IN THE SUPREME COURT OF PENNSYLVANIA

No. 21 EM 2019

The Philadelphia Community Bail Fund, *et al.*,

v.

Arraignment Court Magistrates of the First
Judicial District of Pennsylvania

BRIEF OF *AMICUS CURIAE* UPTURN

On Allowance of Objections to the Report of the Special Master, entered on December 17, 2019, making recommendations pursuant to directions to conduct an inquiry regarding the cash bail system in the First Judicial District.

BY: JIM DAVY
Attorney I.D. No. 321631
2362 E Harold St.
Philadelphia, PA 19125
609-273-5008
jimdavy@gmail.com

Counsel for Amicus Curiae

DATE FILED: January 30, 2020

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Statutes and Rules: **Page(s)**

Other Authorities:

STATEMENT OF INTEREST OF *AMICUS CURIAE*

The *Amicus Curiae* party—Upturn—submits this brief in support of the Philadelphia Community Bail Fund, and to assist the Court in considering particular aspects of the Report of the Special Master that concern algorithmic pretrial risk assessments.

Upturn is a 501(c)(3) nonprofit based in Washington D.C. Upturn advances justice and equity in the design, governance, and use of technology. As part of its mission, Upturn has spent the past four years working alongside civil rights and social justice groups, as well as experts in data science, statistics, and machine learning, to research the use and misuse of pretrial risk assessment instruments. Upturn collaborates with hundreds of civil rights groups, digital justice advocates, and community-based organizations to highlight civil rights concerns with pretrial risk assessments. Upturn staff has also published academic work on the subject, reports documenting the shortcomings of algorithmic risk assessment tools, and recommendations for minimal requirements for any responsible deployment of such tools. Upturn staff also co-chaired the Project on Pretrial Risk Management, which advised the MacArthur Foundation's Safety & Justice Challenge on the use of pretrial risk assessment instruments as part of bail reform efforts nationwide. Upturn submits this amicus brief to address the Special Master's *sua sponte*

recommendation that the First Judicial District adopt a pretrial risk assessment instrument.

Amicus Curiae submits this brief Pursuant to Pa.R.A.P. 531(b)(2), and does not repeat arguments made by the parties. Neither party's counsel authored this brief, or any part of it. Neither party's counsel contributed money to fund any part of the preparation or filing of this brief. The brief was prepared entirely by *Amicus* or its counsel.

In this brief, *Amicus Curiae* urges the Court not to order the implementation of a pretrial risk assessment instrument. Upturn offers an analysis of such instruments focused on racial equity and the protection of Pennsylvanians' civil rights, summarizing our past academic, technical, legal, and policy research that counsels against adopting such a tool. *Amicus Curiae* explains several ways in which pretrial risk assessments, as currently developed and implemented, have technical shortcomings, limited value to pretrial decision-making, and social costs that exceed their merits. Should the Court wish to consider one anyway, Upturn offers a number of suggestions for strong controls and limitations to govern its development and use that would mitigate possible civil rights inequities during its implementation.

SUMMARY OF ARGUMENT

As this Court considers the Report of the Special Master, it should decline to accept the Special Master's *sua sponte* recommendation that the Arraignment Court Magistrates develop and implement a pretrial risk assessment instrument. While this suggestion follows a practice adopted by some jurisdictions in other parts of the country, none of the parties that participated in this mediation suggested it, and for good reason. Implementing a pretrial risk assessment would introduce several problems into the First Judicial District's pretrial justice system, and would have several important deficiencies.

First, pretrial risk assessment instruments are inherently flawed because they learn from, forecast, and reinforce long-standing racial disparities. The Special Master's Report itself acknowledges this, noting that, to be useful, any pretrial risk assessment must "account for bias that may creep into the underlying data." Report of the Special Master, pg. 18. This statement actually understates the scope of the problem of racially biased data driving racially biased results in pretrial risk assessments: Because pretrial risk assessment instruments rely heavily upon historical criminal justice data to project future outcomes, racial bias may be unavoidable.

Second, pretrial risk assessment instruments do not support the purposes of pretrial decision-making because they cannot accurately forecast danger to the

community or accurately differentiate rates of reappearance — let alone do so in an unbiased way. The Report of the Special Master suggests using pretrial risk assessment instruments in part because of the incorrect perception that they "accurately assess whether the defendant presents an [*sic*] community danger." Report of the Special Master, pg. 18. To the contrary, pretrial risk assessment instruments struggle to forecast the likelihood of danger, because it is relatively rare for any defendant to be re-arrested for a violent offense. Pretrial risk assessments similarly fail to reliably predict reappearance, because they cannot distinguish between generalized nonappearance (e.g., lack of transportation) and willful nonappearance (e.g., actual flight). Together, these two problems reveal how pretrial risk assessment instruments, as currently designed and implemented, have flaws that limit their ability to assist courts in considering either of the two acceptable bases for holding a defendant pretrial.

