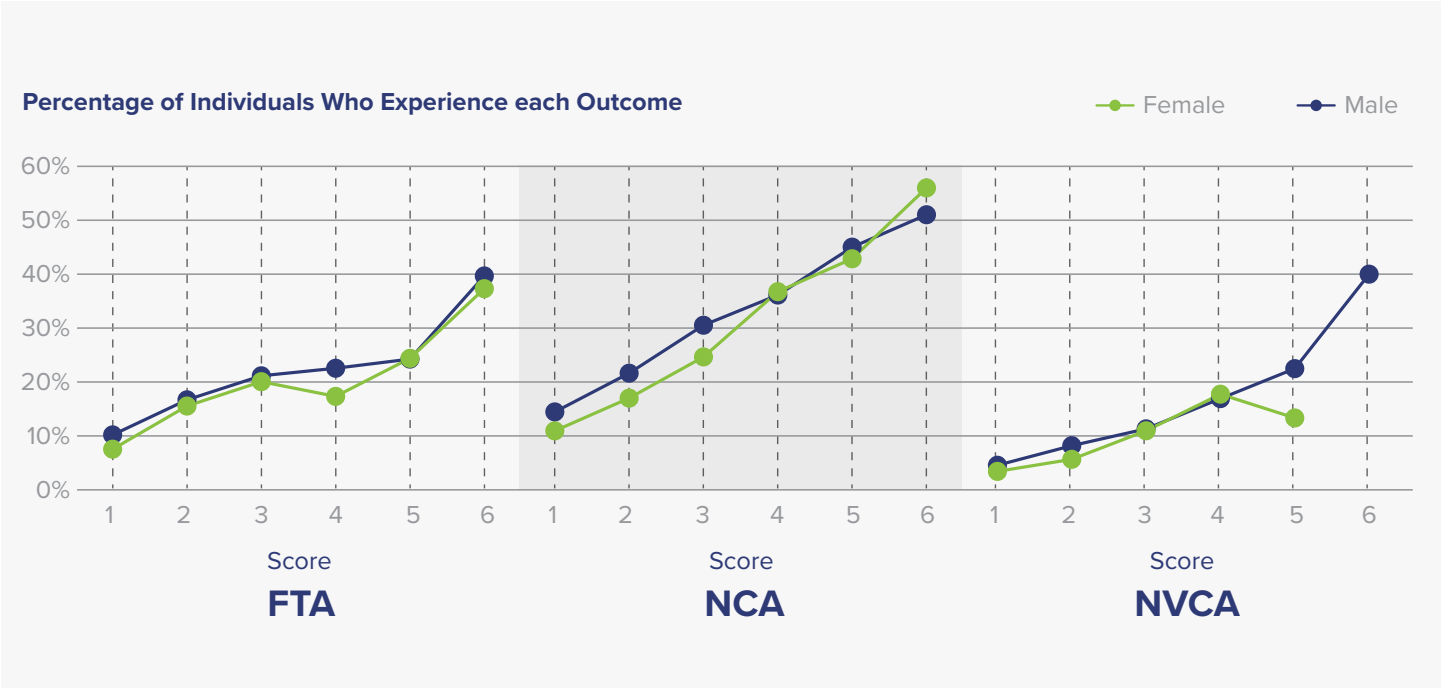


Figure 3 shows the same information as Figure 2 but for male and female individuals. Again, the results suggest that FTAs, NCAs, and NVCAs increase as PSA scores increase across both groups and the relationships between scores and outcomes are similar by sex groups.

Figure 3: PSA Scores and Outcomes by Sex



PSA Probabilities of Pretrial Outcomes for Race and Sex Subgroups

So far, we have shown that the PSA provides good classification by race and sex and that there is little difference in the reported outcomes between race and sex groups. In this section, we share results of statistical analyses that were conducted to determine whether the PSA predicts equal probabilities of each of the outcomes for race and sex subgroups. To address this question, we use logistic regression, which is a statistical procedure that estimates the likelihood or probability of an event happening based on a set of factors or variables. Results from the logistic regression analyses indicate whether the PSA scales predict the pretrial outcomes and whether there are strong relationships between risk, race, sex, and the outcomes that would suggest that the PSA scales are biased. Our approach, essentially, seeks to understand whether the PSA is providing different results for people of color (compared to White individuals) or for females (compared to males).

