December 2019

CIVIL RIGHTS AND PRETRIAL RISK ASSESSMENT INSTRUMENTS

By David G. Robinson and Logan Koepke

Upturn, Inc., Washington, D.C., david@upturn.org, logan@upturn.org



Supported by the John D. and Catherine T. MacArthur Foundation

This report was prepared following a meeting convened by the John D. and Catherine T. MacArthur Foundation as part of the Safety and Justice Challenge, which seeks to reduce over-incarceration by changing the way America thinks about and uses jails. Core to the Challenge is a competition designed to support efforts to improve local criminal justice systems across the country that are working to safely reduce over-reliance on jails, with a particular focus on addressing disproportionate impact on low-income individuals and communities of color.

Acknowledgments

This critical issue brief was commissioned by the Pretrial Risk Management Project of the John D. and Catherine T. MacArthur Foundation. Members of the Pretrial Risk Management Project reviewed and critiqued drafts of the brief, but the authors are solely responsible for its content.

Project Members

Sarah Brayne Alexandra Chouldechova Domingo Corona Khalil Cumberbatch Sarah Desmarais Nneka Jones-Tapia Logan Koepke Kristian Lum Sandra Mayson John Monahan David Robinson Vincent Southerland Elizabeth Thornton Trosch

More information is available at www.SafetyandJusticeChallenge.org.



Supported by the John D. and Catherine T. MacArthur Foundation

In this brief we focus on two questions. First: Why do many in the civil rights community oppose the use of pretrial risk assessment instruments? Second: What concrete reform strategies are available that would avoid risk assessment instruments, or would sharply limit their role?

Civil rights advocates in the United States are among the most important voices calling for an end to mass incarceration and for deep changes to pretrial practice changes designed to keep communities safe while also addressing the enormous human, social, and financial costs associated with jailing accused people who can safely be released. A large and growing number of these advocates and organizations argue that actuarial risk assessment instruments should play no role at all in pretrial administration. They further argue that, where risk assessment instruments remain in use, such instruments be carefully circumscribed in order to be made legally, morally, and practically defensible.

Such arguments are compatible with the best available evidence, which does not clearly establish the impacts of these instruments: Although there is no evidence that they decrease public safety, it remains unclear whether these tools typically cause substantial and lasting reductions in jailing.¹ Moreover, a wide variety of potent, empirically supported interventions are available that can reduce incarceration and mitigate the harms that motivate systemic reform, without the use of an actuarial risk assessment instrument. And in places where actuarial risk assessment tools are adopted, there are both feasible ways and strong reasons to carefully limit their role.

This view, which is sometimes described as an "abolitionist" position on actuarial risk assessment instruments in the pretrial setting, has coalesced over the last three years, alongside a dramatic increase in public interest in risk assessment techniques. It responds in part to a rapid embrace of risk assessment instruments by many stakeholders involved in pretrial reform. It also responds to a widespread perception that such statistical instruments are a "necessary component of a fair pretrial release system[.]"²

The alternatives to actuarial risk assessment that are preferred by civil rights advocates can themselves be

understood as forms of "risk management." Indeed, such advocacy is driven by the desire to manage a wide range of risks that arise in the pretrial context — not only the risks that the accused may commit serious violence, further traumatize victims, or abscond prior to case disposition, but also the risks of assaults and injuries while in jail, lasting harm to children whose parents are jailed, the criminogenic potential of pretrial detention, and other collateral risks of pretrial detention and supervision.³ The risk that people accused of a crime will be unconstitutionally jailed for insufficient reason, as courts have recently found that many people now are, is itself one that civil rights advocates seek to address.

In place of actuarial assessment instruments, the civil rights perspective described here holds that most people accused of crimes should be automatically released on their own recognizance, as a result of simple and categorical rules, and that most of the remainder should be evaluated for possible conditions of release through a careful release hearing. On this view, detention should be possible only for a small subset of the accused. For those inside this small "detention eligibility net," if a prosecutor moves to detain the accused, advocates argue that a thorough and specific detention hearing with heightened procedural safeguards (beyond the already thorough protections at a release conditions hearing) should be required before detention can be imposed.

In jurisdictions that will continue to use a pretrial risk assessment instrument, civil rights advocates have outlined specific steps that can be taken to maximize the extent to which these instruments can be made consistent with civil rights, and with the objective of ending mass incarceration and its attendant human suffering.

Many, perhaps most, participants in pretrial reform efforts would identify the ideal of protecting civil and human rights as an important motivation. Many would also say that advocating for such rights is a practical element

It remains unclear whether these tools typically cause substantial and lasting reduction in jailing, of their daily work — such advocacy happens in a wide range of institutional settings. Stakeholders who press for, design, implement or advocate pretrial risk assessment instruments typically do so out of a deep interest in protecting civil rights. Thus, few, if any, generalizations will apply to everyone who could plausibly be called a civil rights advocate.

The content of this brief is informed by our ongoing conversations with a variety of advocates who are pressing for change in pretrial justice across the United States,⁴ as well as by semi-structured interviews with practitioners in Safety & Justice Challenge implementation sites. One central point of consensus among civil rights critics of risk assessment is that there can be no substitute for direct engagement and partnership with the community wherever a pretrial risk assessment tool is considered for adoption.⁵ This critical issue brief is not a replacement for those vital local discussions.

THE CIVIL RIGHTS OPPOSITION TO RISK ASSESSMENT INSTRUMENTS

CIVIL RIGHTS CONTEXT FOR ASSESSING PRETRIAL RISK ASSESSMENT INSTRUMENTS

Today's pretrial systems evolved out of a long history in which African Americans were categorically denied the equal protection of the laws, first through slavery and then through Jim Crow laws. America's historical treatment of other people of color reflects similar patterns of injustice. The pervasive racial disparities in today's systems cannot be understood apart from this history. Many advocates believe that today's bail systems are unconscionable, unconstitutional, and not appropriate targets for incremental adjustment. Understood from this point of view, pretrial risk assessment instruments often appear to function as a substitute for broader or more fundamental changes. Pretrial risk assessment instruments often appear to function as a substitute for broader or more fundamental changes.