Third, if this Court still wishes to pursue pretrial risk assessments, any attempts to mitigate racial bias would require numerous ongoing policy controls and limitations. While the Report of the Special Master notes that an instrument must be "properly developed, account for bias that may creep into the underlying data … and be routinely tested and calibrated," Report of the Special Master, pg. 18, the lack of detail in the Report on this subject obscures the truth that developing a well-calibrated risk assessment tool risks entrenching bias within the system. Although

routinely testing and recalibrating an instrument meets the minimum baseline for predictive validity, fairness would require many more controls and limitations. These would include expansive transparency mechanisms, community oversight, and other policies that limit potential unintended effects of the instruments.

Ultimately, from observations in other jurisdictions, pretrial risk assessment instruments have not been shown to help reduce incarceration or enhance a bail system's fairness, and raise new problems concerning bias. Current evidence does not support the contention that pretrial risk assessment instruments cause substantial, durable, and racially equitable reductions in jailing. Instead, most research to date regarding pretrial risk assessment instruments has focused on predictive validity. But research on *predictive validity* is not research on *implementation* success.[1] In fact, little methodologically rigorous research has been conducted on the implementation effects of pretrial risk assessment instruments, much less on whether the use of a pretrial risk assessment instrument contributes to reducing racial inequality in pretrial decision-making. The available evidence does not presently support an assumption that implementing a pretrial risk assessment instrument would lead to reductions in jailing, nor does it support an assumption that a pretrial risk assessment would mitigate longstanding racial disparities in pretrial justice.

[1] "Predictive validity" refers to research establishing that an instrument can estimate the probability of failure to appear and/or a re-arrest at a statistically significant rate.

For all of these reasons, *Amicus* urges the Court to reject the *sua sponte* suggestion of the Special Master to design and implement a pretrial risk assessment instrument for use by the First Judicial District.

ARGUMENT

Amicus urges the Court to decline the Special Master's *sua sponte* recommendation to order the development and implementation of a pretrial risk assessment instrument for use in the First Judicial District. First, *Amicus* explains that pretrial risk assessments contain inherent flaws because they rely on a foundation of data that incorporates persistent, historical racial discrimination in the criminal justice system. Second, *Amicus* notes that even the most carefully built existing pretrial risk assessment tools fail to inform pretrial decision-making because their projections do not distinguish risks based on likelihood of events that justify pretrial detention under applicable law. For these reasons, *Amicus* strongly recommends against developing and implementing a pretrial risk assessment. Third, if the Court nevertheless intends to do so, *Amicus* provides several recommendations that might help mitigate some of the racially discriminatory effects of any pretrial risk assessment ordered.

I. Pretrial risk assessment instruments are inherently flawed because they learn from, forecast, and reinforce long-standing racial disparities.

Pretrial risk assessment instruments pose enormous threats to the attempts of Courts, prosecutors, defense attorneys, and other stakeholders to ensure that criminal justice systems do not discriminate against defendants on the basis of race. Many observers, including the Special Master, acknowledge that as precondition for implementing any risk assessment, development of the instrument must "account for

bias that may creep into the underlying data." Report of the Special Master, pg. 18. That caveat, however, understates both the gravity of the problem and its inevitability. Because pretrial risk assessment systems rely upon historical criminal justice data to project future outcomes, the quality of those projections depends on the quality of the criminal justice system data that was used to develop them. As a result, persistent problems with the effects of mass incarceration, racially inequitable policing, racially disparate charging decisions, and other discrimination in the criminal justice system doom pretrial risk assessments to project the same discrimination and bias into the future.

Studies on the effects of pretrial risk assessment have found troubling racial equity problems and have highlighted three primary concerns with pretrial risk assessment. First, research demonstrates that building a tool that is "fair" as a matter of statistics does not necessarily mean that it will treat individuals or groups fairly in practice. Second, research demonstrates that the data that jurisdictions routinely collect are inadequate to the task of projecting future outcomes for individuals. (Taking those issues together, any tool built on existing data that mitigates racial bias might well require interventions to mitigate bias that run afoul of the constitution.) And third, even beyond flaws in the underlying data, studies about the implementation of pretrial risk assessments show pretrial decisions that continue to perpetuate racial bias in troubling ways.

a. Historical data reflect past racial discrimination in the criminal justice system, and doom future forecasts made based on that data.

Biases in historical criminal justice data mean that statistical "fairness" runs counter to justice in practice. Research has shown that divergent re-arrest rates of Black and White people put two important definitions of statistical fairness—predictive parity and equal false positive rates—at odds.[2] Specifically, even if a pretrial risk assessment instrument maintains "predictive parity" for Black and White individuals accused of a crime, then, given differing base rates of re-arrest among those two groups, the likelihood that a pretrial risk assessment instrument incorrectly labels a Black person as a "high risk" will be higher than the chances of making the same mistake for a White person.

