NAPSA

PRETRIAL ASSESSMENTS

An Updated Discussion on Bias and Disparities

A Review of Research

2026
Version 2.0

WEB:
www.napsa.org

EMAIL:
info@napsa.org

Pretrial Assessments

An Updated Discussion on Bias and Disparities

A Review of Research - Updated 2026 (V2)

Prepared By:
National Association of Pretrial Services Agencies (NAPSA)

Purpose:
To examine claims about bias and disparities in pretrial assessment tools, inform stakeholders with broad evidence, and advance pretrial justice.

Original Publication Date:
January 3, 2024

Revision Date:
February 23, 2026

Contact Information:
Info@napsa.org

www.napsa.org

Acknowledgements
NAPSA would like to thank the following individuals who contributed their suggestions and data for this publication. This publication would not have been possible without their contribution and support.

Introduction

Over the years, the National Association of Pretrial Services Agencies (NAPSA) has been closely monitoring claims of disparities and bias in actuarial pretrial assessments. We wrote our first review on this topic in 2023 and this version represents our 2026 updates, including additional contributors and research. This paper does not attempt to redefine fairness or adjudicate philosophical questions about pretrial assessments. Instead, it summarizes the empirical evidence from validation studies to clarify what has—and has not—been found regarding bias and disparities in pretrial assessment tools.

As the debate evolves, NAPSA has noted that a substantial body of research is emerging, providing clarity to this polarizing topic. We have long advocated for pretrial research and have encouraged the exploration of pretrial outcomes for over 50 years. NAPSA continues to promote and encourage research on the validity of pretrial assessment and potential disparities and biases. The evidence itself should lead this debate and must be controlled by rational, objective analyses.

The debate over potential bias in pretrial assessment tools intensified when a 2016 ProPublica article[1] claimed that the COMPAS tool demonstrated racial bias in its risk predictions. In response, COMPAS's developers argued that ProPublica had misinterpreted the data and applied flawed statistical assumptions[2]. Rather than being a simple matter of factual error, subsequent critiques have noted that ProPublica relied on "error rate balance" as its metric—an approach that is not aligned with what most experts regard as the most appropriate standard for fairness: calibration. That is, rather than testing whether predictions were equally accurate across groups (the essence of calibration), ProPublica focused on whether false positive rates were equal. Several scholarly studies have since countered ProPublica's conclusions by showing that its fairness claims depended on this choice of metric and did not reflect actual bias in predictive performance (Chouldechova, 2017; Dieterich et al., 2016; Feller et al., 2016; Flores et al., 2016; Rudin et al., 2020)[3].

Because of the mixed results and ongoing debate surrounding the COMPAS tool—particularly the use of proprietary algorithms and the divergent approaches to measuring bias[4]—NAPSA does not include the original ProPublica research in our presentation of studies examining disparities in pretrial assessments. While ProPublica's challenge was significant in raising public awareness, subsequent

---

[1] Larson, Mattu, Kirchner and Angwin, (2016). How We Analyzed the COMPAS Recidivism Algorithm

[2] The Northpointe Suite - Response to ProPublica (2018): Demonstrating accuracy equity and predictive parity.

[3] Desmarais, S. L., Monahan, J., & Austin, J. (2022). The Empirical Case for Pretrial Risk Assessment Instruments. Criminal Justice and Behavior, 49(6), 807–816.

[4] NAPSA recommends transparent, not-for-profit, and locally validated assessment tools. The COMPAS "is a proprietary instrument that is hidden from public view. Northpointe (now, Equivant), the company that owns the COMPAS, refuses to share the factors, weights, scaling, and detailed validations of the COMPAS. This lack of transparency is problematic." (from DeMichele, Matthew and Baumgartner, Peter and Wenger, Michael and Barrick, Kelle and Comfort, Megan and Misra, Shilpi, The Public Safety Assessment: A Re-Validation and Assessment of Predictive Utility and Differential Prediction by Race and gender in Kentucky (April 25, 2018), page 55.

research highlighted that competing interpretations of bias metrics (e.g., error rate balance vs. calibration) led to differing conclusions. As experts now widely favor calibration as a more appropriate metric, our National Standards on Pretrial Release (2024 Edition)[5] encourage the use of transparent tools and call for ongoing research into potential racial and gender disparities in validation processes. We recognize the value of the broader body of research that has emerged from this discussion.

In 2017, the National Institute of Corrections conducted a review of existing research and concluded that pretrial assessments enhance outcomes and reduce disparities[6]. More recently, researchers published a review of the research in the Criminal Justice and Behavior Journal, stating, *"scientific evidence suggests [pretrial assessments] can be an effective strategy to help achieve pretrial system change, including reducing pretrial detention for people of color and white people, alike, when their results are actually used to inform decision-making."*[7] Even more recently, the updates to the Federal Pretrial Risk Assessment brought the most illuminating potential to pretrial assessments by potentially reducing racial disparities, rather than aggravating them. This is further referenced and discussed later in this paper.

