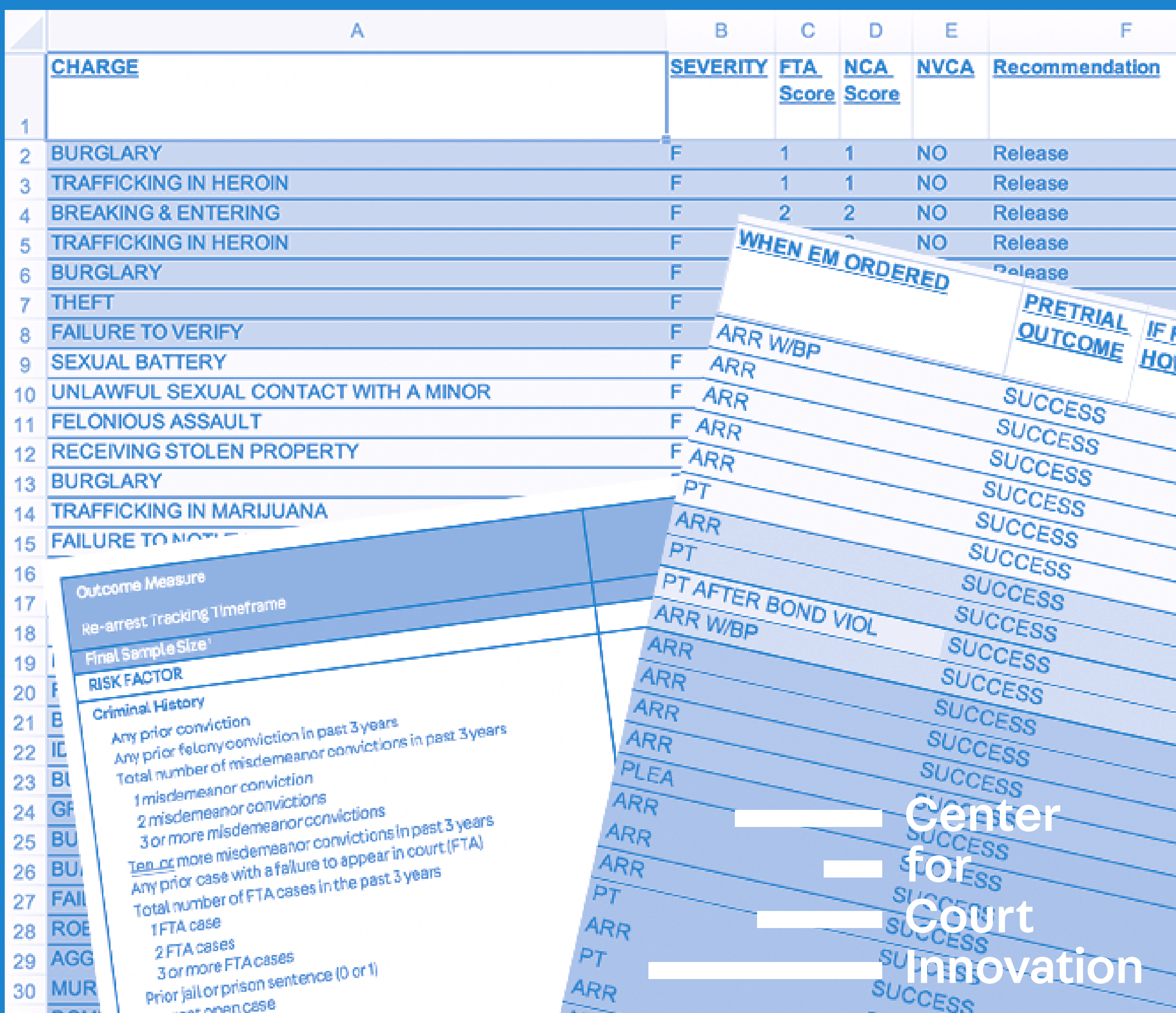
Beyond the Algorithm

Pretrial Reform, Risk Assessment, and Racial Fairness

by Sarah Picard, Matt Watkins, Michael Rempel, and Ashmini Kerodal



Center for Court Innovation

Acknowledgments

The authors are deeply grateful to Arnold Ventures for their thoughtful partnership and support of this project, as well as their ongoing commitment to pretrial justice. We are also indebted to Dr. Alexandra Chouldechova for her numerous intellectual contributions over the life of the project. From the Center for Court Innovation, we thank Michela Lowry for her review of the relevant literature, Julian Adler for his invaluable work in the conceptualization of the project and this publication, Greg Berman for his many editorial contributions, and Samiha Amin Meah and MaaRa Maakheru for designing the publication.

Finally, we thank the New York State Division of Criminal Justice Services (DCJS) for providing the case-level criminal history and administrative data that were the basis of our analyses. The opinions, findings, and conclusions expressed in this publication are those of the authors and not those of DCJS; neither New York State nor DCJS assumes liability for its contents or use thereof. For correspondence, please contact Sarah Picard at picards@courtinnovation.org.

Introduction

Pretrial detention, often resulting from a defendant's inability to afford bail, is one of the primary drivers of incarceration nationwide.[1] The Bureau of Justice Statistics estimates that two out of three people in local jails in 2016 were held while awaiting trial, having not yet been convicted of a crime.[2] Jurisdictions looking to safely reduce their use of bail and pretrial detention have increasingly turned to automated or actuarial risk assessments. These tools employ a mathematical formula, or algorithm, to estimate the probability of a defendant incurring a new arrest or failing to appear in court. Typically, in a risk assessment, defendants' criminal history, criminogenic needs, and/or basic demographic information, such as age and gender, are weighted and combined, generating a score which can be used to group defendants into risk categories ranging from low to high.

With the aid of better information about the defendants who appear before them, judges, in theory, can make more consistent decisions regarding pretrial release and bail. For example, jurisdictions that use risk assessments may be more likely to consider pretrial release for defendants in lower-risk categories, or pretrial supervision in the community for higher-risk defendants. In cases where victim or community safety is a concern, risk assessment may provide guidance regarding the need for bail or detention hearings.