For an assessment instrument to achieve "predictive parity" as a matter of statistics, for "each score value … the proportion of people who actually experience a given outcome (e.g., re-arrested in parole) is the same as the proportion of people predicted to experience that outcome."[3] In context, risk assessment scores should

---

[2] *See, e.g.*, Jon Kleinberg, Sendhil Mullainathan, Manish Raghavan, *Inherent Trade-Offs in the Fair Determination of Risk Scores*, Proceedings of Innovations in Theoretical Computer Science, 2017, *available at*: https://arxiv.org/abs/1609.05807; *see also* Alexandra Chouldechova, *Fair prediction with disparate impact: A study of bias in recidivism prediction instruments*, Proc. FAT/ML, 2016, *available at*: https://arxiv.org/abs/1610.07524.

[3] Richard Berk, et al., *Fairness in Criminal Justice Risk Assessments: The State of the Art*, Sociological Methods & Research, July 2018, *available at*: doi:10.1177/0049124118782533.

mean the same thing for all individuals in the instrument, regardless of their race. For example, in a well-calibrated instrument with predictive parity, among those assessed to have a 30% likelihood of being re-arrested, about 30% of them should in fact be re-arrested. A fair system would, at a minimum, ensure that within each risk category, "the proportion of defendants who [are re-arrested] is approximately the same regardless of race."[4] Most pretrial risk assessment instruments intend to be well-calibrated.

Separately, a fair system should have "false positive" rates that do not vary across racial groups. A pretrial risk assessment system should exhibit similar rates of "false positive" results—where someone who does not go on to be rearrested was nevertheless assessed as high risk—and "false negative" results—where someone who does go on to be rearrested was nevertheless assessed as low risk—across different racial groups. However, research shows that pretrial assessment systems are much more likely to incorrectly label Black defendants as high risk (false positives), and much more likely to incorrectly label White defendants as low risk (false negatives).[5] Critically, this disparity is not the result of faulty statistical

[4] Sam Corbett-Davies, Emma Pierson, Avi Feller, and Sharad Goel, "A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear." Washington Post, Oct. 17, 2016.

[5] Julia Angwin, et al., Machine Bias, *ProPublica*, May 23, 2016.

formulation. Instead it is the "mathematical result of the divergent rates of arrest between the black and white defendants in the underlying dataset . . . so long as the algorithm is also striving to have equal predictive accuracy [calibration] for each racial group."[6]

The racial disparities in arrests mean that pretrial risk assessments cannot simultaneously satisfy multiple important statistical definitions of fairness. So long as the base rate of the outcome the assessment seeks to predict (here, re-arrest) diverges across racial lines, the tool cannot simultaneously achieve predictive parity, parity in the false-positive rate, and parity in the false-negative rate. "It is mathematically impossible to develop a model that will be fair in the sense of having equal predictive value across groups, and fair in the sense of treating members of groups similarly in retrospect."[7] To the extent that pretrial risk assessment instrument scores influence pretrial decision-making, this effect could lead a "well-calibrated" instrument to worsen racial disparities, never mind merely failing to fix

[6] Sandra Mayson, Bias In, Bias Out, 128 YALE L. J. 2218, 2234 (2019).

[7] Laurel Eckhouse, Kristian Lum, Cynthia Conti-Cook, Julie Ciccolini, *Layers of Bias: A Unified Approach for Understanding Problems with Risk Assessment*, 46 Crim. Just. & Behavior 6, Feb. 2019.

them.[8] By contrast, any system that is fair and just across racial lines cannot be well-calibrated to existing data from the past.

b. Pretrial risk assessment system development requires unrealistic assumptions or constitutionally questionable intervention.

The nature of the prediction task at hand itself complicates efforts to "debias" or "account for the bias" in developing a pretrial risk assessment instrument. Pretrial risk assessment instruments rely upon re-arrest as a proxy for public safety—a defendant who gets arrested a second time after having received pretrial release for a first charge, the thinking goes, must have posed a danger to the community. But "the use of arrest as a measure of criminality fundamentally assumes that people who do the same things are arrested at the same rates."[9] Research shows this is simply not the case. Instead, in a disparity that has persisted for decades, Black people are arrested at a higher rate than similarly situated White people for a large

[8] *See* Bo Cowgill, *The Impact of Algorithms on Judicial Discretion, Evidence from Regression Discontinuities*, working paper, 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 'general recidivism' and 'violent recidivism' thresholds, the effect of black defendants passing over the threshold has a much greater magnitude. This is particularly true of 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.").

[9] Eckhouse, *Layers of Bias*, *supra* note 7, at 12.

number of misdemeanor offenses.[10] Decades of research have shown that arrest data primarily document the behavior and decisions of police officers and prosecutors, not the individuals or groups that the data claim to objectively describe.[11] Beyond arrest rates,[12] defendants of different races experience different treatment from

[10] Megan Stevenson, Sandra G. Mayson, *The Scale of Misdemeanor Justice*, 98 B.U. L. REV. 731, 769-770 (2018).