Although this controversy has quieted in the past few years, perhaps due to research reviews like these, NAPSA feels that it is important to continue to inform this topic with the latest evidence. We want to stay out in front of this issue to ensure clarity and guidance, as many jurisdictions have invested tremendous time, resources and effort in pretrial assessments. In the past decade, even some in academia have raised concerns about this[8], prompting us to review more carefully the existing scientific literature on this topic. NAPSA noted that much of what was missing from these claims was specific references to pretrial assessment research. We asked, "What research is actually available that can most objectively guide our perspectives on this?" Many NAPSA board members have undertaken efforts to conduct further research on this topic, thereby enhancing NAPSA's voice in this conversation. As time has progressed, so has the research, and the past several years have yielded numerous additional publications on this topic. NAPSA hopes this research is helpful in informing the field and looks forward to emerging research that will continue to help refine pretrial assessments.

[5] NAPSA Pretrial Standards 2024 https://www.napsa.org/standards

[6] "Myths and Facts: Using Risk and Need Assessments to Enhance Outcomes and Reduce Disparities in the Criminal Justice System", 2017

[7] Desmarais, S. L., Monahan, J., & Austin, J. (2022). The Empirical Case for Pretrial Risk Assessment Instruments. Criminal Justice and Behavior, 49(6), 807–816.

[8] Martha Minow, Jonathan Zittrain, John Bowers (2019).Technical Flaws of Pretrial Risk Assessments Raise Grave Concerns.

Defining Our Terms - Bias and Disparity

Misunderstandings about terms like bias, disparity, and related concepts have contributed to confusion—and, at times, misplaced criticism—of pretrial assessment tools[9]. In public discussions, the existence of group differences in outcomes is sometimes taken as evidence that a tool is inherently biased or discriminatory, without distinguishing between statistical disparities, predictive bias, and broader questions about racism or sexism within the criminal legal system. Clarifying these distinctions is essential for informed decision-making and for the responsible use of pretrial assessments as decision-support tools, consistent with NAPSA's Standards on Pretrial Release.

To support clarity, we offer the following definitions grounded in statistical principles and commonly used in validation research. Importantly, in this report the term *bias* is used in its technical, statistical sense, referring to the accuracy of risk estimation, and not as a claim about broader structural fairness. However, while these findings speak only to the statistical performance of the tools, they may still inform broader inferences about fairness in the pretrial process—particularly when viewed alongside observed outcome patterns and systemic context. These definitions are intended to help practitioners distinguish between observable differences in outcomes and actual inaccuracies in how a tool estimates risk, while reinforcing the importance of local validation, routine revalidation, and careful interpretation of assessment results. Findings—or the absence of findings—regarding predictive bias should not be interpreted as judgments about the existence—or absence—of broader social or historical outcome differences.

Disparity

Technical definition: A disparity is a noticeable and statistically meaningful difference in outcomes or distributions between groups[10].

Pretrial assessment context: A pretrial assessment disparity refers to statistically significant differences in scores, classifications, or observed outcomes across demographic groups (such as sex or race[11]). Disparities describe differences in results, but do not, by themselves, indicate bias or invalidate an assessment tool.

---

[9] As used in this report, the terms disparity and bias are employed for analytic and descriptive purposes only and are not intended to constitute legal determinations or to invoke definitions of discrimination, disparate treatment, or disparate impact under Title VI of the Civil Rights Act of 1964 or other applicable civil rights laws.

[10] "disparity" https://www.merriam-webster.com/dictionary/disparities. 2023.

[11] NAPSA recognizes that race is a social construct, not a biological determinant. While race is commonly tracked in justice system data to monitor disparities, genetic science confirms that human beings do not fall into distinct biological races. This does not mean racism is not real — only that it operates on socially defined categories, not genetic ones. National Human Genome Research Institute. (2020). *DNA, Genes, and Race*. National Institutes of Health.

Bias

Technical definition: Bias is a systematic deviation between a predicted value and the true value it is intended to estimate.[12]

Pretrial assessment context: A pretrial assessment bias occurs when a tool systematically over- or under-estimates risk for one demographic group compared to another group with the same observed likelihood of pretrial success or failure. In practical terms, bias exists if individuals from different groups receive the same risk score but consistently experience different outcomes. When such discrepancies are identified, they are evaluated for both statistical and practical significance and may warrant recalibration or adjustment, consistent with NAPSA's emphasis on continuous improvement and routine revalidation.[13]

Bias vs. Disparities – Simple Pretrial Practitioner Definition

- **Disparity** means there's a difference in outcomes between groups (e.g., men vs. women, or racial groups). This could be due to many factors—not necessarily a flaw in the tool.

