Reprogramming Recidivism: The First Step Act and Algorithmic Prediction of Risk

Amy B. Cyphert[*]

The First Step Act, a seemingly miraculous bipartisan criminal justice reform bill, was signed into law in late 2018. The Act directed the Attorney General to develop a risk and needs assessment tool that would effectively determine who would be eligible for early release based on an algorithmic prediction of recidivism. The resulting tool—PATTERN—was released in the summer of 2019 and quickly updated in January of 2020. It was immediately put to use in an unexpected manner, helping to determine who was eligible for early release during the COVID-19 pandemic. It is now the latest in a growing list of algorithmic recidivism prediction tools, tools that first came to mainstream notice with critical reporting about the COMPAS sentencing algorithm.

This Article evaluates PATTERN, both in its development as well as its still-evolving implementation. In some ways, the PATTERN algorithm represents tentative steps in the right direction on issues like transparency, public input, and use of dynamic factors. But PATTERN, like many algorithmic decision-making tools, will have a disproportionate impact on Black inmates; it provides fewer opportunities for inmates to reduce their risk score than it claims and is still shrouded in some secrecy due to the government's decision to dismiss repeated calls to release more information about it. Perhaps most perplexing, it is unclear whether the tool actually advances accuracy with its predictions. This Article concludes that PATTERN is a decent first step, but it still has a long way to go before it is truly reformative.

[*]Lecturer in Law and Director, ASPIRE, West Virginia University. B.A., 2001, Carnegie Mellon University; J.D., 2005, Harvard Law School. I thank Kenshandra Mitchell for providing excellent research assistance throughout the drafting of this Article. The excellent team at the Seton Hall Law Review greatly improved this Article with their hard work. I am also grateful to my colleagues Amber Brugnoli, Jena Martin, and Joshua Weishart, who provided helpful feedback. As always, thank you to Sam, Bethany, Josh, and Violet for giving me space to write and inspiration to persevere.

I. INTRODUCTION

On December 21, 2018, President Trump signed into law the federal First Step Act,[1] a sweeping piece of criminal justice reform legislation. It has been one of his favorite accomplishments to tout, and he discussed it during the 2020 State of the Union address.[2] The scope of the Act was broad, addressing topics ranging from reducing racial disparities in incarceration to shackling of pregnant inmates during birth. The Act was supported overwhelmingly by Democrats and Republicans and heralded as "almost miraculous"[3] and "a major win for the movement to end mass incarceration."[4] That Congress passed ANY criminal justice reform was, in and of itself, remarkable—it had been eight years since Congress had passed legislation on the topic, and even that was relatively modest.[5] The First Step Act's passage was further surprising given the administration at the time. There had been no suggestion that criminal justice reform was a priority for the Trump Administration,[6] and indeed, the last meaningful criminal justice reform

[1] *See* First Step Act of 2018, Pub. L. No. 115-391, 132 Stat. 5194 (2018).

[2] Matthew Charles, who President Trump described during the State of the Union as "the very first person to be released from prison under the First Step Act," attended the speech as a guest. Mariah Timms, *President Trump: 'Welcome Home' to Matthew Charles, Man Released from Nashville Prison Under First Step Act*, TENNESSEAN (Feb. 5, 2019, 8:37 PM), https://www.tennessean.com/story/news/2019/02/05/donald-trump-state-of-the-union-matthew-charles-alice-johnson-first-step-act-free/2783545002.

[3] Shon Hopwood, *The Effort to Reform the Federal Criminal Justice System*, 128 YALE L.J.F. 791, 795 (2019).

[4] Ames Grawert & Tim Lau, *How the FIRST STEP Act Became Law—And What Happens Next*, BRENNAN CTR. FOR JUST. (Jan. 4, 2019), https://www.brennancenter.org/blog/how-first-step-act-became-law-and-what-happens-next.

[5] The Fair Sentencing Reform Act of 2010 reduced the crack versus powder cocaine disparity by increasing the amount of crack cocaine needed to trigger mandatory minimum sentencing from five grams to twenty-eight grams. Fair Sentencing Act of 2010, Pub. L. No. 111-220, 124 Stat. 2372 (2010) (codified as amended in scattered sections of 21 and 28 U.S.C.). Many commentators have argued that the Fair Sentencing Reform Act did not adequately address the crack-to-powder disparity. *See, e.g.*, Scott R. Hechinger, *Juvenile Life Without Parole: An Antidote to Congress's One-Way Criminal Law Ratchet?*, 35 N.Y.U. REV. L. & SOC. CHANGE 408, 449 (2011) ("After the FSA's passage, some argued that the bill did not go far enough in equalizing the sentences imposed for crack and powder cocaine offenses. Others complained that the bill did not apply retroactively, thus leaving those individuals sentenced for crack-based offenses prior to the FSA's enactment without relief."). The First Step Act addresses the retroactive application in Title IV.

[6] *See, e.g.*, Grawert & Lau, *supra* note 4, at 2 ("[W]hen Donald Trump was elected president in 2016, many worried that sentencing reform would prove impossible for the next four years. Trump's position on criminal justice reform was unclear at best and regressive at worse."); Hopwood, *supra* note 3, at 797 ("After the 2016 election and the appointment of Senator Jeff Sessions as Attorney General, most in the federal reform community believed we would spend the next four years playing defense against the

proposed legislation had been blocked by none other than the Attorney General, Jeff Sessions, while he was in the Senate.[7]

The First Step Act embodied important prison reform goals, including reducing the use of restraints on pregnant inmates;[8] addressing the crack versus cocaine sentencing disparity;[9] improving reentry programs for those leaving prison;[10] and streamlining the process for placing incarcerated persons in prisons near their families.[11] Given the breadth and importance of these initiatives, it is not surprising that many criminal justice reform groups supported the Act's passage.[12] As of December 2019, 7,000 incarcerated persons had been released under this Act.[13]

In the First Step Act, Congress directed the Attorney General to develop a risk and needs assessment tool that could facilitate reducing recidivism and would prioritize the early release of any inmates[14] who were found to pose a "minimum" or "low" risk of reoffending. As a

DOJ's bad policy preferences and a Congress that frequently treats criminal law and punishment as the only way to fix national social problems.").

[7] *See* Ames C. Grawert, *Analysis: Sen. Jeff Sessions's Record on Criminal Justice*, BRENNAN CTR. FOR JUST. (Jan. 6, 2017), https://www.brennancenter.org/our-work/research-reports/analysis-sen-jeff-sessionss-record-criminal-justice (noting that then-Senator Sessions had "personally blocked" a bipartisan criminal justice reform effort in 2016).

[8] First Step Act of 2018, Pub. L. No. 115-391, § 301, 132 Stat. 5194 (2018) (prohibiting the use of restraints on inmates "during the period of pregnancy, labor, and postpartum recovery" unless the inmate is a "flight risk"; poses a threat of harm to herself or others; or the restraints are appropriate for the inmate's medical safety)

[9] *See* First Step Act § 404 (allowing courts to retroactively apply the Fair Sentencing Reform Act of 2010).

[10] First Step Act §§ 502–05 (providing federal grants for the improvement of existing adult and juvenile reentry programs and the development of additional reentry programs).

[11] First Step Act § 601 ("[T]he Bureau of Prisons [shall] place the prisoner as close as practicable to the prisoner's primary residence, and to the extent practicable, in a facility within 500 driving miles of that residence.").

[12] Brandon L. Garrett, *Federal Criminal Risk Assessment*, 41 CARDOZO L. REV. 121, 133 (2019) ("A wide range of civil rights groups such as the National Urban League, the American Civil Liberties Union, #cut50, as well as conservative groups . . . praised the legislation.").

[13] Rachel Anspach, *How the First Step Act Got People Out of Prison and Back with Their Families*, MIC (Dec. 24, 2019), https://www.mic.com/p/how-the-first-step-act-got-people-out-of-prison-back-with-their-families-19629720.

[14] In this Article, I use the terms "inmate" and "prisoner" to refer to people who are currently incarcerated. In past scholarship, I have used the term I prefer, "incarcerated person." *See, e.g.*, Amy B. Cyphert, *Prisoners of Fate: The Challenges of Creating Change for Children of Incarcerated Parents*, 77 MD. L. REV. 385 (2018). Here, I use the terms "inmate" and "prisoner" for the sake of consistency and clarity because they are the terms used throughout both the First Step Act and the subsequent DOJ reports announcing and clarifying the PATTERN tool.

result, the Department of Justice (DOJ) worked with outside consultants to develop an algorithmic assessment tool, PATTERN, to classify each federal inmate as being minimum, low, medium, or high risk of reoffending. This Article examines PATTERN, contextualizing it historically in the evolution of recidivism prediction tools as well as assessing where it represents advances in the field and where it is problematic.

Part II of this Article traces the history of recidivism prediction practices, starting with early clinical models wherein judges and probation officials used their experience to assess the likelihood an individual would reoffend. The historical overview then moves into today's sophisticated algorithmic models, many of which are fueled by machine learning. This new field of algorithms has been attacked as opaque, shielded in secrecy, and racially biased. The Wisconsin Supreme Court's recent decision concerning the recidivism prediction algorithm COMPAS—and the backlash it sparked—offers a lens into how courts treat these tools.

Part III then discusses the First Step Act, and specifically, the requirements therein that a risk and needs assessment tool be developed. Congress was specific about certain features, requiring, for example, that the tool be objective and statistically validated,[15] publicly released, and quickly developed.[16] Congress was silent, however, about other features such as where to set the risk-level cutoffs, how to define recidivism, and how to assess accuracy. Congress's silence on those topics was a de facto delegation of several critical policy decisions to the Attorney General, who in turn delegated those policy decisions to the outside consultants hired to develop PATTERN.

Part IV is the heart of the Article, outlining and dissecting PATTERN's incomplete progress toward better recidivism prediction. This Article attempts to give credit where credit is due, acknowledging the places where DOJ has made attempts at important goals like transparency, public input, and best practices in tool design. But PATTERN still falls short in each of those areas and others, and DOJ needs to do more as it continues to update and calibrate PATTERN. For example, although DOJ has provided some important transparency with respect to PATTERN, the dataset that was used to train the algorithm must be publicly released to allow for independent verification of the DOJ's claims about the tool's accuracy and lack of bias. Further, the Department must continue to listen to the community and other

15 First Step Act § 101(a) (amending 18 U.S.C. § 3632(d)(4)(A)).

16 *Id.*

stakeholders and be responsive to the consistent and informed criticism it has received about things like how to properly evaluate PATTERN's accuracy and how to address racial bias. This is especially important given that PATTERN is being used to determine which inmates are eligible for early release during the COVID-19 pandemic,[17] which has disproportionately impacted Black Americans.[18] Further, the decision to prioritize the risk assessment over the needs assessment reduces PATTERN's ability to reward inmates for participating in programming, which is an evidence-based way to reduce recidivism.

There are things to laud in the First Step Act, and PATTERN's risk predictions may offer some inmates an opportunity for early release and reunification with their families. To ensure this opportunity is fairly apportioned, PATTERN must be rigorously evaluated, and DOJ needs to do more to allow that to happen.

II. HISTORY OF RECIDIVISM PREDICTION

A. *Evolution from Clinical Models to Actuarial Models*

Although the use of algorithms fueled by machine learning may be relatively new to the criminal justice world, predicting how likely a criminal defendant or inmate is to reoffend is not a new phenomenon. Judges, probation professionals, and correctional staff have long used a variety of tools, including their own intuition, to assess the likelihood that a person would reoffend. In the early part of the twentieth century, under what has been termed the "clinical model" of risk assessment,[19] correctional staff and clinical professionals, such as psychiatrists and social workers, played a prominent role in determining "who required

[17] In a March 26, 2020, letter to BOP officials, Attorney General Bill Barr directed the BOP to prioritize the use of granting home confinement to certain inmates in light of the COVID-19 pandemic. Attorney General Barr indicated that decisions about who was eligible for home confinement should include an examination of factors such as inmate age and vulnerability, as well as "[t]he inmate's score under PATTERN, with inmates who have anything above a minimum score not receiving priority treatment." Memorandum from Bill Barr, Attorney General, to Bureau of Prisons Officials (Mar. 26, 2020).

[18] The CDC has acknowledged that early data regarding the COVID-19 pandemic provides increasing evidence that "some racial and ethnic minority groups are being disproportionately affected by COVID-19." *Health Equity Considerations and Racial and Ethnic Minority Groups*, CTRS. FOR DISEASE CONTROL AND PREVENTION, https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html (last updated July 24, 2020).

[19] Alyssa M. Carlson, *The Need for Transparency in the Age of Predictive Sentencing Algorithms*, 103 IOWA L. REV. 303, 305 (2017) ("In the clinical model, assessments to evaluate a defendant are either made by mental health experts or other actors in the criminal justice system, such as judges or parole boards.").

enhanced security and supervision,"[20] and thus who should have a longer carceral sentence.

In the 1970s, "there was a growing recognition that the assessment of risk needed to depend more upon actuarial, evidence-based science and less on professional judgement."[21] Actuarial risk assessment systems were developed to attempt to provide a more objective approach to predicting recidivism. These early actuarial risk assessments examined which factors were most likely to be statistically associated with reoffending—such as a history of substance abuse—and assigned points to inmates who presented with those factors.[22] The higher the point total, the higher, in theory, the risk that a person would reoffend (and therefore the longer the carceral sentence). One major shortcoming was that these early models focused mostly on the types of static, historical data that was easily available to corrections professionals at the time, namely, criminal history information.[23] Because these early models were based on historical data, there was no opportunity for an individual to improve his or her score through activities, such as education courses or substance abuse treatment. Simply put, the early actuarial models did not "account for offenders changing for the better."[24]

In response, researchers began in the late 1970s and early 1980s to develop new actuarial models that took into account dynamic features such as employment while incarcerated and family relationships.[25] These newer tools were called "risk-need assessments" because they looked not only at the risk that a person would reoffend but also at the needs they might have to reduce their likelihood of

[20] James Bonta & D.A. Andrews, *Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation* 2006–07, PUBLIC SAFETY CANADA 3 (2007), https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/rsk-nd-rspnsvty/rsk-nd-rspnsvty-eng.pdf (noting that these professionals were "[g]uided by their own professional training and experience," and that the "assessment of risk was a matter of professional judgement").

[21] *Id.*

[22] *Id.* ("Actuarial risk assessment instruments consider individual items (e.g., history of substance abuse) that have been demonstrated to increase the risk of reoffending and assign these items quantitative scores.").

[23] *Id.* ("The items that create these instruments are chosen simply because they are easily available and show an association with recidivism. The items are not chosen because they are theoretically relevant. Thus, the majority of the items are criminal history items—the type of information that correctional systems are quite efficient at collecting and distributing.").

[24] *Id.* at 4.

[25] *Id.*

reoffending.[26] Having information about inmates' needs helped corrections staff appropriately identify the kinds of interventions inmates required to help reduce their recidivism risk.[27] This risk and needs model is still with us today (and indeed was required by the First Step Act), but advances in technology have changed the tools' level of sophistication and complication.

B. *Modern Risk Assessment Tools*

1. Machine Learning Algorithms

Today's modern risk assessment tools, including PATTERN, continue to address both risk and needs. As of 2015, over sixty different risk assessment tools were used in the sentencing context alone,[28] and more were used for bail determinations and by corrections officials. A 2010 Vera Institute for Justice survey found that "[a]lmost every state uses an assessment tool at one or more points in the criminal justice system," and that "over 60 community supervision agencies in 41 states reported using an actuarial assessment tool."[29]

Many of today's tools are algorithms that are fueled by machine learning, which allows them to consider many more factors than ever before and to also more closely examine the myriad ways in which those factors might interact with each other.[30] "Machine learning is an umbrella term to describe a special subset of algorithms wherein a computer is programmed to revise the code it is using as it works, based

[26] Bonta & Andrews, *supra* note 20, at 4 (noting that these new risk-need models "were sensitive to changes in an offender's circumstances and also provided correctional staff with information as to what needs should be targeted in their interventions").