The appeal of pretrial risk assessment—especially in large, overburdened court systems—is of a fast and objective evaluation, harnessing the power of data to aid decision-making. Research suggests that actuarial risk assessments are more accurate than decisions made by criminal justice officials relying on professional judgment alone.[3] By intervening in a process historically driven by subjective decision-making, risk assessments arguably act as a corrective to a system plagued by bias, as witnessed in the racial disparities long seen in incarceration rates across the country.

That said, important objections have been raised that, far from disrupting racial biases in the criminal justice system, risk assessments unintentionally amplify them, only this time under the guise of science. The debate is still unresolved, but from a justice system practitioner's perspective—let alone that of a defendant—the stakes are urgent.

What follows are the results of an empirical test of racial bias in risk assessment and, based on an original analysis, a consideration of whether there are policy-level solutions that could conserve the benefits of risk assessment, while also addressing valid concerns over racial fairness.

The State of the Debate

The increasingly contentious debate concerning risk assessment and racial bias draws largely upon the findings of a single study of one risk tool implemented in one county: ProPublica's 2016 analysis of the widely-used COMPAS risk algorithm in Broward County, Florida.[4] ProPublica's headline finding was that the risk tool disproportionately labeled black defendants who did *not* go on to be charged with a new crime as high-risk, unfairly exposing them to punitive criminal justice consequences.

The defense mounted by the private company behind the COMPAS assessment and some independent scholars is that the overall predictive accuracy of the COMPAS is similar across racial groups, making the algorithm itself ostensibly unbiased, even where *outcomes* based on the tool—such as who gets detained pretrial—differ systematically by race.[5] We need to pause over this idea that a tool can be unbiased in its overall ability to predict re-arrest. For this claim to be valid indicates three things: first, that the formula does not include race in its calculations; second, that the algorithm performs similarly across racial and ethnic groups in predicting outcomes such as a new arrest or a failure to appear in court; and third, that the factors included in the tool are not so strongly correlated with race that they could be considered racial "proxies."

All risk assessments make mistakes; indeed, they are only assigning probabilities. But the crux of the current debate is about the *kinds* of errors made. Classification errors can have serious real-world consequences. A tool that disproportionately classifies the members of certain groups as high-risk even when they do *not* go on to be re-arrested may unnecessarily expose them to high bail amounts and pretrial detention.

This is precisely what ProPublica found. The ProPublica analysis determined that, while the overall percentage of errors made by the COMPAS was similar for black and white defendants, among those who did *not* go on to be re-arrested, the COMPAS disproportionately misclassified black defendants into the high-risk category. A series of examples demonstrated that this tendency towards high-risk classifications exposed black defendants, in particular, to harmful outcomes such as a high bail amount, pretrial detention, or a longer jail or prison sentence. The authors concluded that the COMPAS in Broward County led to racial bias in pretrial decisions.

The Center for Court Innovation set out to conduct an analysis similar to ProPublica's using a sample of 175,000 anonymized New York City defendants and an assessment tool created solely for the purpose of exploring questions related to risk prediction and pretrial outcomes. While the data is drawn from real defendants, it is important to note that our risk tool was not used to inform pretrial decisions made by New York City courts. Instead, the data was used exclusively for research purposes to illuminate the potential ramifications of applying risk assessment tools to real-world practice.

Our analysis suggests that the racial fairness concerns arising in ProPublica's study of Broward County may well be generalizable to other risk assessment tools and jurisdictions. Specifically, while our risk assessment tool performed similarly across racial and ethnic groups in terms of its overall predictive accuracy, when we looked at the *types* of errors made by our assessment, it was more likely to misclassify black defendants as high-risk when compared to Hispanic or white defendants. If high-risk classification leads to an increased likelihood of high bail or pretrial detention, our tool would potentially foster racially-disparate pretrial outcomes.

In interpreting our findings, we do not, however, argue for eliminating the use of risk assessments. Our principal recommendation—discussed in the concluding pages—is that jurisdictions think "beyond the algorithm." That is, practitioners should take concerns regarding racial fairness seriously and minimize the use of unnecessary incarceration overall. Supporting this conclusion, we found that if pretrial detention was restricted only to defendants who are charged with violent crimes *and* who fall into higher-risk categories, such a policy may both reduce incarceration overall and alleviate racial disparities. These types of targeted risk-informed approaches have not received broad attention in the field, despite the fact they could be of disproportionate benefit to the group that consistently experiences the worst pretrial outcomes: defendants of color.

A Case Study in New York City

Drawing on a case study of defendants arrested in New York City in 2015, our analysis sought to address a challenge facing jurisdictions across the country: *Can risk algorithms be adopted in a manner that maximizes their potential benefits for reducing incarceration, without sacrificing the value of fair and equitable outcomes across racial groups?*

In particular, we sought:

- To examine whether and how the use of a risk assessment tool affects racial disparity in pretrial outcomes; and
- To test a range of approaches to pretrial decision-making in an effort to identify the most effective scenarios for reducing pretrial detention and mitigating racially-disparate outcomes.

Methods

We collected a sample of all arrests made of black, Hispanic, and white individuals in New York City in 2015. The final sample included more than 175,000 defendants, of whom 49% were black (86,227 defendants), 36% Hispanic (64,109 defendants), and 14% white (25,117 defendants).[6] We then applied our risk assessment tool—developed solely for the purposes of research—to this sample of defendants to gain insight into how its use would *likely* affect defendant outcomes in New York City.[7] To that end, we conducted an analysis of the extent to which our tool classified defendants in racially-disparate ways—and of how those classifications could be expected to play out in the real-world under several alternative policy scenarios.