- **Bias** means the tool is systematically inaccurate for one group compared to another. For example, if individuals from one group consistently have *better or worse outcomes* than others with the same risk score, the tool may be biased.

In short: ***Disparities*** mean there are differences in outcomes between groups. ***Bias*** means the tool is working unfairly—if people from different groups get the same score, but one group ends up doing worse or better than the other, that's a sign of bias.

The researchers cited in this review may have used different definitions or statistical thresholds for bias and disparities. NAPSA's definitions are offered as a common-sense framework for understanding these terms across studies.

Review of Pretrial Assessment Studies

The following research summary reveals that the preponderance of empirical research on pretrial risk assessment tools finds no evidence of widespread racial or sex bias. Where isolated disparities have been observed, researchers have not recommended abandoning the instruments; instead, they have recommended improving them through recalibration, validation updates, or jurisdiction-specific adjustments. Taken as a whole, the research supports continued use of these tools with appropriate oversight and periodic revalidation.

---

[12] "bias." https://www.merriam-webster.com/dictionary/bias.

[13] NAPSA 2026 - Approved by NAPSA Board by unanimous vote on February 23, 2026

NAPSA RESEARCH SUMMARY		
Pretrial Assessment name (Alphabetical) with links to validation studies (if available)	**Gender[14] Bias or Disparities Identified (Yes/No)**	**Racial Bias[15] or Disparities Identified (Yes/No)**
Allegheny County, PA - PRA (2020)	No - re-arrest, Yes FTA	No - re-arrest, Yes FTA
Multi-State Analysis - PSA (2024)	No	No
Bernalillo County, NM PSA (2021)	No	No
Bernalillo County, NM PSA (2024)	No	No*
Colorado - CPAT-R (2020)	No	No
El Paso, TX EPPRA-R (2022)	No	No
EL Paso, TX EPPRA-R (2024)	No	No
Federal PTRA (2018)	No	No
Federal PTRA (2024)	Not Evaluated	No
Fulton County, GA (2023)	No	No
Houston, TX PSA (2020)	No	No
Kentucky PSA (2020)	No	No
Kentucky PSA (2021)	No	Not Evaluated
Kentucky PSA (2023)	No	Not Evaluated
Indiana Pooled Study (2020)	No	Yes*

[14] This updated report uses the term “sex” or “gender” to refer to the biological classification of individuals as male or female, consistent with how demographic information is recorded in most criminal justice administrative data and used in the cited validation studies. Gender identity was not captured in the datasets analyzed, unless otherwise noted by the researchers.

[15] Of the four jurisdictions in which predictive bias was identified, all involved outcome-specific or subgroup-specific calibration issues that were subsequently corrected or addressed, and none resulted in recommendations to abandon the assessment tool.

Minnesota MPAT (2021)	Not Evaluated	Yes
Minnesota MPAT (2023)	No	Yes
NYC Criminal Justice Agency (2020)	No	No
Pierce County, WA PSA (2023)	No	No
San Francisco County, CA PSA (2021)	Yes	Yes
Thurston County, WA PSA (2023)	No	No
Virginia VPRAI (2016)	No	No
Volusia County, FL PSA (2021)	No	No
Washington DC RA (2018)	No	No

This review includes publicly available, peer-reviewed, or jurisdiction-validated studies known to NAPSA at the time of publication. Both favorable and unfavorable findings are included.

Allegheny County Pretrial Risk Assessment Instrument Revalidation Study - Pennsylvania (2020)[16] - While the NCA instrument seemed well calibrated, the FTA instrument showed signs of miscalibration with regard to race and sex. A revised model largely corrected the FTA deficiency. The violence flag appeared robust, and the researchers stated that it should be incorporated into pretrial release decision-making. The corrected and revised risk assessment was recommended by the researchers and has since been implemented.

A Multi-Site Evaluation of a Pretrial Risk Assessment Tool (2024) - Amid concerns that the use of assessments may further contribute to the disparities found in the criminal legal system (Eckhouse et al., 2019; Mayson, 2019; Woldgabreal et al., 2020), this study supports the PSA as a valid predictor of pretrial outcomes across the six gender and racial-ethnic groups examined (i.e., White males, White females, Black males, Black females, Hispanic males, Hispanic females). The analyses showed no evidence of differential predictive performance across the six racial-ethnic and gender groupings for the three PSA tools. The regression models failed to produce proof of predictive bias when testing for moderation.[17]

[16] Allegheny County, Pennsylvania - Pretrial Risk Assessment Instrument Revalidation Study Technical Report Avinash Bhati, PhD, Maxarth LLC, 2020.
[17] Demichele, M., Silver, I. A., Labrecque, R. M., Dawes, D., Lattimore, P. K., & Tueller, S. (2024). Testing Predictive Biases at the Intersection of Race-Ethnicity and gender: A Multi-Site Evaluation of a Pretrial Risk Assessment Tool. Criminal Justice and Behavior, 0(0).