[27] *Id.*

[28] Aziz Huq, *Racial Equity in Algorithmic Criminal Justice*, 68 DUKE L.J. 1043, 1075 (2019); *see also* Carlson, *supra* note 19, at 309.

[29] Memorandum from the Vera Inst. of Justice, Ctr. on Sentencing & Corr. to Illinois Risk, Assets and Needs Assessment Task Force 6, 1 (May 27, 2010), https://www2.illinois.gov/idoc/Documents/National_Information_Offender_Assessments_PartII_Memo.pdf; *see also Algorithms in the Criminal Justice System: Risk Assessment Tools*, ELECTRONIC PRIVACY INFO. CTR., https://epic.org/algorithmic-transparency/crim-justice (last visited Aug. 23, 2020) (noting that algorithmic recidivism prediction tools are used in nearly every state).

[30] Danielle Kehl, Priscilla Guo, and Samuel Kessler, Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing 9 (2017) (unpublished student work available through Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School) (on file with Digital Access to Scholarship at Harvard), http://nrs.harvard.edu/urn-3:HUL.InstRepos:33746041 ("A number of modern risk-assessment tools take advantage of machine learning algorithms, which generate risk models based on vast quantities of data. As these algorithms are used over time, their models often dynamically adjust to new data.").

on the results it is generating."[31] Put simply, the machine is programmed to perform certain tasks, "learns" from the results it is seeing, and updates its code accordingly. Because the computer is revising the code, often in real time and perhaps thousands of times, people refer to machine learning algorithms as "black boxes" because the developers cannot easily explain why the algorithm produces the outputs it does, even when they know the inputs.[32]

One of the most commonly used tools today is Correctional Offender Management Profiling for Alternative Sanctions (COMPAS). COMPAS's developers describe it as a "web-based tool designed to assess offenders' criminogenic needs and risk of recidivism."[33] COMPAS makes its predictions based on data it compiles in a 137-question evaluation, with questions ranging from "[d]o you feel that the things you do are boring or dull?" to "[i]f people make me angry or lose my temper, I can be dangerous."[34] The public widely criticized COMPAS, which the software company Northpointe developed, after May 2016 reporting by ProPublica revealed that Black defendants were 77% more likely to be labeled as a higher risk of committing a future violent crime than white defendants.[35]

---

[31] Amy B. Cyphert, Tinker-*ing with Machine Learning: The Legality and Consequences of Online Surveillance of Students*, 20 NEV. L.J. 457, 461 (2020).

[32] *See*, *e.g.*, *Oversight Hearing on the Federal Bureau of Prisons and Implementation of the First Step Act Before the Subcomm. On Crime, Terrorism, and Homeland Sec. of the H. Comm. on the Judiciary*, 116th Cong. (2019) [hereinafter *Oversight Hearing*] (statement of David E. Patton, Exec. Director, Fed. Defenders of New York) ("[A]cross risk assessments in criminal justice, the secrecy that permeates black box instruments causes significant concerns about how reasonable they are in practice."); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 638 (2017) ("Machine learning, one increasingly popular approach to automated decisionmaking, is particularly ill-suited to source code analysis because it involves situations where the decisional rule itself emerges automatically from the specific data under analysis, sometimes in ways that no human can explain."); Han-Wei Liu, Ching-Fu Lin & Yu-Jie Chen, *Beyond* State v. Loomis*: Artificial Intelligence, Government Algorithmization, and Accountability*, 27 INT'L J.L. & INFO. TECH. 122, 135 (2019) (noting that the black box problems cannot be resolved by enacting transparency requirements because "[t]he technical nature of AI techniques is characterized by an inherent lack of transparency," with even the programmers unable to explain why and how certain determinations are made). *But see* Cyphert, *supra* note 31, at 479 (acknowledging that machine learning algorithms can be shrouded in mystery but that there are none the less "several stages in the machine learning process where humans are making decisions and where safeguards can help control against bias.").

[33] Melissa Hamilton, *Risk-Needs Assessment: Constitutional and Ethical Challenges*, 52 AM. CRIM. L. REV. 231, 239 (2015) (quoting NORTHPOINTE, PRACTITIONER'S GUIDE TO COMPAS 1 (2013)).

[34] Carlson, *supra* note 19, at 310–11.

[35] Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. In

2. *State v. Loomis*

In a high-profile and landmark opinion,[36] the Wisconsin Supreme Court examined the use of the COMPAS tool. Eric Loomis pleaded guilty to attempting to flee a traffic officer and operating a motor vehicle without the owner's consent.[37] The court ordered a presentence investigation, and the "Presentence Investigation Report (PSI) included an attached COMPAS risk assessment."[38] Mr. Loomis's COMPAS risk scores "indicated that he presented a high risk of recidivism" across all three of the risk areas COMPAS purports to assess: pretrial recidivism risk, general recidivism risk, and violent recidivism risk.[39] The State relied on the COMPAS scores during its arguments at sentencing,[40] and the sentencing judge relied on the COMPAS scores in ruling out probation for Mr. Loomis.[41] The court sentenced Mr. Loomis to six years of confinement.[42]

After his initial sentencing hearing, Mr. "Loomis filed a motion for post-conviction relief requesting a new sentencing hearing," arguing among other things that the sentencing court's use of COMPAS in crafting his sentence had violated his due process rights.[43] At a subsequent hearing, an expert witness testified that "consideration at sentencing of the risk assessment portions of COMPAS runs a 'tremendous risk of over estimating an individual's risk and . . .

response, Northpointe sent ProPublica a letter, wherein it "criticized ProPublica's methodology and defended the accuracy of its test: 'Northpointe does not agree that the results of your analysis, or the claims being made based upon that analysis, are correct or that they accurately reflect the outcomes from the application of the model.'" *Id*. For more about the fallout from the ProPublica article, see Cyphert, *supra* note 31, at 465–67.

36 *See* Kehl et al., *supra* note 30, at 20 (describing the Loomis opinion as "a landmark decision," and noting that "it was the first time a U.S. court evaluated these algorithms head on").

37 State v. Loomis, 881 N.W.2d 749, 754 (Wis. 2016), *cert. denied*, 137 S. Ct. 2290 (2017).

38 *Id.*

39 *Id.* at 754–55.

40 *Id.* at 755. Specifically, the prosecutor said at sentencing that "the COMPAS report that was completed in this case does show the high risk and the high needs of the defendant. There's a high risk of violence, high risk of recidivism, high pre-trial risk; and so all of these are factors in determining the appropriate sentence." *Id.*

41 *Id.* at 755. The Judge first noted that Mr. Loomis had been identified, "through the COMPAS assessment, as an individual who is at high risk to the community." *Id.* He went on to add that "I'm ruling out probation because of the seriousness of the crime and because your history, your history on supervision, and *the risk assessment tools that have been utilized*, suggest that you're extremely high risk to re-offend." *Id.* (emphasis added).

42 *Id*. at 756 n.18 (noting that Loomis had been sentenced to two years incarceration on one count and four years incarceration on another, to be served consecutively).

43 *Loomis*, 881 N.W.2d at 756.

mistakenly sentencing them or basing their sentence on factors that may not apply.'"[44] Mr. Loomis's expert witness further pointed out that the court was missing key information about COMPAS, including what dataset was used to develop it.[45] The judge denied the post-conviction motion, explaining that he "would have imposed the same sentence regardless of whether [he] considered the COMPAS risk scores."[46]

On appeal to the Wisconsin Supreme Court, Mr. Loomis argued that his due process rights had been violated because the algorithm used by COMPAS was given trade secret protection and so "the proprietary nature of COMPAS prevent[ed] him from assessing its accuracy."[47] In rejecting that argument, the Wisconsin Supreme Court placed great weight on the fact that COMPAS used only publicly available information in formulating a risk score, and therefore Mr. Loomis had an opportunity to verify the information.[48] The fact that Mr. Loomis had no idea how much weight, if any, the COMPAS tool provided to any of those individual factors was not enough, under the court's ruling, to create a due process violation, but many scholars have disagreed with this conclusion.[49] "Simply put, Loomis may have seen the input and output [of COMPAS], but had no idea of [those inputs'] relationship,"[50] information that would have been crucial for him to vigorously challenge his risk classification.[51] The United States Supreme Court declined to hear Mr. Loomis's appeal of the Wisconsin court's decision,[52] so the decision and its problematic holdings remain law in the state of

---

[44] *Id.*

[45] *Id.* at 756–57. The expert testified that "[t]he Court does not know how the COMPAS compares that individual's history with the population that it's comparing them with. The Court doesn't even know whether that population is a Wisconsin population, a New York population, a California population . . . . There's all kinds of information that the court doesn't have, and what we're doing is we're mis-informing the court when we put these graphs in front of them and let them use it for sentence."

[46] *Id.* at 757.

[47] *Id.*

[48] *Id.* at 761. The Court held that "to the extent that Loomis's risk assessment [was] based upon his answers to questions and publicly available data about his criminal history, Loomis had the opportunity to verify that the questions and answers listed on the COMPAS report were accurate." *Id.*

[49] *See, e.g.*, Han-Wei Liu et al., *supra* note 32, at 132 ("While COMPAS algorithms drew on public data and information provided by Loomis, it did not explain the breakdown of each variable, relevant weighting and their correlation.").

[50] *Id.* at 133.

[51] *See* Kehl et al., *supra* note 30, at 28 (noting that many scholars have expressed "broad concerns" about how inputs "are weighted by the algorithm, and . . . whether specific factors (or combinations of factors) may end up serving as proxies for problematic or impermissible variables like race and poverty").

[52] Loomis v. Wisconsin, 137 S. Ct. 2290 (2017).

Wisconsin and remain influential as the most well-known example of a court examining recidivism prediction algorithms.

It is worth noting that the Wisconsin Supreme Court did provide parameters and cautions in the *Loomis* decision that are especially noteworthy given PATTERN's role in determining who is eligible for early release. The *Loomis* court held that that COMPAS scores could not be used: "(1) to determine whether an offender is incarcerated; [ ] (2) to determine the severity of the sentence," or (3) "as the determinative factor in deciding whether an offender can be supervised safely and effectively in the community."[53] In contrast, PATTERN is the only tool that will determine the risk category where an inmate is placed. As is discussed in Part IV, it is unclear whether there will be any meaningful opportunity for an inmate to challenge his or her PATTERN score. Further, unlike COMPAS, PATTERN is not just one factor that is weighed in deciding who is eligible for benefits like early release, it is THE factor. Thus, PATTERN is in certain, important ways more influential than COMPAS in determining how long an inmate is incarcerated.

III. PATTERN: THE RISK AND NEEDS ASSESSMENT TOOL

A. *The First Step Act's Charge to the Attorney General*

In Title I of the First Step Act, Congress delegated to the Attorney General[54] the task of developing and implementing a risk and needs assessment system.[55] The Act's language encouraged the Attorney General to consider algorithmic decision-making, defining the risk and needs system as "an objective and *statistically validated method*" which determines the risk of recidivism.[56] The Act does not require the reinvention of the wheel; in developing the risk and needs assessment system, it allows the Attorney General to use existing risk and needs

[53] State v. Loomis, 881 N.W.2d 749, 769 (Wis. 2016).

[54] The Act provides that the Attorney General work in consultation with several others, namely the Director of the Bureau of Prisons; the Director of the Administrative Office of the United States Courts; the Director of the Office of Probation and Pretrial Services; the Director of the National Institute of Justice; the Director of the National Institute of Corrections; and an Independent Review Committee authorized by the Act. First Step Act § 101(a) (2018).

[55] The Department of Justice has defined risks versus needs as follows: "Whereas risk refers to an inmate's statistical propensity for recidivism or some other adverse outcome, needs represents the areas of intervention that must be addressed to mitigate the risk of recidivism." U.S. DEP'T. OF JUST., OFF. OF THE ATT'Y GEN., THE FIRST STEP ACT OF 2018: RISK AND NEEDS ASSESSMENT SYSTEM 28 (2019) [hereinafter FIRST STEP ACT REPORT] https://nij.ojp.gov/sites/g/files/xyckuh171/files/media/document/the-first-step-act-of-2018-risk-and-needs-assessment-system_1.pdf.

[56] First Step Act § 101(a) (emphasis added).

assessment tools.[57] As its name suggests, the risk and needs assessment tool has twin aims: each inmate is assigned a risk of recidivism category and also assessed for any needs they may have to minimize their risk of recidivism.

With respect to risk, the Act charged the Attorney General with developing a tool to "determine the recidivism risk of each prisoner as part of the intake process, and classify each prisoner as having minimum, low, medium, or high risk for recidivism."[58] It did not provide any guidance on where to set those cutoffs. The stakes of the assessment are high, as inmates assigned a recidivism risk of "minimum" or "low" are eligible to earn additional time credits toward early release[59] as well as home confinement.[60] Further, the Act directed that "prisoners with a similar risk level be grouped together in housing and assignment decisions to the extent practicable."[61] There were also additional incentives for participation in the recidivism reduction programming, including increased phone and visitation privileges,[62] "increased commissary spending limits and product offerings,"[63] and additional "opportunities to access the email system."[64] The Act required that the assessment be dynamic, capable of changing over time as inmates either progress or regress.[65]

---

[57] *Id.*

[58] *Id.* The "recidivism" the system assessed was specifically "the risk of violent or serious misconduct of each prisoner[.]" *Id.* § 3632(a)(2).

[59] *Id.* All eligible inmates are able to earn 10 days of time credits for every 30 days they participate in the recidivism reduction programs; those who are classified as "minimum" or "low" risk for recidivating, earn an additional 5 days of time credits. *Id.*

[60] *Id.* § 602. Section 602 of the First Step Act amends 18 U.S.C. 3624(c)(2) by adding: "The Bureau of Prisons shall, to the extent practicable, place prisoners with lower risk levels and lower needs on home confinement for the maximum amount of time permitted under this paragraph." *Id.*

[61] *Id.* § 101(a).

[62] First Step Act § 101(a).

> A prisoner who is successfully participating in an evidence-based recidivism reduction program shall receive [ ] phone privileges, or, if available, video conferencing privileges, for up to 30 minutes per day, and up to 510 minutes per month; and [ ] additional time for visitation at the prison, as determined by the warden of the prison.

Id. Contact with family while a person is incarcerated has been shown to be one factor in reducing recidivism, especially for incarcerated parents. *See* Amy B. Cyphert, *Prisoners of Fate: The Challenges of Creating Change for Children of Incarcerated Parents*, 77 MD. L. REV. 385, 394–96 (2018)

[63] First Step Act § 101(a).

[64] *Id.*

[65] First Step Act § 101(a) required that the BOP's tool "reassess the recidivism risk of each prisoner periodically, based on factors including indicators of progress, and of regression, that are dynamic and that can reasonably be expected to change while in prison[.]" *Id.*

Congress did not stop at charging the Attorney General with developing a tool to determine and categorize risk. The Act also required a needs assessment for each inmate, which would result in a tailored determination of the recidivism reduction programming and/or productive activities that would best help to reduce that person's likelihood of reoffending upon release from prison.[66] The Act defined "evidence-based recidivism reduction program" as "a group or individual activity that [ ] has been shown by empirical evidence to reduce recidivism or is based on research indicating that it is likely to be effective in reducing recidivism[.]"[67] These programs may include a wide variety of topics and approaches, ranging from "social learning and communication, interpersonal, anti-bullying, rejection response, and other life skills"[68] to vocational training.[69] The Act defined "productive activities" somewhat circularly as activities that allow inmates to remain productive and maintain a low risk of recidivating.[70] The Act required that those labeled with a higher risk of reoffending be given priority access to recidivism reduction programs, and that those classified as lower risk for reoffending be given priority access to productive activities.[71] The Attorney General was also directed to review existing evidence-based recidivism programs and productive activities, identifying which were most effective and coordinating with the Bureau of Prisons (BOP) about them.[72]

---

[66] *Id.* (requiring that the Attorney General "determine the type and amount of evidence-based recidivism reduction programming that is appropriate for each prisoner and assign each prisoner to such programming accordingly, and based on the prisoner's specific criminogenic needs"). With respect to "criminogenic needs," the First Step Act Report announcing the creation of the PATTERN tool notes: "Although research varies about how each of these needs is linked to recidivism risk, criminal thinking, antisocial peers, substance abuse, and education and vocational needs typically are identified as important criminogenic needs." FIRST STEP ACT REPORT, *supra* note 55, at 29.

