Conducting Anti-Racist Research on Pretrial Release Assessments

APPR Research Brief, April 2024
by Megan Comfort, Jenn Rineer, Elizabeth Tibaduiza, and Monica Sheppard

Understanding Actuarial Pretrial Release Assessments

The "pretrial process" refers to the events that happen between the time that one is suspected by law enforcement of violating the law and the time that charges are dismissed, the case is otherwise resolved, or the trial process begins. During the pretrial period, people are considered innocent under the law. The U.S. Supreme Court[1] has stated, "In our society, liberty is the norm, and detention prior to trial or without trial is the carefully limited exception." The only two constitutionally valid reasons for holding someone in jail during the pretrial period are (1) to prevent flight or (2) to prevent harm to people in the community.

Judges make decisions every day about whether to detain or release people going through the pretrial process, as well as about what conditions of release may be needed to help people succeed. Pretrial release assessments are designed to inform their decisions. Unlike assessments that involve a clinician or other professional drawing on their subjective expertise to make a recommendation, *actuarial pretrial release assessments*[a] rely on mathematical processes. Using large data sets with information about people who previously went through the pretrial process, researchers identify factors related to appearing for court hearings and not being arrested again if released. The researchers then create a sequence of instructions for a computer to follow (called an algorithm) that uses these factors to calculate an estimated likelihood that a person will appear in court and remain arrest free while their case is being resolved. This calculation—referred to as a "score"—is provided to the judge as information to consider when making decisions about pretrial release. A person's score is also often provided as information to other courtroom actors, such as prosecutors, defense attorneys, and pretrial services officers.

When thinking about actuarial pretrial release assessments, it is important to understand the history of the criminal legal system in the United States, which is deeply rooted in the legacy of slavery. Read Race and the Criminal Justice System[2] by the Equal Justice Initiative to learn more.

No actuarial pretrial release assessment tool or instrument is considered standard. Numerous assessments have been developed, and they vary in terms of the factors and instructions entered in the algorithm. Some use factors that are available through criminal legal system records, such as whether someone has been arrested before or has previously missed a court date. Others include factors like whether someone has a job, is enrolled in a substance use treatment program, or has a place to live. This information is usually obtained by talking with the person who has been arrested. At the time of this writing, pretrial release assessments use algorithms that are created by humans as opposed to ones that are generated by machine learning or artificial intelligence (AI). It is possible that future assessments will rely on AI, which would raise a different set of issues to consider.

The use of actuarial pretrial release assessments is growing across the United States. Often, they are an element of broader system change aimed at reducing or eliminating the use of cash bonds, which require people to post money to be released from jail. Judges may consider the actuarial pretrial release assessment score when deciding what conditions of release—for instance, electronic monitoring or mandatory check-ins with pretrial services—are appropriate for a person. In systems that retain money bond as a potential release condition, assessments are sometimes used to inform decisions about bond amounts, but the impact on release is lessened if people remain in jail because they cannot afford to pay their way out. Judges may also use the score as part of their decision about whether to keep someone in jail or release them while their case is pending.

---

[a] Actuarial pretrial release assessments are often referred to as "pretrial risk assessments." We use the term "release" as opposed to "risk" because "risk" is potentially stigmatizing and can bias one to think of a person as dangerous.

Some critics of pretrial release assessments express concerns that they may perpetuate racial disparities. This brief addresses these concerns and provides ideas for how to conduct rigorous anti-racist research on how these assessments work and how they are used. It does not argue in favor of or against actuarial pretrial release assessments. Rather, this brief emphasizes the need for research that centers issues of structural racism, unconscious bias, and other forms of discrimination to fully understand the potential impact of assessments and to inform decisions about whether and how to use them.

> The National League of Cities[3] describes anti-racism as "a system in which we create policies, practices, and procedures to promote racial equity. Anti-racism … uplift[s] the innate humanity and individuality of Black, Indigenous, and People of Color." Read more

Equity Considerations about Actuarial Pretrial Release Assessments

Actuarial pretrial release assessments can be controversial.[4] The decision to keep someone in jail has major consequences for that person's life and could result in the loss of housing, income, or employment.[5] Pretrial incarceration also disrupts people's interpersonal relationships and potentially puts them and their loved ones at risk of physical or emotional harm.[6-8] In addition, many scholars have pointed out that actuarial assessment algorithms use information from systems (e.g., health care, social services, criminal legal) that have long histories of structural oppression and that making decisions based on algorithms that rely on these data runs the risk of perpetuating and intensifying racism.[9-11]

Decisions about incarcerating people should

- be made in ways that are fair and understandable,
- focus on the fact that most people who are released from jail show up for their court dates and remain arrest free, and
- consider the safety and well-being of the people who might be jailed as well as the safety and well-being of their families and communities.