Bernalillo County, New Mexico - Public Safety Assessment (PSA) – (2021) - Overall, the key findings of this report indicate that while differences occur by race, they were not statistically significant. Gender differences in scores are observed for New Criminal Activity (NCA) and New Violent Criminal Activity (NVCA) among individuals without the NVCA flag. The predictive validity of the PSA is demonstrably comparable to other jurisdictions, and scores are similar across groups, indicating that the tool is beneficial regardless of the individual's race or sex.[18]

Bernalillo County, New Mexico – Public Safety Assessment (PSA) – (2024)* – The 2024 revalidation found that the PSA continues to exhibit fair-to-good predictive accuracy across demographic groups. Differences by race and gender were observed in base rates but were not statistically significant. The PSA's predictive validity for Failure to Appear (FTA) and New Criminal Activity (NCA) was consistent across racial and gender groups. However, the New Violent Criminal Activity (NVCA) flag showed reduced predictive performance for Black defendants, and Native American defendants had lower-than-expected NCA outcomes at equivalent scores. The authors concluded that these results reflect differences in calibration and accuracy rather than systematic bias or disparity. The PSA remains a reliable and equitable tool for Bernalillo County when locally validated and regularly reviewed. *Minor subgroup calibration and predictive-performance differences were noted for the NVCA flag; not classified as bias or disparity per NAPSA definitions.

Colorado Pretrial Assessment Tool-Revised - CPAT-R (2020): The ACLU raised claims of bias and disparities against the CPAT-R based on research conducted by the University of Northern Colorado (UNC)[19]. This resulted in a few news articles and special-interest opposition[20]. In response, the researchers from UNC reported that "the analyses did not provide support for predictive bias across race/ethnicity, gender, and residential status".[21] The researchers concluded that their study was misinterpreted by the critics.[22]

El Paso County, Texas - EPPRA-R - El Paso Pretrial Research Assessment - Revised (2024) - Researchers from The University of Texas at El Paso conducted an evaluation of predictive utility and bias in the EPPRA-R. Analyses confirmed that there were no significant differences in the tool's risk predictions for defendants in El Paso County by race, ethnicity, gender, or age. When accounting for all

[18] Bernalillo County Public Safety Assessment Validation Study - Prepared for: New Mexico Administrative Office of the Courts and Bernalillo County; Prepared by: Elise Ferguson, M.A., Helen De La Cerda, B.A., Paul Guerin, Ph.D., and Cristopher Moore, Ph.D., June 2021

[19] Dangerous, Misleading and Biased: A Letter on Pretrial Risk Assessment Tools in Colorado. ACLU Colorado, October 29, 2020.

[20] Bias against Black people found in Colorado bail reform tool, John Herrick, 2020

[21] Colorado Pretrial Assessment Tool Validation Project Final Report, Victoria A. Terranova, PhD, Kyle C. Ward, PhD, July 1, 2020 Department of Criminology and Criminal Justice University of Northern Colorado

[22] Response document by the University of Northern Colorado to an ACLU letter disseminated on October 29, 2020, Victoria A. Terranova, PhD, Kyle C. Ward, PhD, November 2020.

defendants (both defendants that were released and detained pretrial), statistical analyses remained consistent – no biases were found for subgroups of defendants in the tool's predicted risk.[23]

El Paso County, Texas - EPPRA-R - El Paso Pretrial Research Assessment - Revised (2025) - A 2025 study of El Paso County's Research Assessment-Revised (EPPRA-R) tool conducted by researchers from The University of Texas at El Paso examined whether the tool performed differently before and after the onset of the COVID-19 pandemic. The evaluation found no evidence of predictive bias across racial or ethnic groups. The EPPRA-R showed similar utility in predicting arrests and FTA, despite fluctuations in pretrial decision-making and release rates. Notably, the assessment continued to perform equitably across groups, with the exception of minor, statistically insignificant disparities that were attributable to changes in release practices rather than to the tool itself. These findings support the conclusion that, when properly validated, actuarial tools can maintain fairness even amid systemic disruptions like COVID-19.[24]