[67] First Step Act § 101(a).

[68] *Id.*

[69] *Id.* Other topics include academic classes, mentoring, and substance abuse treatment. *Id.*

[70] *Id.* The Act defines a "productive activity" as:

> [E]ither a group or individual activity that is designed to allow prisoners determined as having a minimum or low risk of recidivating to remain productive and thereby maintain a minimum or low risk of recidivating, and may include the delivery of the programs described in paragraph (1) to other prisoners.

Id.

[71] The Act amended 18 U.S.C. § 3621(h)(6) as follows: "Priority for participation in recidivism reduction programs shall be given to medium-risk and high-risk prisoners, with access to productive activities given to minimum-risk and low-risk prisoners." First Step Act § 102.

[72] First Step Act § 101(a).

B. *The Development of PATTERN*

The Act had an ambitious timeline, providing only 210 days from the date of its passage until the risk and needs assessment system needed to be developed and publicly released.[73] Further, certain tasks had to be completed *prior* to the release of the system, including the Attorney General's review of existing recidivism programs and corresponding direction to BOP about them.[74] On July 19, 2019, exactly 209 days after the passage of the First Step Act, Attorney General William Barr publicly released the Prisoner Assessment Tool Targeting Estimated Risks and Needs, or "PATTERN." Two outside experts, Dr. Grant Duwe and Dr. Zachary Hamilton, designed PATTERN,[75] both having previously developed risk and needs assessment programs.[76]

Dr. Duwe and Dr. Hamilton participated in three listening sessions the DOJ hosted during PATTERN's development "to enable experts, stakeholders—including organizations representing crime victims—and public interest organizations to comment" on its development.[77] They also participated in conference calls with the constituent groups with whom the Act required the Attorney General to collaborate: the BOP, the National Institute of Justice (NIJ), and the Independent Review Committee.[78] Further, after the original public release of the tool in June of 2019, the DOJ continued to host listening sessions and gather feedback in a 45-day public comment period.[79] On January 15, 2020, the DOJ released a report announcing certain updates to the PATTERN tool,

---

[73] *Id.*

[74] *Id.*

[75] FIRST STEP ACT REPORT, *supra* note 55, at 12.

[76] *Id.* at 13. Dr. Duwe has developed four separate risk assessment instruments, including the Minnesota Screening Tool Assessing Recidivism Risk. *Id.* Dr. Hamilton has developed several risk-need assessment systems for adults and juveniles in the states of Delaware, Iowa, Nebraska, Tennessee, and Washington. *Id*.

[77] *Id.* at 14.

[78] *Id.* at 42. The Act tasked the NIJ with choosing a nonpartisan, nonprofit organization with expertise in risk and needs assessment tools to host the IRC and select its members. First Step Act § 107. The NIJ selected the Hudson Institute as the host of the IRC. FIRST STEP ACT REPORT, *supra* note 55, at 9. The Hudson Institute then selected the IRC's six members: Dr. Patti Butterfield, Dr. James M. Byrne, Dr. Faye S. Taxman, George J. Terwilliger III, John P. Walters, and John E. Wetzel. *Id.* at 10–12; *see also Oversight Hearing*, *supra* note 32 (statement of John P. Walters, Member, First Step Act Independent Review Committee) ("The Independent Review Committee is composed of six experts from a range of fields and with extensive expertise in both research and operations."). For more on the qualifications of the IRC members, see First Step Act sections 107(c) and 107(d); *see also infra* note 225.

[79] *See* U.S. DEP'T. OF JUSTICE, OFFICE OF THE ATTORNEY GEN., THE FIRST STEP ACT OF 2018: RISK AND NEEDS ASSESSMENT SYSTEM—UPDATE 3 (January 2020) [hereinafter UPDATED REPORT].

updates that the Department said were made "as a direct result of the input and suggestions" received during the public comment period.[80] Still, DOJ announced more changes to the tool in June of 2020[81] after ProPublica exposed secret changes made in the wake of the COVID-19 pandemic.[82] These later changes to the tool are discussed at length below.

In developing PATTERN, Dr. Duwe and Dr. Hamilton reviewed BOP data concerning 278,940 former inmates who had been released into the community.[83] The BOP dataset tracked those who were rearrested within three years and excluded those who died during the three-year follow-up period as well as anyone scheduled for deportation.[84] The dataset also included information on factors the DOJ described as "commonly associated with recidivism risk,"[85] such as age at time of release, criminal history, and any misconduct during the time of incarceration.[86]

Dr. Duwe and Dr. Hamilton also reviewed data from BOP's existing recidivism prediction tools, BRAVO and BRAVO-R.[87] BRAVO ("Bureau Risk Assessment Verification and Observation"), created by the BOP in the 1970s, "was designed to predict serious misconduct in prison."[88] BRAVO-R ("Bureau Risk Assessment Verification and Observation-Revised") was "designed to address recidivism in the community," and "to predict misconduct for custody-level classification purposes."[89] Dr.

---

[80] *Id.* at 1.

[81] *See* U.S. DEP'T. OF JUSTICE, FIRST STEP ACT IMPLEMENTATION FISCAL YEAR 2020 90-DAY REPORT 1 (June 2, 2020) [hereinafter JUNE REPORT].

[82] Ian MacDougall, *Bill Barr Promised to Release Prisoners Threatened by Coronavirus—Even as the Feds Secretly Made It Harder for Them to Get Out,* PROPUBLICA (May 26, 2020, 5:00 AM), https://www.propublica.org/article/bill-barr-promised-to-release-prisoners-threatened-by-coronavirus-even-as-the-feds-secretly-made-it-harder-for-them-to-get-out.

[83] The dataset included persons who had been released from BOP facilities from 2009 to 2015. FIRST STEP ACT REPORT, *supra* note 55, at 42. "The vast majority of items within the dataset were drawn from SENTRY, the BOP's centralized inmate management system." *Id*. at 49.

[84] *Id.* at 42–43.

[85] *Id.* at 43.

[86] *Id.*

[87] *Id.* at 12.

[88] *Id*. at 12, 43.

[89] FIRST STEP ACT REPORT, *supra* note 55, at 12, 42. Some advocates have questioned whether PATTERN is different enough from BRAVO or BRAVO-R to add meaningfully to current BOP recidivism prediction efforts. *See* U.S. DEP'T OF JUSTICE, OFF. OF JUSTICE PROGRAMS, NAT'L INST. OF JUSTICE, STAKEHOLDER STATEMENTS SUBMITTED IN RESPONSE TO NIJ'S FIRST STEP ACT LISTENING SESSIONS (September 2019) [hereinafter NIJ LISTENING SESSION STATEMENTS] (written statement of James F. Austin on behalf of the JFA Institute) (noting that since "either the existing BOP BRAVO-R or the proposed PATTERN would be equally

Duwe and Dr. Hamilton used the BOP dataset to develop and "validate" PATTERN;[90] they used the dataset both to build the tool initially and then to test the tool and see what its predictive accuracy would have been with respect to the known data on the three-year rearrest data.[91] Part IV below addresses several outstanding questions regarding the specifics of how Dr. Duwe and Dr. Hamilton developed and validated PATTERN, including whether they used machine learning in developing their algorithm.

C. *Defining Recidivism: What Does PATTERN Purport to Predict?*

PATTERN purports to predict the likelihood that a person will reoffend within the three years following their release from a BOP facility.[92] Of course, as with any recidivism prediction tool, what PATTERN is actually attempting to predict is not the likelihood that an inmate will reoffend but the likelihood that they will be arrested and/or convicted for criminal behavior.[93] PATTERN was developed to predict both "general recidivism" (defined as "any arrest or return to BOP custody following release" within a three-year period post-release) as well as "violent recidivism" (defined as "violent arrests following release" within a three-year period post-release).[94] These definitions of recidivism have been criticized as overly broad for including minor technical violations, like failing to update a residence,[95] though the DOJ

effective [based on their AUC scores, this] raises the question of why the expense and time [was] spent to develop an [sic] 'new' instrument that is already in place").

90 FIRST STEP ACT REPORT, *supra* note 55, at 43.

91 Specifically, they used approximately 2/3 of the dataset as "training data" to train their algorithm with. *Id.* at 49. They then used the remaining approximately 1/3 of the dataset as "test data" to establish the accuracy of the tool. *Id.*

92 *Id.* at 43.

93 *Id.* at 50. The decisions made by recidivism prediction tool developers about "what counts" as recidivism at times "touches on key sentencing policy decisions often left undecided by state actors." Jessica M. Eaglin, *Constructing Recidivism Risk*, 67 EMORY L.J. 59, 78 (2017) (noting that when recidivism is defined solely by arrest, the data will not reflect whether a person was ultimately exonerated because of an overturned conviction or even if the prosecutor declines to press charges). See *infra* Part IV for a discussion of how racial disparities in policing can further bias recidivism prediction when recidivism is defined solely by arrest.

94 FIRST STEP ACT REPORT, *supra* note 55, at 50.

95 *Oversight Hearing*, *supra* note 32, at 9 ("The definition the DOJ chose is unduly broad, sweeping in revocations for minor technical violations such as failure to timely report a change of residence, or failing to timely notify the probation officer of being questioned by police."); *see also Oversight Hearing*, *supra* note 32 (Letter from Ames C. Grawert to David B. Muhlhausen, NIJ Director (Sept. 3, 2019), attached to statement of Ames C. Grawert, Senior Counsel, Justice Program, Brennan Center for Justice at NYU School of Law) [hereinafter Brennan Center letter] ("PATTERN's general tool was designed around a very broad definition [of recidivism]: the risk that an incarcerated

declined to narrow them in the revised PATTERN tool, claiming they lacked the data necessary to do so.[96] The terms have also been criticized as too vague for not including a definition of which infractions are deemed "violent."[97] In the Updated Report, the DOJ clarified that the terms "general recidivism" and "violent recidivism" mirror the definitions used by the Bureau of Justice Statistics and provided more precise definitions and illustrative examples.[98]

In addition to dividing recidivism into "general" and "violent" categories, PATTERN also subdivides depending on the gender of the person being examined. The tool's developers argue that "adding both gender and outcome (i.e., general and violent recidivism) specificity" represents "recent advancements in risk assessment tool construction" over earlier tools such as BRAVO-R.[99] To capture gender-specific risk metrics, men and women were separated into individual samples and a different algorithm was created for each.[100]

---

person would face a 'new arrest **or** return to BOP custody within three years of release.'").

[96] UPDATED REPORT, *supra* note 79, at 13. The DOJ also argued in the Updated Report that it saw benefits in retaining the old definitions, including to allow BOP to "more accurately measure how well a program is working" and to facilitate "comparisons across risk assessment and criminal justice systems." *Id*. at 14.

[97] *Oversight Hearing*, *supra* note 32, Ex. A at 10 ("More information is needed here, as well, regarding what kinds of arrests are considered 'violent.'").

[98] UPDATED REPORT, *supra* note 79, at 12. Specifically, DOJ provided that:

> General recidivism is defined as a return to BOP custody or a re-arrest within three years of release from BOP custody, excluding all traffic offenses except driving under the influence (DUI) and driving while intoxicated (DWI). Violent recidivism is defined as a re-arrest for a suspected act of violence within three years of release from BOP custody. Examples of the violent offenses captured in this definition include, but are not limited to, firearms violations, homicide, child abuse, robbery, sex trafficking, and sexual assault.

Id. at 12–13.

[99] FIRST STEP ACT REPORT, *supra* note 55, at 44. This decision to separate men from women when making recidivism predictions may well improve the accuracy of the tool for both groups, given the disparate rates of violence. "When base rates of the predicted outcome differ across groups, the most accurate algorithm possible will predict that outcome at different rates across groups." Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. 2218, 2249 (2019). Gender-neutral risk assessment tools tend to have an adverse effect on women by overestimating their risk of recidivism. *See* Sharad Goel et al., The Accuracy, Equity, and Jurisprudence of Criminal Risk Assessment 9 (Dec. 26, 2018) (unpublished article), https://ssrn.com/abstract=3306723. PATTERN has been criticized, however, for using different algorithms to assess men and women. *See* UPDATED REPORT, *supra* note 79, at 16 (noting that the DOJ had received feedback that "PATTERN's separate modelling for men and women raises constitutional concerns").

[100] *See* FIRST STEP ACT REPORT, *supra* note 55, at 49 ("Gender-responsive risk metrics are developed by separating males and females into individual samples to produce gender-specific prediction models, which improves both the context and accuracy of

D. *Risk Factors and Risk Levels*

In making its predictions, the PATTERN tool takes into account factors that are dynamic (things an inmate can change, like participation in education classes while incarcerated) as well as static (historical things that are unchangeable, such as the inmate's age at arrest).[101] The use of both dynamic and static factors is considered a best practice,[102] and indeed Congress mandated the use of both in the First Step Act.[103] Examples of dynamic factors that PATTERN relies on in making its predictions include an inmate's participation in drug education and treatment programs, participation in employment, and use of the income earned from that employment for payment toward victim restitution and/or dependents.[104] Examples of static factors included in the original PATTERN tool are age at first conviction, whether the crime was "violent," and whether the inmate was identified as a "sex offender."[105]

PATTERN's inclusion of dynamic factors is noteworthy. As the DOJ's own report remarked, some research suggests that a predictive tool's potential for racial bias is lowered when dynamic factors are included.[106] Prior risk prediction tools utilized the "Burgess method," wherein each predictive factor was equally weighted,[107] but the PATTERN tool analytically weights the dynamic and static factors it uses.[108] It appears that the DOJ has publicly released the weights that

prediction."). Males represented approximately 85% of the training and data samples, and females the remaining 15%. *Id.*

101 *Id.* at 43.

> Static factors are characteristics of inmates that are historical and therefore unchangeable, such as offense severity, age at first arrest, and criminal history at prison entry. By contrast, dynamic factors are variables that may change over time and may reflect more recent inmate behavior, such as prison misconduct or substance abuse.

Id. at 26.

102 *Id.*

103 *See* First Step Act of 2018, Pub. L. No. 115-391, § 101, 132 Stat. 5194 (2018) (requiring that the BOP's tool "reassess the recidivism risk of each prisoner periodically, based on factors including indicators of progress, and of regression, that are dynamic and that can reasonably be expected to change while in prison").

104 FIRST STEP ACT REPORT, *supra* note 55, at 45.

105 *Id.* at 45–46.

106 *Id.* at 60.

107 The "Burgess method," developed by Professor Ernest Burgess in 1927, used 21 factors to determine an individual's likelihood of recidivism while on parole. The "Burgess method" was one of the first risk assessment methods to be developed. Carlson, *supra* note 19, at 308.

108 "Based on findings of previous studies, analytically weighting assessment items improves predictive accuracy, and this was the method adopted for PATTERN." FIRST STEP ACT REPORT, *supra* note 55, at 50.

PATTERN assigns to each of the predictive factors. The First Step Act Report included a Table entitled "Points Assigned in the PATTERN Risk Assessment Models,"[109] and noted that "[l]arger values indicate greater prediction strength for a given model."[110] Several advocates expressed confusion about whether the points in that Table represented the final and sole weighing of the factors.[111] (Understanding the weights assigned to the various predictive factors used in any prediction tool is, of course, important to analyzing it.)[112] In response, the Updated Report clarified that "[e]ach individual's response scores are summed to compute a total score."[113]

The PATTERN developers also had to decide where to place the "cut-off points" that would define, as the Act required, the risk-level categories of "high," "medium," "low," or "minimum."[114] The DOJ has acknowledged that there are "several known methods" for determining such cutoff points, most using statistics.[115] Even though sophisticated statistics may be used to help determine cutoffs in a predictive algorithm like PATTERN, the fact remains that the decisions about where to place the cutoffs are best described as "a matter of policy, not math,"[116] and some scholars have argued that such important policy

---

109 *Id.* at 53–56.