The policy questions surrounding pretrial risk assessment instruments are inseparable from their politics. These tools have, to date, frequently been presented as the foremost or best way to increase release rates and narrow racial disparities in pretrial outcomes — incremental goals that themselves may operate as alternatives to more fundamental changes in the pretrial regime. Many in the civil rights community fear not only that risk assessments may be harmful (or offer little benefit), but also that much of the current wave of reformist energy concerning pretrial justice will be spent implementing these systems (that may yet have to be reformed years later), rather than advancing other changes that would have greater benefit for liberty and racial equity.

Pretrial risk assessment instruments face an inherent legitimacy problem: The world of mass incarceration and racially inequitable criminal law that exists today also provides the data upon which pretrial risk assessment instruments are based.⁶ There is, justifiably, distrust that tools developed on data reflecting this racially inequitable system will avoid perpetuating these patterns, let alone advance substantial reform.

That concern is especially salient given reporting and research on pretrial risk assessment and racial equity.⁷ Research has demonstrated that if risk scores are equally well-calibrated for White and Black individuals accused of a crime, then given differing base rates of arrest between the two groups, the chance of needlessly flagging a Black person as high risk can be higher than the chances of making that same mistake on a White person. That is, the risk that someone who could safely

be released may incorrectly be labelled as "high risk" may be higher for Blacks than for Whites, even if a tool is statistically "well calibrated" for both racial groups.⁸ One central disagreement is whether these seemingly inherent disparities in the human burden of actuarial errors amount to problematic racial bias. Advocates and a number of scholars claim that this outcome does constitute racial bias, while other scholars and many practitioners argue that it does not.⁹

Even if one resolved the disparate burden of errors, there is a deeper problem: Substantial evidence indicates that at every step in the legal system, people of color and White people are treated differently.¹⁰ As a result, the same recorded data (such as a prior arrest) simply does not hold the same meaning for a Black person as for a White person.¹¹ Consider New York City. The city's Criminal Justice Agency is redesigning its risk assessment tool on data from 2009 to 2015.¹² But the city's arrest practices during portions of that period have been held unconstitutional. As a result, the data points should not and cannot be a reliable guide, legally or statistically, to who might come into contact with the criminal law – especially for young men of color who were disproportionately stopped, questioned, and frisked during this period of time. Although the litigation makes New York City's case unusually clear, histories of unconstitutional law enforcement are common across the United States, and evidence that such bias is ongoing can also be found in many places.¹³ As a recent paper crisply explains, ultimately the problem of bias in criminal justice data is such a substantial, serious threat that it calls into question the entire endeavor of data-driven risk assessment.¹⁴

The problem of bias in criminal justice data is such a serious one that it calls into question the entire endeavor of data-driven risk assessment.

SPECIFIC RESERVATIONS ABOUT PRETRIAL RISK ASSESSMENT INSTRUMENTS

Many practitioners today must decide whether the resources available for improving their pretrial system should be spent in part on statistical risk assessment instruments, or should instead be fully devoted to other reform approaches. In this section, we describe specific reservations advocates have about pretrial risk assessment instruments.

RISK ASSESSMENT TOOLS LEARN FROM, PREDICT, AND REINFORCE LONG-STANDING RACIAL DISPARITIES.

Proponents of pretrial risk assessment instruments frequently assert that the tools are race-neutral, or are "free of racial bias."¹⁵ To do so, proponents rely on findings that an instrument's forecasts are equally well *calibrated* for individuals in different racial groups. To be well calibrated means that a given risk score should have the same meaning for those of any race. For example, a Black and a White accused person who are assigned the same risk score should, if all else is held equal, have roughly the same probability of success upon release. The available evidence indicates that several widely used risk assessment instruments do meet this criterion.

Civil rights advocates acknowledge these findings of equal calibration, but deny that statistical calibration can fairly be used as a synonym for "race neutrality" or be used to describe tools as free of racial bias.

The differing underlying rates at which police arrest Black and White people creates an unavoidable, serious problem: *Among those who in fact would succeed if released*, a well-calibrated instrument can still classify more of the Black accused people than the White ones as high-risk. In other words, a Black individual who is ready to succeed on release may be more likely to be deemed high risk than their similarly situated White counterpart. To the extent that risk scores actually influence the outcomes of bail hearings, this effect could lead a wellcalibrated instrument to not only perpetuate, but actually worsen racial disparities.¹⁶ We do not know of any pretrial risk assessment instrument that has acquitted itself of this possibility. In other words, a jurisdiction that implements any pretrial risk assessment instrument today is taking the risk that its instrument may discriminate in this fashion against people of color who are ready to succeed on pretrial release.

It is important here to stress that empirical research about the real-world impact of pretrial risk assessment instruments is as yet insufficient to support generalizations about whether or not, or under which conditions, these tools generally drive lasting reductions in incarceration or racial disparities. Proponents generally believe that, compared to today's state of affairs, risk assessment instruments will describe the accused as safer and will be less biased than the beliefs about risk that judges would act upon in the absence of these tools, leading to less jailing and less racial disparity in outcomes. But, because the implementation of a risk assessment tool is usually simultaneous with other significant changes to the pretrial regime, one cannot confidently infer from a successful (or unsuccessful) overall reform effort that the risk assessment instrument was a helpful (or harmful) factor in advancing reformist goals.

THESE INSTRUMENTS CANNOT FORECAST DANGEROUSNESS OR FLIGHT RISK.

A core concern about pretrial risk assessment instruments is that their forecasts do not address the question judges must answer at arraignment: What, if any, conditions are needed to prevent the accused from hurting other people or from absconding prior to case resolution? Most pretrial risk assessment tools measure other outcomes: the likelihoods of failure to appear at a court appointment, rearrest, rearrest for a violent crime, or pretrial misconduct.

Many pretrial risk assessment tools measure the *composite risk* of generalized "pretrial failure" — that is, the likelihood that someone will either be arrested or fail to appear in court. However, being rearrested and missing an appointment are *distinct* risks that deserve specific interventions.¹⁷ Not only does the law often require judges to consider risks separately, but the decision to show a single composite score may exacerbate perceptions of risk and needlessly stigmatize the accused.¹⁸

No pretrial risk assessment tool measures *flight* risk. Instead, tools can only forecast what the currently available data measures: general risk of missing a court appointment, also known as "Failure to Appear."

