

ISSUE BRIEF

Algorithmic Risk Assessment and COVID-19: Why PATTERN Should Not Be Used.

April 30, 2020



PARTNERSHIP ON AI

Contents

- Introduction 2
- What is PATTERN and what is its intended use? 3
- Why might someone receive a non-minimum score? 4
- Why PATTERN may amplify racial disparities
 - Racial bias in criminal history and re-arrest data5
 - Problems with the inclusion of data on low-level offenses.....5
- PATTERN in the time of COVID.....7
- Bibliography8

Introduction

In an effort to protect the health and safety of inmates and the Bureau of Prisons (BOP) personnel in the wake of the COVID-19 pandemic, Attorney General William Barr issued a memo [14] on March 26, 2020, listing a set of discretionary factors that should be considered when determining which inmates should be transferred from federal prison to home confinement.

Since the release of the memo, BOP has placed 1576 inmates on home confinement [7], and case management staff are actively reviewing all inmates to identify those who meet the criteria established by the memo. These criteria include the inmate's score under a criminal risk assessment tool called "PATTERN." Inmates scoring anything above "minimum," the lowest level of risk, are not to be prioritized for home confinement.

However, PATTERN was not developed with the intent of informing anything akin to COVID-era decisions, nor was it validated on data that reflects the present social and economic landscape. Furthermore, due to both demographic differences between the Black and White federal inmate populations and racial bias in the data used to develop, validate, and score the tool, the proposed use of PATTERN is likely to contribute to significant unjustified racial disparities among the inmates placed in home confinement. This risks making federal prisons yet another place where we see higher infection rates and death rates among Black people.¹

Issues surrounding the validity, bias, and governance of criminal risk assessment tools have been a significant concern of the Partnership on AI (PAI). Last May, PAI released a [report](#) documenting ten minimum requirements for the responsible deployment of risk assessment tools based on the research and expertise of our community, including researchers, civil society organizations, and AI industry practitioners. The finding of the report was that no current tool meets all ten of these requirements, which encompassed concerns around the potential for such tools to perpetuate systemic biases, the lack of transparency surrounding the tools' development, and the lack of governance structures ensuring accountability to communities affected by their deployment.

The current report is part of a recently launched research project focussed on more closely examining questions of racial bias in the arrest data used to develop, validate, and assess risk assessment tools. We begin by providing an overview of PATTERN, its intended use, and inmate risk factors that may preclude receiving a minimum score. We then take a deep dive into reasons why the inclusion of data on criminal justice involvement for low-level offenses as risk factors and targets of the model's prediction may drive significant unjustified racial disparities in PATTERN's risk classifications. Lastly, we discuss why using the output of PATTERN or any other criminal risk assessment tool to make release decisions during the time of COVID is fundamentally unsound.

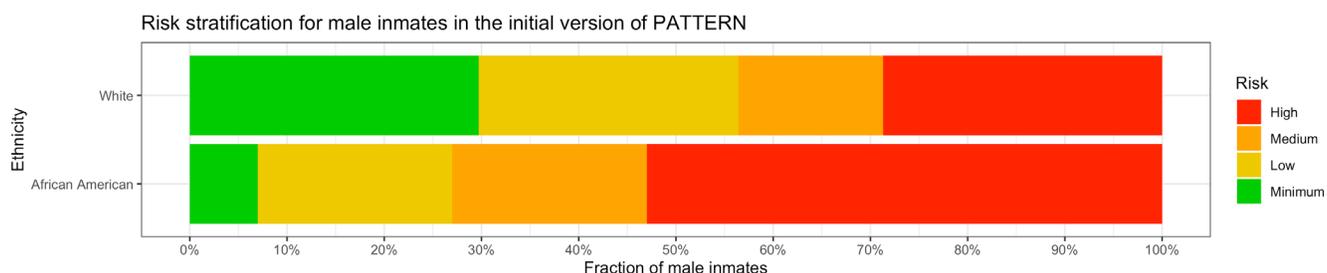
1 As of April 12, 2020, of the over 140,000 federal inmates in BOP custody, 352 have tested positive for COVID-19 and 10 have died.

What is PATTERN and what is its intended use?

