Dear Dr. Muhlhausen,

As researchers in the fields of statistics, machine learning and artificial intelligence, law, and sociology, we submit this comment regarding the First Step Act of 2018: Risk and Needs Assessment System. We have closely watched the development and deployment of risk assessment tools, conducted research, and carefully studied others’ research. We have researched the use of risk assessments in states across the country including California, Colorado, Kentucky, Illinois, Massachusetts, New Jersey, Oregon, South Carolina, Virginia, Washington D.C., and Wisconsin. Drawing from this research, we have written extensively about the harms of actuarial risk assessment as a vehicle for criminal justice reform. We wrote an open letter to MA legislature and an open statement of concern about risk assessments, which was signed by 27 prominent researchers, and we have commented in media outlets including the New York Times.

In this comment, we explain how PATTERN suffers from serious methodological flaws that undermine its accuracy, validity, and effectiveness. PATTERN relies upon data that is deeply flawed and racially biased. As a result, the risk assessment will reproduce these biases and overestimate risk for people of color. PATTERN is also ill-suited for supporting effective risk reduction strategies because the computational methods used to develop the tool are designed to optimize predictive validity rather than diagnostic insight.

Background

On July 19, 2019, the National Institute of Justice (NIJ), a research arm of the U.S Department of Justice’s Office of Justice programs, published a report to accompany the release of PATTERN as a new risk and needs assessment tool that fulfils the FSA’s statutory requirement to assign a “recidivism score” to each incarcerated person which predicts their risk of committing crimes three years after release. The report claims the tool is able to maximize the number of incarcerated people eligible to earn early release, identify those qualified to participate in rehabilitative programming, and to ensure public safety. Specifically, the NIJ claims that PATTERN: (1) “achieves a high level of predictive performance and surpasses” other tools used to evaluate inmate risk; (2) “more appropriately aligns with the goals of the FIRST STEP ACT [by] mak[ing] greater use of dynamic factors; (3) exhibits “predictive performance is unbiased across racial and ethnic classifications”; and (4) itself “encourage(s) risk reducing behavior.”

---

The Data Used to Build and Test PATTERN is Deeply Flawed and Racially Biased

To predict recidivism, PATTERN relies upon criminal history data, including age at first conviction, current incarceration reason, sex offender status, history of violence, history of escape, arrest or return to custody following release, and a criminal history score. PATTERN’s validity rests on the assumption that criminal history data can serve as a reliable and neutral measure of underlying criminal activity, but such records cannot be relied upon for this purpose. Decades of research have shown that, for the same conduct, African-American and Latinx people are more likely to be arrested, prosecuted, convicted and sentenced to harsher punishments than their white counterparts. For decades, communities of color have been arrested at higher rates than white communities, even for crimes that these racial groups engage in at comparable rates. African-Americans are 83% more likely to be arrested for marijuana compared to whites at age 22 and 235% more likely to be arrested at age 27, in spite of similar marijuana usage rates across racial groups. Similarly, African-American drivers are three times as likely as whites to be searched during routine traffic stops, even though police officers generally have a lower “hit rate” for contraband when they search drivers of color. This leads to an overrepresentation of people of color in arrest

---

3 NIJ should describe what instances qualify as a return to BOP custody following release within the training data.


5 Megan Stevenson & Sandra G. Mayson, The Scale of Misdemeanor Justice, 98 B.U. L. Rev. 731, 769-770 (2018). This comprehensive national review of misdemeanor arrest data has shown systemic and persistent racial disparities for most misdemeanor offenses. The study shows that “black arrest rate is at least twice as high as the white arrest rate for disorderly conduct, drug possession, simple assault, theft, vagrancy, and vandalism.” Id. at 759. This study shows that “many misdemeanor offenses criminalize activities that are not universally considered wrongful, and are often symptoms of poverty, mental illness, or addiction.” Id. at 766.


data. Compared to similarly situated white people, African-Americans are also more likely to be convicted\(^8\) and more likely to be sentenced to incarceration.\(^9\)

People of color are treated more harshly than similarly situated white people at each stage of the legal system, which results in serious distortions in the data upon which PATTERN relies. By incorporating this distorted data, PATTERN will produce distorted results and overestimate the risk of people of color.\(^10\) There are no technical fixes for these distortions.

