

Building Community Governance Over Risk Assessment

Hannah Sassaman, Reuben Jones, David Robinson

We are here today to ask for your help.

- **Dozens of criminal justice jurisdictions are considering or have instituted major changes to their bail systems in the US.**
- In many of them, **the opportunity to end cash bail is tied to putting risk assessment algorithms into pretrial decisions** - in between people and their freedom.
- This isn't about the ideal or theoretical role of data scientists: It's about what happens **in practice**, and how you can connect with that.

We are here today to ask for your help (continued).

- In our organizing, we've met hundreds of people with major concerns about how these algorithms will enshrine bias into their predictions, and **who are struggling for power over how court systems will use those predictions.**
- Risk assessment and bail reform are not always decarceral - not always unwinding deep disparities that have caused mass incarceration.
- Our organizing, and much of the work of this community has **changed the conversation** on how communities can engage with algorithms in human decision-making systems.
- **Will you help our communities hold algorithms accountable to us, and to a vision of decarceration and ending racial disparities?**

Goal for today is to begin building community governance of predictive algorithms in pretrial

Urgent need for community governance to happen now in the field. If we don't bring governance to bear on these tools now, risk of embedding larger bias into pretrial decision-making. **We hope to leave today with:**

- A **working definition of community governance over risk assessment** in partnership with data scientists, with good arguments for why it is necessary and possible, and a sense of some of the traps and concerns
- To **start to build a community of practice** from those in the room who want to help build community governance over risk assessment in real time
- Today will **be a participatory exercise** in talking about what that means
- After some **initial level-setting** and sharing examples

Initial level-setting

- Where are we in the **social, legal, and scientific debate** on risk assessments in pretrial decision-making?
- And what about this moment in time brought a **broad coalition of advocates together 18 months ago to develop a major Statement of Concern** about them?

THE USE OF PRETRIAL "RISK ASSESSMENT" INSTRUMENTS:

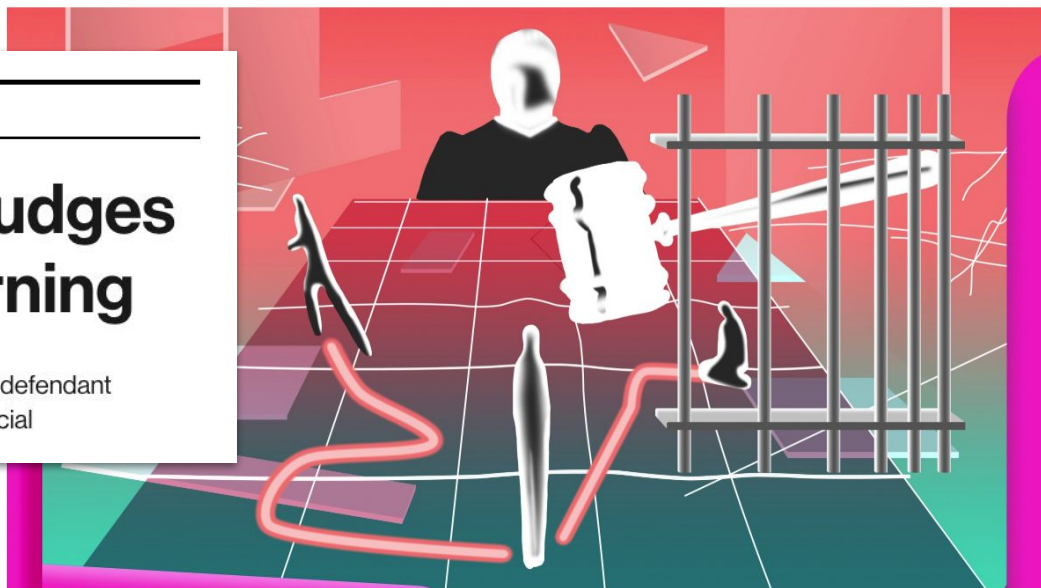
A SHARED STATEMENT OF
CIVIL RIGHTS CONCERNS

Level setting

Intelligent Machines

How to Upgrade Judges with Machine Learning

Software that helps judges decide whether to jail a defendant while they await trial could cut crime and reduce racial disparities amongst prisoners.



Level setting

RELATIONSHIP BETWEEN RISK FACTORS AND PRETRIAL OUTCOMES

Risk Factor	FTA	NCA	NVCA
1. Age at current arrest		X	
2. Current violent offense			X
<i>Current violent offense & 20 years old or younger</i>			X
3. Pending charge at the time of the offense	X	X	X
4. Prior misdemeanor conviction		X	
5. Prior felony conviction		X	
<i>Prior conviction (misdemeanor or felony)</i>	X		X
6. Prior violent conviction		X	X
7. Prior failure to appear in the past two years	X	X	
8. Prior failure to appear older than two years	X		
9. Prior sentence to incarceration		X	

Note: Boxes where an “X” occurs indicate that the presence of a risk factor increases the likelihood of that outcome for a given defendant.