In July 2019, the DOJ released a detailed report describing their work on developing, validating and implementing a criminal risk and needs assessment system for use among the federal prison population [27]. This system, which they call PATTERN (Prisoner Assessment Tool Targeting Estimated Risk and Needs), was developed to help meet the goals of the First Step Act (FSA). As described on the BOP website, PATTERN “provides guidance on the type, amount, and intensity of recidivism reduction programming and productive activities to which each prisoner is assigned, including information on which programs prisoners should participate in based on their criminogenic needs. The system also provides guidance on how to group, to the extent practicable, prisoners with similar risk levels together in recidivism reduction programming and housing assignments” [7]. **We note that PATTERN was not developed to specifically assess recidivism risk during home confinement for inmates transferred from BOP facilities earlier in their sentences.**

PATTERN consists of two different models, one which predicts the risk of any re-arrest within 3 years of release from prison, and another which specifically focuses on re-arrest for a violent offense [27]. Both types of assessment have the same four-level risk scale: minimum, low, medium, and high. Although the models share a common set of input factors, their weighting is different. Consequently, an inmate may receive two different scores and risk classifications. The final classification of risk is taken to be the maximum of the two scales.

Following a public commentary period, BOP made a number of revisions to PATTERN, incorporating additional “dynamic” factors such as participation in mental health treatment programs, and removing elements viewed as likely to drive disparate outcomes in risk classifications, such as the age of first arrest/conviction, and whether the prisoner voluntarily surrendered [28]. While some of these modifications were due in part to concerns over disparities in risk stratification across groups, the report on the revised tool provides no insight into whether the changes served to decrease disparities. **In the initial version of PATTERN, just 7% of African American males received a “minimum” risk classification, compared to 30% of White males [27].**



Why might someone receive a non-minimum score?

While we don't have access to data that would allow us to probe this question directly, we can gain some understanding by looking at how PATTERN is scored and at aggregate statistics about the federal inmate population. This analysis also provides insight into why PATTERN is likely to classify White people as minimum risk at higher rates than African Americans.

Age and criminal history are the most heavily weighed factors in PATTERN score calculation. While the revised version of PATTERN in principle allows even young inmates with poor criminal histories to move to lower risk categories through extensive participation in programs and clean prison records, those initial static factors may be difficult to overcome in practice.² **For one, it is virtually impossible for a male inmate below the age of 30 to receive a minimum risk classification in PATTERN.³ This means that, solely because of age, almost 20% of the current BOP population is not being prioritized under the memorandum [7], even in cases where inmates may have an increased risk of severe COVID-19 complications due to a history of diabetes, lung disease, or other clinical risk factors.**

The matter is further complicated by criminal history. **If a male inmate also has past convictions, he would be unlikely to receive a minimum risk classification unless he is over the age of 40.** The majority of inmates have convictions for prior offenses and three fifths are incarcerated for drug or weapons offenses [7].

Age and criminal history are two major contributors to the racial disparities observed in the risk stratifications produced by PATTERN. Black inmates, who currently form more than one third of the BOP population [7], are on average⁴ younger than their White counterparts [4]. They also tend to be incarcerated for different types of crimes. African American inmates are more likely to have been charged with drug offenses [17], which are crimes that, as we will discuss in greater detail, are more subject to biased enforcement.

2 Although the risk level can be decreased through participation in programs, dynamic factors do not effectively counterbalance age due to their reduced weight. In addition, inmates early in their sentences may not yet have had access to these programs. Among inmates that served one year or more in prisons, two fifths of them do not complete more than one program [28].

3 The scoring system of PATTERN is presented in [28]. A male aged 18-25 would score 35 on the PATTERN risk for re-arrest. Even if they completed the maximum number of programs, participated in work programming, had no necessary drug treatment, and had an high school degree, the final score would be 35-23=12. For a male aged 26-30, the weight relative to age would be 28. Only defendants whose score is lower than 10 are considered at minimum risk.