Federal PTRA (2018) - "This research demonstrates that the PTRA can predict violations irrespective of the defendant's race, ethnicity, and sex. The researchers stated that these findings are supportive of a growing literature showing that risk instruments like the PTRA can be used to assess recidivism risk and inform criminal justice decisions without exacerbating biases in the criminal justice system."[25]

Federal PTRA (2024) – Racial Disparity in Federal Pretrial Detention Recommendations – This study examined how racial disparities in federal pretrial detention recommendations have changed over the past two decades and how those trends relate to the implementation of the Pretrial Risk Assessment (PTRA). Using a nationally representative dataset of more than 650,000 federal pretrial cases from fiscal years 2005 through the first half of 2024, the authors measured monthly relative risk ratios comparing the likelihood that Black and White defendants receive detention recommendations from pretrial services officers. Over the nearly 20-year series, racial disparities decreased substantially—an estimated 75 percent reduction in the relative risk ratio between 2004 and 2024—suggesting a pronounced downward trend in disparity coinciding with the widespread adoption of the PTRA. Interrupted time-series analysis indicated statistically significant reductions in racial disparity following major implementation milestones in 2011 (when PTRA administration reached a systemwide threshold) and continued declines thereafter. These findings support the proposition that structured, empirically

[23] Validation of the El Paso Pretrial Research Assessment Instrument – Revised, May 2024, University of Texas at El Paso, Chelsea Queen, MA, Jennifer Eno Louden, PhD.
[24] Eno Louden, J., & Queen, C. (2025). *The effects of COVID-19 on the utility of the El Paso County Research Assessment–Revised*. The University of Texas at El Paso.
[25] Cohen, Lowencamp, Hicks; (2018) Revalidating the Federal Pretrial Risk Assessment Instrument (PTRA)

grounded instruments like the PTRA may contribute to reducing racial disparities in pretrial detention recommendations, even within a system historically characterized by differential outcomes.[26]

Fulton County, GA - PSA (2023) - "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)". [27]

Houston, Texas - PSA – (2020) – A Harvard study found no substantial evidence that PSA scales performed differently across racial and gender groups. Although some analyses showed statistically significant differences in PSA performance across demographic groups, the differences were substantively small and/or directionally contradictory (i.e., one scale showed higher failure rates for blacks than whites, while another scale showed the opposite).[28]

Indiana Pretrial Assessment Tool (IRAS-PAT) Subgroup Analysis (2020) – This five-county validation found no predictive performance differences or calibration issues by age or sex. However, notable subgroup calibration differences emerged by race. The tool yielded weaker predictive validity estimates for Black defendants, particularly regarding court appearances and arrest outcomes. This performance difference stemmed from high-risk miscalibration; Black defendants classified at high risk exhibited lower actual misconduct rates than similarly classified White defendants (predictive... pp. 2, 28). Item-level analyses indicated this gap was driven by criminal history and residential instability metrics, which may reflect systemic socioeconomic disadvantages rather than individual risk. Because the tool maintained fair utility overall, researchers recommended local calibration adjustments and ongoing validation rather than abandonment (predictive... pp. 2-3, 28).

Kentucky PSA (2020) – Using a statewide data set from Kentucky, researchers found that while the Public Safety Assessment's scores for failure to appear were moderated by race, these differences did not lead to disparate impacts. "We applied a regression modeling approach commonly used to assess bias in test instruments (e.g., cognitive and employment testing), and found some instances of differential prediction by race. These differences suggest that the PSA scores to predict failure to

[26] Lowenkamp, C., Skeem, J. L., & Montoya, L. (2024). *Racial disparity in federal pretrial detention recommendations: Trends over two decades and association with risk assessment implementation*. *Federal Probation*, 88(2).
[27] DeMichele, Tueller, Inkpen, Dawes, Lattimore (2023), Validation and Predictive Bias Testing of the Public Safety Assessment for Fulton County, Georgia Advancing Pretrial Policy and Research, RTI International.
[28] Greiner, Stubenberg, Halen (2020); Validation of the PSA in Harris County, TX - Harvard Law School Access to Justice Lab.

appear (FTA) are moderated by race, with no significant differences found for new crimes and new violent crimes between black and white defendants." [29]

Kentucky PSA (2021) – DeMichele and Baumgartner (2021) examined whether the Public Safety Assessment (PSA) produces racially biased classifications for new criminal arrests by testing error rates for Black and White defendants in Kentucky's statewide pretrial system. The study finds that while false positive and false negative rates differ modestly by race, these differences reflect underlying differences in risk factor distributions and tradeoffs inherent in prediction—not systematic bias in how the PSA estimates risk. Importantly, the PSA demonstrates similar arrest rates for Black and White defendants who receive the same risk scores, supporting equitable risk classification at the score level. The authors emphasize that no single fairness metric can fully capture equity when base rates differ and argue that avoiding unnecessary detention—especially false positives—should be a central policy priority. The study underscores the need for clear standards, local validation, and transparency when using risk assessments as decision-support tools rather than automatic determinants.

