SIX THINGS TO KNOW ABOUT ALGORITHM-BASED DECISION-MAKING TOOLS

1 ALGORITHMIC DECISIONMAKING TOOLS ARE NOT ALWAYS TRANSPARENT – BOTH BECAUSE OF HOW SOME SYSTEMS WORK, AND BECAUSE SOME COMPANIES AREN'T WILLING TO SHOW US HOW THEY WORK.

While some decision-making algorithms used in criminal justice decision-making show the factors they use, and how they are weighted, many tools don't reveal the data that trains them to the public.

Some algorithmic decision-making tools are referred to as a "black box" because developers are unwilling to reveal how the tools are designed or developed due to trade secrecy and other proprietary concerns about the software, or because the ways in which the algorithms make their predictions are inscrutable.[i]

One researcher submitted Freedom of Information Act (FOIA) requests to all 50 states requesting data, source code, validation assessments, etc. related to algorithms used in criminal justice (encompassing parole, probation, bail and pretrial, and sentencing).

Several states refused to provide information because they claimed the algorithms are embedded in software and therefore not subject to FOIA requests, some refused because of contractual obligations, and nine states refused to disclose any information because they said the tools are privately owned.[ii] We should not make significant decisions based on tools that are not designed, calibrated, audited for results, or reviewed by independent parties and the public writ large.[iii]

2 ALGORITHMIC DECISIONMAKING TOOLS DON'T ACTUALLY HELP INFORM BAIL DECISIONS, NOR TELL US WHAT WE WANT TO KNOW.

When making decisions about pretrial release, communities most worry about two things – intentional flight and community safety. But those outcomes are not what most pretrial risk assessment tools generally predict. Instead, pretrial risk assessment tools claim to assess two outcomes: failure to appear and rearrest. Critically, these outcomes are not the same.[i] While rearrest for a violent crime might signal danger to an individual or community, rearrest writ large does not. Similarly, non-appearance is frequently the result of forgetfulness, of childcare issues, of transportation issues, or of misunderstandings, *not* a purposeful desire to flee. Worse, most tools measure *the composite risk* of generalized "pretrial failure" – that is, the likelihood someone will be rearrested or fail to appear in court.[ii] Consequently, "risk scores" from algorithmic risk assessment tools have little bearing on the risks that really matter to communities.[iii]

3 ALGORITHMIC DECISIONMAKING TOOLS ARE NOT VERY ACCURATE.

Multiple studies have examined the accuracy of algorithmic decision-making tools, particularly in the criminal justice context (where they are often referred to as risk assessment tools), and they do not always fare very well.

One study looked at the COMPAS risk assessment tool and found that it was actually less accurate at predicting recidivism than random people given the same data – 65 percent compared to 67 percent, respectively.[i] This discovery came after a ProPublica study in 2016 found that of the people COMPAS flagged as likely to reoffend, only 61 percent actually did so.[ii] Furthermore, algorithmic tools give the statistical probability that something will occur *at this precise moment in time*; the tools must be calibrated frequently as populations and behaviors change. Many risk assessment tools have been implemented without validation and without ensuring that jurisdictions have the resources necessary for the upkeep.[iii]

4 ALGORITHMiC DECISIONMAKING TOOLS CAN PERPETUATE RACIAL BIAS AND DISCRIMINATION.

Algorithmic tools are only as good as the data they analyze and the developers who design them. They are not based on "neutral" data.[i] Even when programmers strive for objectivity and neutrality, these tools reflect the biases of the data upon which they are developed. If inequality is embedded in the data, the computations flowing from that data will reflect and reinforce those inequalities (this is also known as "garbage in, garbage out").[ii] For example, New York City recently announced it will develop a new risk assessment tool based on data from 2009 to 2015.[iii] Of course, the city's arrest practices during part of that period have been held as unconstitutional, meaning that data should not be a reliable legal or statistical guide as to future behavior.[iv]

ProPublica's investigation into n also found that Black defendants were 77 percent more likely to be incorrectly flagged as likely to commit a future violent crime and 45 percent more likely to be incorrectly flagged as committing any type of future offense. Furthermore, Black people were twice as likely to be incorrectly flagged for recidivism as their White counterparts.[v]

5 ALGORITHMIC DECISIONMAKING TOOLS CAN GIVE THE APPEARANCE OF OBJECTIVITY, BUT THEIR RESULTS ARE STILL SUBJECTIVELY INTERPRETED BY DECISION-MAKERS.

In the criminal legal context, it is difficult to determine what encompasses "risk," and it is even more difficult to quantify it with any degree of certainty because of its ill-defined contours. Labeling someone as "high" risk involves an inherent value judgment – one based on moral and political determinations, not scientific ones – under the guise of purportedly neutral data.