In the following section, we highlight several important concerns about actuarial pretrial release assessments and provide ideas for how to conduct rigorous anti-racist research on how they work and how they are used. We also provide resources to help explore these ideas further. Given the increased use of actuarial pretrial release assessments and the impact the decision to release or detain a person has on their life, focusing ongoing research on equity and fairness when studying these assessments is imperative.

Concerns, Contexts, and Recommendations for Anti-Racist Research

Concern: Some actuarial pretrial release assessments are proprietary, meaning that the researchers who develop them do not share what factors they use or how the algorithms work.

Context: *Correctional Offender Management Profiling for Alternative Sanctions* (COMPAS) is an example of a proprietary algorithm that is used widely throughout the criminal legal system to make decisions that have impacts on people's lives. Because the people who developed COMPAS have not provided information about the instructions given to the computer to make calculations, other researchers cannot understand how COMPAS works or assess whether it perpetuates racism and other inequities.[12]

> In its recommendations on AI, the Organisation for Economic Co-operation and Development (OECD)[13] highlights "transparency and explainability" as key principles. Read more

Recommendations: Researchers should share their work, and people have a right to expect transparency. Developers of algorithms should make them public and equip other people—including practitioners and individuals about whom decisions will be made—to understand them and provide feedback. The Public Safety Assessment (PSA)[14] and Virginia Pretrial Risk Assessment Instrument (VPRAI)[15] are two examples of pretrial release assessments that have been made publicly available.

It is important for other researchers, criminal legal system professionals, and community members to be able to access algorithms so they can identify potential sources of bias and work to improve them.[16] For example, by studying a publicly available algorithm that is widely used to identify patients who require extra care, researchers were able to detect and correct an error that was resulting in Black people not receiving the health care that they needed.[17] Fostering this kind of understanding is key to making sure community members and practitioners can fully participate in policy conversations about whether to implement a specific actuarial pretrial release assessment tool in their jurisdiction.

Concern: Actuarial pretrial release assessments use data from systems that reflect societal bias and racism and therefore may perpetuate disparities.

Context: There is ample research showing that systemic racism plays a key role in events such as being arrested, evicted, or denied a job, as well as losing custody of one's children.[18-21] Using data about these events (e.g., history of arrest, having a permanent address) in algorithmic calculations can extend and intensify the impact of systemic racism.[22] As Sandra Mayson[23] writes, "In a racially stratified world, any method of prediction will project the inequalities of the past into the future."

Recommendations: Researchers must actively work to identify and address biased data.[24] One step in this process is to consider whether data reflect more about systems than individuals. For example, students are more likely to be arrested or suspended by school resource officers (SRO) than by security guards.[25] People who attended schools with SROs therefore are more likely to have been arrested for the first time at an earlier age than people who attended schools with security guards, which means an assessment that uses age at first arrest in its calculation will reflect this disparity. Thinking about systemic issues not only mitigates perpetuation of biases but also supports more accurate identification of factors related to negative outcomes, along with solutions to address them.

When data biases are identified, researchers need to use technical approaches to address them in ways that consider fairness and equity.[23] They should include descriptions of what biases they identified and how they addressed them in the documents they share about the algorithm. Researchers also need to ensure that training materials are created to help people understand how biases might affect assessment scores so that practitioners can take this into account when they are making decisions about releasing people from jail. Materials can also be developed to help pretrial services staff better support individuals to appear in court (e.g., providing vouchers for transportation) and to remain arrest free (e.g., providing resources for food, shelter, and other basic needs as well as substance use and mental health treatment).

> "Whereas in a previous era, the intention to deepen racial inequities was more explicit, today coded inequity is perpetuated precisely because those who design and adopt such tools are not thinking carefully about systemic racism."
> — Ruha Benjamin[10] [Read more](#)

Concern: There is a tendency to frame decisions made using algorithmic scores as "objective" and error free.

Context: Pretrial release assessments are often presented as providing information that, because it is based on data, is less prone to bias than a judge's subjective assessment of whether someone is likely to appear in court and remain arrest free. Judicial decision-making has been studied extensively, and there is evidence that many judges draw on years of professional training and experience to make carefully considered decisions.[26] There also is evidence that some judges make decisions that are influenced by racism, sexism, classism—and even how hungry they are.[27; 28] It is important to remember that mathematical scores are also prone to human bias and error: researchers make decisions about what data are included in algorithms, how much weight each factor carries, and how scores are described. Similar to judges, researchers draw on their own expertise and the work of others in the field when creating algorithms, but this is still a subjective process that could be influenced by bias. In addition, both humans and algorithms can make errors when data are unavailable, inaccurate, or otherwise limited.