Kentucky PSA (2023) – This study was notable for identifying equal base rates by gender for failure to appear (FTA), which allowed researchers to isolate and assess potential disparities in error rates independent of underlying outcome differences. The PSA demonstrated predictive validity within acceptable thresholds for both FTA and new arrest outcomes, and the analyses revealed no statistically significant predictive bias by sex across five distinct error rate metrics. These findings contribute to ongoing methodological discussions about how best to define and evaluate bias in risk assessment tools, including whether parity in error rates should be prioritized over other fairness criteria.

Minnesota MNPAT (2021) – Although the researchers found that the current MNPAT is both underestimating risk and overestimating risk of pretrial failure among racial groups, they concluded that "there is no indication that using the MNPAT leads to less favorable outcomes for any racial group than if a risk assessment tool is not used. In fact, research suggests that using a tool improves pretrial outcomes…The Committee is confident, based on the preliminary data and available research, that use of the current tool results in better outcomes for the majority of cases than not using an assessment tool." After reviewing these findings, the committee decided to seek improvements to the MNPAT rather than maintaining the status quo.[30]

Minnesota MNPAT (2023) – The MPAT found similar predictive performance between males and females but identified racial bias in the data. Although the MNPAT was found to be predictive for White, Asian/Native Hawaiian, Hispanic, and "Other", it was not found to be as predictive for Black and Native American people, and the levels may have misrepresented the actual risk. Adjustments were made to

[29] DeMichele M., Baumgartner P., Wenger M., Barrick K., Comfort M., 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.

[30] Minnesota Pretrial Release Evaluation Form and Assessment Tool (MNPAT) Validation Study, Minnesota Judicial Branch, 2021.

the tool, and a Validation Committee considered three models. After extensive review, discussion, and analysis of other potential risk assessments, the Validation Committee chose a model that was more predictive than the current MNPAT and did not exhibit racial or gender bias.[31] The Committee found that their own tool was superior to existing tools.

New York City Criminal Justice Agency Release Assessment Update (June 2020) – Using the largest sample size in any study of over 1.6 million… "One of the major objectives of updating the Release Assessment was to address concerns about disparities. The updated model was designed to promote greater equity while maintaining predictive accuracy… Importantly, the updated Release Assessment reduces disparities in recommendations by race, ethnicity, and gender. The racial and ethnic disparity in release recommendations decreased by approximately half—from 9.4 percentage points under the prior tool to 4.3 percentage points under the updated tool...The gender disparity also decreased substantially."[32]

Pierce County, WA, PSA (2023) – The researchers "did not find any evidence of predictive bias by race or gender for the PSA in Pierce County. The predictions from the PSA for FTA, NCA, and NVCA are not different for White individuals and people of color or males and females in the historical Pierce County data."[33]

San Francisco, CA - PSA (2021) - The analysis showed predictive bias in the FTA, NCA, and NVCA scales. The results of the moderator regressions indicate that the FTA scale is calibrated by gender but not by race and ethnicity for the population studied. Further, the NCA scale is calibrated by race and ethnicity, but not by gender. Finally, the NVCA flag is not calibrated by race; rather, it is calibrated by ethnicity and gender. Overall, these results suggest that, for each of the tool's scales, the relationship between the tool's score and the outcome differs across demographic groups. The researchers recommended working closely with the tool developer to ensure that any modifications to the scales do not invalidate the tool altogether, and to re-validate every three years.[34]

Thurston County, WA - PSA (2023) - The researchers reported no evidence that people of color or women are being scored higher than their actual outcome rates (i.e., no overprediction). These results were based on a historical cohort.[35]

---

[31] Medhanie, Ph.D. Hoheisel, M.P.P., Ogunleye, M.P.H., Welter, M.A. (2023). Minnesota Pretrial Assessment Tool Validation Study.

[32] Luminosity & New York City Criminal Justice Agency. (2020). *Updating the NYC Criminal Justice Agency Release Assessment.* New York, NY: NYC Criminal Justice Agency. Retrieved from https://www.nycja.org/assets/Updating-the-NYC-Criminal-Justice-Agency-Release-Assessment-Final-Report-June-2020.pdf

[33] DeMichele, Tueller, Burtch, Dawes, Lattimore, RTI International, (2023); Validation and Predictive Bias Testing of the Public Safety Assessment for Pierce County, Washington Advancing Pretrial Policy and Research.