Even where pretrial risk assessment tools do clearly differentiate types of risk, the outcomes they forecast are seen as imperfect proxies at best and dangerous standins at worst for the important questions that judges need to answer. Advocates report that community concerns mirror the outcomes of legal significance: The risk of violent crime while on pretrial release, the possible intimidation of witnesses or victims, and the possibility that an accused person will flee the jurisdiction. These three outcomes are quite rare, and most pretrial risk assessment tools do not address them.

No pretrial risk assessment tool measures *flight* risk. Instead, tools can only forecast what the currently available data measures: general risk of missing a court appointment, also known as "Failure to Appear" (FTA). This difference — which current risk assessment tools elide — is a crucial one. A consensus among stakeholders is that very few of the people who miss court appointments actually flee the jurisdiction or otherwise willfully fail to appear. Instead, most failures to appear are much more mundane: they are the result of missed busses, inability to arrange childcare, inability to take time off of work, court system malfunctions, or a simple misunderstanding of the schedule. A series of research studies suggests that simple reminders — either delivered by postcard, voicemail, or text message – can dramatically increase the rate of appearance at court dates. As one scholar aptly summarizes, "these tools ignore both longstanding" doctrinal and statutory emphasis on concerns about flight risk and clear, practical policy needs for a more nuanced understanding of the problems of nonappearance."19

Similarly, pretrial risk assessment tools usually use generalized rearrest as an outcome to measure public safety.²⁰ Here, there are several distinct but overlapping concerns.

First, pretrial risk assessment tools often forecast rearrests of all kinds combined, but only a minority of rearrests involve allegations of violence. Second, research shows that Black people are arrested at higher rates compared to their similarly situated White counterparts for a large number of misdemeanors offenses, a decades long, consistent disparity.²¹ "[T]he use of arrest as a measure of criminality fundamentally assumes that people who do the same things are arrested at the same rates,"22 and yet they are not.²³ In other words, pretrial risk assessment instruments ask us to assume that arrest is not a racially biased measure of criminality. But arrest statistics document the behavior and decisions of police officers and prosecutors, which does not consistently correspond to the behavior of the people arrested. Arrest statistics are greatly influenced by what citizens choose to report and what police departments and officers choose to enforce, both of which can vary widely from one neighborhood or precinct to another.²⁴

Arnold Ventures' Public Safety Assessment (PSA) and COMPAS do forecast the likelihood of arrest for a crime that has been designated as violent. Some nonetheless see violent crime as "so infrequent among pretrial individuals that 'risk assessment' tools cannot produce meaningful conclusions" that someone is a purported "high risk."²⁵ In fact, those forecast with the highest risk of new arrest for a violent crime only have about an 8 percent chance of doing so within roughly six months.²⁶

PRETRIAL RISK ASSESSMENT INSTRUMENTS INSULATE MORAL CHOICES FROM PUBLIC INPUT AND SCRUTINY.

What numeric threshold should be counted as "moderate" or "high" risk pretrial — and who gets to make that decision? Ultimately, these are political and moral questions, not scientific or technical ones. Yet judgments of this kind "are sometimes incorporated into risk assessment instruments without careful attention."²⁷

Probabilistic estimates from risk assessment tools are generally (but not always) translated into a categorical label. An instrument might characterize the accused as "low," "medium/moderate," or "high" risk, or output a point score along a numbered scale – perhaps one to six, or one to ten. Someone must decide what threshold numerical level of risk will be mapped to each label or category, and what course of action will be recommended for people with each label. "In other words, it requires a normative decision about what fraction of defendants should receive the stigmatic high-risk label versus the more benign low-risk label."²⁸ The question of who is at the table when these decisions are made is of utmost concern. Where risk assessment designers may see a statistical exercise and look for logical cut points based on "worsening" outcome data, advocates (correctly) see an inherently political decision.

In many jurisdictions, the substantive decisions actually reached on these questions, often through a non-public process, are problematic — and can easily be changed after reforms have been implemented, potentially undermining the decarceral influence of the tools. The same tool used in two different jurisdictions might have different cutoffs for who is deemed high risk, or recommended for burdensome pretrial conditions.

Advocates also consider the labels themselves misleading: For instance, the probability of *success* for people that today's tools consider the highest risk group ranges from 83.5% likelihood of success under the Federal Pretrial Risk Assessment Instrument, to 74% under the Public Safety Assessment, to 57.9% under COMPAS.²⁹ In other words, for each of these tools, the *majority* of people who are labeled as being in the highest risk group will not be rearrested if released pretrial. For COMPAS and PSA (two of the most studied tools) those with the highest risk label or a violence flag have less than a 13 percent chance of being arrested for a new violent crime.³⁰ One may fairly ask whether it makes sense to deem those with a 75% chance of success a "high" risk of failure. It is a further question whether judges who are advised of such a label accurately understand that the "high risk" person before them may have a 75% chance of avoiding rearrest pretrial.

Generally speaking, pretrial risk assessment tools frame their predictions harshly — that is, the likelihood that some *negative* event will occur.

Of course, if the tool is well-calibrated then those in the high risk group will be *relatively* less likely to succeed on release than others. But the act of describing an entire group of individuals as "high risk" can easily prompt stakeholders to overestimate the underlying level of absolute risk. For example, a recent study found that people vastly overestimate the recidivism rate for the accused who are labeled as "moderate-high" or "high" risk.³¹ Compounding the labeling problem, a recent Pew survey found that twothirds or more of Americans, when given the underlying numbers about likelihood of success, support the release of those who are often labeled as "moderate or high risk" by risk assessment tools.³² In other words, those likely subject to harshest scrutiny are the same people many Americans want to see released.

Widespread confusion about the nature of statistical risk estimates is a further problem, and amplifies potential harm from the numbers themselves. Pretrial risk assessment tools use data about groups of people - typically about people who have been arrested — to assess the probability of future behavior. However, the resulting statistical forecasts can easily be interpreted in the courtroom not as a statement about people like the accused, but rather as a specific diagnosis of the accused individual – for instance that they personally have a certain (often not quantified) "high" likelihood of being rearrested. Here, pretrial risk assessment instruments intersect with a long-standing academic debate about how to understand, interpret, and communicate probability estimates. "[R]easoning from the group to an individual case presents considerable challenges,"33 which are beyond the scope of this brief. From the civil rights perspective, whichever way risk estimates are framed (as a probability, frequency, or categorically) they must express a likelihood of success.