4 This refers to the aggregate population of inmates including both state and federal prisons.

Why PATTERN may amplify racial disparities

Racial bias in criminal history and re-arrest data

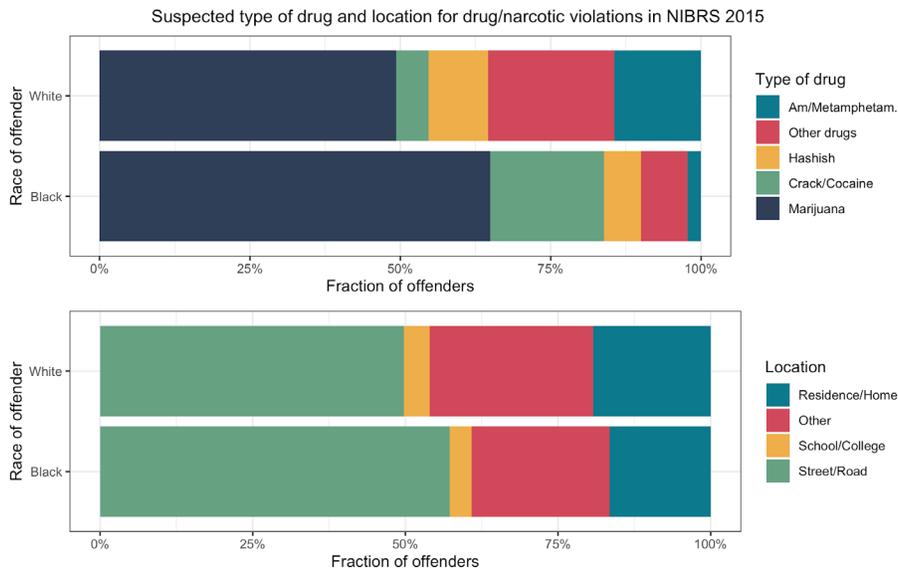
There are two interrelated sources of bias that are of concern in criminal risk assessment. The first is bias in the inputs, such as criminal history information, used to produce scores. The other is bias in the outcome that the tool is developed to predict. Like many other tools, PATTERN has been trained and validated on re-arrest data. This means that, to the extent that the tool is validated, it tells us how predictive the tool is of re-arrest, not of new criminal activity. Arrests coincide with offending only to the extent to which offenses result in the apprehension of the offender. Yet only a small fraction of all offenses is known to the police [16] and an even smaller fraction is cleared by arrest [5].

The uncertainty arising from measurement bias in the data should be taken into account in the evaluation of risk assessment tools, particularly when those tools result in disparate risk stratifications. As part of our work, we have developed methods for assessing the predictive bias properties of risk assessment tools as predictors of re-offense that account for uncertainty and racial disparities in the likelihood of arrest for crimes committed [8]. Our analysis of two risk assessment tools shows that even small racial disparities in the likelihood of arrest may be sufficient to ensure that an instrument that is a racially equitable predictor of re-arrest is a racially biased predictor of re-offending.

In addition, over the past few months, we have undertaken an extensive review of the literature and conducted in-depth analysis of nationwide data on arrests, crime reporting, and victimization to better understand racial disparities in the likelihood of arrest. We have found that factors such as whether and how police are informed of the incident, cooperation of victims and offenders, availability of resources, and the willingness and discretion of the police all affect the likelihood that an offender will be apprehended. Race plays a key role in shaping the relationship between these factors.

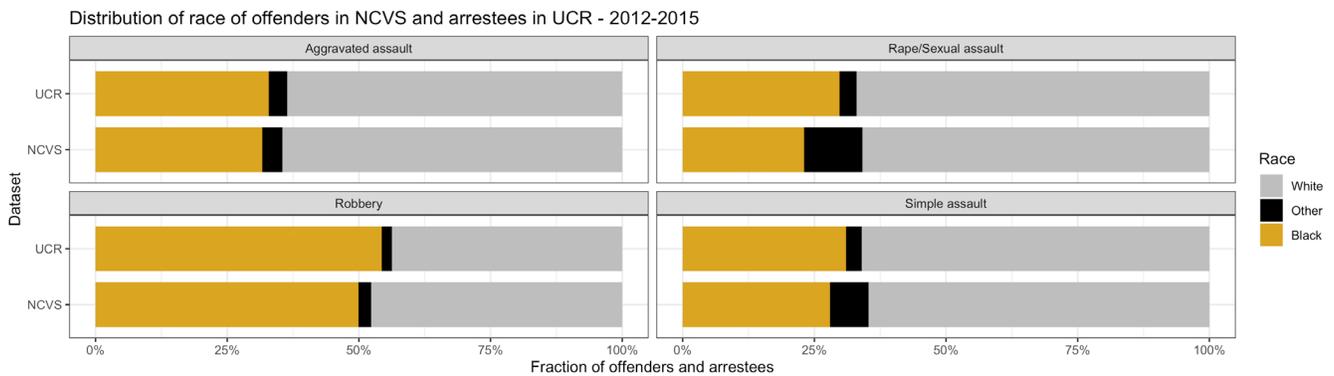
Problems with the inclusion of data on low-level offenses