[11] Carl B. Klockars, *Some Really Cheap Ways of Measuring What Really Matters, in Measuring What Matters: Proceedings. from the Police Research Institute Meetings*, 201, U.S. Dept. of Justice, Office of Justice Programs, 1999, *available at*: https://www.ncjrs.gov/pdffiles1/nij/170610.pdf ("It has been known for more than 30 years that, in general, police statistics are poor measures of true levels of crime. This is in part because citizens exercise an extraordinary degree of discretion in deciding what crimes to report to police, and police exercise an extraordinary degree of discretion in deciding what to report as crimes. ... In addition, both crime and crime clearance rates can be manipulated dramatically by any police agency with a will to do so. It is also absolutely axiomatic that for certain types of crime (drug offenses, prostitution, corruption, illegal gambling, receiving stolen property, driving under the influence, etc.), police statistics are in no way reflective of the level of that type of crime or of the rise and fall of it, but they are reflective of the level of police agency resources dedicated to its detection.").

[12] Brad Heath, *Racial Gap in U.S. Arrest Rates: 'Staggering Disparity,'* USA Today, Nov. 18, 2014, *available at*: https://www.usatoday.com/ story/news/nation/2014/11/18/ferguson-black- arrest-rates/19043207 ("Blacks are more likely than others to be arrested in almost every city for almost every type of crime.").

police officers,[13] during plea bargaining,[14] in ability to prepare or wait for trial,[15] and at sentencing,[16] among other points. Because pretrial risk assessment tools ask

[13] *See, e.g.*, Rob Vogt et al., *Language from Police Body Camera Footage Shows Racial Disparities in Officer Respect*, 114 Proceedings of the Nat'l Acad. Sci. 6521, 6521 (2017) ("We find that officers speak with con- sistently less respect toward black versus white community members, even after controlling for the race of the officer, the severity of the infraction, the location of the stop, and the outcome of the stop.").

[14] *See* Carlos Berdejó, *Criminalizing Race: Racial Disparities in Plea Bargaining*, 59 B.C. L. REV. 1187 (2018) (finding in Wisconsin state courts that "[w]hite defendants are twenty-five percent more likely than black defendants to have their principal initial charge dropped or reduced to a lesser crime," making whites who face felony charges less likely to be convicted of felonies, and that "white defendants initially charged with misdemeanors are more likely than black defendants either to be convicted for crimes carrying no possible incarceration, or not to be convicted at all.").

[15] Kristian Lum & Mike Baiocchi, *The Causal Impact of Bail on Case Outcomes for Indigent Defendants*, Proceedings of 4th Workshop on Fairness, Accountability & Transparency in Machine Learning 1, 4, Aug. 2017, *available at*: https://arxiv.org/pdf/1707.04666.pdf ("We find a strong causal relationship between setting bail and the outcome of a case. . . . [F]or cases for which different judges could come to different decisions regarding whether bail should be set, setting bail results in a 34 percent increase in the chances that they will be found guilty.").

[16] *See* U.S. Sentencing Comm'n, *Demographic Differences in Sentencing: An Update to the 2012 Booker Report*, 2 (Nov. 2017) (finding that from 2012 to 2016, "Black male offenders received sentences on average 19.1 percent longer than similarly situated White male offenders"); *see also* Jill K. Doerner & Stephen Demuth, *The Independent and Joint Effects of Race/Ethnicity, Gender, and Age on Sentencing Outcomes in U.S. Federal Courts*, 27 Justice Quarterly 1 (2010) ("We find that Hispanics and blacks, males, and younger defendants receive harsher sentences than whites, females, and older defendants after controlling for important legal and contextual factors.").

participants in the system to pretend these disparities do not exist, "bias in the [criminal justice] data is a serious threat to the entire endeavor of data-driven risk assessment."[17]

In fact, attempts to create a tool that equalizes false positive rates could run afoul of the Constitution. Equalizing false-positive and false-negative rates for Black defendants and similarly situated White defendants would likely require race-specific thresholds for each risk class. Such thresholds—and treating some similarly-situated defendants differently depending on their racial background—are governmental classifications on the basis of race that, if challenged, need stand up to strict scrutiny.[18] While equity in pretrial risk assessment might amount to a compelling government interest, the narrow tailoring analysis poses more

[17] Eckhouse, *Layers of Bias*, *supra* note 9, at 13.

[18] Mayson, *Bias In, Bias Out*, *supra* note 6; *see also* Aziz Huq, Sam Corbett-Davies, Emma Pierson, Avi Feller & Sharad Goel, *Algorithmic Decision Making and the Cost of Fairness*, 1 Proceedings of the 23rd International Conference on Knowledge Discovery and Data Mining 797, 2017, *available at*: https://arxiv.org/pdf/1701.08230.pdf (observing that "with race-specific thresholds … would likely trigger strict scrutiny, the most stringent standard of judicial review used by U.S. courts under the Equal Protection Clause of the Fourteenth Amendment."); *see also* Aziz Huq, *Racial Equity in Algorithmic Criminal Justice*, 68 DUKE L. J. 1043, 1133 (2019) (arguing that "a multiple threshold rule for different racial groups runs headlong into the anticlassification rule of equal protection doctrine. At a minimum, it would receive strict scrutiny.").

challenges.[19] And as no case has addressed this issue, this Court would step into unprecedented territory if it adopts such a system.

c. Studies suggest that even a perfectly designed pretrial risk assessment system could exacerbate racial inequity during implementation.