BY EMPHASIZING THE PROSPECT OF FAILURE UPON RELEASE — RATHER THAN THE MUCH MORE LIKELY PROSPECT OF SUCCESS — THESE TOOLS ERODE THE PRESUMPTION OF INNOCENCE.

Generally speaking, pretrial risk assessment tools frame their predictions harshly — that is, the likelihood that some *negative* event will occur. For example, a pretrial risk assessment tool might suggest that someone is at 20 percent risk of new criminal activity. Given a legal framework which strongly favors broad release and the presumption of innocence,³⁴ advocates see the very question pretrial risk assessment tools pose as poorly framed.

To the extent pretrial risk assessment tools are used, advocates want pretrial risk assessment tools to communicate the likelihood of success upon release in clear terms.³⁵ Framing a pretrial risk assessment instrument's forecast as the likelihood the accused would fail to appear or be rearrested upon release may unnecessarily and unjustly lead pretrial decisionmakers to perceive and treat the accused more punitively.

IN MANY CASES, IT IS UNCLEAR HOW A TOOL WAS DEVELOPED OR IS BEING USED.

Woefully little publicly available information describes how a pretrial risk assessment instrument was developed, what data was used to develop it, or how the instrument is actually used. In particular, public defenders worry that the accused's constitutional due process rights are jeopardized without sufficient answers to these questions.

Here, transparency includes, and extends beyond, a strong objection to developer claims that some or all aspects of a tool's construction are trade secrets or otherwise unsuitable for public disclosure. The civil rights interest is in a comprehensive understanding of how a risk assessment tool was built, determining what factors a tool considers, how it weighs those factors, what data was used to develop and validate the instrument, and what criteria were used to determine its validation. Robust transparency measures that detail how an instrument was developed and validated, and how it operates can help pave the way for broader public engagement, but cannot themselves resolve other civil rights concerns.³⁶ A related but distinct gap in public information concerns how a risk assessment tool is actually used in practice. There is shockingly little information on this point: How many people are assigned to each category of risk? How many people are placed on what kinds of conditions? How are decisionmaking frameworks established? For whom and in what cases do judges diverge from a recommendation based on a decision-making framework or release matrix?³⁷ This information is critical.

THE BENEFITS OF SUPPORTIVE SERVICES ARE LARGELY IGNORED IN TODAY'S PRETRIAL RISK ESTIMATES.

Historically, few accused people have received the benefit of the kinds of simple and effective supports that many jurisdictions are now implementing — supports such as text message reminders for upcoming court dates, or assistance with transportation. Pretrial risk assessment tools built on past patterns in a world where these new supports were not offered forecast what might happen *if supportive services are not provided*. Given that, tools may overestimate actual risk and deserve heightened scrutiny.³⁸

The tools make no attempt to forecast the outcome that needs to guide judges: What is the accused person's likelihood of success *after* available steps are taken to maximize the likelihood of successful appearance and nonarrest throughout the pretrial period? This matters because there is good reason to believe that supportive services and, in some cases, monitoring conditions can and do materially reduce risk. For instance, a series of studies suggests that reminding the accused person of an upcoming court date can reduce failure to appear by as much as one-third.³⁹

Today's risk assessment tools provide fairly limited information: the likelihood of certain outcomes — such as rearrest for any reason or failure to appear — under the condition that *no new assistance is provided to the accused*. But, under Supreme Court precedent, state constitutional provisions, and many court rules, the threshold question for pretrial detention in most cases is whether there are any conditions that could reasonably assure an individual's appearance or the safety of the community, and if so, what these conditions are.⁴⁰ Only when no such conditions are possible does detention become constitutionally defensible.

APPROACHES THAT AVOID PRETRIAL RISK ASSESSMENT INSTRUMENTS

PRETRIAL REFORMS WITHOUT RISK ASSESSMENT INSTRUMENTS

Last year, over 110 civil rights, social justice, and grassroots groups signed onto a Shared Statement of Civil Rights Concerns regarding the use of pretrial risk assessment instruments, arguing that "jurisdictions *should not use risk assessment instruments* in pretrial decisionmaking, and instead [should] move to end secured money bail and decarcerate most accused people pretrial . . . Jurisdictions can — and should — abolish systems of monetary bail, combat mass incarceration, make meaningful investments in communities, and pursue pretrial fairness and justice without adopting such tools." More recently, a group of academics argued that risk assessment instruments "suffer from serious technical flaws that . . . cannot be resolved with technical fixes."⁴¹ Accordingly, they "strongly recommend turning to other reforms."

MANY PEOPLE HAVE RESPONDED BY ASKING: IF NOT RISK ASSESSMENT, THEN WHAT?

In this section, we describe the core features of emergent, alternative visions of pretrial justice that civil rights organizations and advocates are pursuing across the nation in lieu of risk assessment and in opposition to the for-profit bail industry. Like any intervention, these measures could

Last year, over 110 civil rights, social justice, and grassroots groups signed onto a Shared Statement of Civil Rights Concerns regarding the use of pretrial risk assessment instruments. produce paradoxical or unintended results. But unlike the decision to adopt a pretrial risk assessment instrument, these policies respond directly to the challenge of mass incarceration by compelling release in more cases, and by requiring a careful hearing with a high burden before any deprivation of pretrial liberty. Our purpose in this section is not to necessarily persuade you of the correctness of such views. It is, instead, to make clear that concrete and powerful alternatives to both money bail *and* groupbased actuarial assessment of risk are available, and need to be considered. Understanding this vision may help all stakeholders better understand the resistance to risk assessment tools.

ENSURE THAT THE VAST MAJORITY OF ACCUSED PEOPLE ARE ELIGIBLE FOR PRESUMPTIVE OR AUTOMATIC RELEASE WITH NO CONDITIONS.

A central tenet of today's bail reform movement, like the bail reform movement in the 1960s, is that pretrial liberty should be assured in all but narrow circumstances.⁴² For too long, pretrial justice systems have failed to live up to the Supreme Court's holding that "[i]n our society, liberty is the norm, and detention prior to trial or without trial is the carefully limited exception." Given that the vast majority of people are likely to show up to their court dates and pose no significant public safety threat, the majority of those who are arrested should be released without conditions, or else cited in lieu of arrest.