110 *Id.* at 53. The Report also explained that "boosted regression models were computed for each of the four models, with selected items and weights created from model coefficients. Coefficient values for each model were converted (multiplied by 100) to whole numbers to improve ease of risk scoring." *Id.* at 66.

111 *See, e.g.*, *Oversight Hearing*, *supra* note 32, Ex. A at 9 (letter from David E. Patton, Exec. Director, Fed. Defenders of New York to David Muhlhausen, Director of Nat'l Inst. of Justice, attached to statement of David E. Patton, Executive Director, Federal Defenders of New York) ("The DOJ Report indicates that PATTERN involves 'analytically weighting assessment items,' but more information is needed on whether the weights are assigned solely through the points identified for each of the factors included in Table 2, or are somehow reweighted in an algorithm not discussed in the report.").

112 Eaglin, *supra* note 93, at 73. Tool creators generally use one of two statistical methods to weigh assessment factors. One statistical method assigns every factor equal weight. The other statistical method assigns different weight to assessment factors depending on how predictive each factor may be. *Id.* at 81.

113 UPDATED REPORT, *supra* note 79, at 37.

114 First Step Act of 2018, Pub. L. No. 115-391, § 101, 132 Stat. 5194 (2018).

115 *See* FIRST STEP ACT REPORT, *supra* note 55, at 50.

116 Chelsea Barabas et al., *An Open Letter to the Members of the Massachusetts Legislature Regarding the Adoption of Actuarial Risk Assessment Tools in the Criminal Justice System*, BERKMAN KLEIN CTR. FOR INTERNET & SOC'Y 1, 4 (Nov. 9, 2017), http://nrs.harvard.edu/urn-3:HUL.InstRepos:34372582 ("By way of example, the classification of a risk category applicable to a particular criminal defendant with respect to a given risk score (e.g., high risk, medium risk, or low risk) is a matter of policy, not math.").

decisions should not be left to predictive tool developers.[117] PATTERN's developers set the cutoffs for its risk-level categories by basing them, in part, on BOP base rates for general and violent recidivism,[118] a decision that has been criticized as "arbitrary."[119]

In May of 2020, ProPublica published an article reporting that the DOJ had quietly lowered the PATTERN risk-level cutoffs without making any public acknowledgment of doing so, making it harder for inmates to be classified as "minimum" risk, and had used the new risk-level cutoffs in determining which inmates were eligible for early release during the COVID-19 pandemic.[120] A week later, the DOJ released a fiscal report where it acknowledged the new, lower risk-level cutoffs.[121] Ultimately, to be classified as "low" or "minimum" risk and therefore be eligible for early release and other benefits, an inmate must be classified as "low" or "minimum" risk in *both* the general and violent recidivism prediction models.[122]

---

117 *See*, *e.g.*, Eaglin, *supra* note 93, at 59 ("Government agencies should always decide the cut-off points."). Indeed, the First Step Act Report acknowledged that the choice of cutoff points is in some ways a policy decision.

> The current cut points endeavor to set the appropriate balance between maximizing the number of inmates eligible to earn early release time credits and to participate in evidence-based programming that would reduce their recidivism risk to a low or minimum category, while also considering public safety and the risk of recidivism upon release.

FIRST STEP ACT REPORT, *supra* note 55, at 51.

118 *Id.*

> Specifically, the low risk threshold was set at roughly half the base rate (or 24 percent) for general recidivism and just over two-thirds the base rate for violent recidivism (or 12 percent). The minimum risk category was set at just under one-quarter of the base rate (or 10 percent) for general recidivism and one-third of the base rate for violent recidivism (or 5 percent). The high risk category was set at roughly two-thirds above the base rate (or 80 percent) for general recidivism, and just over twice the base rate (or 33 percent) for violent recidivism.

Id. Anyone not identified as minimum, low, or high risk was, by default, classified as medium risk. *Id.* "This specific set of cut-off points was one of nearly a dozen" cutoff points that were tested by the PATTERN developers. *Id.*

119 *Oversight Hearing*, *supra* note 32, Ex. A at 16 (letter attached to statement of David E. Patton, Exec. Director, Fed. Defenders of New York) ("For example, the decisions on the cut-points, which necessarily impact fairness measures such as false positive rates and positive predictive values, appear to have been made by the researchers and based on arbitrary fractions or multiples of the recidivism rates.").

120 MacDougall, *supra* note 82.

121 JUNE REPORT, *supra* note 81, at 2.

122 FIRST STEP ACT REPORT, *supra* note 55, at 51.

> The four categories were created for both general and violent recidivism risk scores. A final set of categories was created where an

PATTERN's developers "analyzed" the impact of the initial risk-level cutoff points based on an inmate's race and/or ethnicity by creating a relative risk index. This index compares the proportion of individuals from racial or ethnic groups across the sample to the proportion of individuals from those same groups within the risk levels.[123] The First Step Act Report provides the results of the relative risk index at Tables 8, 9, and 10, but does not discuss them.[124] Scholars who have examined those tables report that they "reflect[] racial disparity."[125] No information was provided in the June Report concerning the new, lower risk-level cutoff points or any racial disparities.

E. *Assessing PATTERN's Accuracy*

Having developed their tool using the BOP data, PATTERN's developers tested it to determine its predictive validity. They chose to use a statistical method known as "area under curve" or "AUC."[126] At the risk of oversimplifying what is an "an inherently difficult concept,"[127] AUC is essentially a measure of how well a predictive tool performs when it makes the classifications it has been programmed to make. For a recidivism prediction tool like PATTERN, AUC is best thought of as a measure of how often the tool correctly assigned a higher risk level to someone who reoffended than it did to someone who did not

> individual must be identified as minimum risk of both general and violent recidivism to be classified as minimum in the final RLCs. An individual that was identified as lower than medium risk in both the general and violent models was labeled as low risk in the final RLC. Those individuals identified as high risk in either the general or violent models were classified as high risk in the final RLCs. Finally, those not classified as minimum, low, or high risk were identified as medium risk in the final RLCs.

Id.

123 *Id.* at 53. The results of this relative risk index are displayed in Tables 8–10 of the FIRST STEP ACT REPORT. *Id.* at 62.

124 *Id.* at 62.

125 *Oversight Hearing*, *supra* note 32, at 12 (statement of Prof. Melissa Hamilton). Professor Hamilton notes that "[a]n RRI over 1.0 indicates disparity between groups," and that Table 8 of the First Step Act Report shows an RRI of 1.54 for males. *Id.* Put another way, "[n]on-Whites are one-and-a-half times more likely to be assessed as medium/high risk than Whites" under the original PATTERN tool. *Id.*

126 *See* FIRST STEP ACT REPORT, *supra* note 55, at 50 ("Relying on the AUC as the primary metric for evaluating predictive validity, the research team analyzed how the PATTERN instruments performed in predicting recidivism on the test set . . . .").

127 Melissa Hamilton, *Adventures in Risk: Predicting Violent and Sexual Recidivism in Sentencing Law*, 47 ARIZ. ST. L.J. 1, 27 (2015).

reoffend.[128] In other words, if person A reoffended and person B did not, and PATTERN classified person A as "high risk" and person B as "medium risk," that would be counted as "accuracy" for AUC purposes. If the tool is correct in this classification half the time, the AUC score is .5.[129] If the tool is correct in this classification all of the time, the AUC score is 1.[130] Most of the recidivism prediction tools that are used today have AUC values that fall between .60 to .80.[131]

PATTERN's creators report that, based on its AUC value, PATTERN is "15% more predictive than other [recidivism prediction] tools"![132] It is worth noting here that although AUC is a widely used, popular method to assess the predictive accuracy of risk assessment tools,[133] the First Step Act Report itself acknowledged that this method "can provide an overly optimistic estimate of predictive discrimination for imbalanced datasets."[134] According to the DOJ, PATTERN's AUC values, which range from .77 to .80,[135] represent a "consistent, modest improvement" over the predictive reliability of the BRAVO-R tool.[136] PATTERN's developers also analyzed the tool and its risk-level cut points via odds ratios. The developers calculated that, for general recidivism, inmates labeled as minimum or low risk had an 86% reduced chance of recidivating as compared to those labeled as medium or high risk.[137] For violent recidivism, those labeled as minimum or low risk had an 88% reduced

128 *Id.* at 25*; see also* Charlotte Hopkinson, *Using Daubert to Evaluate Evidence-Based Sentencing*, 103 CORNELL L. REV. 723, 741 (using AUC to measure "how well the predictive test can differentiate those who experienced the outcome of interest (labeled as likely to recidivate and did recidivate) versus those who did not (labeled as likely to recidivate but did not recidivate)").

129 FIRST STEP ACT REPORT, *supra* note 55, at 56 ("[A]n AUC of .5 means there is a 50 percent probability that the result is accurate, which is the same as a coin toss.").

130 *Id.* ("An AUC of 1.0 means there is a 100 percent probability in the result, which is essentially perfection."); *see also* Hamilton, *supra* note 127, at 25 ("AUC values lie between 0 and 1, with .5 indicating discriminatory ability no better than chance and 1 indicating perfect discrimination.").

131 Eaglin, *supra* note 93, at n.180.

132 FIRST STEP ACT REPORT, *supra* note 55, at 57 fig.3.

133 *Id.* at 37; *see also* Eaglin, *supra* note 93, at 90 ("[A] popular method to assess predictive accuracy measures the area under the curve, or the AUC value."); Hamilton, *supra* note 127, at 25 (noting that "[s]everal statistical measures of discrimination for actuarial tools are available," but that AUC "has come to dominate the relevant literature").

134 FIRST STEP ACT REPORT, *supra* note 55, at 66 n.24. The Report goes on to note that AUC "is relatively robust across different recidivism base rates and selection ratios." *Id.*

135 Specifically, the AUC scores were calculated as follows: female violent recidivism at .77; female general recidivism at .79; male violent recidivism at .78; and male general recidivism at .8. *Id.* at 57 tbl.3.

136 *Id.* at 57.

137 *Id.* at 58.

chance of recidivating, as compared to those labeled medium or high risk.[138] The First Step Act Report concluded that these findings "identify a substantial demarcation in recidivism rates between early-release-eligible (minimum and low) and non-eligible (medium and high) risk categories."[139]

DOJ also reported AUC scores for PATTERN that were broken down by the race/ethnicity of the inmate. They used the four categories the BOP data used: white, African American, Hispanic, and other.[140] The First Step Act Report is internally inconsistent on the presence of any racial bias in PATTERN. At one point, the report concludes that the AUC value findings "suggest" "minimal" racial or ethnic disparity with respect to PATTERN's predictive value,[141] and just a few pages later the report says that "the PATTERN instrument's predictive performance is unbiased across racial and ethnic classifications."[142] In any event, the AUC calculations relied on to demonstrate racial neutrality actually show that in both categories for male inmates (male general recidivism and male violent recidivism), the AUC value calculated for white inmates is higher than the AUC value for African American and Hispanic inmates.[143] This means that, accepting the DOJ's own test for predictive validity, PATTERN is less accurate in predicting recidivism for men of color than for white men. The AUC values for white women were lower than the AUC values for African American women but higher than those of Hispanic women.[144] The presence of racial bias in the PATTERN tool and how to minimize it is addressed in Part IV.

---

138 *Id.*

139 *Id.*

140 FIRST STEP ACT REPORT, *supra* note 55, at 60.

141 *Id.* ("Overall, these findings indicate strong and comparable prediction strength for PATTERN models across all race/ethnicity categories, suggesting minimal racial/ethnic disparity for PATTERN's prediction strength.").

142 *Id.* at 63.

143 *See id.* at 60 tbl.7. The AUC value for general recidivism among white men was calculated to be .81, as opposed to .78 for both African American and Hispanic men. *Id.* The AUC value for violent recidivism among white men was calculated to be .80, as opposed to .75 for African American men and .78 for Hispanic men. *Id.*

144 *Id.* The AUC value for general recidivism among African American women was calculated to be .80, as opposed to .79 for white women and .76 for Hispanic women. *Id.* The AUC value for violent recidivism among African American women was calculated to be .80, as opposed to .75 for both white and Hispanic women. *Id.* Professor Melissa Hamilton notes that even though demographic parity is better for women than men under PATTERN, it is still problematic and means that "African-American females benefit less often from the First Step Act's incentives and rewards." *Oversight Hearing*, *supra* note 32, at 13 (statement of Prof. Melissa Hamilton).

IV. PATTERN'S INCOMPLETE PROGRESS

There are many unanswered questions about PATTERN and several problems with its development and the assumptions that underlie it. Still, it is important to note that there are some laudable aspects to PATTERN, even as each is not yet sufficient. Specifically, PATTERN's development involves a level of transparency greater than that of private software companies (although, as is discussed in the next Section, still greater transparency is required); PATTERN's commitment to using dynamic factors, at least in theory, is a step in the right direction (though its practical application is currently in doubt); the PATTERN developers reached out to community stakeholders and held public comment sessions in creating the tool; and PATTERN's commitment to recalibration and continued study is an important best practice. Ultimately, the decision of the PATTERN developers to prioritize accuracy over other important policy goals, and their decision to use AUC as a proxy for accuracy, remain problematic.

A. *Transparency*

With respect to transparency, it should be acknowledged that the DOJ has provided more information about PATTERN than many private software companies provide about their recidivism prediction tools. Northpointe's decision to claim trade secret protection over its COMPAS algorithm in the *Loomis* case is a troubling but not isolated example. Indeed, one reason some academics have advocated for government-developed algorithms is precisely to avoid such trade secret and proprietary claims, in the hope that government-funded software would be available to the public at large to inspect.[145] One academic noted that

[145] *See, e.g.*, Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 2025–26 (2017) (citing CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 27 (2016)) (advocating that "courts and legislatures should consider requiring that software used in criminal trials and sentencings be publicly designed and open-source," and noting that "[p]ublic models would have the benefit of being 'transparent' and 'continuously updated, with both the assumptions and the conclusions clear for all to see'"); *see also* Kehl et al., *supra* note 30, at 28 (noting that "[a]cademic researchers and governments, [as opposed to private software companies,] . . . tend to have more incentives to make the details of their algorithms publicly available and ensure that they are subject to appropriate scrutiny and oversight"). It is worth noting that the DOJ and other government agencies *have* claimed trade secret protection over government-developed software programs. *See, e.g.*, Rebecca Wexler, *Life, Liberty and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1346 (2018) ("[T]he federal government claimed that trade secret interests should shield details about how a cybercrime investigative software program operates . . . ."). Further, "New York City's Office of Chief Medical Examiner (OCME) argued, repeatedly and successfully, that the source code for a forensic software program

even providing the "minimal disclosures" of "information about the specific origin of the data set underlying a tool" and "the selection of risk factors"—both items the DOJ appears to have provided—would "go beyond what many developers currently provide."[146]

1. Congress Mandated Certain Disclosures in the Act

Congress mandated in the First Step Act that the Attorney General not only create a risk and needs assessment but further that he *publicly release* that system.[147] Accordingly, the First Step Act Report provides the following about PATTERN: (1) some information about what the dataset used to develop the tool was;[148] (2) what the numeric risk-level cutoffs are[149] (though, of course, as noted above, those cutoff levels appear to have been changed without a timely public acknowledgment); (3) the factors that the tool uses in calculating an inmate's risk score;[150] (4) the relative weight of each of those factors;[151] and (5) some information about the relative accuracy of the tool, at least as measured by one technique as against other BOP recidivism prediction tools.[152] The Updated Report also offers additional information about PATTERN, namely which factors the PATTERN developers consider "dynamic" and which they consider "static."[153]

Congress also mandated in the First Step Act a certain level of ongoing reporting about PATTERN. For example, the Act requires the Director of the Bureau of Justice Statistics to annually submit to Congress statistics on "[t]he breakdown of prisoners classified at each risk level by demographic characteristics."[154] The Act also requires that the Independent Review Committee submits to Congress a report addressing the demographic percentages of inmates ineligible to

developed in-house using taxpayer funds should be protected from subpoena by criminal defendants." *Id.* at 1397.

