



Pretrial Risk Assessment in Connecticut: Assessment of Current Instrument and Recommendations Moving Forward

October 30, 2019

Summary and Background

The Public Safety Risk Assessment Clearinghouse (PSRAC) at the Urban Institute (Urban) has prepared this report in response to a request from the Connecticut Sentencing Commission (CSC), in partnership with the Judicial Branch Connecticut Court Services Support Division (CSSD), to review the Pretrial Risk Assessment Point Scale (Point Scale) currently in use and provide recommendations for improvement. PSRAC was further asked that its review address:

1. Whether separate tools that predict failure to appear (FTA) and rearrest during the pretrial release period would be preferable to a single tool that predicts both outcomes.
2. How the tool and/or the decision-making framework building on the tool might need to change in a no-money bail system.¹
3. Whether elements in the current tool are potentially discriminatory and should be modified or removed to avoid exacerbating racial and ethnic disparities in pretrial detention.

In response to this request, the PSRAC team conducted a review using publicly-available material on the development, validation, and use of the Point Scale. This report examines the current iteration of the Point Scale and associated decision-making framework, focusing on the three points named above relative to best practice and the extant research evidence regarding pretrial risk assessment and risk-based pretrial decision-making. It concludes by providing recommendations for analyses to conduct to test tool performance that might lead to modifications and optimal implementation of assessment-informed pretrial decision-making.

¹ We do not presume to know the form that such a policy change might take in Connecticut and therefore keep our observations and recommendations at a general level for a pretrial decision-making environment in which money bond has no role.

THE PUBLIC SAFETY RISK ASSESSMENT CLEARINGHOUSE

The Bureau of Justice Assistance (BJA) partnered with the Urban Institute to create the content for the Public Safety Risk Assessment Clearinghouse (PSRAC), a comprehensive online source for current and objective information on criminal justice risk and needs assessments. PSRAC brings together a wide range of resources from research and practice and presents this information in a way that is easy to understand and use. PSRAC also provides training and technical assistance on issues related to the development, validation, adoption, implementation, and use of risk and needs assessments.

BJA is a component of the U.S. Department of Justice's Office of Justice Programs working to support local, state, and tribal justice policy development and capacity building through grant administration, training, and technical assistance. Urban conducts research and evaluation to create evidence-based criminal justice policy and practice at the federal, state, and local levels. The development of PSRAC content and technical assistance reports was supported by Grant No. 2015-ZB-BX-K004 from BJA.

Pretrial Risk Assessment: An Overview

Pretrial risk assessment is a decision aid that sorts defendants into categories based on the predicted probability of pretrial failure. Pretrial failure is generally conceptualized in two ways: failure to appear in court (FTA) and reoffending during a period of pretrial release, with the latter almost always measured by rearrest. Risk assessment tools are developed by analyzing data sets to identify correlations between outcomes like rearrest and FTA and various individual characteristics such as age, offense, prior convictions, and prior failure to appear (Stevenson 2018). These characteristics can be risk factors (making pretrial failure more likely) or protective factors (making pretrial success more likely).

Pretrial risk assessment results inform judicial decision-making about pretrial detention, release, and supervision based on an individual's assessed risk category. Proponents of pretrial risk assessment argue that objective, empirically-driven risk assessment tools make more accurate predictions of pretrial success or failure or do than do individual decision-makers, and have the potential to reduce pretrial detention rates without jeopardizing public safety. However, pretrial risk assessment has faced criticism for potentially contributing to racial and ethnic disparities in the criminal justice system (Leadership Conference on Civil and Human Rights 2018).

Studies on tool performance often use the Area under the Curve (AUC) statistic to assess predictive validity. The AUC statistic measures predictive discrimination, or the degree to which a tool separates failures from successes. Within the context of pretrial risk assessment, the AUC statistic measures the degree to which a tool meaningfully differentiates levels of risk between individuals who experience pretrial failure (reoffending or not appearing in court) and those who succeed. Ranging from 0 to 1, the value of the AUC statistic for pretrial risk assessment can be interpreted as the probability that a randomly selected individual who does not succeed will score higher on a risk assessment tool than a randomly selected individual who succeeds. While most existing risk assessment tools have an AUC statistic in the high 0.60s or low 0.70s, an AUC

between 0.90 and 1.00 is considered excellent, between 0.80 and 0.89 is good, between 0.70 and 0.79 is fair, between 0.60 and 0.69 is poor, and between 0.50 and 0.59 indicates a failure to achieve predictive discrimination (Baird et al. 2013; Thornton and Laws 2009). Although other measures exist to assess the performance of risk assessment tools, the AUC statistic enables researchers, practitioners, and policymakers to compare the predictive validity of different tools across different states and jurisdictions, despite differing base rates of misconduct and success.