Several studies on the effects of risk assessment instruments during implementation have found troubling discrimination on the basis of race. Even if the recommendations reflected racially equitable data about arrests and other factors untainted by historical bias, pretrial decision-makers may re-impose bias in considering the recommendations. One study, for example, found that Kentucky judges responded to pretrial risk assessment scores differently based on the race of the individual under consideration.[20] Specifically, judges were more likely to override recommendations in a punitive way for Black defendants compared to

[19] *See, e.g.*, Mayson, *Bias In, Bias Out, supra* note 18, at 2271 (noting that the intervention to equalize false-positive rates and false-negative rates might not reduce the *net burden* of errors and that the interventions will likely have a substantial cost in inaccuracy, leading to even more errors).

[20] Alex Albright, *If You Give a Judge a Risk Score: Evidence from Kentucky Bail Decisions*, working paper, Sept. 3, 2019, *available at*: https://thelittledataset.com/about_files/albright_judge_score.pdf ("[E]ven within judge and time, black moderate risk defendants are treated more harshly than similar white moderate risk defendants after but not before HB463. After HB463, judges are 10% more likely to deviate from the non-financial bond recommendation for moderate black rather than similar moderate white defendants. (The same is not true for low risk defendants.) This is suggestive evidence that judges interpret risk score levels differently based on race.").

similarly situated White defendants. Overall, the study found that "the introduction of risk score recommendations can widen racial disparities for individuals who share the same predicted risk level."[21] And although New Jersey began using the PSA in 2017, the state's Judiciary noted in a 2018 report that, despite reductions in the pretrial population, "the jail population studies found that … the racial makeup of defendants in New Jersey remained similar."[22]

This result manifests at a statistically significant level at multiple stages in the criminal process. Another study, for example, found that racial disparities increased in the subset of courts that appeared to rely most heavily on risk assessment information at the criminal sentencing stage. Specifically, the study found that Virginia's Nonviolent Offender Risk Assessment not only did not contribute to a reduction of the state's incarceration rate, but that racial disparities increased among the state's circuits that appeared to use the risk assessment instrument the most.[23]

---

[21] *Id.*

[22] Glenn A. Grant, *Report to the Governor and the Legislature*, New Jersey Judiciary, 7, April 2019, *available at*: https://njcourts.gov/courts/assets/criminal/2018cjrannual.pdf?c=taP

[23] Megan T. Stevenson, Jennifer L. Doleac, *Algorithmic Risk Assessment in the Hands of Humans*, working paper, Nov. 18, 2019, *available at*: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3489440 ("The most striking difference … is that here we see an increase in racial disparities. The probability of incarceration for black defendants increased by about four percentage points ($P < 0.10$) relative to white defendants, and the length of the sentence increased by approximately 17% ($P < 0.05$). This increase is partially due to racial disparities in

Yet another study found that Black and Hispanic defendants were, respectively, two and one-and-a-half times more likely than White defendants to be made ineligible for New York City's supervised release based on the supervised release risk assessment instrument.[24]

***

Today's pretrial risk assessment instrument predictions follow yesterday's patterns. Pretrial risk assessment instruments cannot escape those patterns. At worst, pretrial risk assessments can exacerbate disparities rather than merely perpetuate them. And to the extent that the Report of the Special Master and these entire proceedings seek to solve these persistent problems, adopting a pretrial risk assessment offers little value in that process.

II. Pretrial risk assessment instruments do not support the purposes of pretrial decision-making, because they cannot forecast danger to the community or accurately differentiate rates of appearance.

Pretrial decision-making, at its root, seeks to protect the community from risks posed by a small number of defendants and to ensure that defendants appear at future proceedings in their cases. *See, e.g.*, Pa. Rule Crim. P. 523 ("Release Criteria").

---

the risk score, but partially due to the fact that, as in the full sample, judge are more likely to deviate downward for white defendants with high risk scores than black.")

[24] Kristian Lum, Tarak Shah, *Measures of Fairness for New York City's Supervised Release Risk Assessment Tool*, Oct. 1, 2019, *available at* https://hrdag.org/wp-content/uploads/2019/09/2019-HRDAG-measures-of-fairness-CJA-1.pdf.

Pretrial risk assessments do not serve either of these purposes. Pretrial risk assessments struggle to forecast the likelihood that releasing a defendant pretrial would pose a danger to the community because of the relative rarity of re-arrest for violent offenses among any defendants. Similarly, pretrial risk assessments do not aid decision-making because they struggle to distinguish between inadvertent non-appearance and willful non-appearance including flight risk. Indeed, because pretrial risk assessment instruments rely upon available data, shortcomings in that data lead to a profound gap between outcomes that pretrial risk assessment instruments can predict, and potential outcomes that courts and magistrates seek to assess. Worse still, these shortcomings are not balanced by countervailing improvements to bail system fairness or overall reduction in incarceration rates.

a. Pretrial risk assessment instruments do not aid in making pretrial decisions about community danger.