The risk algorithm we developed for this study drew exclusively on criminal history and demographic factors—the factors generally proven to be the most predictive of a future arrest. While the inclusion of gender as a factor in risk assessment tools remains controversial generally, its inclusion in the current tool improves the overall accuracy of the risk algorithm and mitigates the tendency of the tool to over-classify female defendants as high-risk. In other words, female defendants in our sample have substantially lower actual rates of re-arrest than male defendants, even after controlling for criminal history. The tool did not explicitly use race or ethnicity in calculating risk scores.

Specifically, our assessment relied on the nine risk factors listed below to estimate the probability of a new arrest over a two-year tracking period. Although studies of pretrial risk often limit their analysis specifically to the pretrial period, longer tracking periods can improve the stability of algorithms for predicting outcomes of interest.[8] In 2015, 87% of criminal cases in New York City were disposed within one year of arrest, and 36% of those defendants who were re-arrested in our sample were re-arrested prior to the disposition of their case. Thus, the tracking period selected for the current analysis covers both pretrial and post-disposition periods for the vast majority of defendants.

Criminal History[9]

1. Prior convictions
2. Prior jail or prison sentence
3. Prior failure to appear in court
4. Probation status

Current Case Characteristics

5. Charge type
6. Charge severity
7. Concurrent open cases

Demographic Characteristics

8. Age
9. Gender

Details regarding the specific items, weights, and predictive performance of the tool can be found in Appendix A.

Tool developers typically measure predictive accuracy using area-under-the-curve ("AUC") statistics, with AUCs above 0.700 indicating good predictive accuracy by current industry standards. Our tool had strong predictive accuracy for defendants as a whole (AUC=.745). Moreover, the five risk categories produced by the tool (minimal, low, moderate, moderate-high, and high-risk) clearly differentiated among defendants with varying rates of re-arrest over a two-year follow-up. For example, only one out of 10 defendants labeled minimal-risk went on to be re-arrested compared to more than seven out of 10 defendants labeled high-risk.

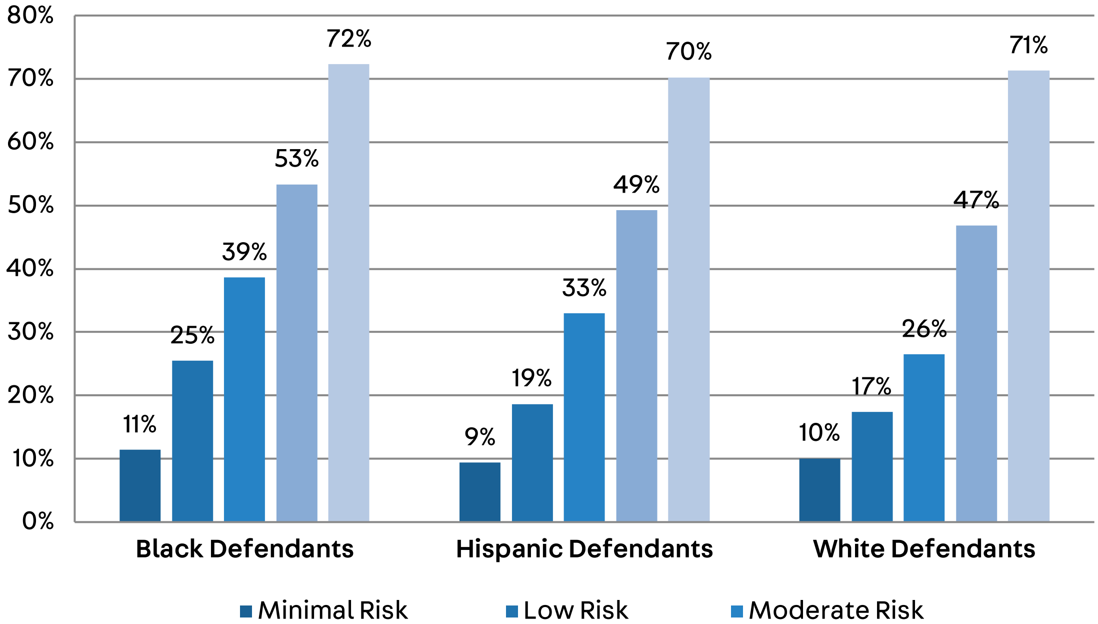
Findings

Predictive Accuracy by Race

A core question for the current study was whether our tool would perform well in making predictions, regardless of race or ethnicity. This question is explored in Exhibit 1 below. As shown across all groups, re-arrest rates increased progressively, in near-lockstep, as risk categories move from minimal to high. Moreover, rates of re-arrest were similar, though not equal, for black, Hispanic, and white defendants in each risk category. For example, re-arrest rates for defendants who the tool classified as high-risk are 72% for blacks, 71% for Hispanics, and 70% for whites. At the other end of the spectrum, we saw similarly consistent, and much lower, rates of re-arrest across the same three racial and ethnic groups in the minimal-risk category (11%, 9%, and 10%, respectively). More substantial differences in re-arrest can be seen in the low- and moderate-risk categories in our tool. For example, black defendants classified as moderate risk by our tool were re-arrested at a higher rate than Hispanic or white defendants classified in the same category (39%, 33%, and 26% respectively). Some differences in re-arrest rates within categories are inevitable in any "colorblind" risk assessment tool that is developed in a jurisdiction where actual re-arrest rates differ between racial groups. Indeed, in large part, the desire to understand the impact of such differences on pretrial decisions and outcomes by race was the motivation for the current study.

Exhibit 1. Two Year Re-arrest Rates by Risk Category and Race

New York City Defendants, 2015



Note: Differences in re-arrest rates within categories reflect real differences in overall re-arrest rates between racial groups. Overall two-year re-arrest rates are 43% for black defendants, 34% for Hispanic defendants and 26% for white defendants.

Perhaps a more important measure of predictive accuracy is the rate at which our tool correctly classifies defendants with a higher probability of new arrest into higher-risk categories (as represented by AUC statistics). As shown in Exhibit 2, AUC statistics exceeded .700 for black, Hispanic, and white defendants alike, suggesting that our tool effectively classifies risk irrespective of race.