> "When the field of AI believes it is neutral, it both fails to notice biased data and builds systems that sanctify the status quo and advance the interests of the powerful. What is needed is a field that exposes and critiques systems that concentrate power, while co-creating new systems with impacted communities: AI by and for the people." — Pratyusha Kalluri[30]
> [Read more](#)

It also is possible that errors are introduced when information is entered into databases.[12] County-level criminal legal system data are often messy and difficult to decipher, and the employees tasked with data entry may be overworked, under-trained, or rushed. Information about an individual is not necessarily "true" just because it appears in an official database.[29]

Recommendations: It is important for researchers to acknowledge that algorithmic scores potentially reflect subjectivity, bias, and error. When designing algorithms, researchers should consider who has conducted, participated in, and funded the prior work informing the algorithm, as well as how power, privilege, and systemic inequities have influenced that work. In striving to reduce algorithmic bias, researchers can collaborate with a wide variety of experts, including people who have been processed through pretrial systems and their families, as a way of challenging assumptions, considering different views, and thinking through how various decisions could favor majority groups and harm other groups. By encouraging discussions of questions, doubts, and alternatives, researchers can produce rigorous work that moves past simply claiming that it is "objective" or "neutral."

Researchers also should also highlight the possibility that humans make errors in inputting data, that data can be collected and recorded inconsistently, and, as discussed above, that even correctly entered data can reflect systemic racism. People need to be provided with free and accessible opportunities to contest, contextualize, and correct data about themselves and receive support in understanding how data are used to calculate scores that inform decisions about their freedom. Researchers can also support jurisdictions that adopt a new actuarial assessment in obtaining technical assistance and setting up quality assurance practices to make sure that data are being entered correctly and that the assessment is being scored and implemented as intended, with ongoing efforts to detect and address bias.

Concern: Algorithms do not consider the circumstances in people's lives that explain their behavior or might facilitate better outcomes in the future.

Context: Pretrial release assessments attempt to use selected information to predict what someone will do. Contextual information cannot be factored into an algorithm. For example, someone may have missed prior court appointments because of a lack of childcare but recently enrolled their child in a pre-kindergarten program. Another person may have experienced a substance use relapse in the wake of losing a loved one but has since been connected to grief counseling. Contextual information might, however, be communicated to a judge through other means, such as through pretrial services staff or a defense attorney.

> "Hypotheses [about people's motives for missing court dates] reflect our cultural assumptions: Either the people summonsed by the state—largely low-income Black and brown residents—are deserving of compassion or leniency because they are oblivious, or they are undeserving because they are obstinate. But they are not excused on the same terms many of us frequently invoke: We are busy, our time and resources are overtaxed, and we sometimes fail to discharge some of our obligations because we are just emotionally overwhelmed."
>
> – Issa Kohler-Hausmann[31] [Read more](#)

Recommendations: Researchers need to create documentation that clearly explains what the scores generated by actuarial pretrial release assessments indicate (e.g., a prediction of future behavior based on a limited amount of information about a person). This documentation can also note what information is not included in the algorithm that might facilitate a change in behavior or contextualize the factors that are in the algorithm. For example, a score indicating that someone is likely to miss their upcoming court date does not reflect any information about why that might happen. Indeed, that person might be much more likely to appear if they receive a reminder about the appointment, a transportation voucher, information about on-site childcare, or the option to participate by videoconferencing. Documentation should emphasize that practitioners may choose to use actuarial pretrial release assessments as one component of a comprehensive approach to avoid keeping people in jail when there are less detrimental alternatives.[32]

Conclusion

The above is by no means an exhaustive list of considerations about the impact of actuarial pretrial release assessments on racial equity in the criminal legal system. Indeed, a crucial first step is for researchers to acknowledge that if these tools are to be equitable and unbiased, they need to be designed with those goals in mind; it will not happen by default. Deliberate steps should be taken to provide documentation to help make algorithms and the research conducted on them understandable to a broad audience. Researchers also should invite questions and debate from the people who use assessments and those whose lives are profoundly affected by them. These conversations can help bring transparency to the research process, which lays the groundwork for researchers' commitment to anti-racism.

Anti-racist research on actuarial pretrial release assessments can help develop the knowledge base to inform critical thinking about these tools: how to develop, whether to use, and ways to improve them. It is important to also acknowledge that problems may exist in the absence of using such assessments (i.e., with full judicial discretion).[32] As the field moves forward, the entire range of tools, practices, strategies, and supports should be drawn upon in an anti-racist approach to improving equity, fairness, and justice in the criminal legal system.