RESULTS FOR RACIAL SUBGROUPS

Table 9 provides the results of a set of logistic regression models estimated with the Fulton County data.¹⁴ The dependent variable in each model is the zero-one indicator of whether the event was observed (e.g., for the FTA model, the dependent variable (FTA) equals 1 if the individual had an FTA or zero if the individual did not have an FTA). Four models are shown for each of the three PSA scales (FTA, NCA, NVCA). Model 1 includes only race (White = 1; people of color = 0) as a covariate—showing the direct relationship between race and the observed outcome. Model 2 includes only the relevant scale score (i.e., FTA score for the FTA model, NCA score for the NCA model, or NVCA score for the NVCA model) as a covariate—showing the direct relationship between the score and the observed outcome. Model 3 includes both the race and scale score—testing the relationship of both variables to the outcome simultaneously. Model 4 includes race, the scale score, and an interaction term (White*Score) that tests whether there is a differential effect of the score by race.¹⁵

Results from Model 1 show statistically significant relationships between race and each of the outcomes (e.g., an odds ratio of 0.822 for White for the FTA Scale Model 1). For all three models, White individuals are predicted to have a smaller likelihood of the outcomes ($p < .001$). This result is not surprising as it reflects the differences in base rates that were observed and discussed in table 9.

14 The values in Table 11 are odds ratios which provide an indication of the direction of a relationship. An odds ratio less than 1 means higher values for that variable are associated with less risk (i.e., less likelihood) of the outcome occurring. An odds ratio Greater than 1 means higher values for that variable are associated with more risk (i.e., greater likelihood) of the outcome occurring. The p-value indicates whether the odds ratio is significantly different from 1. If the odds ratio isn't significantly different from 1 then there is no relationship between the variable and the outcome. As before, because of the size of the dataset, we are using a p-value of ≤ 0.001 to indicate statistical significance.

15 This approach is used by several researchers: DeMichele, M., Baumgartner, P., Wenger, M., Barrick, K., Comfort, M., and Misra, S. (2020). The Public Safety Assessment: A re-validation and assessment of predictive utility and differential prediction by race and gender in Kentucky. *Criminology & Public Policy*, 19(2), 409–431. Flores, A., Bechtel, K., and Lowenkamp, C. (2016). False positives, false negatives, and false analyses: A rejoinders 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. Skeem, J. L., and Lowenkamp, C. T. (2016). Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology*, 54, 680–712.

Table 9: Logistic Regression Results Testing Racial Bias

Variable	Model 1		Model 2		Model 3		Model 4	
	Odds Ratio	p-value	Odds Ratio	p-value	Odds Ratio	p-value	Odds Ratio	p-value
FTA Scale								
White	0.822	<0.001			0.937	0.243	0.742	0.009
FTA Score			1.344	<0.001	1.342	<0.001	1.326	<0.001
White X FTA Score							1.098	0.017
Intercept	0.201	<0.001	0.091	<0.001	0.092	<0.001	0.095	<0.001
NCA Scale								
White	0.760	<0.001			0.947	0.269	0.665	<0.001
NCA Score			1.509	<0.001	1.506	<0.001	1.482	<0.001
White X NCA Score							1.158	<0.001
Intercept	0.337	<0.001	0.109	<0.001	0.111	<0.001	0.116	<0.001
NVCA Scale								
White	0.586	<0.001			0.671	<0.001	0.481	<0.001
NVCA Score			1.626	<0.001	1.610	<0.001	1.588	<0.001
White X NVCA Score							1.166	0.051
Intercept	0.086	<0.001	0.029	<0.001	0.031	<0.001	0.032	<0.001