Empirical evidence suggests that the overrepresentation of Black offenders is highest in the case of arrests for lower-level crimes. A large majority of such apprehensions is the result of proactive policing strategies where considerable discretion is afforded to police officers. In some proactive policing strategies such as stop-and-frisk [9,10,19] or traffic stops [20,24], Black people have been found to be disproportionately targeted by police. Black people have also been shown to be substantially more likely than White people to be stopped and arrested for drug offenses [2,3,15,18]. According to one hypothesis [26], differences in the nature of the offending (e.g. location [21]) may make Black people an easier target for the police. Consistent with this hypothesis, our analysis on official crime data has shown that Black offenders are more likely than White people to be apprehended outside of their residence and with drugs such as crack cocaine or marijuana.



The nature of violent crimes is different. These offenses are typically reported to the police either by victims or third parties rather than being witnessed by police directly. This greatly reduces the level of discretion police have in their response to the criminal activity. Comparisons of victimization and official incident reporting data have found little discrepancy between the fraction of Black people in reported offenders and arrestees [1,23,25]. Studies on official data have also found that Black offenders may be equally or less likely than White people to be apprehended for certain types of crimes [6], suggesting that arrest data for violent crimes may exhibit less bias.⁵ Our initial analysis of recent victimization and official data largely agrees with these conclusions.

Of course, all of this empirical analysis speaks to patterns in offending and enforcement in the pre-COVID era. As will be discussed below, tools validated on pre-COVID data fail to capture the unique dynamics of the COVID era.



The initial version of PATTERN tool produced highly racially disparate risk stratifications, in part due to its reliance on racially biased data of low-level offenses, and we expect similar disparities to persist for the revised tool. Relying on PATTERN in making release decisions thus risks making federal prisons yet another place where we see higher infection rates and death rates among Black people.

⁵ This is not to say that arrest data for violent crimes is unbiased. There have been numerous reports on the issue of racially-biased wrongful arrests and convictions for violent crimes. [13] Moreover, there are many limitations with official criminal justice data sources, including notable cases of data manipulation by police departments [22].

PATTERN in the time of COVID

In developing, deploying, and implementing PATTERN, the BOP laid out four principles of best practice [27]. The first two among them is that the tool should be developed using an appropriate inmate population, and validated on an appropriate inmate population. This is why, for instance, BOP considered only federal inmates throughout their analysis. Yet these principles are also a reason why using PATTERN to guide decisions in the time of COVID is an unsound, potentially perilous, way forward: **PATTERN was neither developed nor validated on an inmate population during the COVID crisis.** Fundamentally, it was not developed to assess in any meaningful way the public safety risk that transferring an inmate to home confinement might present.

While the tool may turn out to have utility in serving its intended purpose of guiding recidivism risk reduction programming for inmates in BOP custody during standard operations, this is not the type of decisions case management staff are tasked with making today. In times of crisis there is an understandable desire to repurpose existing tools, be they pharmaceuticals or risk models, in confronting the challenges of the day. But not all tools are fit for purpose. The experts, practitioners, policy makers and other stakeholders engaged in the process of developing PATTERN were not asked to weigh in on a tool that would be used in making potentially life-or-death transfer decisions. Had this been the question put to the group, the resulting tool would likely look very different.

The question now is no longer one of which strategy best reflects the correctional goals of retribution, incapacitation, deterrence and rehabilitation. No initial sentence imposed prior to the US outbreak of the COVID pandemic considered the risk of infection and death among federal inmates. And no reasonable release strategy should prioritize extrapolated assessments of risk for largely low-crimes over considerations of the health and safety of vulnerable individuals.