Policies that directly ensure the automatic or presumptive release of broad categories of the accused are a natural means to accomplish this. As one example, policymakers could specify that for certain categories of criminal charges - like misdemeanors - the accused are automatically released. This could help many accused to avoid pretrial incarceration, and substantially reduce pretrial caseload, enabling judges to conduct more thoroughly individualized hearings where necessary. For other criminal charges, policies could establish a presumption of release with no conditions and create specific and distinct "release hearings" to individually consider the possible imposition of release conditions. To the extent that conditions of release that restrict liberty – like curfews and GPS monitors – are contemplated, courts can be required first to find that

these are the least restrictive means necessary to mitigate a specific, legally relevant risk, and should not impose their financial cost on the accused person.

In many places, local advocates seek to begin this transformation by completely eliminating money bail systems and ending the for-profit bail industry. To the extent money may still play a court-administered role as a potential least restrictive condition of release, existing systems will have to be sharply cabined to account for ability to pay.

SIGNIFICANTLY NARROW WHO IS ELIGIBLE TO BE JAILED BEFORE TRIAL, AND ENSURE ROBUST HEARINGS SOON AFTER ARREST.

Dramatically limiting the population that can be preventively detained is core to this vision. Policymakers can ensure that pretrial detention is limited to a narrow set of serious charges, and that for people who are so charged, a specific and thorough hearing is required before detention of more than a few days can be imposed. Grounds for detention in such hearings could be, for instance, that there is evidence that establishes the accused has made credible, serious, and articulable threats of violence to specific persons or the community or will flee to avoid prosecution.⁴³ Such hearings could also be designed to include robust due process protection, including the right to counsel, access to discovery materials, and the ability to cross-examine witnesses. By design, such a process not only protects the rights of the accused, but also imposes an appropriately high cost on system actors before detention can be imposed.

PROVIDE NEW SUPPORTIVE SERVICES THAT MINIMIZE THE RISKS CREATED BY THE PRETRIAL PROCESS.

A fundamental principle of this pretrial vision is to reduce the harm that the pretrial system can inflict on people waiting for their release, as well as on their families and communities. Accordingly, policies would maximize new kinds of supportive services that help ensure individuals appear to court. These include court date reminders for all those released, new mechanisms that allow people to reschedule court dates without burdensome procedures, basic transportation assistance, and flexible, on-site childcare opportunities. As we described at the outset, those who believe that risk assessment instruments should not be used pretrial may nonetheless, on balance, be willing to accept a package of reforms that includes such an instrument.

CONTROLS AND REFORMS THAT LIMIT PRETRIAL RISK ASSESSMENT INSTRUMENTS

As we described at the outset, those who believe that risk assessment instruments should not be used pretrial may nonetheless, on balance, be willing to accept a package of reforms that includes such an instrument. And in any case, a large number of jurisdictions have already rolled out such instruments or are planning on doing so soon. In such circumstances, what controls can be imposed to maximize compatibility between these instruments and broadly shared civil rights goals?

In this section, we describe the core features of policies that civil rights organizations and advocates have described as necessary controls on the use of pretrial risk assessment instruments in places where these tools are adopted. Many of these controls are not only highlighted in the Shared Statement of Concern,⁴⁴ but were also recently described by a separate group of technical experts on bias in machine learning as "minimum requirements for the responsible deployment of criminal justice risk assessment" instruments.⁴⁵

EXPANSIVE TRANSPARENCY — THROUGHOUT DESIGN AND IMPLEMENTATION — IS SEEN AS MINIMALLY NECESSARY.

Risk assessment models used in the courtroom pretrial and the process used to develop and test them — must be public by default. Assertions of trade secrecy or that certain processes are proprietary are seen as illegitimate.

Generally, advocates want policies to make all data, records, and information used to *develop* and *validate* the pretrial

risk assessment system available in an easily accessible manner and format. Other relevant information advocates want disclosed includes:

- A complete description of the design and testing process, reflecting community input.
- A list of factors that the tool uses and how it weighs them.
- The thresholds and data used to determine labels for risk scores, where applicable.
- The outcome data used to develop and validate the tool at an aggregate and privacy-protecting level, disclosing breakdown of rearrests by charge, severity of charge, failures to appear, age, race, and gender.
- Clear definitions of what an instrument forecasts and for what time period.

But advocates do not just want transparency as to the design of an instrument. Their vision of expansive transparency also requires policy assurances that make the implementation and effects of pretrial risk assessment instruments transparent to the community. Here, it is critical that policymakers establish clear processes and mechanisms that allow the community to meaningfully and regularly evaluate the system for decarceral and racially equitable goals.

Similarly, jurisdictions must make it easy for community members to monitor *how* an instrument is performing by regularly comparing predictions against actual outcomes of interest. Without such data, there is no way to know whether the risk assessment data is systematically wrong about the risks posed by individuals. Such regular monitoring would allow jurisdictions to evaluate how well their risk assessment tool classifies risk and track how reform efforts may be changing risk levels.⁴⁶

Generally, advocates want policies to make all data, records, and information used to *develop* and *validate* the pretrial risk assessment system available in an easily accessible manner and format.

COMMUNITY OVERSIGHT AND GOVERNANCE ARE CRITICAL TO ENSURE REFORMS PROMOTE DECARCERAL AND RACIALLY EQUITABLE OUTCOMES.

Ultimately, advocates are not only looking for ways to ensure they have a seat at the metaphorical table, but are also looking for policies and procedures that vest those impacted by the pretrial legal system — alongside policymakers and other relevant stakeholders — with some form of oversight and governance over how pretrial risk assessment instruments are chosen, implemented, and evaluated. Core to this vision is a widespread recognition that the lived experiences of people and families who have been directly impacted by mass incarceration, as well as by violence, must be carefully understood and reflected in decisionmaking processes surrounding pretrial reform.

The exact mechanisms and structures for establishing robust community governance may, rightfully, vary widely from community to community. As one potential approach, the Shared Statement calls for the creation of a funded and staffed "community advisory board" that aids in the development and revalidation of instruments.⁴⁷ Members of this community advisory board would include those directly impacted by the criminal justice system, survivors of harm and violence, criminal justice stakeholders like public defenders, judges, prosecutors, and community groups focused on racial and economic justice. This board would also oversee implementation of the pretrial risk assessment instrument in order to ensure that, at the outset. jurisdictions build processes so community members can meaningfully and regularly evaluate the system for decarceral and racially equitable results.48

ENSURE DECISIONS TO DETAIN ARE RARE, DELIBERATE, AND NOT DEPENDENT SOLELY ON PRETRIAL RISK ASSESSMENT INSTRUMENTS.