[34] Skog & Lacoe (2021). Validation of the PSA in San Francisco (capolicylab.org), Page 34, California Policy Lab,

[35] DeMichele, Tueller, Janda, Dawes, Lattimore, RTI International, (2023); Validation and Predictive Bias Testing of the Public Safety Assessment for Thurston County, Washington Advancing Pretrial Policy and Research.

Virginia VPRAI (2016): A revalidation of the Virginia Pretrial Risk Assessment Instrument (VPRAI) found that race was not a significant predictor of missed court dates or new arrests pretrial. Race also did not impact the relationship between individual or combined factors and any type of pretrial failure. The VPRAI predicted outcomes that were statistically significantly better for Whites than for Persons of Color. However, when the assessment's risk factors were weighted, summed, and collapsed into levels, the differences in predictive value were not statistically significant. Taken as a whole, the analyses support the VPRAI's neutrality in classifying People of Color and Whites by likelihood of pretrial failure.[36]

Volusia County, Florida - PSA (2021) - This peer-reviewed study found no racial disparities in the PSA assessment tool. There were some mixed outcomes between black and white defendants, where one group was arrested more than the other in certain areas, but overall, no significant differences were found.[37]

Washington, D.C. Pretrial Risk Assessment (2018) - The Pretrial Services Agency for the District of Columbia contracted with Maxarth, LLC, to complete an independent revalidation of the Agency's pretrial assessment instrument. The evaluation found that, although scores and misconduct rates vary by race, the relationship between scores and observed misconduct remained relatively stable across races. Where disparities were detected, the authors noted them as minimal and distributed evenly among all groups.[38]

Observations:

Although the vast majority of pretrial assessment research has not found evidence of disparities or bias, the instances where there were disparities highlighted the reason this research is important. The emerging standards and trends of researching this area is an advancement in the pretrial world and this body of evidence clarifies the state of the pretrial field on this topic. This 2026 NAPSA synthesis revealed that where subgroup differences or calibration issues arose, models were updated accordingly or the disparities were found not to indicate systemic bias.

Some critics contend that risk assessments provide a veneer of scientific objectivity to biased data, effectively "baking in" historical racial disparities into modern pretrial decisions through their reliance on arrest and conviction histories. This critique does not allege measurement error, but rather that the tools may reflect broader social conditions embedded in the justice system. Validation studies, however, are specifically designed to test whether the tool systematically misestimates risk across demographic groups. Across jurisdictions, the majority of the available evidence does not show consistent differences in outcomes among individuals with the same risk scores by race or ethnicity—

[36] Danner, M.J.E., VanNostrand, M., and Spruance, M.S. (2016). *Race and gender Neutral Pretrial Risk Assessment, Release Recommendations, and Supervision: VPRAI and PRAXIS Revised.* Luminosity: Ft. Lauderdale, FL.
[37] Brittain, B. J., Georges, L., & Martin, J. (2021). Examining the predictive validity of the Public Safety Assessment. Criminal Justice and Behavior, 48(10), 1431-1449.
[38] Pretrial Services Agency for the District of Columbia (2022). PSA's Risk Assessment Ensures Fair Administration of Pretrial Justice in the District of Columbia

results that would be expected if predictive bias were inherent to the tool itself. In other words, disaggregated empirical studies allow us to test whether risk tools actually propagate bias. The vast majority of validations show that calibrated assessments do not yield disparate predictive performance and often outperform unstructured judicial discretion. While a few jurisdictions have identified and corrected disparities, most peer-reviewed validations show either no increase in disparity—or measurable reductions. A 2024 federal analysis of the PTRA found a 75% reduction in racial disparities in detention recommendations over two decades. NAPSA does not seek to downplay findings of bias or disparity where they exist, but emphasizes that researchers have consistently recommended adjustments and recalibration—not abandonment—of these tools.

It is also important to recognize that the causes for the differences in arrest and justice system contact rates are not uniform across all demographic groups and reflect a complex interaction of social, economic, community, and behavioral factors—such as peer networks, substance use patterns, family structure[39], neighborhood conditions, and prior system exposure—rather than a single explanatory cause.[40] For example, nationally aggregated arrest data show that while Black individuals often have higher arrest rates per capita compared to Whites, other groups such as Asians tend to have lower arrest rates, and patterns may vary substantially by offense type and jurisdiction[41]. In some studies, Whites and Hispanics have higher rates in certain crimes than blacks, and vice versa.

Behavioral experiments with lay participants, like Green & Chen (2020)[42], do highlight risks of naïve application. Their simulation showed a 1.9% increase in disparity with lay participants under a specific cognitive bias scenario. However, such artificial settings do not reflect actual judicial practice, where judges apply their own expertise and experience in conjunction with the results of risk assessments to make decisions. By contrast, real-world studies of pretrial implementation—especially when paired with policy safeguards and training—offer stronger evidence that validated tools enhance consistency and fairness.