The most common type of research and analysis regarding pretrial risk assessment focuses on the predictive performance of such tools. Substantial literature establishes that such tools can successfully separate pretrial populations into categories with meaningfully different likelihoods of pretrial success (Desmarais and Lowder 2019). While actuarial tools predict outcomes more reliably than professional judgments alone (Gottfredson and Moriarty 2006), these tools do not always meet the AUC criterion for an excellent, good, or even fair tool in achieving predictive discrimination. Moreover, even well-validated risk assessment tools can perform poorly if they are not used or interpreted correctly (Stevenson 2018).

Evaluations of the impact of implementing pretrial risk assessment tools on release and pretrial success outcomes at the state and local levels have yielded mixed results and cannot be generalized to draw conclusions about overall outcomes of pretrial decision-making nationwide (Desmarais and Lowder 2019). Pretrial risk assessment in practice may have varied impacts in a single jurisdiction. For example, a 2018 study of Kentucky's mandatory pretrial risk assessment tool (Stevenson 2018) found that it led to a trivial increase in pretrial release (with a major increase in release for low-risk individuals, a moderate increase for moderate-risk individuals, and a decrease for high-risk individuals), an increase in failure to appear and pretrial crime rates, and minimal and county-dependent changes in racial disparities. These effects also eroded over time; within a couple of years, the pretrial release rate was lower than before the tool became mandatory.

Moreover, there are few rigorous studies on how risk assessment tools interact with racial and ethnic disparities in pretrial decision-making and outcomes (Desmarais and Lowder 2019). Different ways to measure algorithmic fairness (whether risk assessment tools differ in predictive validity across subgroups such as race, ethnicity, and gender) have been proposed and are often incompatible with each other (Berk et al. 2017; Corbett-Davies et al. 2017).

Current Pretrial Risk Assessment Practice in Connecticut

Pretrial practice in Connecticut was examined and comprehensively described in the CSC's report to the Governor and General Assembly on Pretrial Release and Detention in Connecticut (2017). Therefore, we will avoid treating it in great detail here in favor of summarizing some key points of particular importance to consideration of the Point Scale.

Connecticut has employed a pretrial risk assessment since the Judicial Branch implemented a risk assessment point scale in 1982, with the most recent revalidation and recalibration of the tool scoring conducted in 2015 (Hedlund 2015). While the details of the pretrial release and bond-setting processes in Connecticut have changed in various ways since then, there are two main decision points informed by the tool. First, CSSD bail staff use it to inform their independent recommendation regarding whether an individual should be considered for financial bond or non-

financial release based on likelihood to appear in court and likelihood of rearrest. Next, CSSD bail staff use the tool to inform the recommended bond amount for cases of financial bond.

The Point Scale categories include charge, marital status, living situation, means of support, length at employer, total years of education, substance/mental health, character references, prior failure to appear, number of convictions, pending charges, safety risk convictions, safety risk pending, and use of a dangerous instrument (Connecticut Sentencing Commission 2017). The current tool items and scoring weights are presented in table 1. As a general guideline, clients with scores of zero or more are prioritized for non-financial release; those with negative point scores are considered more appropriate for financial bond (Connecticut Sentencing Commission 2017; Hedlund and Bantley 2009).

For those individuals for whom the court decides to set a financial bond, the Point Scale results are also factored into the decision for the recommended bond amount. Hedlund and Bantley (2009) developed a bond-setting decision support aid, the Financial Bond Guidelines, to systematically incorporate Point Scale results into bond setting after analyses found that correlations between Point Scale scores and imposed bond amounts were weak. The Guidelines are a matrix with charge type on one axis and a scale on the other axis combining the Point Scale score with mitigating or aggravating factors related to the most serious charge faced by the individual (Hedlund and Bantley 2009). While the relationship between Point Scale scores and pretrial outcomes have been subject to analysis several times as discussed previously, the relationship between the bond amounts and failure to appear in court or rearrest has not been empirically validated (Connecticut Sentencing Commission 2017).