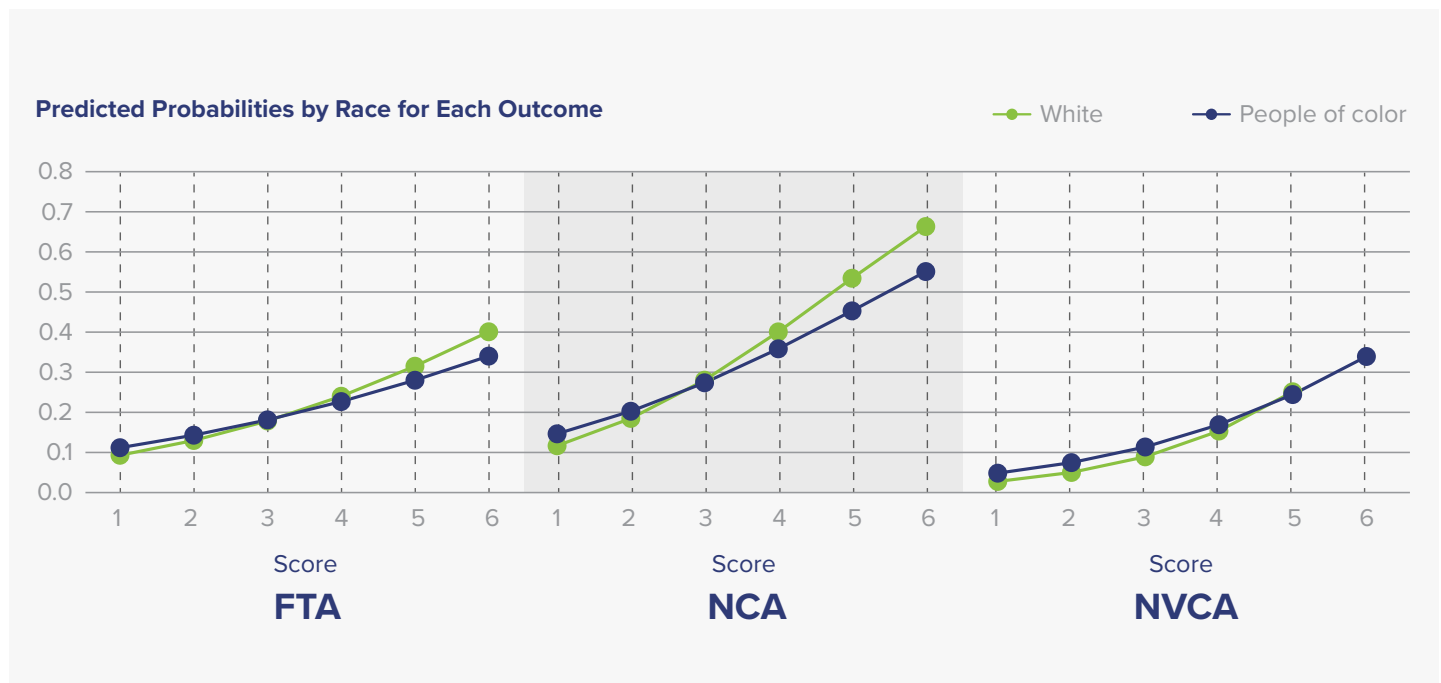
Results from Model 2 estimate whether the PSA scales perform as expected—in other words, higher scale scores are predictive of greater likelihood of the outcome. The results from the three models confirm that increases in scores are associated with statistically significant increases in the likelihood of the outcome. Specifically, the odds ratios are 1.34 (FTA), 1.51 (NCA), and 1.63 (NVCA) and all are statistically different from 1 as the associated p-values are ≤ 0.001 . These results mean that for each point increase in the FTA, NCA, and NVCA score there is a 34%, 51%, and 63% increase in the odds of those outcomes occurring, respectively. Higher scores are related to significantly greater likelihood that someone will miss court or be rearrested for any crime or a violent crime during their pretrial release.