“Frameworks” — political and moral judgment

	NCA 1	NCA 2	NCA 3	NCA 4	NCA 5	NCA 6
FTA 1	ROR 11.8% of population	ROR 7.7% of population				
FTA 2	ROR 0.5% of population	ROR 6.9% of population	PML 1 11.7% of population	PML 2 6.4% of population	PML 3 0.1% of population	
FTA 3		PML 1 2.0% of population	PML 1 8.8% of population	PML 2 6.0% of population	PML 3 2.9% of population	Release Not Recommended 0.5% of population
FTA 4		PML 1 0.6% of population	PML 1 1.5% of population	PML 2 3.5% of population	PML 3 4.8% of population	Release Not Recommended 0.9% of population
FTA 5		PML 2 0.0% of population	PML 2 0.4% of population	PML 3 2.0% of population	PML 3 + EM/HD 2.4% of population	Release Not Recommended 1.5% of population
FTA 6				Release Not Recommended 0.1% of population	Release Not Recommended 0.3% of population	Release Not Recommended 1.8% of population

Marie VanNostrand, INTRODUCTION TO THE PUBLIC SAFETY ASSESSMENT AND DECISION MAKING FRAMEWORK: NEW JERSEY (on file with the authors).
Reformatted for this report.



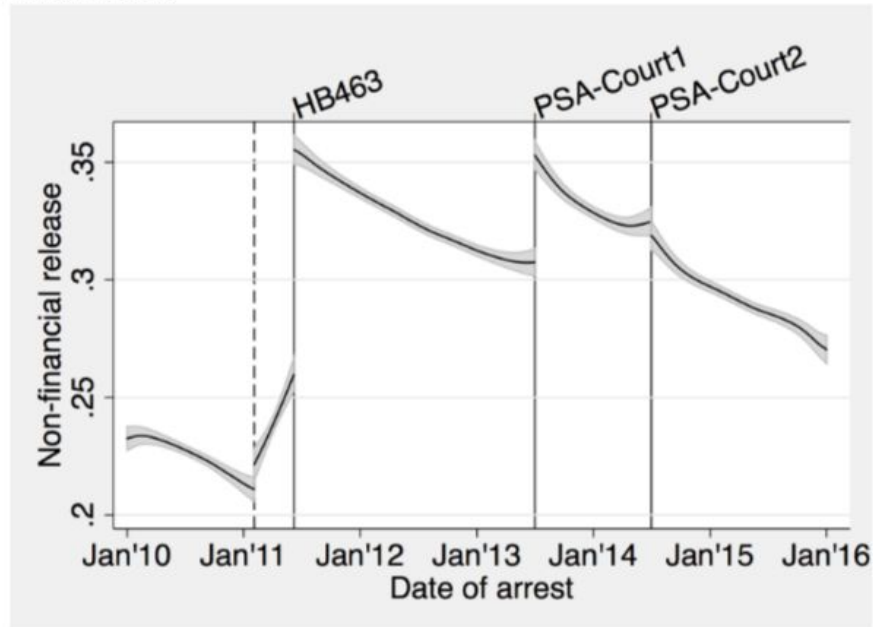
What we know: Two kinds of research

- **Statistical validation:** Do the tools' forecasts come true?
 - Some data here. Forecasts of missed appointments and rearrest can be generally accurate
 - **But** the burden of error is racially disparate

- **Impact on Decision-making:** What happens to people, once tools are introduced?
 - *Very scant research here*

What we know - risk assessment isn't always used decarcerally

Figure 4 - Non-financial release rates over time for all defendants



Note: This figure shows the fraction of defendants who are granted non-financial release over time. From left to right, the vertical lines indicate the date HB 463 was introduced as legislation, the date it was implemented as law, the date the PSA was adopted, and the date it was modified.

THE PRESENCE OF JUSTICE

Innocence Is Irrelevant

This is the age of the plea bargain—and millions of Americans are suffering the consequences.