Empirical validations—especially those disaggregated by race and ethnicity—provide a means to test whether such bias propagation occurs in practice. Most studies find that well-calibrated tools do not systematically yield disparate predictive performance and often improve accuracy and consistency relative to unstructured decision-making. Therefore, this concern frequently conflates base-rate differences—such as varying arrest rates across groups—with predictive bias, which would mean the tool inaccurately estimates risk for one group relative to another. Validation studies are designed to detect this problem by testing calibration and predictive accuracy across racial and ethnic groups. While a small

---

[39] J. Kroese, *Growing up in single-parent families and the criminal involvement of adolescents* (2021)
[40] Lauren N. Gase *et al.*, "Understanding Racial and Ethnic Disparities in Arrest: The Role of Individual, Home, School, and Community Characteristics," *Race and Social Problems* 8, no. 4 (2016): 296–312
[41] Federal Bureau of Investigation, *Crime in the United States, 2019: Table 43 – Arrests by Race and Ethnicity*, U.S. Department of Justice, 2020, https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/tables/table-43
[42] Ben Green and Yiling Chen, "Algorithmic Risk Assessments Can Alter Human Decision-Making Processes in High-Stakes Government Contexts," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2 (2021): Article 457.

number of studies have found and corrected for predictive bias—such as in Allegheny County—the majority of peer-reviewed validations have not found meaningful disparities in tool performance.[43] Moreover, a 2024 systematic review of risk assessment instruments found that most studies reported risk tools did not increase racial or ethnic disparities—and in many cases were associated with significant reductions in disparate outcomes across settings.[44]

NAPSA's Position

In the 2024 NAPSA Pretrial Standards, the Association refined its guidance on pretrial assessments to emphasize greater clarity, transparency, and consistency in their development and use, while reaffirming the core principles established in earlier editions. The updated Standards make clear that validated pretrial assessments are intended to inform, not replace, individualized judicial decision-making. They are not designed to determine guilt, culpability or specific conditions. They should not be used as the sole or primary driver of pretrial detention nor to automatically impose predetermined, sweeping, or overly restrictive conditions of release.[45] Instead, pretrial assessment results are intended to support careful, individualized consideration of pretrial decisions consistent with the presumption of release and the least restrictive conditions necessary to promote court appearance and public safety.

The 2024 Standards further underscore the importance of local validation and routine revalidation to ensure that pretrial assessments remain accurate, fair, and responsive to the populations they serve, including ongoing examination of outcomes across demographic groups. Observed differences in outcomes are intended to prompt empirical analysis, transparency, and, when appropriate, recalibration or adjustment—rather than assumptions of bias or calls for abandonment. Consistent with the broader body of empirical research, NAPSA continues to support the appropriate use of pretrial assessments as a valuable evidence-based tool that can improve consistency, inform targeted support, and enhance pretrial success when properly implemented, while recognizing that assessments are only one component of an enhanced pretrial justice system.

To further confirm these conclusions, the most notable update to this paper is the 2024 longitudinal study on the Federal PTRA, in which the researchers found that implementation and widespread use of the tool was associated with a substantial reduction in racial disparities in detention recommendations—by as much as 75% over a 20-year period[46]. This evidence reinforces the

---

[43] *Desmarais, S.L., Monahan, J., & Austin, J. (2022). The Empirical Case for Pretrial Risk Assessment Instruments. Criminal Justice and Behavior, 49(6), 807–816.* https://doi.org/10.1177/00938548211041651

[44] Lawson, S. G., Narkewicz, E. L., & Vincent, G. M. (2024). *Disparate impact of risk assessment instruments: A systematic review*. *Law and Human Behavior*.

[45] NAPSA Standards 1.4 and 2.9 (2024)

[46] Christopher T. Lowenkamp, Jennifer L. Skeem & Lina Montoya, *Racial Disparity in Federal Pretrial Detention Recommendations: Trends Over Two Decades and Association with Risk Assessment Implementation* (Oct. 15, 2024), *Federal Probation Journal* (2024),(finding a 75 percent decrease in racial disparities in detention recommendations over 2004–2024 coinciding with implementation of the Pretrial Risk Assessment).

position that properly implemented pretrial assessments can support better outcomes and reduce bias and disparities, rather than contribute to them. Findings like this should be recognized and celebrated as a sign of responsible, data-driven improvement.

NAPSA hopes that this information helps inform stakeholders, practitioners, the public, and the media. Please let us know about research that may emerge or be missing from this review. We are committed to continued updates to this paper as evidence continues to emerge.

NAPSA would like to thank all jurisdictions that are tirelessly researching pretrial justice to better inform and enhance the field.