146 Eaglin, *supra* note 93, at 111.

147 First Step Act § 3632(a) ("Not later than 180 days after the date of the enactment of the FIRST STEP Act, the Attorney General shall develop and release a risk and needs assessment system[.]")

148 FIRST STEP ACT REPORT, *supra* note 55, at 43 ("Hamilton and Duwe used the BOP datasets to develop and validate" PATTERN).

149 *Id.* at 58.

150 *Id.* at 53–56. The Updated Report also provides the factors for the updated tool. UPDATED REPORT, *supra* note 79, at 37–39.

151 FIRST STEP ACT REPORT, *supra* note 55, at 53–56.

152 *Id.* at 57.

153 UPDATED REPORT, *supra* note 79, at 10–11.

154 *See* First Step Act of 2018, Pub. L. No. 115-391, § 610(a)(26), 132 Stat. 5194 (2018) (requiring the annual breakdown of prisoners to include the following demographic characteristics: "age, sex, race, and the length of the sentence imposed").

receive and apply time credits, including by age, race, and sex.[155] These reports, provided they are publicly released, should add to the overall transparency of PATTERN by allowing researchers to examine them for indicators of bias.

2. Additional Transparency Is Critical

Although these steps toward transparency are laudable, they are wholly insufficient. There are several important steps that the DOJ should take to provide more transparency about PATTERN—transparency that will allow advocates to truly assess the fairness of the system. "Greater transparency will have little impact if outside researchers do not have access to the data and tools to evaluate and test the algorithms for bias."[156]

i. The Dataset Used to Create PATTERN Should Be Publicly Released

First, the BOP dataset used to develop PATTERN should be publicly released. Several advocates and scholars—ranging across the ideological spectrum from the Brennan Center to the Charles Koch Institute—have already made this request in Congressional testimony[157] and at NIJ listening sessions,[158] and the National Association of Criminal Defense Lawyers (NACDL) has submitted a FOIA request for the dataset.[159] The dataset is already complete, has already

155 *See* First Step Act § 107(g)(4) (requiring the Independent Review Committee to submit "recommendations regarding recidivism reduction" with its report); *see also id.* § 103(8) (requiring the Comptroller General to conduct an audit of the use of the risk and needs assessment system every two years, which must include an analysis of "[t]he rates of recidivism among similarly classified prisoners to identify any unwarranted disparities, including disparities among similarly classified prisoners of different demographic groups, in such rates").

156 Kehl et al., *supra* note 30, at 33.

157 *See, e.g.*, *Oversight Hearing*, *supra* note 32 (statement of David E. Patton, Executive Director, Federal Defenders of New York) ("Full transparency requires the DOJ to release the same dataset used by Grant Duwe, Ph.D., and Zachary Hamilton, Ph.D., to create PATTERN."); *Oversight Hearing*, *supra* note 32 (statement of Prof. Melissa Hamilton) ("The datasets should be public released.").

158 *See* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Jeremiah Mosteller on behalf of the Charles Koch Institute) ("[W]e strongly recommend that the Department of Justice allow a group of independent, external researchers to analyze the data used to develop the PATTERN and confirm its predictive validity before full implementation by the Bureau of Prisons.").

159 *See Oversight Hearing*, *supra* note 32 (written Statement of Norman Reimer, Executive Director of the National Association of Criminal Defense Lawyers). The FOIA request, attached to the written statement, requested the dataset used to develop PATTERN along with an array of other information, including the model used to train

been provided to PATTERN's developers, and should be relatively easy to produce.[160] Without access to the dataset, researchers and advocates have to accept at face value the DOJ's claims about PATTERN's validity, accuracy, and bias.[161] Further, access to the dataset would allow researchers to calculate other methods of accuracy beyond AUC, a measure that is problematic for the reasons discussed below.[162]

Despite these compelling and uniform calls for release of the dataset, the DOJ declined in the Updated Report to do so, arguing that the data "is restricted because [it] . . . includes arrest and conviction information provided directly to DOJ by state and local jurisdictions, who have a significant interest in protecting their data."[163] Putting aside the question of how much of the PATTERN data actually comes from states and localities as opposed to how much is already in the underlying BOP dataset, as the initial report provided, the argument about data privacy does not withstand scrutiny. The DOJ is essentially arguing that it cannot release data that is overwhelmingly publicly available[164] and was compiled with taxpayer dollars because of data privacy concerns. Further, the DOJ claims that the information is subject to "sharing agreements" that prohibit the release of the data.[165] The DOJ does not provide whether states insisted on these sharing agreements or if the DOJ offered them as a matter of course. The Updated Report even claims that anyone who accesses the data would have to have a background check—without providing *any* source for that

PATTERN, clarification about certain risk factors, and information about how certain risk factors were weighted. *Id.*

160 *See Oversight Hearing*, *supra* note 32 (statement of Prof. Melissa Hamilton) (noting that any datasets "have already been anonymized and delivered to the external PATTERN developers. Hence, there is no obvious burden to personnel or resources for the BOP and NIJ to publicly release these datasets.").

161 *See, e.g.*, *id.* (noting that with the public release of the PATTERN dataset, researchers would "be able to calculate a host of measures that are absent in the DOJ Report yet are relevant to a more holistic analysis of the validity, reliability, and equity of the PATTERN tool as it exists"); *Oversight Hearing*, *supra* note 32 (statement of David E. Patton, Executive Director, Federal Defenders of New York) ("Access to the data would allow independent researchers to isolate individual factors and determine which contributed to any disparate impact.").

162 *See infra* Section IV.D.; *see also Oversight Hearing*, *supra* note 32 (statement of Prof. Melissa Hamilton) ("[T]he AUC has serious limitations and thus cannot present a holistic portrait of a tool's abilities.").

163 UPDATED REPORT, *supra* note 79, at 14.

164 *See* Eldar Haber, *Digital Expungement*, 77 MD. L. REV. 337, 351 (2018) (noting that criminal records "are generally considered public information" and that digital data practices make them more publicly available than ever before).

165 *See* UPDATED REPORT, *supra* note 79, at 14 ("The retrieval, disclosure, and redistribution of that criminal history data is prohibited by the sharing agreements used to acquire the underlying data.").

requirement—and that the expense and delay associated with background checks make sharing the data widely "not a feasible solution to increasing transparency with outside stakeholders."[166] The DOJ does claim it will allow an unknown number of researchers an opportunity to apply for the chance to conduct a five-year review of PATTERN's data—subject to "appropriate background investigations, and permission from relevant justice agencies."[167] Ultimately, the DOJ is offering its own version of Northpointe's "trade secret" defense, and it falls flat here as it should have in *Loomis*.[168]

ii. The DOJ Should Release Additional Information About the Development of PATTERN

Even releasing the dataset, reticent as the DOJ is to do so, is not enough to allow for a full and independent evaluation of PATTERN. The DOJ needs to provide some additional information about how PATTERN was developed. The NACDL FOIA request seeks additional information beyond the dataset,[169] and for good reason: this additional information is critical to fully assessing PATTERN. "Full transparency" for algorithmic assessment purposes requires the release of items such as any source code.[170] Further, we know that PATTERN's developers tested and rejected numerous risk-level cutoffs and static and dynamic factors.[171] Decisions about where to set risk-level cutoffs are "highly

166 *Id.* at 15.

167 *Id.*

168 The Wisconsin court's decision to recognize any trade secret protection in *Loomis* has been widely attacked. *See, e.g.*, TAYLOR R. MOORE, CTR. FOR DEMOCRACY & TECH., TRADE SECRETS & ALGORITHMS AS BARRIERS TO SOCIAL JUSTICE (2017) (arguing that trade secret law is not properly calibrated to balance social good in situations like risk assessment and "that a social justice framework should be incorporated into trade secret protection when applied to risk-assessment algorithms").

169 *See supra* note 159.

170 Kroll et al., *supra* note 32, at 641 ("[W]ithout full transparency—including source code, input data, and the full operating environment of the software—even the disclosure of audit logs showing what a program did while it was running provides no guarantee that the disclosed information actually reflects a computer system's behavior.").

171 *See* FIRST STEP ACT REPORT, *supra* note 55, at 51 (noting that PATTERN's developers tested nearly a dozen different cut points for the risk levels before selecting one); *see also* UPDATED REPORT, *supra* note 79, at 11 ("The Department analyzed more than a hundred different iterations and variables in order to find those factors most predictive of the risk of recidivism. The process of weighting the variables was based on scientific research and analysis.").

subjective,"[172] and are matters of policy, not math.[173] The decisions surrounding which factors to test and which to ultimately include in PATTERN are important policy decisions that should be made transparently,[174] and researchers and advocates should be able to probe the decisions. The DOJ needs to share information about the factor and risk-level testing process and why it reached the decisions it made. This is especially important given the troubling decision to lower the risk-level cutoffs in March 2020 with no public acknowledgment that the cutoffs were lowered until June 2020 and *after* investigative reporting by ProPublica. The June Report does not provide any explanation for this decision beyond noting that "changes were made to raw scores to ensure that inmates were accurately placed in the appropriate Risk Level Category and the tool achieved the same high-level of predictive accuracy" that it had previously purportedly attained.[175]

The DOJ also needs to disclose whether machine learning was used to develop the risk classification algorithm. Although the question of whether machine learning was used to develop PATTERN was raised in Congressional testimony, there has so far been no answer from the DOJ, including in the Updated Report it released in January of 2020.[176] If

172 *See* Eaglin, *supra* note 93, at 87 ("Translating tool outcomes into risk categories is a highly subjective, policy- oriented process."). The risk-level cutoff point decisions for PATTERN have been questioned. *See* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Professor Christopher Slobogin) ("[W]hy are the high, medium, and low risk cut-scores in the PATTERN set relative to the population sample rather than the absolute probability of recidivism? The PATTERN's method of setting cut scores means that individuals with a 12% probability of committing a violent crime are considered medium risk (see p. 51 of the report). A probability level that low is, at best, only justifiable if 'violent crime' is defined narrowly.").

173 *See Oversight Hearing*, *supra* note 32 (letter attached to statement of David E. Patton, Executive Director, Federal Defenders of New York) ("Risk assessments are not simply math. Every risk assessment involves moral choices and tradeoffs.").

174 *See* Eaglin, *supra* note 93, at 79 ("Developers make judgment calls about what factors to study in a data set."); *see also* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (written statement of Brandon L. Garrett and Megan T. Stevenson) ("The DOJ Report that announced the development of the PATTERN describes how the experts who developed it made certain key decisions concerning the risk thresholds that separate these groups. However, no information is provided in that Report about how those thresholds were set. Determining how many individuals are rated minimum or low risk, and therefore get the many benefits associated with this designation, is one of the most influential decisions pertaining to the risk assessment tool.").

175 JUNE REPORT, *supra* note 81, at 2.

176 *See Oversight Hearing*, *supra* note 32 (letter attached to statement of David E. Patton, Executive Director, Federal Defenders of New York) ("The DOJ Report provides so few details on weighting, it is unclear what type(s) of models were used (such as regressions) and/or whether any type of machine learning (supervised or unsupervised) was employed. If the former, more information is needed regarding whether and how stepwise procedures were used, data on intercorrelations, and if multicollinearity exists. If the algorithm was developed with any form of machine

machine learning was used to develop PATTERN (and many researchers theorize that most new recidivism prediction tools will take advantage of it),[177] it could be quite difficult for even the tools' developers to explain why the factors were ultimately weighed in the way they were.[178] Ultimately, if PATTERN's developers chose to use the more opaque machine learning process over the more easily explainable process of standard regression analysis in developing PATTERN,[179] that is a decision they should acknowledge, because "the choice to field such a system instead of one which can be interpreted and governed is itself a decision about the system's design."[180] Further, they should be willing to explain why any losses in explainability that occur as a result of using machine learning are offset by gains in areas such as accuracy. In sum, PATTERN's developers need to say much more about how they developed the tool and the policy decisions they codified in it.

B. *Use of Dynamic Factors*

As discussed above, it is a best practice for recidivism prediction tools to use dynamic as well as static factors in classifying someone's risk of reoffending, and DOJ officials touted PATTERN's use of dynamic factors in testimony before Congress.[181] PATTERN's use of dynamic factors is problematic, however, because it undervalues them in the algorithm compared to static factors and because the BOP cannot presently support all of the programs in which an inmate would need to participate to meaningfully reduce his or her score.

learning, this more 'black box' method has different and profound implications on transparency of the developmental procedures.").

177 *See, e.g.*, Huq, *supra* note 28, at 1067–68 ("Not all such tools use machine learning or deep learning. But it is only a question of time before these powerful instruments crowd out simpler models. . . . Even if machine-learning and deep-learning tools are not now omnipresent, they will be soon.").

178 Eaglin, *supra* note 93, at 119 (noting that the use of machine learning in risk tools causes interpretability issues because "the developers creating the tools cannot explain what factors a tool uses to predict recidivism risk").

179 Because a computer may independently update the code algorithm in the machine learning process, it is not enough to provide the original source code. "Online machine learning systems update their decision rules after every query, meaning that any disclosure will be obsolete as soon as it is made." Kroll et al., *supra* note 32, at 660.

180 *Id.* at 637 n.9.

181 *See Oversight Hearing*, *supra* note 32 (statement of Antoinette T. Bacon, Associate Deputy Attorney General) ("Consistent with the Act, PATTERN incorporates dynamic risk factors—things that an inmate can change over time. Such dynamic factors include, among others, an inmate's infractions, beneficial programs, and vocational courses.").