EMILY YOFFE SEPTEMBER 2017 ISSUE



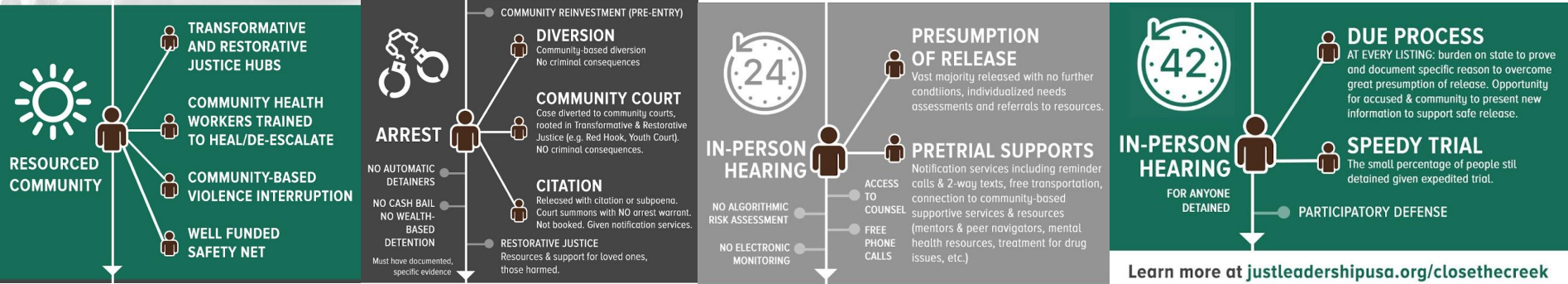
With all this in mind - how communities came to abolition





COMMUNITY DEMANDS for PRETRIAL JUSTICE

Divert from the system. Reinvest in the people.



Learn more at justleadershipusa.org/closethecreek

THE USE OF PRETRIAL "RISK ASSESSMENT" INSTRUMENTS:

A SHARED STATEMENT OF CIVIL RIGHTS CONCERNS

We believe that jurisdictions should work to end secured money bail and decarcerate most accused people pretrial, without the use of "risk assessment" instruments.

The extraordinary measure of pretrial detention should be treated as a last resort and should only be imposed upon an accused person after they've received a thorough, adversarial hearing that observes rigorous procedural safeguards respecting individual rights, liberties, and the presumption of innocence.

In light of the concerns raised in this document, we urge jurisdictions to reconsider their use of risk assessment tools. Pretrial risk assessment instruments – although they may seem objective or neutral – threaten to further intensify unwarranted discrepancies in the justice system and to provide a misleading and undeserved imprimatur of impartiality for an institution that desperately needs fundamental change.

Where these tools are used, in order to reduce the harm they can cause we urge the following:

1

Pretrial risk assessment instruments must be designed and implemented in ways that reduce and ultimately eliminate unwarranted racial disparities across the criminal justice system. Those engaged in the design, implementation, or use of risk assessment instruments should also test ways to reduce the racial disparities that result from using historical criminal justice data, which may reflect a pattern of bias or unfairness.

2

Pretrial risk assessment instruments must be developed with community input, revalidated regularly by independent data scientists with that input in mind, and subjected to regular, meaningful oversight by the community. The particular pretrial risk assessment instrument chosen should be trained by, or at least cross-checked with, local data and should be evaluated for decarceration and anti-racist results on a regular basis by the local community, including people impacted by harm and violence, and people impacted by mass incarceration, and their advocates.

3

Pretrial risk assessment instruments must never recommend detention; instead, when a tool does not recommend immediate release, it must recommend a pretrial release hearing that observes rigorous procedural safeguards. Such tools must only be used to significantly increase rates of pretrial release and, where possible, to ascertain and meet the needs of accused persons before trial, in combination with individualized assessments of those persons. Risk assessment instruments must automatically cause or affirmatively recommend release on recognizance in most cases, because the U.S. Constitution guarantees a presumption of innocence for persons accused of crimes and a strong presumption of release pre-trial.

4

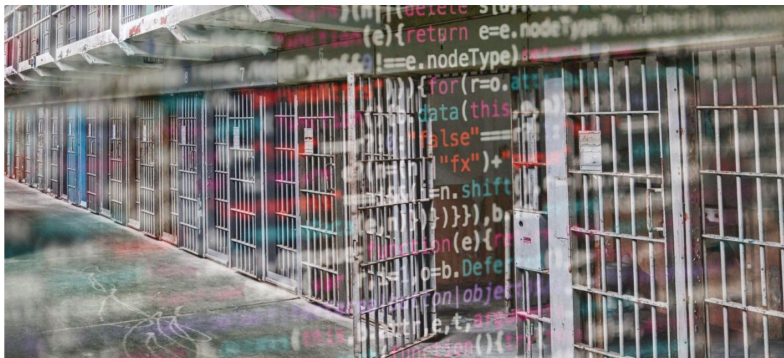
Neither pretrial detention nor conditions of supervision should ever be imposed, except through an individualized, adversarial hearing. The hearing must be held promptly to determine whether the accused person presents a substantial and identifiable risk of flight or (in places where such an inquiry is required by law) specific, credible danger to specifically identified individuals in the community. The prosecution must be required to demonstrate these specific circumstances, and the court must find sufficient facts to establish at least clear and convincing evidence of a substantial and identifiable risk of flight or significant danger to the alleged victim (or to others where required by law) before the exceptional step of detention of a presumptively innocent person, or other onerous supervisory conditions can be imposed. All conditions short of detention must be the least restrictive necessary to reasonably achieve the government's interests of mitigating risks of intentional flight or of a specifically identified, credible danger to others. Any person detained pretrial must have a right to expedited appellate review of the detention decision.