Model 3 estimates the relationship between the race variable and the scale score with the outcomes simultaneously. The odds ratios are similar for each variable in size and direction to those from the simpler models, however the race variables do not reach the required level of significance in the FTA and NCA models, which suggest racial differences in these models are not present once we account for the PSA factor scores.

The final set of models tests whether predicted outcomes for a score are the same for each race group. This is tested by including the interaction term (White*Score) in the models.¹⁶ Specifically, if the resulting odds ratio is significantly different from 1 (i.e., if the p-value is ≤ 0.001) then the results suggest that the scale is providing different results for different racial groups. Results from Model 4, in Table 10 show that the odds ratios for the interaction term (White*Score) for the FTA and NVCA models are not significantly different from 1. In other words, the predictions of the PSA for appearing in court (FTA) and experiencing a new violent criminal arrest (NVCA) are not different for White individuals and people of color. Results for the NCA model, however, suggest that race moderates the relationship between the NCA score and the prediction of a new criminal arrest. Specifically, the odds ratio of 1.158 suggests that the NCA scale overpredicts new arrests for White individuals compared to people of color by about 16%.

Figure 4 provides a visualization of the analyses. Figure 4 shows the plots of the predicted probabilities for the outcomes by race for each scale score using the results from Model 4. The vertical y-axis shows the predicted probability that a person with the score indicated on the horizontal x-axis will have each of the outcomes. These probabilities represent the proportions of individuals predicted to have each outcome. The red lines provide values for people of color and the blue lines provide values for White individuals. As before, the shaded areas show the 99% confidence interval for the point estimate and overlap of the red and blue areas indicates that there is not a statistically significant difference between the point estimates for the two race groups. In addition to observing whether the lines are close together or highly divergent, the placement of the two lines suggests whether there is a consistent pattern in which one group's line consistently is above the other's which would indicate that one group consistently has a higher predicted outcome for a scale's scores.

Figure 4: Predicted PSA Outcomes by Score and Race



¹⁶ This type of analysis is referred to as moderation analysis and is used to determine whether scale scores have different meanings among different races or sexes.

The FTA plot in Figure 4 indicates that there are nearly identical probabilities of an FTA for people of color and White individuals who have scores of 1 to 4 (e.g., about 25% of both White individuals and people of color with an FTA score of 4 have predicted probabilities of not appearing in court). For scores of 5 and 6, White individuals have slightly higher predicted FTA rates, however, these differences are not statistically significant. The NCA plot results are more complicated because we find evidence of predictive bias consistent with the Model 4 results. The lines are similar for scores 1 through 3. However, for scores of 4 or greater White individuals have higher predicted probabilities of a new criminal arrest compared to people of color. For instance, the NCA plot shows that about 45% of people of color and 50% of White individuals with a score 5 will have an NCA, whereas about 50% of people of color and 60% of White individuals with a score of 6 will have an NCA. Thus, race does moderate the relationship between the risk of a new criminal arrest and the NCA score. In this case, the evidence of bias suggests that the NCA scale overpredicts new arrests for White individuals. The NVCA plots in Figure 4 does not show evidence of predictive bias. Similar to the FTA plot, scores of 1 through 5 have nearly identical predicted probabilities of an NVCA for both racial groups. For scores of 6, although White individuals have a slightly higher predicted probability of a new violent arrest the difference is not statistically significant.

Overall, these results demonstrate that, although there are some differences between people of color and White individuals, these differences do not produce predictive bias such that people of color are disadvantaged, which is the central concern for pretrial assessments.

RESULTS FOR SEX SUBGROUPS

Although concern about predictive bias has often focused on race, there are related concerns that assessments may incorrectly score female individuals as higher risk than observed.¹⁷ There are consistent relationships between sex and the outcomes such that female individuals have significantly lower likelihood for each of the pretrial outcomes compared to male individuals. Table 1 earlier showed that female individuals are less likely to experience an FTA, an NCA, or NVCA and, as noted earlier, these differences in base rates are not evidence of predictive bias.