References

1. National Conference of State Legislatures. (2020). Legal presumptions to guide courts making pretrial determinations. https://www.ncsl.org/civil-and-criminal-justice/legal-presumptions-to-guide-courts-making-pretrial-determinations

2. Equal Justice Initiative. (2024). Race and the criminal justice system. https://eji.org/news/history-racial-injustice-race-and-criminal-justice/

3. National League of Cities. (2024). What does it mean to be an anti-racist? https://www.nlc.org/article/2020/07/21/what-does-it-mean-to-be-an-anti-racist/

4. Courtland, R. (2018). Bias detectives: The researchers striving to make algorithms fair. *Nature, 558*, 357–360. https://doi.org/10.1038/d41586-018-05469-3

5. Wakefield, S., & Andersen, L. H. (2020). Pretrial detention and the costs of system overreach for employment and family life. *Sociological Science, 7,* 342-366.

6. Pelvin, H. (2017). *Doing uncertain time: Understanding the experiences of punishment in pre-trial custody* (Publication No. 10635082) [Doctoral dissertation, University of Toronto]. ProQuest Dissertation & Theses Global.

7. Walker, M. L. (2022). *Indefinite: Doing time in jail.* Oxford University Press.

8. The Bail Project. (2024). *Out of pocket: The high cost of pretrial incarceration.* https://bailproject.org/wp-content/uploads/2024/04/the_bail_project_out_of_pocket.pdf

9. Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica.* https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

10. Benjamin, R. (2019). Assessing risk, automating racism. *Science, 336*(6464), 421–422. https://doi.org/10.1126/science.aaz3873

11. Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism.* New York University Press.

12. Rudin, C., Wang, C., & Coker, B. (2020). The age of secrecy and unfairness in recidivism prediction. *Harvard Data Science Review, 2*(1). https://doi.org/10.1162/99608f92.6ed64b30

13. Organisation for Economic Co-operation and Development. (2024). *Recommendation of the Council on Artificial Intelligence.* https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

14. Laura and John Arnold Foundation. (2024). About the Public Safety Assessment. https://advancingpretrial.org/psa/about/

15. National Institute of Corrections. (2009). *Virginia Pretrial Risk Assessment Instrument (VPRAI).* https://nicic.gov/resources/nic-library/all-library-items/virginia-pretrial-risk-assessment-instrument-vprai

16. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

17. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

18. Bronner, L. (2020, June 25). Why statistics don't capture the full extent of the systemic bias in policing. *FiveThirtyEight.* https://fivethirtyeight.com/features/why-statistics-dont-capture-the-full-extent-of-the-systemic-bias-in-policing/

19. Roberts, D. (2002). *Shattered bonds: The color of child welfare.* Basic Civitas Books.

20. Desmond, M. (2012). Eviction and the reproduction of urban poverty. *American Journal of Sociology, 118*(1), 88–133.

21. Pager, D. (2007). *Marked: Race, crime, and finding work in an era of mass incarceration.* University of Chicago Press.

22. Grunwald, B. (2023). Racial bias in criminal records. *Journal of Quantitative Criminology, 39.* https://doi.org/10.1007/s10940-023-09575-y

23. Mayson, S. G. (2018). Bias in, bias out. *Yale Law Journal, 128*(8), 2218–2300.

24. Hadley, E. (2020). 5 steps to take as an antiracist data scientist. *Towards Data Science.* https://towardsdatascience.com/5-steps-to-take-as-an-antiracist-data-scientist-89712877c214

25. Vernon, K., & Curran, F. C. (2023). On guard but not sworn: The relationship between school security guards, school resource officers, and student behavior, discipline, and arrests. *Criminology & Public Policy.* Advance online publication. https://doi.org/10.1111/1745-9133.12653

26. Dhami, M. K., Belton, I., & Goodman-Delahunty, J. (2015). Quasirational models of sentencing. *Journal of Applied Research in Memory and Cognition, 4*(3), 239–247.

27. Chew, P. K., & Kelley, R. E. (2008). Myth of the color-blind judge: An empirical analysis of racial harassment cases. *Washington University Law Review, 86*(5), 1117–1166.

28. Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences, 108*(17), 6889–6892.

29. Lageson, S. (2020). *Digital punishment: Privacy, stigma, and the harms of data-driven criminal justice.* Oxford University Press.

30. Kalluri, P. (2020). Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature, 583,* 169. https://doi.org/10.1038/d41586-020-02003-2

31. Kohler-Hausmann, I. (2020). Nudging people to court. *Science, 370*(6517), 658–659.

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