5

Pretrial risk assessment instruments must communicate the likelihood of success upon release in clear, concrete terms. In accordance with basic concepts of fairness, the presumption of innocence, and due process, pretrial risk assessment instruments must frame their predictions in terms of success upon release, not failure. Further, such tools should only predict events during the length of the trial or case – not after the resolution of the open case.

6

Pretrial risk assessment instruments must be transparent, independently validated, and open to challenge by an accused person's counsel. At minimum, the public, the accused person, and the accused person's counsel must all be given a meaningful opportunity to inspect how a pretrial risk assessment instrument works. The accused person's counsel must also be given an opportunity to inspect the specific inputs that were used to calculate their client's particular categorization or risk score, along with an opportunity to challenge any part – including non-neutral value judgments and data that reflects institutional racism and classism – of that calculation.



Pretrial risk assessment instruments must be developed with community input, revalidated regularly by independent data scientists with that input in mind, and subjected to regular, meaningful oversight by the community. The particular pretrial risk assessment instrument chosen should be trained by, or at least cross-checked with, local data and should be evaluated for decarceral and anti-racist results on a regular basis by the local community, including people impacted by harm and violence, and people impacted by mass incarceration, and their advocates.

<https://leadershipconferenceedfund.org/pretrial-risk-assessment/>

Some examples of what community validation could include:

- **Data justice**
 - explaining what you are doing to impacted people, survivors of harm and survivors of mass incarceration, and listening to what they have to say about their experiences of the system
- **Clearly explaining what tools are factoring in and what they aren't**
 - data scientists explaining the *limits* of numbers & encouraging consideration of factors not systematically quantified, like a person's community ties and how risk can be mitigated by meeting someone's often quite simple needs
- **Working with communities to make transparent and accountable HOW these predictions will be used - not what the numbers "say", but what they "do"**
 - data scientists pushing systems to involve communities in deciding how the system RESPONDS to an assignation of risk. (How about meet someone's needs?)
- **Revalidating based on input from the community**
 - Not just testing to say "the tool is accurate in what it was programmed to predict" but "the tool is being used in ways that reduce the harms of mass incarceration as defined by the community"

Some examples of what community validation could include (part 2):

- **Watching for the results that communities care about:** not just reductions in FTA and NCA/NCVA, but are jails getting smaller, are jails getting less racist - and building needs assessment algorithms
- **Using risk assessment to watch decision-makers and people with power:** algorithms to watch police or judges
- **Fight bias in framing of questions:** example: how questions asked about contact w/father for prisoner intake (Reuben)
- **Fight to make sure these tools are only used decarcerally:** only to send people home, preventing judges from using group predictions to punish individuals. A high-risk label should only trigger a hearing, not a decision to incarcerate

Here's how we're gonna get everyone's input:

For the next 20 minutes:

- Attendees and organizers together will mock up up, from a data scientist's perspective, **how communities judged by risk assessment tools can be a part of a design, auditing, validation, or revalidation process.**
- We will ask people to get up and write on the wall **possible processes that would work and possible hurdles to this working**
- Then we will have people indicate reactions visually to the proposed steps, leading into a whole group discussion. **STICKERS!**
- Finally, **we will identify people interested in continuing to work to develop, refine, challenge and implement these ideas** — paving the way for continued momentum after the tutorial.

Questions to keep in mind:

- (a) How can we do community validation in practice? The governance structures are hard to build and hard to explain.
- (b) How can we track successes and challenges of community control in a national way?
- (c) What continued dialog and learning is necessary between designers of these tools and impacted communities?

Questions?

- reuben@justleadershipusa.org, 267-414-4764
- hannah@mediamobilizingproject.org, 267-970-4007
- david.robinson@cornell.edu, 202-657-9892