1. Inmates' Theoretical Ability to Reduce Risk Score

In theory, the original PATTERN tool allowed an inmate to reduce his or her recidivism risk rather dramatically through dynamic factors such as job training and drug education. Theoretically, the original PATTERN tool would have allowed a male inmate to reduce his general recidivism score by as many as nineteen points[182] (though there is some dispute about which factors in that tool were truly dynamic and several ways to approach this calculation).[183] To contextualize what this nineteen-point reduction could mean, imagine a male inmate to whom PATTERN assigned an initial general recidivism risk score of fifty. This initial score was due to static factors—factors that are historical and/or outside of the inmate's current control—namely, being age twenty-three (+ thirty points); having been age twenty-two at the time of his first conviction (+ eight points); and having a BRAVO criminal history

[182] According to Table 2 of the First Step Act Report, a hypothetical male inmate could receive the following reductions to his general recidivism score: twelve points for completing eleven or more programs; four points for completing residential drug treatment; two points for taking more than one technical or vocational courses; and one point for receiving drug education while incarcerated. (Table 2 confusingly states that an inmate would receive a two-point deduction if they take zero technical or vocational courses but a zero-point deduction for taking two or more courses. This appears to be an error and is treated as such here. *See* Brennan Center letter, supra note 95 ("We believe that this [increase in risk-level point for completing vocational courses] may be a typographical error"). The point allocation in the Updated Tool makes clear that completing programs reduces an inmate's score. *See* UPDATED REPORT, supra note 79, at 38.) In the original PATTERN tool, female inmates received smaller deductions for those categories but could have additional reductions for participating in employment opportunities and complying with any financial responsibilities. FIRST STEP ACT REPORT, *supra* note 55, at 53–55. Additionally, I do not classify the factor "no need for drug treatment," which deducts six points, as a "dynamic" factor because an inmate's absence of the need for drug treatment is presumably historical by the time they are assessed for PATTERN purposes. But in an attempt to acknowledge all dynamic deductions available under PATTERN, I have included the four-point deduction for an inmate who *did* need drug treatment and who received residential treatment, a more traditionally dynamic feature. *Id.*

[183] I did not classify as dynamic the factors "infraction convictions," "history of violence," "history of escape," or "voluntary surrender," because each is historical as of the time of the initial assessment and can only be changed by an inmate if he or she attempts an escape or has an infraction. The developers of PATTERN, however, have classified "infraction convictions," "history of violence," and "history of escape" as dynamic factors. *See* UPDATED REPORT, *supra* note 79, at 7. Other advocates have pointed out that PATTERN developers are "mistaken" in their classification of certain factors as dynamic. *See* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Sarah Anderson on behalf of FreedomWorks) ("[W]e believe that PATTERN mistakenly treats some factors as 'dynamic' that, although they can reasonably be expected to change while in prison, are not truly dynamic in nature."). Of course, an inmate would also reduce his or her PATTERN score on a subsequent rescoring simply by virtue of aging and being placed in a different band for the "age at time of assessment" factor. Again, this is not a truly "dynamic" factor, and so I do not include it here.

score of six (+ twelve points).[184] This hypothetical inmate's score of fifty would place him in the "high" risk level under PATTERN,[185] and thus preclude him from earning early release credits and other benefits, even though he has never done things that increase a PATTERN score, such as attempting escape or having "infraction violations" while incarcerated. But if he was to earn the maximum deduction of nineteen points through dynamic factors like participation in available programming,[186] his revised score would become thirty-one, thus placing him in the "low" risk level under PATTERN's original risk-level cutoffs, and thereby allowing him to earn early release.[187]

The revised PATTERN tool was purportedly designed to include additional dynamic factors that provide an inmate additional opportunities to earn deductions to his or her score.[188] The opportunity to "earn" reductions is, at least for our hypothetical male inmate, not substantially increased with the revised PATTERN tool, even as the point "penalties" for factors like age at time of assessment and criminal history score are increased relative to the original PATTERN tool.[189] Imagine the same hypothetical male inmate. His new general recidivism score under the revised PATTERN tool will have increased from fifty to fifty-one,[190] even though nothing about him has changed. Just as his initial score has increased by one point, the theoretical reduction he can earn through dynamic factors has likewise increased by one point. He can now theoretically deduct six points for completing residential drug treatment, eight points for completing at least eleven programs, two

184 All of these point totals are drawn from FIRST STEP ACT REPORT, *supra* note 55, at 53–55 tbl.2.

185 PATTERN places the "medium" risk-level cutoff for general recidivism at forty-five; scores above that are labeled "high." *Id*. at 58 tbl.4.

186 See FIRST STEP ACT REPORT, *supra* note 55, at 53–55 for a full breakdown of the math and assumptions that go into the nineteen-point deduction calculation.

187 The PATTERN tool originally places the "low" risk-level cutoff for general recidivism at thirty-three. FIRST STEP ACT REPORT, *supra* note 55, at 58 tbl.4.

188 UPDATED REPORT, *supra* note 79, at 8 (noting that the updated PATTERN tool includes "an additional dynamic measure of offender's 'infraction free' period during his or her current term of incarceration").

189 Under the original PATTERN tool, the points assigned for age at time of assessment ranged from zero for male inmates over age sixty to thirty for inmates aged eighteen to twenty-five. FIRST STEP ACT REPORT, *supra* note 55, at 54. In the revised PATTERN tool, the points assigned for age at time of assessment ranged from zero for inmates over age sixty to thirty-five for inmates under aged eighteen to twenty-five. UPDATED REPORT, *supra* note 79, at 37. Because the tool no longer assigns points for age at first conviction, some recalibration of the other remaining static factors is to be expected.

190 The static factors now give him an initial score as follows: being aged twenty-three (+ thirty-five points); and having a BRAVO criminal history score of six (+ sixteen). UPDATED REPORT, *supra* note 79, at 37–38.

points for participating in work programming, and four points if he earns his high school degree or GED.[191] Thus, his total possible reduction is now twenty points, where it was previously nineteen points. With his new base score of fifty-one and a reduction of twenty points, he has a score of thirty-one, the same as he did under the original PATTERN tool.[192] This is because the new "dynamic factor" included in the revised PATTERN tool—a measure of how long the inmate had been "infraction free"—does not allow inmates to deduct points from their score but instead builds additional points onto the base score.[193] Further, because of the change to the risk-level cutoff scores announced in June of 2020, a score of thirty-one now places our hypothetical inmate in the *medium* risk-level category,[194] ineligible for early release, where before a score of thirty-one placed him in the *low* risk-level category, eligible for early release. So, for an inmate who had long been infraction free, he or she may well be worse off under the revised PATTERN tool. Indeed, the June Report acknowledges that under the revised tool, 4,430 inmates' scores did increase, from low to medium.[195]

2. Inmates' Actual Ability to Reduce Risk Score

The decision to reduce the total points inmates can deduct for participation in certain programming in the revised PATTERN tool[196] is all the more troubling when you consider how unlikely it is that any inmate can actually participate in all of that programming to begin with.

---

191 *See id.* at 38–39. As before, I have not classified the factor "no need for drug treatment," which would now deduct nine points, as a "dynamic" factor, given that an inmate's lack of need for drug treatment is presumably historical by the time he or she is assessed for PATTERN purposes. As noted, inmates who did need drug treatment and received residential treatment can deduct six points under the revised PATTERN tool.

192 The Updated Report does not provide an explanation for why the points deducted or added under PATTERN changed in the updated version, but the risk-level cutoffs remained at the same level as in the original PATTERN tool.

193 UPDATED REPORT, *supra* note 79, at 37. If an inmate has been infraction free for more than twelve months, he has zero points added to his PATTERN score. If he has had an infraction in the past seven to twelve months, he has one point added. If he has had an infraction in the past three to six months, he has two points added. If he has had an infraction in the past three months, he has three points added. There is no opportunity to actually deduct points under this factor.

194 JUNE REPORT, *supra* note 81, at 2.

195 *Id.* at 2–3. The same report notes that there were also thousands of inmates whose scores decreased. Almost no information is given about the demographics of these inmates, however, or why the change to the risk-level cutoff scores was necessary, nor why it was not publicly released for months.

196 For example, under the original tool, inmates could deduct twelve points for completing eleven or more programs. FIRST STEP ACT REPORT, *supra* note 55, at 54. In the revised tool, they can deduct only eight points for participating in eleven or more programs. UPDATED REPORT, *supra* note 79, at 38.

The preceding paragraphs use the terms "in theory," "theoretically," and "hypothetical," and they use them with good reason. Although PATTERN's inclusion of dynamic factors is a step in the right direction, it is at present a hollow gesture in light of current inmates' actual ability (or, more accurately, nonability) to earn deductions. Advocates have criticized the BOP for its lack of educational and rehabilitative programming, and even when the programming exists, inmates face "long wait lists" before they can access it.[197] Indeed, in the NACDL's FOIA request about PATTERN, they seek specific information on BOP waitlists, "including numbers of offenders on each waitlist and time periods remaining on such waitlist by program and location."[198] As David Patton, the Executive Director of the Federal Defenders of New York, expressed to Congress in his testimony, even though the First Step Act "relies heavily on the BOP offering substantially increased programming and productive activities for incarcerated individuals . . . the BOP has failed to provide adequate programming to meet current needs, much less the increased demand that will be required to make the FSA a success."[199] The DOJ chose to develop the risk assessment portion of the risk and needs assessment tool before developing the needs assessment portion. That is a decision that vastly undercuts the tool's ability to properly weigh dynamic factors.

When the DOJ publicly released the information about the PATTERN tool, it acknowledged that it fulfilled the "risk" portion of the risk and needs assessment but had not yet fulfilled the "needs" portion.[200] The DOJ has reported that it will be relying initially on the

---

[197] Rachel Anspach, *How the First Step Act Got People Out of Prison and Back With Their Families*, MIC (Dec. 24, 2019), https://www.mic.com/p/how-the-first-step-act-got-people-out-of-prison-back-with-their-families-19629720 (quoting David Booth, deputy executive director of Black & Pink, an LGBTQ prison abolition organization, as saying, "[t]he BOP historically has long wait lists and inadequate programming . . . I'm afraid [PATTERN is] going to create a lot more problems and hurdles down the road").

[198] *See Oversight Hearing*, *supra* note 32 (statement of Norman Reimer, Executive Director of the National Association of Criminal Defense Lawyers).

[199] *See Oversight Hearing*, *supra* note 32 (statement of David E. Patton, Executive Director, Federal Defenders of New York). Patton went on to note that "[t]he true extent of the deficit is not known because the BOP has not been transparent about the number of programs offered, the capacity of these programs, and the length of the waitlists for these programs." *Id.*

[200] *See* FIRST STEP ACT REPORT, *supra* note 55, at 64 n.3 ("As presently developed, PATTERN does not currently include a needs assessment component. . . . [T]he BOP will continue to use its current needs assessment system, which is being modified and enhanced at the time of this writing. The BOP's needs assessment system will continue to be developed and enhanced over the coming months."); *see also Oversight Hearing*, *supra* note 32 (statement of Antoinette T. Bacon, Associate Deputy Attorney General) ("The prototype needs assessment system is expected to be available for testing by the second quarter of 2020.").

existing needs assessment the BOP implements, but that assessment is inadequate by nearly any measure.[201] For example, according to a 2019 report from the BOP, the wait list for literacy programs is approximately 16,000 people long.[202] When Congress first considered the First Step Act, the BOP and government labor unions expressed concern that facilities were already understaffed and, therefore, lacked adequate programming.[203] The DOJ said that it anticipated having additional needs programming by the second quarter of 2020.[204] The First Step Act, however, required that the DOJ complete the initial risk-level assessment for each inmate in a BOP facility *before* that, by January 2020.[205] Therefore, inmates were assessed by a PATTERN tool that took into account things like the employment and education opportunities they take advantage of, even though many of those inmates were held in facilities with limited to no education or employment opportunities.[206]

---

[201] *See* Julie Samuels, Nancy La Vigne, & Chelsea Thomson, *Next Steps in Federal Corrections Reform: Implementing and Building on the First Step Act*, 32 FED. SENT. REP. 92, 92 (2019) (noting that the BOP has been "characterized by dangerous overcrowding, escalating costs, and insufficient programming and services to prepare people for law abiding lives after release"); *see also* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Laura Mate on behalf of the Sentencing Resource Counsel Project, Federal Public and Community Defenders) ("There is significant evidence that current programs are inadequate to satisfy the congressional mandates in the First Step Act."). The DOJ appears to dispute this conclusion, at least in the First Step Act Report. *See* FIRST STEP ACT REPORT, *supra* note 55, at 61 ("BOP has a strong needs assessment process in place to match inmates with programs to address their criminogenic needs.").

[202] NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Antoine Prince Albert III on behalf of The Leadership Conference on Civil and Human Rights and The Leadership Conference Education Fund) (citing U.S. DEP'T. OF JUSTICE, FEDERAL PRISON SYSTEM FY 2019 PERFORMANCE BUDGET, CONGRESSIONAL SUBMISSION, SALARIES AND EXPENSES 27 (2019), https://www.justice.gov/jmd/page/file/1034421/download). *But see* UPDATED REPORT, *supra* note 79, at 18 (noting that the term "waitlist" may not mean the same thing in BOP reports and that "[t]he vast majority of Bureau programs do not have capacity-related constraints").

[203] *See* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Brandon L. Garrett and Megan T. Stevenson) ("During the consideration of the Act, the American Federation of Government Employees, AFL-CIO, and Council of Prison Locals called for sentencing reforms, but raised the concern as to the use of risk assessment, that federal prisoners are suffering from 'sustained cuts' to staffing levels, which reduces access to rehabilitative programming.").

[204] *Oversight Hearing*, *supra* note 32 (statement of Antoinette T. Bacon, Associate Deputy Attorney General) ("The prototype needs assessment system is expected to be available for testing by the second quarter of 2020."); FIRST STEP ACT REPORT, *supra* note 55, at 78 ("The prototype system addressing [the needs assessment] is expected to be available for testing by the second quarter of 2020.").

[205] FIRST STEP ACT REPORT, *supra* note 55, at 83 ("Importantly, the Department will meet the FSA's requirement to assess all BOP prisoners by January 2020.").

[206] Several commentators have criticized the amount of policymaking that the First Step Act delegated to the BOP. *See, e.g., Oversight Hearing*, *supra* note 32 (statement of David E. Patton, Executive Director, Federal Defenders of New York) ("As a result of the

Further, even once DOJ develops the needs assessment, the BOP only plans on rolling it out for the first year in eight facilities.[207] Only after assessing this rollout and validating any results will the needs assessment be introduced to the remaining 106 BOP facilities, with no definite timetable given for completion of that process.[208] In the Updated Report, the DOJ claims that the BOP has already begun the process of expanding its programming and hiring staff, and notes that the First Step Act calls for $75 million per year in implementation funding.[209] It remains to be seen if the BOP will be willing and able to prioritize that money on recidivism reduction programming.[210]

3. Weight of Static versus Dynamic Factors in PATTERN

Even assuming the BOP could properly fund and implement such programming, PATTERN still disproportionately weighs static features over dynamic ones.[211] In the Updated Report, the DOJ said that it had considered "assigning more weight to dynamic factors" but did not ultimately do so because it tested the updated PATTERN tool "to ensure that individuals could successfully move from a higher to a lower risk score."[212] The DOJ provided several examples in Attachment A to the Updated Report that it claimed demonstrated "how different types of inmates can use the dynamic factors in the updated PATTERN to lower their risk score."[213] What the eight examples in Attachment A actually show, however, is something different.

Act, the BOP will now establish and implement a risk and needs assessment system that will directly determine how long tens of thousands of people serve in prison. If not done wisely, there are countless ways the system will result in unfair, biased, and overly punitive outcomes. With history as a guide, this committee should be very concerned about whether the BOP will rise to the challenge of these new responsibilities.").

207 FIRST STEP ACT REPORT, *supra* note 55, at 79 ("To collect a sufficient amount of data for evaluation purposes, identify issues and trends, and conduct a preliminary assessment of effectiveness, BOP plans to test the system using eight facilities, across four security levels and six regions, for twelve months.").

208 *Id.* ("After testing has been conducted and the results validated, BOP will implement the final system and begin rolling it out to the remaining 114 BOP sites.").

209 UPDATED REPORT, *supra* note 79, at 21–22.

210 The June Report notes that "April 6, 2020, DOJ's Office of Justice Programs announced nearly $70 million in
grants to support successful reentry." JUNE REPORT, *supra* note 81, at 5. If these grants have been awarded as of the date of this Article, that information does not appear to be public, and so this Article is unable to assess their likely impact.

211 *See* Brennan Center letter, *supra* note 95, describing the original PATTERN tool ("PATTERN appears to under-value dynamic factors related to rehabilitation . . . . Taken together, the balance of factors seems to reflect an over-reliance on static factors and an under-valuing of rehabilitative factors.").

212 UPDATED REPORT, *supra* note 79, at 11–12.

213 *Id*.

The eight examples look at a male and female version of four inmate archetypes: "young drug offender taking multiple programs," "'average' released inmate," "aging career criminal," and "white collar offender."[214] For each of the examples, the Updated Report assesses how the inmate's risk score would change under three different scenarios: first, if they were a "model" inmate (one who took advantage of nearly all programmatic and educational opportunities and avoided any infractions while incarcerated); second, if they were a "non-compliant" inmate (one who did not take advantage of programmatic and educational opportunities and who had infractions while incarcerated); and third, if they were a "mixed" inmate (participated in some programming but not all that is offered and had some infractions, though less serious ones).[215]

The point of the examples is obviously to illustrate that dynamic factors have weight and inmates control their destinies under PATTERN—if you are a model inmate, you will have the ability to earn an earlier release. But it really shows that the factor that makes the most difference in reducing most inmates' scores under PATTERN is that they will inevitably age, rather than anything that inmates themselves can control. For example, in the very first example provided in the Updated Report, a twenty-five-year-old male drug offender who is a "model" inmate and takes advantage of every single educational, rehabilitative, or professional program offered can reduce his score from a fifty-seven to a twenty-one, a reduction of thirty-six points.[216] But this reduction is not due solely to dynamic factors. Instead, fourteen points of the thirty-six, nearly forty percent, are awarded *simply because the inmate aged from one bracket into another*.