Table 10 shows the results of logistic regression models that mirror the analyses in Table 11 but use sex as opposed to race as a control variable. Results for Model 1 show statistically significant relationships between sex and the outcome variables. In this instance, females are less likely to experience any of the outcomes compared to males. The results for Model 2 are identical to Table 10, as the scale score is the only variable included in the model. As before, Model 3 shows how sex and the scale scores relate to the outcomes when accounting for both in the model. In these models, both odds ratios are significantly different from 1, which suggests that sex and the scale scores are both significant predictors of the outcomes. Model 4 shows odds ratios for the interaction term between sex and the scale score (Female*scale). Like the models focusing on racial differences, the interactions between sex and the FTA and NVCA scale scores are not statistically different from 1. Similar to the race models, the interaction term between sex and the NCA scale is significant, suggesting that sex moderates the relationship between the NCA scale and a new criminal arrest. In this case, the odds ratio of 1.113 indicates that the NCA scale is overpredicting new criminal arrest for females by roughly 11% compared to males.

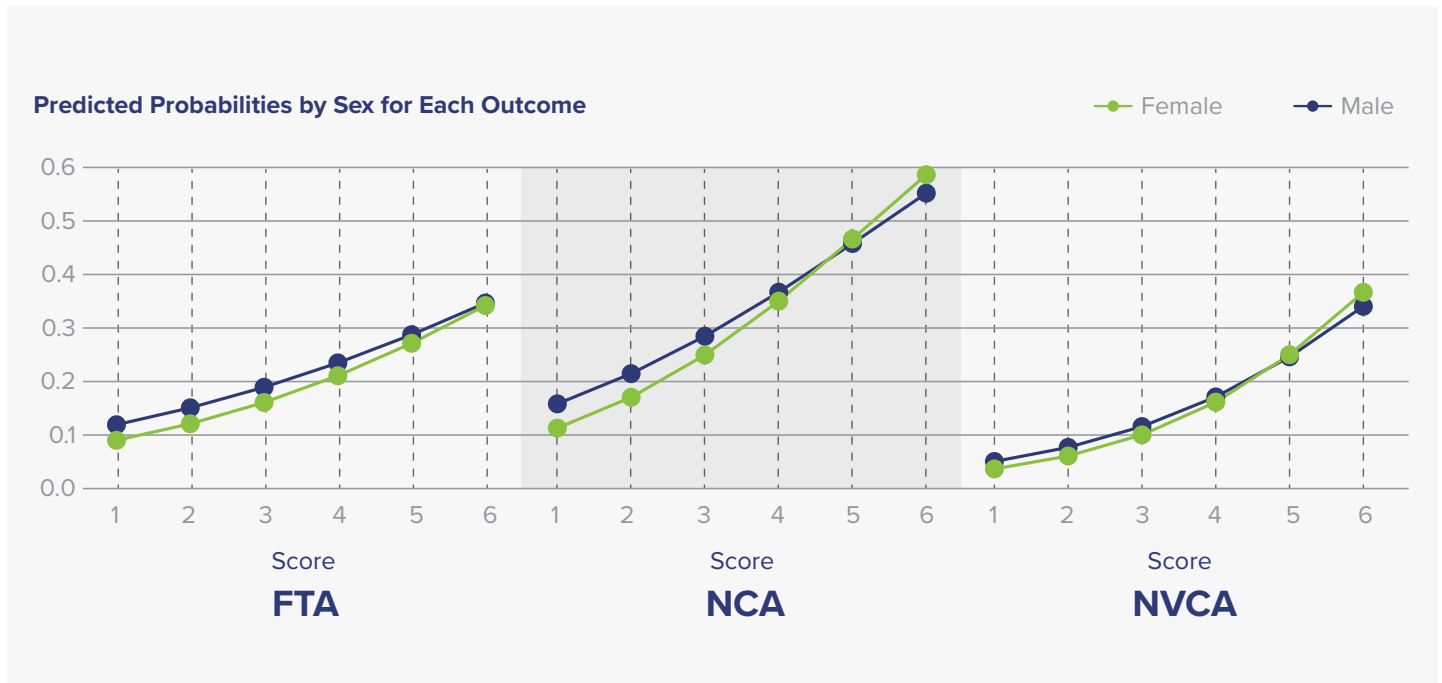
¹⁷ When one group is consistently scored higher than their true risk level this is referred to as “overprediction.” Overprediction can happen when one group appears more frequently in the data and the scores of the larger group are routinely higher than the other. In this case, males are about 75% of the sample and they have higher average (mean) scores than females (released Male NCA scale mean = 2.64, released Female NCA scale mean = 2.01).