Civil rights advocates believe that pretrial risk assessment tools cannot responsibly be used to support the detention of an accused person. As we observed above, risk assessment tools provide fairly limited information: the likelihood of certain outcomes under the condition that no assistance is provided to the accused. But, the question for detention decisions is whether there are any possible conditions that could reasonably assure an individual's appearance or the safety of the community, and if so, what these conditions are. Despite this, several jurisdictions have, worryingly, explicitly tied pretrial risk assessment scores to detention schemes.⁴⁹

Civil rights advocates are not alone in their concern. Arnold Ventures, the developer of the Public Safety Assessment — likely the most used pretrial risk assessment tool — agrees: "Risk assessment should not be used as the basis to detain someone, only to inform release conditions. Detention decisions should be reserved for a legal process as outlined by the Supreme Court in *U.S. v. Salerno.*"⁵⁰ In other words, the absolute most that risk assessment scores can suggest is a potential hearing, which must meet substantial requirements. Pretrial risk assessment scores cannot be tied to "release not recommended" or detention recommendations, which may prejudicially disadvantage the accused.⁵¹

CONCLUSION

At the very moment jurisdictions across the United States explore and implement pretrial risk assessment instruments, many in the civil rights advocacy community argue that such instruments should play no role at all in pretrial administration. Where instruments remain in use, advocates have described policies and controls to circumscribe their use. This brief aims to help you better understand this viewpoint and the context for it. As the first critical issue brief in this series observed, "a well-validated tool may not produce the intended results of more accurate, decarceral, and racially and ethnically equitable decisions."⁵² Understanding civil rights critiques of pretrial risk assessment tools may help all participants better engage community partners and align bail reform policies with the twin goals of decarceration and greater racial equity.

ENDNOTES

- 1 Such effects are extremely hard to measure. A recent systematic review found that "[a]lthough some researchers and policymakers have hypothesized that the adoption of tools might reduce rates of incarceration ... we found tenuous results. [T]he overall strength of evidence that tools reduce placements is low." See Jodi L. Viljoen, et al., Impact of Risk Assessment Instruments on Rates of Pretrial Detention, Postconviction Placements, and Release: A Systematic Review and Meta-Analysis, 43 Law and Human Behavior 397, Aug. 2019. Jurisdictions that adopt a risk assessment instrument are often making many simultaneous changes to their culture and practices, so that it is unclear what positive or negative contribution the instrument itself made to subsequent pretrial outcomes. A recent report on the changes in Mecklenburg County, North Carolina, is typical: "As is the case in many jurisdictions that implement the PSA [Public Safety Assessment], when Mecklenburg County adopted the tool it was undergoing broad cultural and policy shifts that also included training for court staff members, magistrates, and judges in the best practices of pretrial release and detention and in how to use risk assessment to help determine release conditions. Changes in leadership also occurred shortly after the PSA was adopted that led to additional shifts in policy and practice. This study is assessing the effect of all of these changes that occurred around the same time[.]" See Cindy Redcross et al., Evaluation of Pretrial Justice System Reforms That Use the Public Safety Assessment: Effects in Mecklenburg County, North Carolina, MDRC, Mar. 2019.
- 2 Gideon's Promise, National Legal Aid & Defenders Association, National Association for Public Defense, National Association of Criminal Defense Lawyers, "Joint Statement in Support of the Use of Pretrial Risk Assessment Instruments," May 10, 2017, *available at* https://perma.cc/7FE5-NA7R.
- 3 See, e.g., Kristian Lum, Mike Baiocchi, The Causal Impact of Bail on Case Outcomes for Indigent Defendants (2017); Paul Heaton, Sandra Mayson, Megan Stevenson, The Downstream Consequences of Misdemeanor Pretrial Detention, 69 Stan. L. Rev. (2017); Christopher T. Lowenkamp, Marie VanNostrand & Alexander Holsinger, Laura and John Arnold Found., The Hidden Costs of Pretrial Detention (2013); Arpit Gupta, Christopher Hansman & Ethan Frenchman, The Heavy Costs of High Bail: Evidence from Judge Randomization, 45 J. Leg. Stud. 1, 3 (2016).
- 4 This brief also reflects our years of direct experience working on risk assessment policy issues, including our work helping research and develop a July 30, 2018 shared statement of civil rights concerns on pretrial risk assessment instruments, which was endorsed by more than one hundred organizations. When we use the term "civil rights advocates" below, we are referring not to every person or group that does or could seek to protect civil rights, but instead to the large community of civil rights organizations, advocacy staff and grassroots organizations that are opposed to the use of pretrial risk assessment instruments.
- 5 Advocates argue that it is especially important that policymakers hear directly from the people who have the most at stake in pretrial proceedings, including people who have been accused of crimes (both those who were and who were not detained), their loved ones, and community leaders from the neighborhoods where arrests leading to arraignment are most common, in addition to crime victims (indeed, these identities often overlap). Many cities and towns have organizations that are deeply engaged in these issues and can provide this kind of input, and every state has grassroots organizations that can bring a perspective that is informed by local community experiences and also familiar with state law.