Indeed, the fourteen-point reduction awarded for "aging" is the largest deduction the inmate earns, dwarfing points deducted for earning a GED (four points), completing drug treatment (six points), completing a work program (two points), avoiding any violent infractions (two points), and completing eleven or more programs (eight points).[217] In order to get the fourteen-point aging reduction, our twenty-five-year-old inmate, who only had a five-year sentence to begin with, would have had to reach the minimum age of twenty-nine and therefore be incarcerated for at least four years of his five-year sentence.[218] Indeed, because the revised PATTERN tool's "age at time of

214 *Id.* at 27–34.

215 *Id.*

216 *Id.* at 27.

217 *Id.*

218 UPDATED REPORT, *supra* note 79, at App. II, p.37.

assessment" factor has age brackets that range up to ten years,[219] many inmates could serve their entire sentences without ever "aging out" of the age bracket they entered PATTERN in and thus never receive any "aging" deductions. Despite this, the fourteen-point deduction for "aging" is the largest in any of the eight examples provided in the Updated Report. Because aging is not a proper "dynamic factor," these examples do not provide evidence for the argument that PATTERN properly weighs dynamic factors.

Including dynamic factors in the PATTERN tool was a good first step. But for this inclusion to be meaningful, PATTERNS's developers should reconsider the weighting of the dynamic factors or, at the very least, should explain and justify them, and much more work needs to be done to give inmates a real chance to access and complete anti-recidivism programming.

C. *Seeking Input from Experts and Stakeholder Groups*

In developing PATTERN, the DOJ and the PATTERN developers sought the input of experts in the field of risk assessment, stakeholder organizations, and the public at large. The DOJ announced a commitment to refining and recalibrating PATTERN over time, both as a result of public feedback as well as the result of the initial classification of BOP inmates. For example, one DOJ official testified before Congress in October 2019 that the Department has continued to consult with the Independent Review Committee (IRC) outside experts "to consider ways to improve PATTERN," noting that this improvement "is an ongoing process, and [the DOJ] looks forward to refining the [risk and needs assessment system] over time."[220] Further, in the First Step Act Report, the DOJ committed to data review and tool reassessment before using PATTERN to assess any individual inmate.[221] The DOJ has also requested bids from outside experts who would assess PATTERN and

[219] *Id.* at 37 (including age ranges between 40–50 and 50–60).

[220] *Oversight Hearing*, *supra* note 32 (statement of Antoinette T. Bacon, Associate Deputy Attorney General).

[221] FIRST STEP ACT REPORT, *supra* note 55, at 70 ("After reviewing the input, we will reassess the tool, make any appropriate changes, and begin the process of assessing each individual inmate.").

offer recommendations.[222] Recalibration is a best practice[223] and therefore a laudable goal. Seeking public input is also an important best practice, and the DOJ deserves to be commended for it. Of course, seeking public input is one thing; embracing it and implementing it is another.

1. Statutory Requirements

As explained above, in developing PATTERN, the DOJ contracted with two researchers with prior experience in risk assessment as well as published scholarship on the topic, Dr. Duwe and Dr. Hamilton.[224] As required by the First Step Act, the NIJ named a six-person IRC that included members with peer-reviewed scholarship about risk prediction and members with prior experience implementing and assessing risk assessment systems.[225] It appears that the IRC took a relatively robust role in the review of PATTERN, as the DOJ claims that the tool developers spent "more than 100 hours addressing questions and concerns raised by the IRC" after PATTERN's initial release.[226] The DOJ claimed that the IRC's recommendations "helped confirm DOJ's confidence in the accuracy of PATTERN,"[227] though it also acknowledged

[222] The June Report notes that "On February 10, 2020, the Department released a competitive funding opportunity to hire outside consultants to review and revalidate PATTERN on an annual basis for up to five years. The deadline for submitting submissions was April 24, 2020. The peer review process is underway." JUNE REPORT, *supra* note 81, at 3.

[223] *See* Kehl et al., *supra* note 30, at 33 ("States should require regular repetition of validity studies and develop procedures to make appropriate alterations based on any changes in the population or new information that emerges about these tools.").

[224] *See Oversight Hearing*, *supra* note 32 (Statement of Antoinette T. Bacon, Associate Deputy Attorney General) ("In developing PATTERN, the Department contracted with two national experts in risk assessment systems—Dr. Grant Duwe, the Director of Research for the Minnesota Department of Corrections, and Dr. Zachary Hamilton, an Associate Professor of Criminal Justice and Criminology and the Director of the Washington State Institute for Criminal Justice.").

[225] The Act was fairly prescriptive about the composition of the IRC, requiring that the NIJ "appoint not fewer than 6 members to the Independent Review Committee," and that "all [IRC members] have expertise in risk and needs assessment systems and shall include (1) 2 individuals who have published peer-reviewed scholarship about risk and needs assessments in both corrections and community settings; (2) 2 corrections practitioners who have developed and implemented a risk assessment tool in a corrections system or in a community supervision setting, including 1 with prior experience working within the Bureau of Prisons; and (3) 1 individual with expertise in assessing risk assessment implementation." First Step Act of 2018, Pub. L. No. 115-391, § 107(c)–(d), 132 Stat. 5194 (2018).

[226] UPDATED REPORT, *supra* note 79, at 2.

[227] *Id.*

that it did not adopt all of the IRC's recommendations for updating the tool.[228]

Per the Act's requirements, the NIJ also selected a "nonpartisan and nonprofit organization with expertise in the study and development of risk and needs assessment tools to host the Independent Review Committee,"[229] the nonpartisan Hudson Institute.[230] DOJ officials testified before Congress that the Hudson Institute "took no institutional position on the First Step Act at any time, and . . . [that the NIJ] select[ed] . . . the most qualified group of criminal justice experts . . . to advise the Attorney General."[231]

2. Listening Sessions and Congressional Testimony

The DOJ also sought feedback from stakeholder organizations as well as the public at large during PATTERN's development.[232] The Department held three listening sessions in April and May of 2019.[233] Those who submitted written and/or oral comments included attorneys representing the Federal Defenders office,[234] members of organizations ranging from the Heritage Foundation to the ACLU,[235] and criminal justice reform organizations.[236]

After the PATTERN tool was announced, the DOJ continued to seek input on the tool, commencing a forty-five-day "public study period" during which "the public [would] be able to review the System and

---

228 *Id.*

229 First Step Act § 107(b).

230 *See Oversight Hearing*, *supra* note 32 (statement of Dr. John P. Walters, Member, First Step Act Independent Review Committee) ("In April 2019, the National Institute of Justice awarded Hudson Institute a contract to serve as host organization for the Independent Review Committee specified by Title I of the First Step Act of 2018."); *see also* FIRST STEP ACT REPORT, *supra* note 55, at 9–10.

231 *Oversight Hearing*, *supra* note 32 (statement of Dr. John P. Walters, Member, First Step Act Independent Review Committee).

232 *Id.* (noting that June of 2019, "the IRC solicited recommendations, concerns, and priorities for further research from nearly eighty-five expert, external organizations and individuals concerned with criminal justice, victims' rights, and law enforcement.").

233 *Oversight Hearing*, *supra* note 32 (statement of Antoinette T. Bacon, Associate Deputy Attorney General) ("The Department also held three listening sessions in April and May of 2019 to allow stakeholders to provide input regarding the [risk and needs assessment system].").

234 FIRST STEP ACT REPORT, *supra* note 55, at 15.

235 *Id.* at 16.

236 *Id.* at 15–16; *see also Oversight Hearing*, *supra* note 32 (statement of Antoinette T. Bacon, Associate Deputy Attorney General) ("The Department received written and in-person statements from 27 individuals representing a variety of communities, including legal experts, law enforcement, criminal justice advocates, academics, victims' rights advocates, and others.").

consider ways in which it may be improved."[237] The DOJ specifically invited and said it would welcome comment on the dynamic factors chosen for inclusion in the PATTERN tool.[238] As part of that public comment period, NIJ hosted two more listening sessions where eight "stakeholder organizations" provided input and feedback and another eight submitted written comments.[239] At the end of that public study period, DOJ officials familiar with the creation of PATTERN testified before the House Oversight Committee, alongside representatives from a variety of organizations and academics.

In seeking input from interested organizations, the DOJ took a step toward one of the advocates' goals: public participation in the development of recidivism prediction tools. One academic has argued that for recidivism risk prediction tools to be more fair, these tools "must reflect the values of the communities where the tools are applied,"[240] and that "[o]nly the communities affected by the tools can voice those values."[241] She concludes that "public notice and comment on normative decisions throughout the development process [of recidivism prediction tools] would resolve many" issues such as fairness, bias, and accountability.[242] In the written statement the ABA submitted as part of NIJ's listening sessions on PATTERN, President Judy Perry Martinez applauded the DOJ's "commitment to public engagement," noting that such engagement "may strengthen public confidence in the assessments that PATTERN renders."[243]

---

237 FIRST STEP ACT REPORT, *supra* note 55, at 70; *see also Oversight Hearing*, *supra* note 32 (Statement of Antoinette T. Bacon, Associate Deputy Attorney General) ("In announcing PATTERN, the Department made clear that it was only the first step in implementing the Act. Indeed, as part of PATTERN's announcement, the Department immediately began the process of considering how to improve it. That process included a 45-day public study period that recently concluded.").

238 *See* FIRST STEP ACT REPORT, *supra* note 55, at 64 n.10 ("During the 45-day public study period discussed in Chapter 4, the Department welcomes input on what changes to these variables may increase predictability.").

239 UPDATED REPORT, *supra* note 79, at 3. The NIJ compiled all of the statements from the September 2019 listening sessions into a report. *See* NIJ LISTENING SESSION STATEMENTS, *supra* note 89. The eight "in-person" comments came from academics and representatives of organizations including the Sentencing Project, the Federal Public and Community Defenders, and the Charles Koch Institute. *Id.* Written statements came from, among others, a victims' rights group, Deloitte Consulting LLP, and the American Bar Association. *Id.*

240 Eaglin, *supra* note 93, at 105.

241 *Id.* at 108.

242 *Id.* at 109.

243 NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (Aug. 16, 2019 letter from Judy Perry Martinez, President of the ABA, to Attorney General William Barr).

Of course, for public participation to be meaningful, the DOJ had to not only solicit the feedback on PATTERN but also take it seriously. In the Updated Report the DOJ issued in January of 2020 announcing certain changes to PATTERN, the Department claimed that it had "taken the feedback [it received] seriously and considered the various and, at times, competing views presented."[244] Notwithstanding these assurances, many stakeholder groups have expressed frustration that the DOJ has not taken their input seriously enough. For example, victims' rights groups expressed frustration that PATTERN reflected a "lack of regard for victims" as well as a failure to listen to the input they had previously provided.[245]

Certain criticisms of PATTERN have made for strange bedfellows. In his written statement to the NIJ after their September 2019 listening session, the General Counsel for the American Conservative Union Foundation remarked upon the diversity of the organizations calling for changes to PATTERN:

> It is worth noting that groups on the right (FreedomWorks and Stand Together), left (ACLU, the Leadership Conference on Civil and Human Rights, and #Cut50), and center (US Justice Action Network and National Association of Criminal Defense Lawyers) voiced similar concerns about the current iteration of PATTERN. In an era of polarized politics, the fact that these voices are unified underscores the compelling need for changes to the risk assessment tool.[246]

But there is also evidence that the PATTERN developers did listen to some of the public comments they received and have already reengineered the tool to reflect that feedback. For example, the Brennan Center made a compelling argument about the lack of emphasis on dynamic factors in the original PATTERN tool by pointing out the relative weights that the tool placed on whether someone voluntarily

---

244 UPDATED REPORT, *supra* note 79, at 1.

245 NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Prof. Mary Graw Leary on behalf of the Victims Advisory Group, United States Sentencing Commission) ("[T]he V.A.G. was disappointed with the lack of regard for victims and a failure to utilize the input previously provided by victim groups who addressed this body in May."); *see also* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Bridgette Stumpf on behalf of Network for Victim Recovery of DC) ("[W]e are gravely disappointed by the absence of any substantive mention or consideration of crime victims and their rights" in the First Step Act Report.). In its Updated Report, the DOJ reaffirmed its commitment to victims' rights. *See* UPDATED REPORT, *supra* note 79, at 17 ("DOJ will continue to engage and communicate with organizations that represent crime victims as it works to fully implement PATTERN, the forthcoming needs assessment, and other provisions of the FSA.").

246 NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of David H. Safavian on behalf of American Conservative Union Foundation).

surrendered versus on someone's participation in recidivism reduction programming:

> Tellingly, for men, whether the person surrendered to federal custody, a static factor, counts for 12 points—as much as completing more than 10 recidivism reduction programs. Postsentencing voluntary surrenders, the only type that PATTERN appears to score, are rare, occurring in only 25 percent of the diagnostic sample. Worse, our understanding is that such surrenders occur most frequently in cases presenting special circumstances or involving affluent defendants. PATTERN should value rehabilitation above access and privilege.[247]

In January of 2020, the DOJ announced that the PATTERN tool would no longer consider whether an inmate had voluntarily surrendered because it "removed or changed certain measures that might be associated with bias, especially racial bias, in order to implement the most fair and predictive tool possible."[248] The DOJ also removed the factor that considered an inmate's age at their first arrest/conviction for the same reason.[249]

These changes, and the announcement that they are made as a direct result of feedback, would be more welcome if PATTERN had not been simultaneously changed, without explanation, to make it even harder for certain inmates to initially score in the low or minimum risk level.[250] Further, there is a long list of issues raised and recommendations made by advocates after the initial release of the tool that the Updated Report dismisses. For example, despite receiving multiple requests to release the dataset they used to develop PATTERN and to make public additional information about the tool's development, the Updated Report declines to do so, instead saying they will give

---

[247] Brennan Center letter, *supra* note 95, at 6. The Brennan Center further pointed out that in the original PATTERN tool, "[p]articipation in prison education programs, which is proven to reduce recidivism, is barely scored, and not scored at all for men." *Id.* The updated PATTERN tool does allow men to earn point deductions for educational attainment, though those deductions remain smaller than the deductions available for female inmates. *See* UPDATED REPORT, *supra* note 79, at 39.

[248] UPDATED REPORT, *supra* note 79, at 9.

[249] *Id. But see supra* note 189, explaining that PATTERN was revised to increase the points assigned for other static factors and to decrease the points that could be earned as deductions.

[250] As explained above, the Updated Report removed voluntary surrender and age at first arrest as factors at the same time that the tool was recalibrated to increase the total points for static features like age at time of assessment and decrease the point deductions for dynamic factors like participation in prison programming. Further, the risk-level cutoff points were lowered without any meaningful explanation.

certain researchers a limited opportunity to review the data.[251] Likewise, despite several calls to narrow the broad definition of recidivism they adopted, the DOJ defended its original definition.[252] Finally, despite numerous criticism of the use of AUC to assess accuracy, much of which is discussed below, the DOJ does not even address that issue in its Updated Report.

D. *Issues with AUC and an Overall Focus on "Accuracy"*

In developing PATTERN, the DOJ has seemed especially preoccupied with establishing that the tool is "accurate" and has favorable predictive reliability.[253] The PATTERN developers repeatedly tout the original tool's AUC score in the First Step Act Report, and their willingness to recalibrate the tool seems to be constrained in part by a fear of making any changes that result in too great a reduction to the AUC score. For example, in the Updated Report, the DOJ announced that removing certain factors from PATTERN "reduce[d] PATTERN's predictive accuracy by approximately one percent," but that it "viewed this decrease as acceptable, if it prevents the actual or perceived perpetuation of any bias."[254] Further, the only explanation given for the decision to lower the risk-level cutoff levels was that doing so "ensure[d] that inmates were accurately placed in the appropriate Risk Level Category and the tool achieved the same high-level of predictive accuracy."[255]

1. Accuracy Is Not a Straightforward Goal

Although this hyper-focus on accuracy is understandable,[256] it is problematic for several reasons, especially given the conflation in the PATTERN reports of accuracy as meaning a higher AUC score. First, accuracy cannot be the only goal of a recidivism prediction tool, and "may need to yield to other important goals, such as differential validity, group fairness, and individual rights."[257] Further, accuracy is not a

251 *See supra* notes 157, 160 and 161 and accompanying text.