Exhibit 2. Risk Classifications by Race

New York City Defendants, 2015

	AUC Statistic
Black Defendants	0.719
Hispanic Defendants	0.735
White Defendants	0.724

Note: While AUC statistics are similar in magnitude across racial groups, the predictive accuracy of the tool is systematically better for Hispanic defendants, when compared to black or white defendants (p<.001).

Disparate Pretrial Outcomes by Race

Demonstrating that a risk tool performs similarly across racial and ethnic groups does not by itself resolve questions of racial disparity. Indeed, even tools that appear to perform well across groups may, nonetheless, foster disparate pretrial outcomes in practice. How is this possible? Many risk assessment tools produce scores based primarily on criminal history factors (such as prior convictions or warrants). Yet, in jurisdictions across the country, people of color are far more likely to accumulate such histories.[10] The reasons for this vary and may range from deep historical inequalities adversely affecting communities of color, to the disproportionate policing of these communities, to racial bias in criminal justice decision-making, but the end result is the same: people of color are likely to average longer criminal histories, increasing their average risk score. For example, a recent analysis of a risk tool used to inform sentencing decisions for defendants at the federal level revealed systematic correlations between race and criminal history items included in the assessment (i.e., black defendants in the sample had higher rates of prior arrests and convictions and consequently fell into higher risk categories). Ultimately, the authors warn that these differences might lead black defendants to be disproportionately exposed to harsher sentencing practices.[11]

The existence of racial differences in criminal justice system contact is not, on its own, evidence of biased decision-making by law enforcement, judges, or other criminal justice actors. However, such differences help explain why black defendants may inevitably be more frequently classified as high-risk by any assessment tool relying heavily on criminal history in its algorithm. If, because of the number of criminal history factors used as “inputs” in a tool, black defendants are classified disproportionately as high-risk, it follows that there will also be a greater total number of black defendants available to be *mis*classified as high-risk. With more potential high-risk black defendants to draw upon, an algorithm will automatically assign more of them to this category, regardless of whether this assignment proves correct.

This concern became more visible following ProPublica’s 2016 study, which showed that black defendants who were *not* ultimately arrested on a new charge were still twice as likely to have been classified as high-risk by the COMPAS.[12] In a situation where risk assessment is used to inform pretrial decisions, the disproportionate misclassification of people of color into high-risk categories can directly lead to racial disparities in pretrial detention.

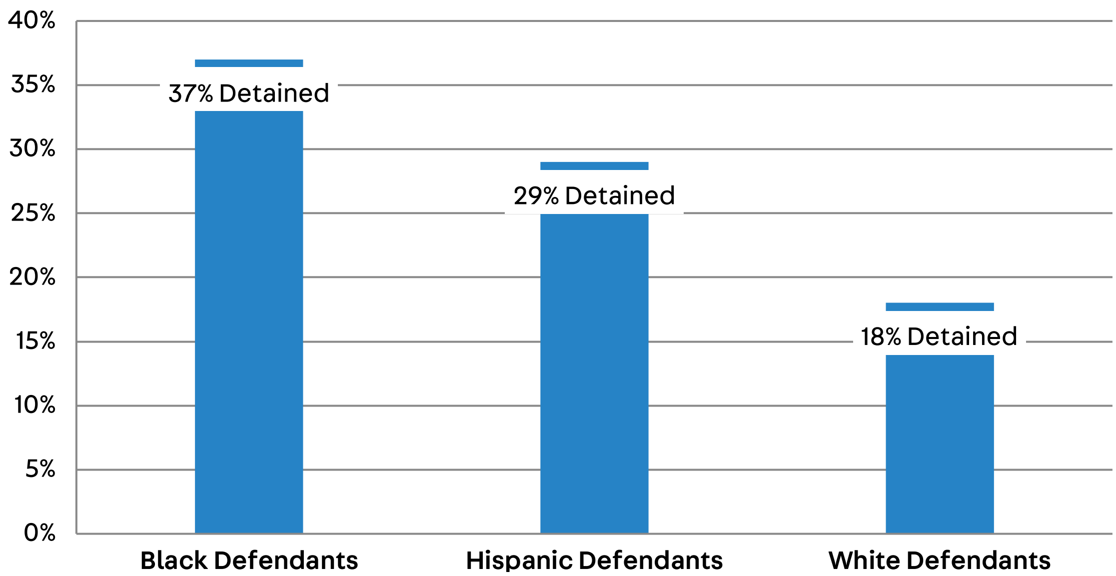
Unpacking the “High-Risk” Label and Its Impact on Pretrial Decisions

What is really meant by “high-risk”? And what happens if jurisdictions base pretrial decisions primarily on this categorization? Like many tools currently in use, the assessment created for this study was designed to classify defendants into a range of risk categories from minimal to high. However, actual decisions drawing on risk assessments often rely on a simpler calculus: a risk threshold is established above which defendants are typically not considered appropriate candidates for straight release. Depending on the norms of the specific jurisdiction, defendants above this threshold are more likely to be recommended for pretrial supervision, bail, or a pretrial detention hearing, which increases their likelihood of detention.[13] For research purposes, we collapsed the top two categories produced by our assessment (initially labeled “moderate-high” and “high” risk) into a combined “high-risk” group on the assumption that, in a typical risk-informed decision-making scenario, these defendants would face the most onerous pretrial conditions in a given jurisdiction.