Table 10: Logistic Regression Results Testing Sex Bias

Variable	Model 1		Model 2		Model 3		Model 4	
	Odds Ratio	p-value	Odds Ratio	p-value	Odds Ratio	p-value	Odds Ratio	p-value
FTA Scale								
Female	0.726	<0.001			0.803	<0.001	0.686	<0.001
FTA Score			1.344	<0.001	1.336	<0.001	1.317	<0.001
Female X FTA Score							1.061	0.051
Intercept	0.211	<0.001	0.091	<0.001	0.097	<0.001	0.101	<0.001
NCA Scale								
Female	0.618	<0.001			0.788	<0.001	0.606	<0.001
NCA Score			1.509	<0.001	1.487	<0.001	1.458	<0.001
Female X NCA Score							1.113	<0.001
Intercept	0.364	<0.001	0.109	<0.001	0.120	<0.001	0.127	<0.001
NVCA Scale								
Female	0.698	<0.001			0.789	<0.001	0.642	0.002
NVCA Score			1.626	<0.001	1.612	<0.001	1.585	<0.001
Female X NVCA Score							1.098	0.107
Intercept	0.087	<0.001	0.029	<0.001	0.031	<0.001	0.032	<0.001

Figure 5 provides the results from the moderation analyses for sex, the PSA scores, and the pretrial outcomes. These figures can be interpreted the same way as those used for Figure 4, but here the red lines are females, and the blue lines are for males.

Figure 5: Predicted PSA Outcomes by Score and Sex



The FTA plots show little difference between the predicted probabilities of an FTA for males and females. Males have slightly higher probabilities for scores between 1 and 4 and similar probabilities for scores of 5 and 6.

The NCA plots illustrate the challenge in summarizing predictive bias for the PSA because it involves explaining the trends across the risk continuum. Here, males have slightly higher predicted rates of a new criminal arrest for scores of 1 through 3, similar predicted rates for scores of 4 and 5, and lower predicted rates for a score of 6. These differences resulted in a significant moderation effect. Similar to the NCA score comparing White individuals and people of color, these analyses show a difference in how the NCA score predicts new criminal activity for males compared to females. The NVCA plots illustrate that the PSA provides nearly identical predicted probabilities for a new violent criminal arrest for males and females. Other than the slightly higher score for females with a score 6, the probabilities are the same across the risk scale.

Discussion

The results described here suggest that, overall, PSA assessment scores do not exhibit a substantial problem with predictive bias with respect to race and sex. One concern with pretrial release assessment instruments is that people of color and females could be scored too high relative to their actual risk of missing a court date or committing a new crime. If conditions of release are predicated on risk scores, people of color or females may be disadvantaged. We do not find evidence that people of color are being scored higher than their actual outcome rates (i.e., no overprediction), but we do find overprediction for females for NCAs.

A primary concern with pretrial release assessments is that structural racism in criminal legal system practices has resulted in people of color being arrested at earlier ages, charged with more serious crimes, and convicted more often than White individuals. Since criminal history contributes to the PSA scores, these factors will result in higher scores for people of color, who may face additional conditions for pretrial release (or less likelihood of release) based on these inflated scores. The analyses conducted for this study did not find that people of color were being overpredicted (i.e., scored higher than their actual outcomes). Rather, we found that White individuals were being overpredicted for NCAs. We can look at the actual outcome rates (Figure 2) and the predicted outcomes rates (Figure 4) by PSA score for each race group and see there are few differences by race for NCA scale scores below 4.