252 *See* UPDATED REPORT, *supra* note 79, at 13; *see also supra* note 96.

253 "In crafting the System, the Department worked to make the benefits of the FSA as widely available as possible without compromising predictive reliability." FIRST STEP ACT REPORT, *supra* note 55 (Letter from William Barr, Attorney General).

254 UPDATED REPORT, *supra* note 79, at 9.

255 June Report, *supra* note 81, at 2.

256 "Risk tool developers have a natural incentive to focus on overall accuracy." *Oversight Hearing*, *supra* note 32 (letter attached to statement of David E. Patton, Executive Director, Federal Defenders of New York).

257 *Id.* Mr. Patton goes on to note that "[s]electing the right tradeoff between these sometimes competing goals are more rightly within the power of policymakers and stakeholders."

simple target, and developers make important decisions in even defining what accuracy means. For example, is it more important that PATTERN have high specificity (which measures the algorithm's accuracy among the true negatives—people "who are not ultimately rearrested") or that it have high sensitivity (which "measures the algorithm's accuracy among the true positives—people who are ultimately rearrested")? Put another way, do we care more about avoiding false positives (falsely classifying an inmate as likely to reoffend) or false negatives (falsely classifying an inmate as unlikely to reoffend)?[258] Algorithms like PATTERN will have to contend with Blackstone's ratio—"it is better that ten guilty persons escape than that one innocent suffer"[259]—just as judges throughout time have. Algorithms can be designed to achieve one definition of accuracy or parity, but not all of them at the same time.[260]

2. AUC Is a Weak Proxy for Accuracy

Even if accuracy were a simple or straightforward goal, using AUC to assess it is not without risk and criticism. As noted above, AUC for PATTERN is best thought of as a measure of how often someone who reoffended was placed in a higher risk category than someone who did not. Although that information is valuable, it is not a complete picture of "accuracy" by any means.[261] For example, an AUC score cannot tell us if the chosen risk-level cutoffs are well calibrated because "[p]redictive accuracy measures like the AUC value provide no insight into whether the cut-off points located between high, medium, and low recidivism risk categories accurately calibrate with actual outcomes in the real world."[262] In other words, while the AUC can tell us something about how well PATTERN distinguishes recidivists from non-recidivists, it

258 For a more fulsome discussion of specificity versus sensitivity, *see* Mayson, *supra* note 99, at 2243–45.

259 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 1743 (William Draper Lewis ed., Rees Welsh & Co. 1902) (1765).

260 Mayson, *supra* note 99, at 2248 ("An algorithm can be designed to achieve any one of the above metrics of output equality, but not *all* of them together. That is, an algorithm cannot be designed to achieve 'total fairness.'"); *see also* Richard Berk et al., *Fairness in Criminal Justice Risk Assessments: The State of the Art* 1 (Univ. Pa. Working Paper No. 2017-1.0, 2017) (arguing that "there are at least six kinds of fairness, some of which are incompatible with one another and with accuracy").

261 *See, e.g.*, *Oversight Hearing*, *supra* note 32, at 4 (statement of Prof. Melissa Hamilton) ("[T]he AUC has serious limitations and thus cannot present a holistic portrait of a tool's abilities.").

262 Eaglin, *supra* note 93, at 91; *see also* Hopkinson, *supra* note 128, at 742 ("[I]n statistics, there is no consensus as to which AUC scores represent small, moderate, or large effect sizes.").

cannot tell us very much at all about how accurately the tool will be able to predict future recidivism.[263] An AUC score cannot even tell us whether the majority of errors PATTERN makes are false positives or false negatives.[264]

With all of the potential problems involved in using AUC as a proxy for accuracy, it is not surprising that even Dr. Hamilton and Dr. Duwe, the outside consultants hired to develop PATTERN, have themselves written about AUC's limitations in their prior publications.[265] This makes the decision to declare the tool "validated" by using AUC values all the more puzzling.[266] Because AUC can be so confusing, many laypeople (and even some researchers) incorrectly conclude that AUC is a measure of how well a tool predicts recidivism, and thus that an AUC score of .8 means that a tool is right 80% of the time about whether or not someone reoffends. It does not mean anything remotely close to that. Rather, as explained above, it means that 80% of the time, the tool will classify someone who does end up reoffending as being in a higher risk category than someone who does not.

3. Racial Disparities in PATTERN

The DOJ does not just use AUC to support its argument that PATTERN is "validated"; it also uses it to counter arguments that PATTERN is racially biased. As noted above, PATTERN's developers used the calculated AUC scores to bolster the conclusion that there is "minimal" racial or ethnic disparity with respect to PATTERN's predictive value,[267] and that "the PATTERN instrument's predictive performance is unbiased across racial and ethnic classifications."[268] Several advocates, however, have challenged that conclusion. In a letter to the NIJ on behalf of the Brennan Center, Ames Grawert pointed out that

---

263 For more on the issue of discrimination (distinguishing those who reoffend from those who do not) versus calibration ("how accurate the tool statistically estimates recidivism"), and why AUC is a better measure of discrimination, see *Oversight Hearing*, *supra* note 32, at 4 (statement of Prof. Melissa Hamilton).

264 *Id.*

265 *See id.* at 5 (citing to previous studies published by Drs. Duwe and Hamilton that noted that AUC validation is "rather easily achieved" and not something that should be given great weight or solely relied upon to assess accuracy).

266 FIRST STEP ACT REPORT, *supra* note 55, at 84 ("PATTERN was developed using a seven-year dataset of BOP releases, and it was validated as an effective predictor of recidivism over the inmates' subsequent three-year period in the community.").

267 *Id.* at 60 ("Overall, these findings indicate strong and comparable prediction strength for PATTERN models across all race/ethnicity categories, suggesting minimal racial/ethnic disparity for PATTERN's prediction strength.").

268 *Id.* at 63.

> PATTERN identified more than half of all Black men (53 percent) in the diagnostic sample as having a high risk of recidivism, compared to 29 percent of white men. Indeed, the plurality of white men (30 percent) were classified as "minimum risk"; just 7 percent of Black men received the same classification.[269]

Racial disparities like this are not necessarily surprising in light of the fact that PATTERN was developed from a BOP dataset, and therefore any historical racial disparities in our criminal justice system. For instance, issues like Black defendants being given longer sentences for the same underlying crime than white defendants[270] or Black people having higher arrest rates for drug use than white people—despite having similar underlying drug usage rates[271]—are codified in the data.[272] Even if PATTERN was designed to be racially neutral, the fact that it relies on data that is not racially neutral means it will produce racially disproportionate results.

After the original release of the PATTERN tool, advocates suggested ameliorating this racial bias by eliminating or reducing PATTERN's reliance on an inmate's criminal history in calculating a risk score.[273] In response to these criticisms and suggestions, the Updated

---

269 Brennan Center letter, *supra* note 95.

270 *See, e.g.*, UNITED STATES SENTENCING COMMISSION, DEMOGRAPHIC DIFFERENCES IN SENTENCING:
AN UPDATE TO THE 2012 BOOKER REPORT (Nov. 2017) (concluding that Black men received sentences on average 19.1 percent longer than similarly situated white men during the four-year period that was studied).

271 Brennan Center letter, *supra* note 95 ("Due to historical discrimination and enforcement patterns, Black men and women may have longer criminal records than their white counterparts despite similar offending patterns. Blacks are disproportionately arrested for drug offenses, for example, despite using drugs at rates similar to whites."); *see also* NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Antoine Prince Albert III on behalf of The Leadership Conference on Civil and Human Rights) ("Fourth, NIJ's claim that PATTERN's 'predictive performance is unbiased across racial and ethnic classifications' ignores historical and enduring patterns of racial bias and discrimination that infect the data upon which PATTERN relies.").

272 *See, e.g.*, Mayson, *supra* note 99, at 2251 ("Any form of prediction that relies on data about the past will produce racial disparity if the past data shows the event that we aspire to predict—the target variable—occurring with unequal frequency across racial groups.").

273 *See, e.g.*, Brennan Center letter, *supra* note 95 ("PATTERN could be designed to incorporate criminal history in a novel way, by discounting the impact of drug convictions. Better yet, it could exclude arrest records entirely when 'scoring' someone's criminal history, for the reasons stated in the previous section, and focus solely on a person's history of conviction and incarceration."); *Oversight Hearing*, *supra* note 32 (statement of Prof. Melissa Hamilton) (suggesting that researchers use the PATTERN dataset to determine if the criminal history factor accounts for most of the racial disparity, and if so, that "modifying criminal history in risk-sensitive ways may improve the tool and its equitable outcomes.").

Report points once again to PATTERN's AUC scores and claims that these establish that PATTERN is a "neutral assessment tool."[274] The DOJ further says, without explaining why, that "[t]he risk assessment tool cannot correct for any outside biases that lead to higher recidivism."[275]

4. More Accurate Than What?

A more proper way of assessing the accuracy and fairness of PATTERN may not be to ask, "Is it fair?" but to ask, "Is it fair as opposed to the alternatives?" Is PATTERN better at predicting recidivism, and doing so in a fair way, than are judges or probation officers? There is some evidence that algorithmic recidivism prediction tools do outperform the "unaided clinical judgment" of people, even trained professionals.[276] Some academics have argued that even where a recidivism prediction tool like PATTERN may produce biased results, "the default alternative—subjective risk assessment—is very likely to be worse."[277] Nonetheless, if PATTERN can be improved in ways that minimize racial bias without reducing its accuracy, as many commentators have suggested it can,[278] the DOJ must consider those improvements.

E. *Unclear Opportunities to Appeal/Challenge*

It is not entirely clear at this point where the information the PATTERN tool uses to assess each new inmate will come from, but the most likely source seems to be the Presentence Investigation Report ("PSR") that the United States Probation Office completes for each federal criminal defendant who has pleaded or been found guilty.[279]

274 UPDATED REPORT, *supra* note 79, at 9.

275 *Id.*

276 *See, e.g.*, Goel et al., *supra* note 99, at 2–4 (noting that algorithms tend to outperform human judgment, even criminal justice professionals' judgment, when it comes to predicting recidivism); Kehl et al., *supra* note 30, at 11–12 (discussing how many academics "argue that actuarial assessment, which is at work in risk-assessment algorithms, is preferable to clinical assessment," and that "studies have generally credited greater accuracy and predictive validity to the objectivity of actuarial tools compared to the theoretical nature professional clinical judgment").

277 Mayson, *supra* note 99, at 2277–78 n.208–09.

278 *See supra* note 273.

279 The First Step Act Report explains how a PSR is created. *See* FIRST STEP ACT REPORT, *supra* note 55, at 91 n.15 ("Prior to sentencing, the probation officer will interview the defendant and conduct an investigation to provide the sentencing judge with pertinent information relevant for sentencing. This information is captured in a Presentence Investigation Report and includes details of the defendant's family history, community ties, education background, employment history and physical and mental health."). Other portions of that Report suggest that the PSR is where much of the PATTERN data would come from. *See id.* at 75 ("BOP has held preliminary discussions with the United

Although this may be an obvious choice for populating data for PATTERN, it is not one without problems. First, if the information in the PSR is incorrect, the inmate will face "great difficulty" trying to have it corrected.[280] This is because "[t]he sentencing court does not have jurisdiction to correct the PSR after sentencing, thus creating a 'jurisdictional obstacle' for the inmate."[281] Of course, an inmate might not even know that the information in his or her PSR was incorrect, as some parts may be considered "confidential" and kept from the defendant and his or her counsel.[282] Finally, sentencing information like the PSR is admitted into evidence under the lower "preponderance of the evidence" standard, and so the PSR can include hearsay.[283]

Wherever the PATTERN information comes from, it appears that an inmate has only one opportunity to challenge or "appeal" if the risk score the algorithm assigns to him or her places them outside of the "minimal" or "low" range: through a warden override.[284] There are several reasons a warden may decline to provide such an override. First,

States Probation Office about supplementing the Presentence Investigation Report with more detailed education information about learning needs."). Deloitte Consulting LLP, which has done previous consulting work for the DOJ, recommended at one of the listening sessions that prison staff "use information in the presentence report to populate data in PATTERN and the needs assessment." NIJ LISTENING SESSION STATEMENTS, *supra* note 89 (statement of Deloitte Consulting LLP).

280 Gregory W. Carman & Tamar Harutunian, *Fairness at the Time of Sentencing: The Accuracy of the*
Presentence Report, 78 ST. JOHN'S L. REV. 1, 7 (2004).

281 *Id.*

282 Kehl et al., *supra* note 30, at 15 ("Although the information in the PSI is generally made available to the defendant or his counsel as well, certain information or parts of the report the report may be considered confidential and kept from the defendant.").

283 Noel L. Hillman, *The Use of Artificial Intelligence in Gauging the Risk of Recidivism*, 58 JUDGES J. 36, 37 (2019) (noting that "while the process [of creating the PSR] is largely transparent with input from the defendant and counsel, the officer obtains the information through ex parte interviews of third parties and a review of documents often containing multiple layers of hearsay").

284 *See* First Step Act of 2018, Pub. L. No. 115-391, § 102(g)(1)(D), 132 Stat. 5194 (2018), which provides that an inmate, even one who has not been labeled low or minimal risk by PATTERN, may petition the warden "to be transferred to prerelease custody or supervised release" and that the warden may approve such a petition "after the warden's determination that—the prisoner would not be a danger to society if transferred to prerelease custody or supervised release; the prisoner has made a good faith effort to lower their recidivism risk through participation in recidivism reduction programs or productive activities; and the prisoner is unlikely to recidivate[.]" Advocates have urged the DOJ and BOP to consider having another, independent appeals process for inmates to challenge their PATTERN score. *See, e.g.*, *Oversight Hearing*, *supra* note 32 (statement of Prof. Melissa Hamilton) ("A stand-alone mechanism for disputing risk scores must be established. The current plan appears to be to simply apply the current prisoner grievance system. This is insufficient and inapplicable. Algorithmic risk assessment practices require their own processes to challenge.").

the "anchoring effect" of knowing that a seemingly neutral and objective algorithm has declared an inmate at a higher likelihood of recidivism may give many wardens pause.[285] Second, any individual warden may be hesitant to approve an early release for fear of negative publicity if the inmate does reoffend. Finally, wardens may introduce their own biases into the system through such an override[286]—biases that algorithmic systems are theoretically designed to protect against.

V. CONCLUSION

The First Step Act is just that—a first step, one which is necessary but not sufficient. The Act carries with it the opportunity for truly meaningful criminal justice reform, and it has made some exciting progress. The PATTERN tool includes certain best practices in recidivism prediction, and its developers have made a good faith effort to engage advocates and scholars about the tool's development. But much remains to be done if PATTERN is to truly represent an advance in algorithmic recidivism prediction. The DOJ must release more information about PATTERN and its underlying datasets so that scholars, advocates, and community members can truly assess the tool and offer responses for improving it.

[285] For more on "anchoring effect" research with respect to recidivism prediction tools, see Liu et al., *supra* note 32, at 130 (noting that there is a "psychological 'anchoring effect' for courts using scientific and technological tools," and that "numerous studies have demonstrated how judges (and human individuals) are submissive to computer-generated numbers and results that may further frame and condition the view of judges"); *see also* Cyphert, *supra* note 31 ("Laypeople without technical expertise can be especially vulnerable to placing too much faith in algorithmic outcomes.").

[286] Goel et al., *supra* note 99, at 15 ("[A]s illustrated by Part I's discussion of how professional 'overrides' of actuarial estimates can backfire, it could well reintroduce the bias that instruments are designed to prevent.").