Exhibit 3 shows that, in a hypothetical scenario in which New York City judges strictly adhered to the high-risk threshold established above—and used that threshold to make decisions on who to detain—we would observe substantially different rates of pretrial detention by race. Specifically, black defendants would be detained at twice the rate of white defendants and nearly 1.25 times the rate of Hispanic defendants. To be clear, this scenario simply illustrates that our tool more often classified black defendants into the two highest risk categories and—if strictly followed—would lead them to be more frequently exposed to pretrial detention. It does not suggest racial bias in the risk assessment itself. However, as recent research from the field of data science demonstrates, such racial disparities in the distribution of the risk scores and categories will inevitably emerge in a context—such as New York City's—where the overall probability of re-arrest differs by race (this is known as the "base rate" problem).[14]

Exhibit 3. High Risk Categorization and Hypothetical Pretrial Outcomes

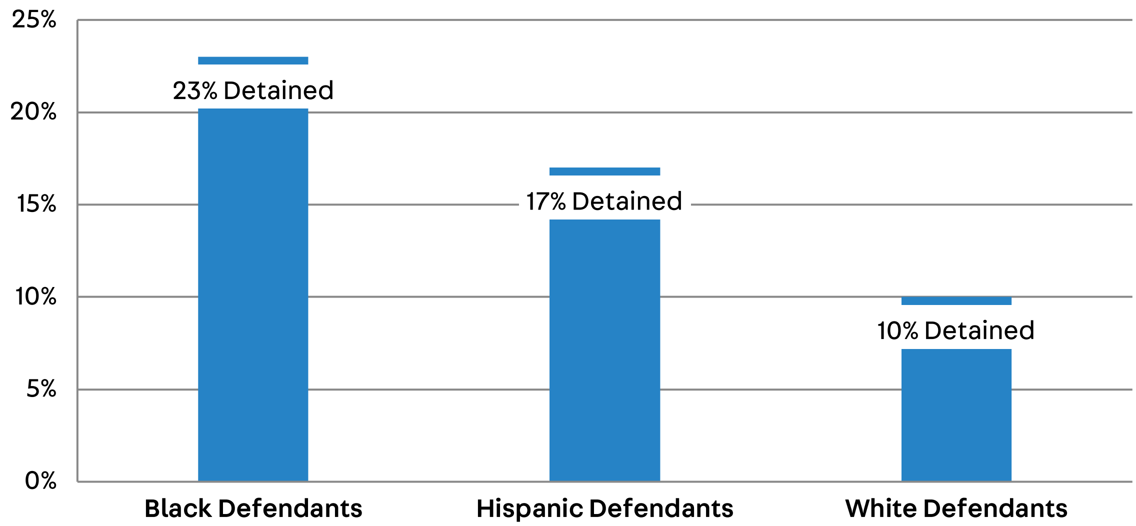
New York City Defendants, 2015



False Positives. To what extent might racial differences in high-risk classification influence the fairness of pretrial outcomes by inappropriately labeling black or Hispanic defendants as high-risk? In Exhibit 4, we examine the rate of false positives—that is, the rate of high-risk classifications among those individuals who were *not* in fact re-arrested on a new charge. As shown, among those who were not re-arrested, 23% of black defendants were nonetheless classified as high-risk and flagged for detention, compared with 17% of Hispanic defendants, and only 10% of white defendants.[15] In short, our findings regarding false positives suggest that black and Hispanic defendants would be substantially more likely to be exposed to unwarranted pretrial detention if the risk-based decision-making approach described above was actually employed in New York City.

Exhibit 4. False Positives and Hypothetical Pretrial Outcomes

New York City Defendants, 2015



What Is to Be Done?

Thus far our findings underscore a critical lesson for the field: even tools with similar predictive accuracy for all groups can ultimately lead to disparate negative outcomes for black and Hispanic defendants. One by-product of risk algorithms is that the members of whichever groups have more frequent contact with the justice system will, as a matter of course, be more frequently classified—and also *misclassified*—as high-risk. One potential solution would be to explicitly tailor risk algorithms or high-risk thresholds by race with the goal of reducing disparities in false positive rates.[16] However, for reasons practical, ethical, and likely constitutional, race-specific predictive models would be a cure worse than the disease. Imagine, for example, the challenges of defining "race" in a courtroom in order to accord a defendant their corresponding race-tailored algorithm. Moreover, the idea of judicial decision-making based on a race-specific assessment would appear to run squarely counter to the Fourteenth Amendment's guarantee of "equal protection," leaving any such decisions subject to legal challenge.[17]

Despite the amount of controversy risk assessments have engendered, they are an increasing part of pretrial practice in jurisdictions across the country. Given this trend, there has been surprisingly little concrete work on how risk assessment may fit into larger strategies to promote racial fairness in pretrial outcomes. To help jumpstart the conversation on how to reduce pretrial detention and racial disparities therein, we used our New York City data to compare the hypothetical outcomes of black, Hispanic, and white defendants under three distinct decision-making scenarios.

- **Scenario 1. Business as Usual**: This scenario uses status-quo decision practices in New York City to simulate business as usual in many jurisdictions. Like many jurisdictions, in New York City, pretrial decision-making is largely subjective, relying primarily on judicial discretion as informed by arguments from attorneys. While New York City judges are provided with the results of a tool assessing defendants' likelihood for failure-to-appear, prior research indicates judicial decisions often deviate from the recommendations of this assessment in the status quo.[18]

- **Scenario 2. Risk-Based Approach (Adjusted High-Risk Threshold)**: In this scenario, pretrial decisions rely solely on the risk assessment tool. However, unlike the assumption made in our original analysis that defendants in the top *two* risk categories would likely be detained, in this scenario, we adjusted the high-risk threshold so that only defendants classified in the *highest* of the five risk categories were flagged for pretrial detention. This scenario reduced the proportion of all defendants whose risk classification exposes them to pretrial detention.

- **Scenario 3. Hybrid Charge- and Risk-Based Approach**: In this final policy scenario, pretrial detention was reserved exclusively for defendants charged with a violent felony or a domestic violence offense who also fell into the top *two* risk categories on our risk assessment tool. This final scenario presumes that most misdemeanor and non-violent defendants are not appropriate candidates for bail or detention consideration, regardless of risk level. At the same time, it recognizes that charge alone is not a good proxy for risk, and that some individuals with violent charges can be safely supervised in the community.