The overprediction is likely related to the small number of White individuals that have a new NCA compared to people of color with NCA scores of 5 or 6. Although the percentage of White individuals and people of color with a scale score of 5 and 6 do not differ, there are fewer White individuals that have an NCA with these scores. Specifically, a total of 44% of people have an NCA score of 5, but there are only 37 White individuals with an NCA, compared to people of color ($n = 506$). There is a similar result for NCA scores of 6 with little difference in the percentage (around 50%) with an NCA, but large differences in the number of White individuals ($n = 13$) compared to people of color ($n = 228$) with an NCA. The statistical techniques used to assess predictive bias, for the most part, are more accurate when we have more cases in each group. The small numbers of White individuals with an NCA score 5 or 6 contributes to the large error bars (wide blue shading) at these higher scores in Figure 4.

Different concerns arise for the predictive bias tests for females and the NCA scale. As females are less likely to commit a new crime or miss court compared to males, the PSA scores overpredict the outcomes for females with lower scores. Females with NCA scale scores of 1-3 have significantly lower NCA rates than males. Specifically, females with an NCA score of 1-3 have NCA rates of 11%, 17%, and 25% whereas for the same scores males have rates of 15%, 22%, and 31%. Interestingly, NCA rates do not differ between males and females for NCA scores of 4-6. There is a similar degree of association between the PSA and NCAs for males ($AUC = 0.64$) and females ($AUC = 0.65$). NCA rates increase as the NCA scale increases for both males and females such that the NCA scale is predictive for both males and females. The reality is that females with lower scores (1-3) on the NCA scale are rearrested in Fulton County at lower rates than males. Females have lower criminal involvement evidenced by having fewer factors on the NCA scale including pending charge (15% v. 22%), prior misdemeanor conviction (39% v. 61%), prior felony conviction (18% v. 44%), prior violent conviction (15% v. 36%), and prior sentence to incarceration for more than 14 days (14% v. 36%). These differences alone are not indicative of predictive bias, but along with higher overall NCA rates for males this suggests there are some differences in NCAs between lower scoring males and females.

Conclusion

The PSA was created through investments made by the Arnold Ventures using a large database drawn from several jurisdictions to examine the predictive validity of hundreds of risk factors.¹⁸ The PSA was developed to identify the strongest predictors of failure to appear (FTA), new criminal activity (NCA), and new violent criminal activity (NVCA). The PSA leaves out demographic factors related to race/ethnicity and sex as well as socioeconomic variables such as residential stability, educational attainment, and employment. The PSA is completed with court and criminal history data and does not require an interview. These items were excluded to reduce potential for predictive bias for the poor and communities of color. The PSA is available to the public and jurisdictions can use the PSA for free.

The results demonstrate that the PSA meets standards of predictive validity for criminal legal system assessments.¹⁹ For the three scales, we found that the Area Under the Curve (AUC) values are in the fair (FTA, AUC = 0.62) and good (NVCA, AUC = 0.65 and NCA, AUC = 0.65) ranges,²⁰ there are consistent increases in pretrial outcomes as scores increase, and significant increases in the predicted likelihood of missed court appearance and new arrests as scores increase across a series of regression models. The report shows that the PSA meets validity standards used for criminal justice assessments, and the report includes tests for predictive bias. The results show that for each point increase in the FTA, NCA, and NVCA score there is a 34%, 51%, and 63% increase in the odds of those outcomes occurring, respectively. Higher scores are related to significantly greater likelihood that someone will miss court or be rearrested for any crime or a violent crime during their pretrial release.