Pretrial risk assessments are not useful in aiding pretrial decision-making in even the Special Master's formulation because they do not "accurately assess whether the defendant presents a community danger." Report of the Special Master, pg. 18. Most of the pretrial risk assessment instruments available today forecast something much less useful to pretrial decision-making than danger to the community: likelihood of generalized re-arrest, or, even more broadly, generalized pretrial failure. That is, most pretrial risk assessments project either the likelihood of the defendant either being re-arrested on any basis at all, or, even less helpfully,

the combined likelihood that a defendant released pretrial will either be rearrested on any basis or will fail to appear for any reason. Because of the wide range of arrestable conduct in Pennsylvania's criminal code, a system that predicts generalized re-arrest—without distinguishing between re-arrest for intentional murder and disrupting a public procession—cannot usefully inform a magistrate's consideration of potential community danger. *Compare* 18 Pa. C.S.A. § 2501 ("Criminal homicide") *with* 18 Pa. C.S.A. § 5508 ("Disrupting meetings and processions"). Worse, failing to distinguish between violent crimes and misdemeanors compounds racial disparities in prediction, because arrests for low-level offenses and drug offenses exhibit the biggest racial disparities.[25]

Even the few pretrial risk assessment tools that distinguish between re-arrest for violent offenses and generalized re-arrest rates pose problems. Those that make such a distinction provide limited predictive value at best, while also prompting stakeholders to overestimate the actual risk posed by individuals in the highest-risk categories. Two of the most common pretrial risk assessment tools—the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) and the Public Safety Assessment (PSA)—exhibit these problems. In those tools,

[25] *See, e.g.*, Heath, *Racial Gap in U.S. Arrest Rates*, *supra* note 12 ("Blacks are more likely than others to be arrested in almost every city for almost every type of crime.").

defendants flagged with the highest risk of violence have less than a 13 percent chance of being rearrested for a new violent offense.[26] For COMPAS, defendants with the highest risk of new arrest for a violent offense only have about an eight percent rate of re-arrest for a violent offense within six months.[27]

To the contrary, most individuals across all risk categories successfully avoid re-arrest and make their subsequent criminal process appearances in court. Research shows that even the individuals labeled the highest risk for re-arrest by popular pretrial risk assessment tools have objectively high rates of successfully avoiding re-arrest (on suspicion of *any* offense) when released pretrial. Among those classified as highest risk, 83% of those assessed by the Federal Pretrial Risk Assessment Instrument succeed; 74% so categorized under the PSA succeed; and about 58% under COMPAS succeed.[28] The vast majority of individuals assessed as the highest

---

[26] *See* Matthew DeMichele, et al., *Public Safety Assessment: Predictive Utility and Differential Prediction by Race in Kentucky*, Crim. & Pub. Pol (forthcoming); *see also* Thomas Blomberg, et al., *Validation of the COMPAS Risk Assessment Classification*, Sept. 2010, at 52.

[27] Thomas Blomberg, et al., *Validation of the COMPAS Risk Assessment Classification*, Broward County Sheriff's Office, 51-52, Sept. 2010, *available at*: http://criminology.fsu.edu/wp-content/uploads/Validation-of-the-COMPAS-Risk-Assessment-Classification-Instrument.pdf.

[28] *See* Thomas Cohen, et al., *Revalidating the Federal Pretrial Risk Assessment Instrument (PTRA): A Research Summary*, 82 Federal Probation 2, 26, Sept. 2018, *available at*: https://www.uscourts.gov/sites/default/files/82_2_3_0.pdf; *see also* DeMichele, *Public Safety Assessment*, *supra* note 26, at 28; *see also* Blomberg, et al., *Validation of the COMPAS Risk Assessment Classification*, *supra* note 27, at 36.

risk are not re-arrested for anything if released pretrial, much less for violent offenses. Labeling them as "high risk" disserves pretrial decision-making and the policy-making underlying it.[29]

b. **Pretrial risk assessment instruments do not aid in making pretrial decisions about reappearance for future proceedings.**

Pretrial risk assessments do not aid in making pretrial decisions about reappearance for future court proceedings because the available data does not help identify willful non-appearance at future proceedings. Because available data simply catalogues whether a defendant reappears in court or not, pretrial risk assessments simply forecast generalized non-appearance in future proceedings rather than distinguishing generalized non-appearance and willful non-appearance including flight. In fact, *no* existing pretrial risk assessment tool measures actual flight risk.[30]

---

[29] Pretrial risk assessments may also confuse lay audiences to the detriment of pretrial justice. A recent Pew survey found that two-thirds or more of Americans, when given the underlying numbers about likelihood of success upon release, support the release of individuals often labeled as "moderate or high risk" by pretrial risk assessment instruments. *See* Pew Charitable Trusts, *Americans Favor Expanded Pretrial Release*, *Limited Use of Jail*, Figure 8 at 9, Nov. 2018, *available at*: https://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2018/11/americans-favor-expanded-pretrial-release-limited-use-of-jail.