Results

In our analysis, we examined how each scenario would affect both pretrial detention rates and false positive rates—as well as racial disparities—for defendants in New York City. Exhibit 5 presents our results for pretrial detention rates—both overall, and separately, for black, Hispanic, and white defendants. Exhibit 6 presents our results for false positive rates.

- **Scenario 1. Business as Usual**: To begin with, our analysis demonstrated that real-world differences in pretrial detention rates exist by race and ethnicity under the current approach to pretrial decisions used in New York City. In 2015, for example, bail decisions at arraignment led to the detention of 26% of defendants, including 31% of black, 25% of Hispanic, and 22% of white defendants (see Exhibit 5). Further, false positive rates were relatively high for all groups in the status quo, and higher for black and Hispanic defendants when compared with similarly situated white defendants. Specifically, among those defendants in our sample who were not ultimately arrested on a new charge, one out of five were initially detained pretrial, including 21% of black, 19% of Hispanic, and 17% of white defendants (Exhibit 6).[19]

- **Scenario 2. Risk-Based Approach (Adjusted High-Risk Threshold)**: A risk-based approach in which defendants classified as moderate-high *or* high-risk are candidates for bail and detention created disparities in both pretrial detention and false positive rates, at least in our New York City example (see Exhibits 3 and 4 above). But what if a more restrained approach to pure risk-based decision-making was implemented? By adjusting our risk threshold so that only defendants in the *highest* risk category were candidates for bail or detention, we reduced the raw numbers of people detained and the disproportionate impact of pretrial detention on black and Hispanic defendants. As shown in Exhibit 5, this approach reduced the overall detention rate by nine percentage points when compared to business as usual. Moreover, this stricter risk-based scenario also improved the accuracy of pretrial detention decisions, reducing the overall false positive rate to less than 10 percent. That said, even as false positives declined for everyone, exposing fewer individuals of all races to misclassification, we still observed a gap in false positive rates between black, white, and Hispanic defendants (Exhibit 6).

- **Scenario 3. Hybrid Charge- and Risk-Based Approach**: Finally, we considered an approach that attempted to restrict detention by both charge and risk level, specifically limiting the use of pretrial detention to defendants charged with a violent felony or a domestic violence offense who also fall into the moderate-high and high-risk categories. At the same time, it takes pretrial detention off the table for *all* defendants charged with a misdemeanor or non-violent felony (except where domestic violence was involved). Would this approach reduce overall detention rates *and* lessen the racial disparities found in our prior scenarios? Our findings suggest it would. In New York City, such an approach would cut overall pretrial detention by 51% compared to business as usual and nearly eliminate disparities in detention, with black and white defendants both detained at a rate of 13 percent, compared to 14 percent for Hispanic defendants (Exhibit 5). Moreover, racial disparities in false positives would also be largely alleviated in this scenario (see Exhibit 6).[20]

Exhibit 5. Pretrial Detention by Race Under Three Decision-Making Scenarios

New York City Defendants, 2015

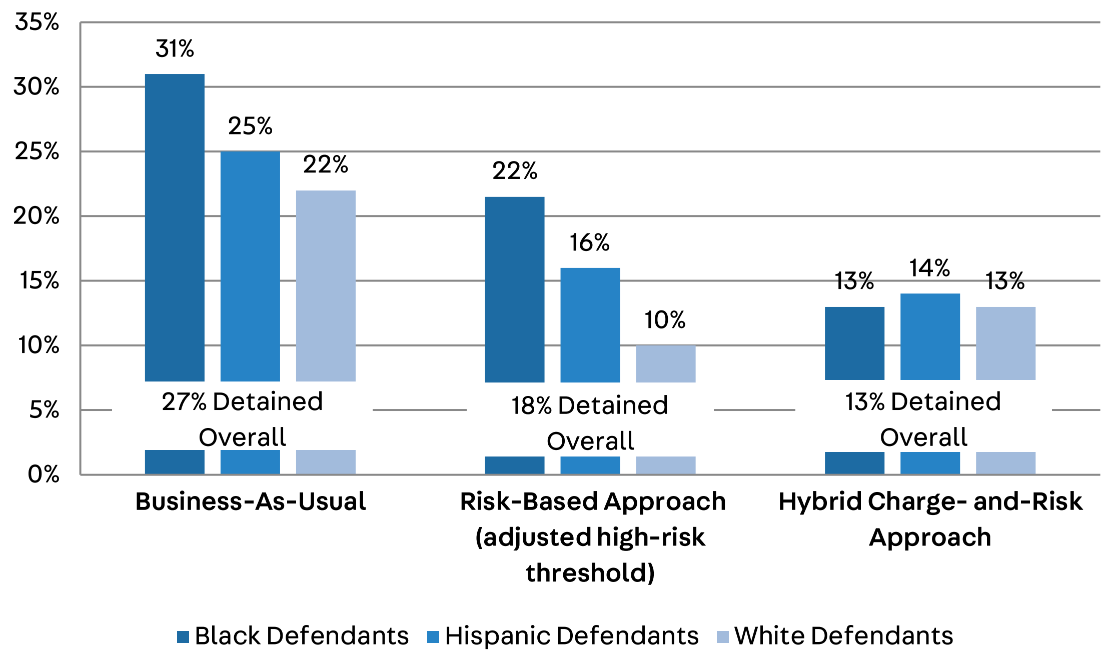
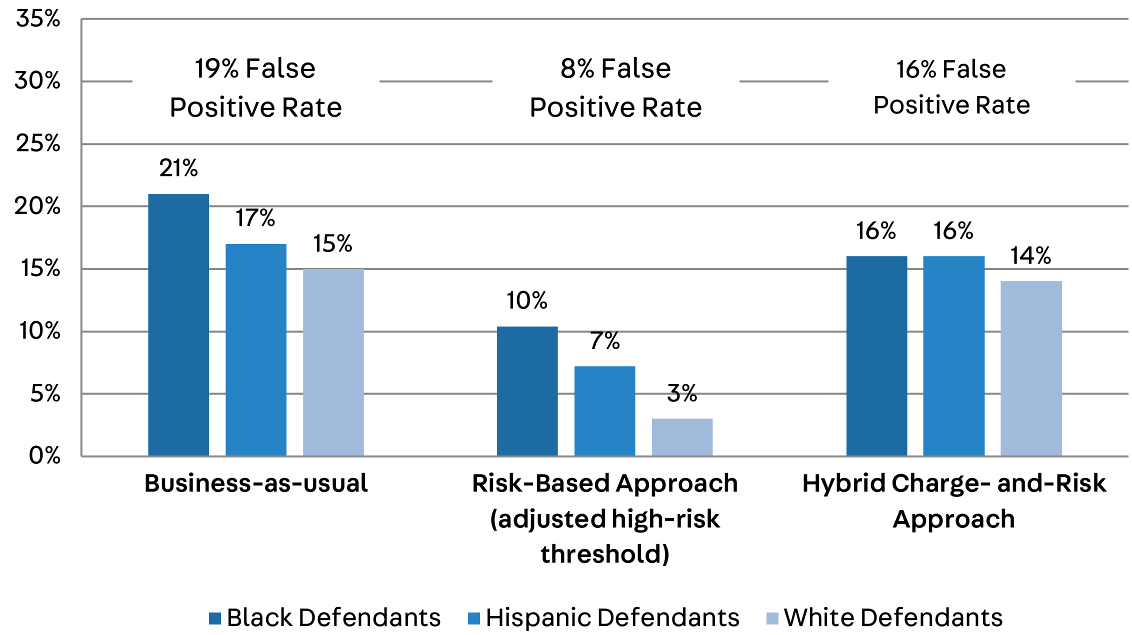