- 6 See Gil Rothschild-Elyassi, Johann Koehler, Jonathan Simon, Actuarial Justice, in The Handbook of Social Control (Mathieu Deflem ed., 2019) at 203 (noting, along similar lines, a "special twist to actuarial justice's legitimacy puzzle: the very system that produced an outcome as illegitimate as mass incarceration must simultaneously be treated as legitimate in producing the data on which actuarial justice-based reform efforts must rely.").
- 7 Julia Angwin et al., *Machine Bias*, ProPublica May 23, 2016.
- 8 See Jon Kleinberg, et al., Inherent Tradeoffs in the Fair Determination of Risk Scores. Here, misclassification is defined as a situation where the instrument assesses a person as high risk, but where that person would not in fact be rearrested if released.
- 9 William Dietrich, et al., Northpointe Inc. Research Dept. "COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity," (2016); Anthony W. Flores, Kristin Bechtel, Christopher T. Lowenkamp, False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks." 80 Fed. Prob. 2 (Sept. 2016).
- 10 See David G. Robinson, Hannah Jane Sassaman, Megan Stevenson, Pretrial Risk Assessments: A Practical Guide for Judges, 57 Judges' Journal 3, at 9-10 ("Nonwhite people are more likely to face hostility from the police, and more likely be arrested, than similarly situated whites. Those who can't afford bail are more likely to plead guilty (resulting, at times, in lengthy sentences) than those who can afford to post bond. Black people are less likely to have their charges dropped, dismissed, or reduced, and when sentenced to incarceration, they receive harsher sentences than similarly situated whites.") (internal citations omitted).
- See Sandra Mayson, Bias In, Bias Out, 128 Yale L.J. 2218 (2019) ("In New Orleans, when I worked there as a public defender, the significance of arrest varied by race. If a black man had three arrests in his past, it suggested only that he had been living in New Orleans. Black men were arrested all the time for trivial things. If a white man had three past arrests, on the other hand, it suggested that he was really bad news! White men were hardly ever arrested; three past arrests indicated a highly unusual tendency to attract law-enforcement attention. A race-blind algorithm would be blind to this difference. It would treat the two men as posing an identical risk. The algorithm could not consider the arrests in the context of disparate policing patterns and recognize that they were a much less significant indicator of risk for the black man than for the white."
- 12 New York City Criminal Justice Agency, Annual Report 2016, at 2.
- 13 Id.
- 14 Laurel Eckhouse, Kristian Lum, Cynthia Conti-Cook, Julie Ciccolini, *Layers* of Bias: A Unified Approach for Understanding Problems with Risk Assessment, Criminal Justice And Behavior, Nov. 23, 2018.
- 15 *See, e.g.*, Pretrial Justice Institute, "Pretrial Risk Assessment Can Produce Race-Neutral Results," *available at* https://university.pretrial. org/HigherLogic/System/DownloadDocumentFile.ashx?Document-FileKey=5cebc2e7-dfa4-65b2-13cd-300b81a6ad7a.

- 16 One recent paper argues that algorithmic guidance does impact bail decisions in a racially inequitable way. Bo Cowgill, *The Impact of Algorithms on Judicial Discretion: Evidence from Regression Discontinuities*, Dec. 5, 2018, available at http://www.columbia.edu/~bc2656/papers/RecidAlgo.pdf. (Finding that "[c]rossing the 'general recidivism' low/medium threshold [under COMPAS] causes an increase in detention of around two weeks, while crossing the same threshold for violent recidivism has a treatment effect of almost a month of additional jail time ... For both the "general recidivism" and "violent recidivism" thresholds, the effect of black defendants passing over the threshold has a much greater magnitude. This is particularly true for the violence threshold, where black defendants crossing the threshold receive extra penalty of two months. The equivalent penalty for white defendants is not statistically significant from zero.")
- 17 Lauryn P. Gouldin, Disentangling Flight Risk from Dangerousness, 2016 BYU L Rev 837, 893. ("The final policy-based reason that flight risk should be analyzed separately from dangerousness is because of differences in conditions of release: many of the conditions that judges can impose to manage or mitigate risk are suited to flight risk or dangerousness, but not to both.")
- 18 Andrea Woods, Brandon Buskey, "Making Sense of Pretrial Risk Assessments," *The Champion*, June 2018. ("Especially where the tool collapses risk categories or measures only the risk of re-arrest for any offense, this confrontation might include asserting that the tool's risk nomenclature is unduly prejudicial, or even an encroachment on the right to an individualized bail determination.")
- 19 Lauryn P. Gouldin, *Defining Flight Risk*, 85 U. Chi. L. Rev. 677, 713 (2018).
- 20 Sandra G. Mayson, *Dangerous Defendants*, 127 Yale L.J. 490 (2018) at 127, Tb. 2.
- 21 Megan Stevenson, Sandra G. Mayson, *The Scale of Misdemeanor Justice*, 98 B.U. L. Rev. 731, 769-770 (2018).
- 22 Laurel Eckhouse, Kristian Lum, Cynthia Conti-Cook, Julie Ciccolini, Layers of Bias: A Unified Approach for Understanding Problems with Risk Assessment, Criminal Justice And Behavior, Nov. 23, 2018.
- 23 See Stevenson and Mayson, supra note 21.
- 24 See, e.g., Jennifer L. Eberhardt, Biased: Uncovering the Hidden Prejudice that Shapes What We See, Think, and Do (Viking 2019).
- 25 Civil Rights Corps, "Risk Assessment" Policy Statement, *available at* https://www.civilrightscorps.org/resources/7B7FS7wRTSKnWMX-SqR9Y.
- 26 This problem is discussed later in the brief.
- 27 David Robinson, Hannah Sassaman, Megan Stevenson, Pretrial Risk Assessments: A Practical Guide for Judges, Judges Journal, Vol. 57, No. 3 at 9 (2018).
- 28 Megan Stevenson, Assessing Risk Assessment in Action, 103 Minn. L. Rev. 303, 335 (2019).
- 29 See: Thomas Cohen, et al., Revalidating the Federal Pretrial Risk Assessment Instrument (PTRA): A Research Summary, 82 Fed. Prob. (Sep. 2018) at 26; Matthew DeMichele, et al., Public Safety Assessment: Predictive Utility and Differential Prediction by Race in Kentucky. Criminology and Public Policy (forthcoming) at 28; Thomas Blomberg, et al., Validation of the COMPAS Risk Assessment Classification, Sept. 2010, at 36 (6 months).
- 30 See Blomberg, et al., Validation of the COMPAS, 52; DeMichele et al., Public Safety Assessment, 28.