A risk assessment score doesn't mean anything on its own, nor can an algorithm determine what level of risk warrants detention or release. Instead, the scores are often translated into proposed courses of action in decision-making frameworks, and then interpreted by judges, bail magistrates, or other pretrial officers who often have discretion in how they can use that information.[i]

To that end, there have been reports of jurisdictions in California manipulating the risk categorizations up or down depending on whether too many or too few people are being released,[ii] making any objectivity illusory. Additionally, one study found that more often than not, judges in Kentucky disregarded the action-directive associated with risk assessment scores in favor of their own judgment.[iii]

6

ALGORITHMIC DECISIONMAKING TOOLS MEASURE GROUP RISK, NOT INDIVIDUALIZED RISK.

Risk assessment tools calculate the statistical likelihood that an event will occur in an individual's life by examining data points gathered from the lives of groups of people. Risk tools do not factor in circumstances relevant to a particular case or individual, but rather lump people into broad categories.

For example, if someone who is facing a bail determination by a judge lives in a highly policed neighborhood that yields high arrest rates, the likelihood of them being arrested, whether warranted or not, could be much higher than someone who lives in a community with a lower police presence. An algorithm that weighs that fact – where someone lives – as a component of the risk calculation will naturally flag a person who lives in a community with a high arrest rate as a high risk, even if that individual has not been arrested themselves. The presumption that someone presents a high risk of arrest by virtue of where they live flies in the face of the presumption of innocence and the individualized determination to which all people in the criminal legal system are entitled.

---

[1] *See, e.g.* Robert Brauneis, & Ellen P. Goodman, *Algorithmic Transparency for the Smart City* 20 Yale J.L. & Tech. 103 (2018); Trey Popp, *Black Box Justice*, Pa. Gazette, Aug. 28, 2017 at 38, http://thepenngazette.com/black-box-justice.

[1] *See* Michael Kassner, *Decision-Making Algorithms: Is Anyone Making Sure They're Right?*, TechRepublic (July 18, 2016, 1:53 PM), https://www.techrepublic.com/article/decision-making-algorithms-is-anyone-making-sure-theyre-right/.

[1] *See, e.g.*, Assoc. for Computing Mach. U.S. Pub. Policy Council, Statement on Algorithmic Transparency and Accountability, (2017).

[1] *See, e.g.*, John Logan Koepke, & David G. Robinson, *Danger Ahead: Risk Assessment and the Future of Bail Reform* Wash. L. Rev. (forthcoming 2018).

[1] *See, e.g.*, Lauryn Gouldin, *Disentangling Flight Risk from Dangerousness*, 2016 BYU L. Rev. 837; Sandra Mayson, *Dangerous Defendants* 127 Yale L.J. 490 (2018).

[1] *See, e.g.,* Lauryn Gouldin, *Defining Flight Risk* 85 U. Chi. L. Rev. 677 (2018); Gouldin, *supra* note 5; Shima Baradaran Baughman & Frank McIntyre, *Predicting Violence*, 90 Tex. L. Rev. 497 (2012); Megan Stevenson & Sandra G. Mayson, *Pretrial Detention and Bail*, *in* 3 Reforming Criminal Justice (Erik Luna ed., 2017); Jessica Eaglin, *Constructing Recidivism Risk* 67 Emory L.J. 59 (2017).

[1] Rowan Walrath, *Software Used to Make 'Life-Altering' Decisions Is No Better Than Random People at Predicting Recidivism*, Mother Jones (Jan. 17, 2018, 3:15 PM) https://www.motherjones.com/crime-justice/2018/01/compas-software-racial-bias-inaccurate-predicting-recidivism/.

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

[1] *See* Cecelia Klingele, *The Promises and Perils of Evidence-Based Corrections*, 91 Notre Dame L. Rev. 537 (2015).

[1] *See Terms of Use*, Public Safety Assessment, https://psapretrial.org/terms/non-commercial (last visited July 24, 2018) ("The PSA uses neutral, reliable data to produce two risk scores.")

[1] *See* Lee Rainie & Janna Anderson, *Code-Dependent: Pros and Cons of the Algorithm Age* Pew Res. Ctr. (Feb. 8, 2017) http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/.

[1] Remarks at Measuring Justice: Redesigning New York City's Pretrial Risk Assessment and Recommendation System (Sep. 18, 2017).

[1] *See, e.g.*, Floyd v. City of New York, 959 F. Supp. 2d 540 (S.D.N.Y. 2013).

[1] Angwin et. al., *supra* note 8.

[1] *See, e.g.*, Mayson, *supra* note 5; Koepke & Robinson, *supra* note 4.

[1] *See* Hum. Rts. Watch, Not in it for Justice: How California's Pretrial Detention and Bail System Unfairly Punishes Poor People 8 (2017), https://www.hrw.org/report/2017/04/11/not-it-justice/how-californias-pretrial-detention-and-bail-system-unfairly ("For example, in Santa Cruz County, the tool was adjusted to double the number of people released under conditions of supervision.").

[1] *See* Megan Stevenson, *Assessing Risk Assessment in Action*, 103 Minn. L. Rev., (forthcoming 2018).

SIGN ONTO THIS STATEMENT OF CONCERN HERE
FOR QUESTIONS OR COMMENTS, CONTACT pretrialjustice@civilrights.org