Exhibit 6. False Positive Rates by Race Under Three Decision-Making Scenarios
New York City Defendants, 2015



Moving Forward

Risk algorithms used to inform pretrial release decisions have shown promise for driving efforts to reduce pretrial incarceration but have also come under increasing fire. Critics argue that risk assessments that rely on factors such as criminal history will inevitably produce unfair outcomes, for black defendants in particular.

In this project, our explicit goal was to reach practitioners and policymakers dedicated to reducing the use of pretrial incarceration who are confused or alarmed by the debate over risk assessments and race. After reexamining the issue with data drawn from New York City, we advance three key conclusions for the field.

The first is that current "business-as-usual" approaches to pretrial decision-making fall short of achieving the goals of pretrial reformers, whether in terms of accurately assessing risk or improving racial fairness. The second is that concerns regarding the potential for risk assessments to perpetuate racial disparities are real, regardless of whether the tool in question is deemed unbiased in its algorithmic construction. Moreover, disparities are likely to prove especially wide in jurisdictions where black, Hispanic, or other racial or ethnic groups have disproportionate contact with the justice system. Unfortunately, even where the use of risk assessment tools lends itself to a reduction in pretrial detention rates for all groups, this dynamic may still result in persistent racial disparities in pretrial outcomes.

Our final conclusion is that, while the persistence of disparities is concerning, it is not an argument for abandoning the use of risk assessments in pretrial decision-making. We show that targeted risk-based pretrial strategies—specifically, a strategy of reserving pretrial detention only for defendants facing serious, violent charges and using risk-based decision-making only with those charges—holds significant potential for reducing both unnecessary detention and reducing racial disparities. Indeed, to the extent that risk assessments are thoughtfully applied to promote decarceration and alternatives to bail, in many jurisdictions they will, as a matter of course, be of particular benefit to defendants of color.

Too often the debate over risk assessments portrays them as either a technological panacea, or as evidence of the false promise of machine learning. The reality is they are neither. Risk assessments are tools with the *potential* to improve pretrial decision-making and enhance fairness. To realize this potential, the onus is on practitioners to consider a deliberate and modest approach to risk assessment, vigilantly gauging the technology's effects on both racial fairness and incarceration along the way.

Appendix A. Risk Assessment Tool: Risk Factors, Weights, Scoring, and Performance

Outcome Measure	Any Re-Arrest
Re-arrest Tracking Timeframe	Two Years
Final Sample Size [1]	**177,753**
RISK FACTOR	**Weight**[2]
Criminal History	
Any prior conviction	1
Any prior felony conviction in past 3 years	1
Total number of misdemeanor convictions in past 3 years	
1 misdemeanor conviction	*2*
2 misdemeanor convictions	*4*
3 or more misdemeanor convictions	*6*
Ten or more misdemeanor convictions in past 3 years	5
Any prior case with a failure to appear in court (FTA)	4
Total number of FTA cases in the past 3 years	
1 FTA case	*1*
2 FTA cases	*2*
3 or more FTA cases	*3*
Prior jail or prison sentence (0 or 1)	2
Current open case	3
Currently on probation	2
Demographics	
Age Category	
Up to 19 years old	*12*
20-24 years old	*10*
25-29 years old	*8*
30-39 years old	*6*
40-49 years old	*4*
50-59 years old	*2*
Under 25 years old	2
Male sex	2
Current Charge	
Current charge: Misdemeanor Property	1
Current charge: Felony drug possession	1
Current charge: Felony drug sales	1
Current charge: Felony weapons possession	1
Scoring	
Total Risk Score	0-43
Risk Categories	
Minimal Risk	0-6
Low Risk	7-12
Moderate	13-16
Moderate-high risk	17-19
High risk	20-43
Performance	
Area-under the Curve (AUC) Statistics	
Raw Risk Score	0.745
Risk Categories	0.733
Two Year Re-arrrest Rates by Risk Category	
Minimal Risk	10%
Low Risk	21%
Moderate	35%
Moderate-high risk	51%
High risk	72%

[1] Includes all arrests made of black, white or Hispanic defendants in New York City between January 1, 2015 and December 31, 2015.

[2] Weights represent the number of risk points associated with each risk factor.