Bibliography

- [1] Beck, Allen J., and Alfred Blumstein. "Racial disproportionality in US state prisons: Accounting for the effects of racial and ethnic differences in criminal involvement, arrests, sentencing, and time served." *Journal of Quantitative Criminology* 34.3 (2018): 853-883.
- [2] Beckett, Katherine, et al. "Drug use, drug possession arrests, and the question of race: Lessons from Seattle." *Social Problems* 52.3 (2005): 419-441.
- [3] Beckett, Katherine, Kris Nyrop, and Lori Pfingst. "Race, drugs, and policing: Understanding disparities in drug delivery arrests." *Criminology* 44.1 (2006): 105-137.
- [4] Bronson, Jennifer, and E. Ann Carson. "Prisoners in 2017." *Age* 500 (2019): 400.
- [5] "Crime in the U.S. 2018." FBI, FBI, 22 July 2019, ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018. Accessed 16 Apr. 2020.
- [6] D'Alessio, Stewart J., and Lisa Stolzenberg. "Race and the probability of arrest." *Social forces* 81.4 (2003): 1381-1397.
- [7] "Federal Bureau of Prisons." BOP, www.bop.gov/. Accessed 16 Apr. 2020.
- [8] Fogliato, Riccardo, Max G'Sell, and Alexandra Chouldechova. "Fairness Evaluation in Presence of Biased Noisy Labels." arXiv preprint arXiv:2003.13808 (2020).
- [9] Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. "An analysis of the New York City police department's "stop-and-frisk" policy in the context of claims of racial bias." *Journal of the American statistical association* 102.479 (2007): 813-823.
- [10] Goel, Sharad, Justin M. Rao, and Ravi Shroff. "Precinct or prejudice? Understanding racial disparities in New York City's stop-and-frisk policy." *The Annals of Applied Statistics* 10.1 (2016): 365-394.
- [11] Golub, Andrew, Bruce D. Johnson, and Eloise Dunlap. "The race/ethnicity disparity in misdemeanor marijuana arrests in New York City." *Criminology & public policy* 6.1 (2007): 131-164.
- [12] James, Nathan. "The First Step Act of 2018: An overview." (2019).
- [13] Loeffler, Charles E., Jordan Hyatt, and Greg Ridgeway. "Measuring self-reported wrongful convictions among prisoners." *Journal of Quantitative Criminology* 35.2 (2019): 259-286.
- [14] "Memorandum." 26 Mar. 2020, <https://perma.cc/F4VV-3JEN> Accessed 16 Apr. 2020.

- [15] Mitchell, Ojmarrh, and Michael S. Caudy. "Examining racial disparities in drug arrests." *Justice Quarterly* 32.2 (2015): 288-313.
- [16] Morgan, Rachel E., and Grace Kena. "Criminal victimization, 2018." Bureau of Justice Statistics (2019).
- [17] Motivans, Mark. Federal justice statistics, 2015-2016. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, 2019.
- [18] Nguyen, Holly, and Peter Reuter. "How risky is marijuana possession? Considering the role of age, race, and gender." *Crime & Delinquency* 58.6 (2012): 879-910.
- [19] Pierson, Emma, Sam Corbett-Davies, and Sharad Goel. "Fast threshold tests for detecting discrimination." arXiv preprint arXiv:1702.08536 (2017).
- [20] Pierson, Emma, et al. "A large-scale analysis of racial disparities in police stops across the United States." arXiv preprint arXiv:1706.05678 (2017).
- [21] Ramchand, Rajeev, Rosalie Liccardo Pacula, and Martin Y. Iguchi. "Racial differences in marijuana-users' risk of arrest in the United States." *Drug and alcohol dependence* 84.3 (2006): 264-272.
- [22] Richardson, Rashida, Jason Schultz, and Kate Crawford. "Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice." *New York University Law Review Online*, Forthcoming (2019).
- [23] Sampson, Robert J., and Janet L. Lauritsen. "Racial and ethnic disparities in crime and criminal justice in the United States." *Crime and justice* 21 (1997): 311-374.
- [24] Simoiu, Camelia, Sam Corbett-Davies, and Sharad Goel. "The problem of infra-marginality in outcome tests for discrimination." *The Annals of Applied Statistics* 11.3 (2017): 1193-1216.
- [25] Tonry, Michael. *Malign neglect: Race, crime, and punishment in America*. Oxford University Press, 1995.
- [26] Tonry, Michael, and Matthew Melewski. "The malign effects of drug and crime control policies on black Americans." *Crime and justice* 37.1 (2008): 1-44.
- [27] U.S. Department of Justice, Office of the Attorney General. "The First Step Act of 2018: Risk and Needs Assessment System." July, 2019.
- [28] U.S. Department of Justice, Office of the Attorney General. "The First Step Act of 2018: Risk and Needs Assessment System - UPDATE". January, 2020.



PartnershipOnAI.org
115 Sansome Street, Suite 1200
San Francisco, CA 94104
© 2020 Partnership on AI