[30] *See* Lauryn P. Gouldin, *Defining Flight Risk*, 85 U. CHI. L. REV. 677 (2018) (Arguing that there are three "subcategories" of "nonappearing defendants": true flight, local absconders (those who *remain* in the jurisdiction but persistently and actively avoid court dates), and low-cost appearances (those who remain in the jurisdiction but whose failures to appear are easily preventable, addressable, and/or non-willful)).

This matters because magistrates making pretrial decisions about setting bail, allowing release, or allowing release with conditions could save money and promote justice by distinguishing between people who might actively flee and those who would reappear with as low-impact (and inexpensive) interventions as text message reminders of court dates.

As with community danger, pretrial risk assessment instrument forecasts also fail to aid decision-making because they overstate risk. In this context, pretrial risk assessments make the risk of nonappearance appear greater than it is. Even beyond failing to distinguish why people do not reappear, pretrial risk assessment instruments overstate nonappearance by forecasting the probability that an individual will fail to appear for even *one* court date. Most individuals have many court dates, and missing any one of them does not necessarily connote a person who willfully rejects or flees the criminal process. As a result, many pretrial risk assessment instruments likely overstate the underlying risk to magistrates and other decision-makers, and deter them from embracing the same simple and inexpensive solutions referenced above that would promote reappearance for people who have no active intent to avoid process.

c. Pretrial risk assessments do not make up for other shortcomings by reducing incarceration or promoting bail system fairness.

Current evidence does not support the contention that pretrial risk assessment instruments cause substantial, durable, and racially equitable reductions in jailing.

Notably, hypothesized reductions have not yet manifested in statistically significant ways in other jurisdictions. Studies about deployed systems often contain methodological shortcomings, if anyone has undertaken them at all. Far from "deny[ing] the utility of social science," as the Special Master cautioned against, Report of the Special Master, pg. 18, *Amicus* notes that the utility of social science depends in part on testing hypotheses through rigorous studies—and that part of using social science and statistics for good policy-making means acknowledging the divergence between speculative and proven capabilities of new statistical tools.

Most research to date regarding pretrial risk assessment instruments has focused on predictive validity. This line of research generally suggests that instruments can distinguish between those individuals who present different risk of re-arrest while on release pretrial, and different propensities to fail to reappear for court dates.[31] But this body of research is limited in important ways. First, the methods and statistics used in those studies "often fail to meet the standards of practice in the field of risk assessment and the standards for educational and psychological testing more generally."[32] Second, "there has been no independent

[31] Sarah Desmarais, Evan Lowder, *Pretrial Risks Assessment Tools: A Primer for Judges Prosecutors, and Defense Attorneys*, Safety + Justice Challenge, Feb. 2019, *available at*: http://www.safetyandjusticechallenge.org/wp-content/uploads/2019/02/Pretrial-Risk-Assessment-Primer-February-2019.pdf.

[32] *Id.*

evaluation or synthesis of this research, limiting definitive conclusions regarding the predictive validity of pretrial risk assessment tools overall."[33]

Beyond predictive validity, little methodologically-rigorous research has been conducted on the implementation effects of pretrial risk assessment instruments. Among the few studies into implementation that exist, even fewer rigorous studies investigate whether the use of a pretrial risk assessment instrument contributes to reductions in racial and ethnic inequalities in pretrial decision-making. The nascent research on the broader impacts and effects of pretrial risk assessment instruments is mixed, at best. A recent systematic review found that "[a]though some researchers and policymakers have hypothesized that the adoption of tools might reduce rates of incarceration . . . we found tenuous results. The overall strength of evidence that tools reduce placements is low."[34] Although there is no evidence that the use of these instruments definitively *decreases* public safety, based on the evidence base available today, it cannot (and should not) be assumed that the implementation of a pretrial risk assessment instrument will necessarily *increase* safety or lead to reductions in jailing.

***

[33] *Id.*

[34] 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 5, 397, Aug. 2019.

As the Report of the Special Master rightfully notes, "[w]hat risks are counted and how those risks are assessed makes the difference between a tool that is generally useful in applying the law and one that distorts the application of the law." Report of the Special Master, pg. 18. Based on how pretrial risk assessments are designed and implemented today, the tools do more to distort the application of the law than to help courts and magistrates apply the law, and this Court should reject the recommendation to incorporate them into the First Judicial District's pretrial decision-making.

III. While unbiased pretrial risk assessments do not exist, any attempts to mitigate racial bias would require several necessary policy controls and limitations.