**PATTERN is Ill-suited as a Diagnostic Tool to Inform Risk Reduction Programming.**

The NIJ report claims that PATTERN’s use of “dynamic risk factors” can lead to a reduction in risky behavior over time. But PATTERN is ill-suited for supporting effective risk reduction strategies – even with the inclusion of dynamic risk factors – because the computational methods used to develop the tool (boosted regression) are designed to optimize predictive validity, rather than diagnostic insight.

Today risk assessments are used for two primary purposes, which practitioners have deemed “prediction-oriented” and “reduction-oriented” approaches to risk assessment.\(^11\) Prediction-oriented assessments are used to facilitate prediction of key outcomes, such as future recidivism, while reduction-oriented tools are intended to inform treatment and supervision plans.

According to the NIJ report, PATTERN is designed to achieve both predictive and diagnostic goals. However, PATTERN was built using boosted regression modeling, which is ill-suited for the purpose of effective diagnosis and intervention on criminogenic needs. Assessments designed for intervention require computational techniques, such as causal inference, which can identify causal relationships between risk factors and future crime. In contrast, regression analysis is widely used for purposes of forecasting future events. The main goal of regression is to identify a set of variables that are predictive of a given outcome variable. This is achieved by determining the optimal weights for a given set of covariates, ones that are best predictive of the outcome variable of interest. This is

---


\(^10\) There have been attempts to solve this problem on the back end by mitigating outcome disparities in risk assessment predictions, but they overlook and do not address the fundamental distortions outlined above.

done through processes called model checking and selection,\textsuperscript{12} where by statistical tests are run on each covariate to see how significantly predictive they are of the outcome. Covariates that are identified through this process are correlational, not causal, of the outcome of interest.

Regression based models, such as PATTERN, are not well-equipped to answer causal questions on interventional covariates. Tacking on interventional covariates to a list of risk covariates that were surfaced using regression models suffers from the following drawbacks:

First, the set of dynamic factors considered in PATTERN are constrained to a narrow set of variables, ones which are amenable to individualized treatment plans and are informed by particular psychological and normative theories of re-offending.\textsuperscript{13} Very little consideration has been given to broader social or structural drivers of crime. As Hannah-Moffat has argued, in Risk-Needs-Responsivity (RNR) assessments an offender’s needs are stripped of their broader social context, and are framed largely in terms of poor choices or moral and psychological deficiencies that can be treated through “re-education” at the individual level.\textsuperscript{14}

The narrow theoretical focus of PATTERN stems in part from the fact that regression is ill-suited to test and differentiate competing models of criminal behavior. This has led to a fairly narrow conceptualization of criminogenic need that is limited to attributes which demonstrate a statistically significant relationship to future recidivism rates in a regression model. The NIJ report acknowledges that criminogenic needs include a wide range of factors which directly or indirectly contribute to adverse outcomes. Yet, the tool only incorporates needs factors that are \textit{predictive} of recidivism.\textsuperscript{15} Other common sense needs factors for which a statistically significant relationship could not be established, such as mental health, have been excluded.

Second, it is challenging to differentiate intermediate outcomes from covariates in a regression based model without a causal inference framework.\textsuperscript{16} Because PATTERN does not establish a causal relationship between a given covariate to the risk of recidivism, the instruments risks misinterpreting an intermediate outcome as a causal driver of crime. This is deeply significant if we are to use PATTERN as the basis for identifying effective interventions to address the underlying drivers of...


\textsuperscript{14} Supra note 5.

\textsuperscript{15} P. 28 NIJ report

crime. To use an example from the healthcare context, lung cancer might be a variable that is highly predictive of mortality even though the root cause is smoking. If one wanted to intervene on increased mortality rates in a population that includes a large population of smokers, one would need to characterize lung cancer as an intermediate outcome of smoking. Similarly in PATTERN, arrest rates are taken as a covariate because of their predictive power, but this practice hinders our ability to understand the core causal drivers of crime, while simultaneously running the risk of exacerbating racial disparities in the system.

If PATTERN is to be conceived of as a diagnostic tool that can be used to inform effective risk reduction strategies, then causal inference offers the best framework for pursuing these goals. More work should be done to examine how these statistical methods could be incorporated into the existing PATTERN framework.

**Conclusion**

This comment specifically addresses fundamental, technical problems with PATTERN. Beyond the technical flaws outlined in this comment, a broader and growing body of research questions the validity, ethics, and efficacy of actuarial risk assessments. The PATTERN risk assessment will produce distorted results and is ill-suited as a diagnostic tool to inform risk reduction programming.