- 31 Daniel A., Krauss, Gabriel I. Cook, Lukas Klapatch, *Risk Assessment Communication Difficulties: An Empirical Examination of the Effects of Categorical Versus Probabilistic Risk Communication in Sexually Violent Predator Decisions*, Behavioral Sciences & Law (2018). (Finding that respondents vastly overestimated the true recidivism rate for those assessed as moderate-high risk category — it was less than half of what participants thought it was.)
- 32 Pew Charitable Trusts, Americans Favor Expanded Pretrial Release, Limited Use of Jail, Figure 8 at 9, Nov. 2018, *available at* https://www. pewtrusts.org/-/media/assets/2018/11/pspp_public_favors_expansion. pdf.
- 33 David Faigman, John Monahan, Christopher Slobogin, Group to Individual (G2i) Inference in Scientific Expert Testimony, 81 U. Chi. L. Rev. 417 (2014).
- 34 The United States and all fifty states prohibit excessive bail and nearly all states have a constitutional or statutory presumption in favor of release that applies in nearly all cases. The presumption of innocence – while more generally applicable in the context of trials and burden of proof – is inextricably attached to the presumption of pretrial release. As the Supreme Court explained in the 1951 case *Stack v. Boyle* 342 U.S. 1 (1951), a robust right to pretrial release is critical to preserving the presumption of innocence.
- 35 The Leadership Conference on Civil and Human Rights, "The Use of Pretrial 'Risk Assessment' Instruments: A Shared Statement of Civil Rights Concerns," 2018, *available at* http://civilrightsdocs.info/pdf/ criminal-justice/Pretrial-Risk-Assessment-Full.pdf ("For example, a tool could say that, in the past, people similar to the accused person have had an 80 percent likelihood of appearing for all of their court dates over 18 months of hearings.")
- 36 Recently enacted laws in both Idaho and New York offer helpful examples of fairly robust transparency requirements that many in the civil rights community see as a first step in the right direction. New York's law requires underlying data from validation studies to be made public upon request, while Idaho's law goes further, requiring all "documents, data, records, and information" used to build or validate a risk assessment tool to be made public for auditing and testing. *See* https://legislation.nysenate.gov/pdf/bills/2019/S1509C and https://legislature. idaho.gov/sessioninfo/billbookmark/?yr=2019&bn=H0118
- 37 Though not strictly related to the transparency as to *how* risk assessment tools are used, it is important to note that advocates also want to continuously monitor who is on what kind of conditions of release, broken down by age, race, and gender.
- 38 See John Logan Koepke, David G. Robinson, Danger Ahead: Risk Assessment and the Future of Bail Reform, 93 Wash. L. Rev. 1725, 1757 (2018). ("Simply put, risk-mitigating policies will likely change the risks a defendant faces upon release, just like a change in economic conditions or in time can. Overall, using historic, pre-reform outcome data to predict future risks within a jurisdiction that significantly reforms its bail system deserves heightened, continued scrutiny.")
- 39 *See Danger Ahead* at 1786-1787 for a summary of available empirical findings on pretrial reminders.

- 40 United States v. Salerno, 481 U.S. 739 (1987). ("In a full-blown adversary hearing, the Government must convince a neutral decisionmaker by clear and convincing evidence that no conditions of release can reasonably assure the safety of the community or any person.") See also, American Bar Association, Standards of Pretrial Release, Part V., 10-5.10(f) ("In pretrial detention proceedings, the prosecutor should bear the burden of establishing by clear and convincing evidence that no condition or combination of conditions of release will reasonably ensure the defendant's appearance in court and protect the safety of the community or any person.")
- 41 Technical Flaws of Pretrial Risk Assessments Raise Grave Concerns, July 2019, available at https://dam-prod.media.mit.edu/x/2019/07/16/ TechnicalFlawsOfPretrial_ML%20site.pdf.
- 42 For a time capsule of bail reform advocacy from that era, see then-Attorney General Robert F. Kennedy's testimony in favor of a proposed federal law: https://www.justice.gov/sites/default/files/ag/legacy/2011/01/20/08-04-1964.pdf.
- 43 Senate Bill 10 offers a particular example of what advocates do not want to see as it pertains to the detention eligibility net. The (now-suspended) law authorized a prosecutor to file a motion seeking detention pending a trial if they believed that there was "substantial reason to believe that no nonmonetary condition or combination of conditions of pretrial supervision will reasonably assure protection of the public or a victim, or the appearance of the defendant in court as required." To advocates, this standard is highly discretionary, if not unlimited, and unnecessarily expands the net of who is eligible to be detained. This prosecutorial prerogative was absent from the earlier versions of the bill that advocates supported.
- 44 The Shared Statement observes, "understanding that these tools are already used in communities across the country ... we have developed the following principles that, in total, provide tools and guidance for reducing the harm that these assessments can impose. In order to ameliorate the strong dangers and risks we see in the implementation of risk assessment instruments in pretrial decisionmaking, all of the standards described in our principles below should and must be met by a jurisdiction."
- 45 Partnership on Al, *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System*, April 2019, *available at* https://www.partnershiponai.org/report-on-machine-learning-in-risk-assessment-tools-in-the-u-s-criminal-justice-system/.
- 46 It is also important for someone's risk assessment prediction, their subsequent outcome, and case file to be linked. This would make it possible to analyze how predictions or outcomes correlate with other features, such as a defendant's race, socioeconomic status, recent rearrests, or type of pretrial monitoring. Doing so would help ensure that risk assessment tools lead to more equitable outcomes across race, gender, and socioeconomic status.
- 47 Here, the quality of technical assistance is critical. This is why the Shared Statement calls for the board to include independent data scientists who can help communities best articulate their priorities at a technical level.
- 48 Importantly, in order to realize this vision, jurisdictions would have to make key outcomes — like risk scores, release rates, release conditions, technical violations or revocations, detention motions decisions, and other relevant metrics broken down by race, age and gender — easily discernible.

- 49 For example, in New Jersey, an initial court rule allowed judges to consider a high pretrial risk assessment score as "prima facie evidence sufficient to overcome the presumption of release," though this has now changed; Similarly, in California, Senate Bill 10 (which is currently suspended and subject to a voter referendum) would have explicitly tied a risk score to the determination of whether to release the accused; And in Yakima, Washington, release is not recommended for many of those with high risk scores, despite the fact that many of those assessed are not, under Washington law, legally eligible for detention.
- 50 Arnold Ventures, "Statement of Principles on Pretrial Justice and Use of Pretrial Risk Assessment," available at https://craftmediabucket.s3.amazonaws.com/uploads/AV-Statement-of-Principles-on-Pretrial-Justice.pdf 9.
- 51 In fact, such a recommendation may contravene state law which frequently defines the group of individuals who can be preventively detained quite narrowly.
- 52 Sarah L. Desmarais, Evan M. Lowder, "Pretrial Risk Assessment Tools: A Primer for Judges, Prosecutors, and Defense Attorneys," (Feb. 2019) *available at* http://www.safetyandjusticechallenge.org/wp-content/uploads/2019/02/Pretrial-Risk-Assessment-Primer-February-2019.pdf.



Supported by the John D. and Catherine T. MacArthur Foundation

www.SafetyandJusticeChallenge.org