If this Court wishes to proceed with pretrial risk assessments despite the likelihood that they will exacerbate racial disparities and inhibit informed decision-making by overstating risks of violence and flight, *Amicus* offers several recommendations that might help mitigate some of the racial bias inherent to such an instrument. Accounting for bias requires controls and limitations to mitigate potential harms, as even the Special Master acknowledged. Report of the Special Master, pg. 18. Those controls involve both law and policy subject to constitutional

analysis, *see* section I, *supra*, as well as other necessary controls and limitations described here.[35]

First, development and implementation of any pretrial risk assessment instrument should both include expansive transparency requirements. Transparency in this context means publicly disclosing several kinds of information: a complete description of the design and testing process, including engagement of community stakeholders; a list of factors that the tool uses and how it weighs them; clear definitions of what the instrument forecasts, and over what time period it forecasts those defined outcomes; the outcome data—stripped of personally identifiable information—used to develop and validate the instrument, including re-arrest by charge, severity of charge, failures to appear, age, race, gender, and more; the decision-making framework or release conditions matrix used, as well as the basis for translating the estimate of risk into a proposed course of action for a court or magistrate; and the thresholds and data used to determine labels or categories for risk scores, where necessary. Importantly, transparency must also include a requirement to compare predictions against outcomes on a regular basis—the only

[35] *See also, e.g.*, Partnership on ai, *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System*, April 2019; 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.

way to know whether the instrument systematically succeeds or fails at predicting risks posed by individuals.

Second, in light of the inherent shortcomings of predictive tools, any pretrial risk assessment instrument should never recommend detention to a magistrate or other decision-maker. The most specific observation that a risk assessment tool can make capably is that an individual falls within a group that collectively has a higher risk (of, e.g., non-appearance). Such a finding does not address, much less answer, the question of whether interests in public safety and reappearance require detaining that individual pretrial. Where individual liberty is at stake, a model must at most inform an individualized analysis at a robust hearing, rather than dictate a decision to detain.

Third, any pretrial risk assessment that attempts to minimize bias must rely on the most recent possible data, including future data gathered only after other implementation of other pretrial reforms. As described in section I, *supra*, pretrial risk assessment instruments developed solely from historical data that predates the enactment of significant risk-mitigating reforms will perpetuate the biases that exist in those data. Waiting to develop a tool until reforms in the First Judicial District have gone into effect, and worked long enough to manifest in collected data, provides a better chance the tool will assist efforts at reform rather than hinder

them.[36] Such reforms include many of those the parties have already agreed to in the case, *see* Report of the Special Master, pg. 12-13, which will likely result in significant changes to defendants' odds of success on release. Those changes should inform any potential pretrial risk assessment tool.

Fourth, beyond transparency, any potential pretrial risk assessment tool should include structures for community oversight. While such structures might take different forms, oversight from the community subjected to decision-making by the tool is critical because the community has the greatest interest in ensuring that risk assessment forecasts promote racially equitable outcomes and do not inadvertently increase incarceration. An oversight body should have the power to audit the pretrial risk assessment instrument, and raise alarms about its ongoing use depending on how it works in practice.

Fifth, any potential pretrial risk assessment adopted here should frame its forecasts of potential outcomes in clear, affirmative language. Most pretrial risk assessment instruments forecast the likelihood that a negative event — like a failure to appear or re-arrest — might occur. For example, an individual might be assessed with a 20 percent likelihood of failing to appear for at least one court date. This framing not only subjects defendants to potential prejudice at the pretrial decision-

[36] *See generally,* John Logan Koepke, David G. Robinson, *Danger Ahead: Risk Assessment and the Future of Bail Reform*, 93 WASH. L. REV. 1725 (2018).

making stage by undermining the presumption of innocence, but also subtly primes magistrates against pretrial release by highlighting failure rather than the high likelihood of success for even individuals in the highest risk categories. Those individuals, for example, have an 80% likelihood of appearance to their court dates. Fulfilling the promises of this litigation and the purposes of pretrial decision-making requires framing the language of the risk assessment tool in affirmative, clear language that accurately reflects the low risk posed by the majority of individuals across all risk categories.

CONCLUSION

For the foregoing reasons, *Amicus Curiae* urges this Court to reject the Special Master's *sua sponte* suggestion to incorporate a pretrial risk assessment into any set of reforms that emerge from this litigation.

Respectfully submitted,

/s/ Jim Davy
Jim Davy, Esq.
Attorney I.D. No. 321631
2362 E Harold St.
Philadelphia, PA 19125
609-273-5008
jimdavy@gmail.com

Counsel for Amicus Curiae

CERTIFICATES OF COMPLIANCE

I certify that the foregoing brief complies with the word count limitation of Rule 2135 of the Pennsylvania Rules of Appellate Procedure. This brief contains 6,782 words exclusive of exempted portions in Rule 2135(b). In preparing this certificate, I relied on the word count feature of Microsoft Word.

I certify, pursuant to Pa.R.A.P. 127, that this filing complies with the provisions of the Public Access Policy of the Unified Judicial System of Pennsylvania: Case Records of the Appellate and Trial Courts that require filing confidential information and documents differently than non-confidential information and documents.

Dated: January 30, 2020

/s/ Jim Davy

Jim Davy, Esq.

CERTIFICATE OF SERVICE

I hereby certify that on this 30th day of January, 2020, a true and correct copy of the foregoing BRIEF OF AMICUS CURIAE was served on all parties of interest via the Commonwealth's PACfile system.

/s/ Jim Davy

Jim Davy, Esq.