We found that many lower scoring individuals were detained, and many higher scoring individuals were released (table 4). Nearly 40% (N=5,208) of the detained individuals (N=13,695) score 1-2 on the FTA scale, and about 10% (N=2,176) of the released individuals score 5-6 on the FTA scale.

This report provides an initial step for Fulton County to enhance their empirical approach to pretrial decision making. Although there are many similarities across criminal legal systems, each jurisdiction has unique features, institutions, and cultures that requires conducting localized research. Pretrial systems are highly variable in their release rates, outcome rates, and population characteristics. Pretrial assessment research is ongoing and requires using ‘fresh, local data’ to compare the current findings to future outcomes post-PSA implementation to avoid misstating the likelihood of outcomes.²¹ Future research is needed because Fulton County’s criminal legal system is embedded within an ever-changing environment such that they are implementing risk-reducing strategies that will change the likelihood of outcomes. The current research provides a snapshot of how the PSA is likely to perform in Fulton County.

18 VanNostrand, M., & Lowenkamp, C. T. (2013). Assessing pretrial risk without a defendant interview. Houston: Laura and John Arnold Foundation. <https://nicic.gov/assessing-pretrial-risk-without-defendant-interview>

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

20 Rice, M. E., and Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, 29, 615–620.

21 Koepke, Logan, J., & Robinson, D. G. (2018). Danger Ahead: Risk Assessment and the Future of Bail Reform, Retrieved from <https://ssrn.com/abstract=3041622> or <https://doi.org/10.2139/ssrn.3041622>

Appendix—A

Table 1. Fulton County, Georgia, Violent Crime List

Statute Number	Georgia Statute Title
16-5-1	Murder
16-5-2	Voluntary Manslaughter
16-5-3	Involuntary Manslaughter
16-5-20	Simple Assault
16-5-21	Aggravated Assault
16-5-23	Simple Battery
16-5-23.1	Battery
16-5-24	Aggravated Battery
16-5-27	Female Genital Mutilation
16-5-28	Assault of Unborn Child
16-5-29	Battery of an Unborn Child
16-5-40	Kidnapping
16-5-42	False Imprisonment Under Color of Legal Process
16-5-44	Aircraft Hijacking
16-5-44.1	Motor Vehicle Hijacking
16-5-46	Trafficking a Person for Labor or Sexual Servitude
16-5-70	Cruelty to Children*
16-5-90	Stalking
16-5-91	Aggravated Stalking
16-5-102	Punishment for any Person Who Exploits, Threatens, Intimidates, or Attempts to Intimidate a Disabled Adult Elder Person

Statute Number	Georgia Statute Title
16-5-102(a)	Exploitation and intimidation of disabled adults, elder persons, and residents; obstruction of investigation
16-5-102.1(b)	Trafficking of a disabled adult, elder person, or resident*
16-6-1	Rape
16-6-2(a)(2)	Aggravated sodomy
16-6-3	Statutory Rape
16-6-4	Child Molestation; Aggravated Child Molestation
16-6-5.1	Sexual Assault - by persons with supervisory or disciplinary authority*
16-6-14	Pandering by Compulsion
16-6-22.1	Sexual Battery
16-6-22.2	Aggravated Sexual Battery
16-7-1	Burglary*
16-7-5	Home Invasion; Degrees
16-7-60	Arson in the First Degree*
16-7-88	Possession, Transportation, Receipt, or Use of Destructive Device or Explosive with Intent to Kill, Injure or Intimidate or to Destroy Any Public Building
16-8-40	Robbery
16-8-41	Armed Robbery; Robbery by Intimidation
16-10-32	Attempted Killing or Threats to Prevent Information on Criminal Activity From Being Communicated to Law Enforcement or to the Courts
16-10-56	Riot in a Penal Institution
16-11-30	Riot
16-11-37	Terroristic Threats and Acts
16-11-102	Pointing Gun or Pistol at Another
16-12-100	Sexual Exploitation of Children*
16-12-123	Bus or Rail Vehicle Hijacking; Boarding Aircraft, Bus, or Rail Vehicle with Certain Items