Endnotes

[1] Stevenson, M.T. (2018). Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes. *The Journal of Law, Economics, and Organization*, 34 (4): 511–542.
[2] Zeng, Z. (2018). *Jail Inmates in 2016.* Washington, D.C.: U.S. Bureau of Justice Statistics.
[3] Dawes, R. M., Faust, D. & Meehl, P. E. (1989). Clinical vs. actuarial judgment. *Science*, *243*, 1668–1674; Gendreau, P., Little, T. & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34,* 575-607; Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology*, *98*(6), 1060.
[4] Angwin, J., Kirchner, L., Larson, J., & Surya, M. (2016). Machine Bias. *ProPublica*. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
[5] See, for example, Dieterich, W., Mendoza, C. and Brennan, T. (2016). COMPAS risk scales: Demonstrating accuracy equity and predictive parity. Northpointe, Inc. Retrieved from: http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf; Flores, A.W., Bechtel, K., & Lowenkamp, C. T. (2016). False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks. *Federal Probation, 80* (2): 38-46.
[6] As in many jurisdictions across the country, black and Hispanic defendants are disproportionately represented in the population of those arrested in New York City, when compared to the general population of the city. As of 2017, the U.S. Census reports New York City's population as 32% white, 29% Hispanic, and 24% black (see https://www.census.gov/quickfacts/newyorkcitynewyork).
[7] The risk assessment used for this study draws on a model developed in using a full sample of defendants arrested in New York City in 2012, as one component of an in-depth study of the jail population in New Yok City (see Rempel, M., Kerodal, A., Spadafore, J., and Mai, C. (2017). *Jail in New York City: Evidence-Based Opportunities for Reform*. Retrieved from http://www.courtinnovation.org/sites/default/files/documents/NYC_Path_Analysis_Final%20Report.pdf). For the current study, risk factor weights that were originally developed to predict recidivism separately among defendants with a current felony or current misdemeanor charge were revised to optimize predictive accuracy for a mixed population of felony and misdemeanor defendants arrested in 2014. (Misdemeanor weights were retained where they performed equally well for the mixed misdemeanor and felony population, or were reweighted toward the felony weight where appropriate). Finally, a split-sample methodology was used to both confirm the accuracy and validate the revised tool, which has been applied to the 2015 arrestee population for the current study.
[8] Isolating the pretrial period (prior to a case disposition) for each case or controlling for length of pretrial detention for those cases detained was outside the scope of the current study However, the same risk algorithm was previously validated for a shorter six-month timeframe that more closely approximates the typical length of the pretrial period (see Rempel et al., 2017, Op Cit.)
[9] Our analyses found statistically significant correlations between defendant race and many of the individual criminal history factors utilized in our tool. In particular, when compared to white and Hispanic defendants, black defendants in the sample were more likely to have a prior conviction ($r=.122$, $p<.001$), a prior sentence to jail or prison ($r=.150$, $p<.001$), and a prior warrant for failure to appear ($r=.146$, $p<.001$) on their record. As discussed at length in our findings and in other recent research on risk assessment and racial bias (e.g., see note 11 below), these types of correlations contribute to higher average risk scores and risk classifications for black defendants, and may disproportionately expose them to harsher consequences.
[10] For a more in-depth discussion of the antecedent causes of racial disparities in the criminal justice system, see Hinton, E., Henderson, L. & Reed, E. (May 2018). An Unjust Burden: The Disparate Treatment of Black Americans in the Criminal Justice System. New York: Vera Institute for Justice. Retrieved from https://www.vera.org/publications/for-the-record-unjust-burden
[11] Skeem, J. & Lowenkamp, C. (2016). Risk, Race, and Recidivism: Predictive Bias and Disparate Impact. *Criminology, 54*(4), 680-712.
[12] Angwin, J., Kirchner, L., Larson, J., & Surya, M. (2016). Machine Bias. *ProPublica*. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
[13] The logic used to create risk-based decision matrices will vary by jurisdiction and tool, but always will require the collapse of score ranges into categories associated with particular recommended outcomes (e.g., bail, supervised release, detention hearing).
[14] E.g., see Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, *5*(2), 153-163.
[15] Specifically, false positive results affected 11,460 black defendants; 7,466 Hispanic defendants; and 1,822 white defendants.
[16] The development of algorithmic adjustments that reduce tool bias or disparate outcomes is the subject of ongoing study in the scholarly literature, but applications of these adjustments to practice may face significant legal, practical and ethical obstacles. For an accessible discussion of some of these issues, see Drosser, C. (2017, December 22). In Order Not to Discriminate, We Might Have to Discriminate. *Simons Institute: Theory of Computing.* Retrieved from https://simons.berkeley.edu/news/algorithms-discrimination.
[17] Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017, August). Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 797-806). ACM.
[18] See New York City Criminal Justice Agency. (2018). *Annual Report 2016*. New York, NY: Criminal Justice Agency. Additionally, in the status quo, New York State law prevents judges from explicitly considering consider public safety risk in their decision-making.
[19] Not shown in our exhibits: In a scenario where New York City used domestic violence or violent felony charge as a proxy for risk without recourse to risk assessment or attorney recommendation, overall detention rates would decrease modestly compared to business-as-usual (19% overall; 21% among black defendants; 19% among Hispanic defendants; and 15% among white defendants). False positive rates would resemble those in the business-as-usual scenario (21% among black defendants; 20% among Hispanic defendants; and 15% among white defendants).
[20] Not shown in our exhibits: In a scenario where detention was reserved for only those defendants with a violent charge in the *highest* risk category, we observe a significant reduction in pretrial detention (3.1% overall) and false positive rates (2% overall), as well as an alleviation in racial disparities in the area of false positives similar to that observed in Scenario 3 (2.4% among black defendants; 1.5% among Hispanic defendants; and .4% among white defendants).

Center
for
Court
Innovation

520 Eighth Avenue
New York, NY 10018
p. 646.386.3100
courtinnovation.org