TABLE 1
Connecticut Pretrial Risk Assessment Point Scale (2015)

Charge (Most Serious)	-20 = Capital Felony -10 = Class A Felony -9 = Class B Felony -8 = Class C Felony -7 = Class D Felony -6 = Class E / Unclassified Felony	-5 = Class A Misdemeanor -4 = Class B Misdemeanor -3 = Class C Misdemeanor -2 = Class D Misdemeanor -1 = Unclassified Misdemeanor 0 = Motor Vehicle Violation
Marital Status	0 = Not Married (includes separated, divorced, and widowed) +3 = Married	
Lives with	0 = Alone +3 = Nonimmediate family or roommate +6 = Immediate family	
Means of Support	0 = None or Incarcerated +2 = Reliance on others (includes government support) +5 = Self-reliance (part-time, seasonal, and full-time employment)	
Length at Employer	0 = Less than one year at current job +3 = One year but less than two years at current job +5 = Two or more years at current job	
Total Years (Education)	0 = Less than High School +2 = High School or equivalent +5 = More than High School	
Substance/Mental Health	+3 = No 0 = Yes	
Verifiable References	0 = No +4 = Yes	

Prior Failure to Appear*	0 = No prior failure to appear -7 = Prior FTA for a misdemeanor charge -8 = Prior FTA for a felony charge
Number of Convictions	0 = No convictions -1 = One or two convictions -5 = More than two convictions
Pending Charges	0 = No pending charges -4 = Pending charges
Safety Risk Convictions	0 = Not charged with a Safety Risk Offense and does not have a Safety Risk Offense conviction -2 = Charged with a Safety Risk Offense and has a Safety Risk Offense Conviction
Safety Risk Pending	0 = Not charged with a Safety Risk Offense and does not have a Safety Risk Offense pending -2 = Charged with a Safety Risk Offense and has a Safety Risk Offense pending
Dangerous Instrument	0 = No Dangerous Instrument Involved -2 = Dangerous Instrument Involved
TOTAL POINTS	Below zero: Surety or 10% Bond Zero or more: Nonfinancial form of release

Source: Hedlund, Jennifer. 2015. "Bail Revalidation: An Analysis of Weighted Release Criteria and Bail Points Final Report." Rep. Wethersfield, CT: Submitted to the Judicial Branch Court Support Services Division.

Notes: *Count pending or convicted FTA charges.

The Connecticut Pretrial Point Scale in Comparative Context

One basic way of examining the Point Scale is to compare it with other commonly-used pretrial risk assessment tools that we believe are representative of best (or at least standard) practice in pretrial risk assessment. While there are many jurisdictions like Connecticut that have developed their own tool locally, there are several tools in the public domain that have been implemented and validated in multiple jurisdictions that can serve as reference points for considering the Point Scale: the Public Safety Assessment (PSA), the Virginia Pretrial Risk Assessment Tool (VPRAI), and the Pretrial Assessment Tool (PAT) component of the Ohio Risk Assessment System (ORAS).² Table 2 summarizes the attributes of these tools alongside those of the Point Scale.

In terms of basic tool attributes and risk factors used, the most obvious distinction between the Point Scale and the other three tools is that the Point Scale includes more items, and in particular several sociodemographic factors that are not included in the other tools. Pretrial risk assessment developers often put a premium on factors that are readily available from administrative data systems for two reasons. First, staff time to conduct pretrial interviews is a limited resource in most places. The fewer items in a tool requiring an interview to obtain, the more feasible it is to implement pretrial risk assessment for a large number of individuals. The PSA was explicitly developed as a pretrial risk assessment model that would not require an interview for this reason. Second, when initially developing a tool, using administrative data makes it possible to create a model using historical data, whereas interview-based data must be collected prospectively, adding to the tool development time. Connecticut has the benefit of the staffing necessary to conduct a large volume of pretrial interviews, and the Point Scale development reflects this.

² For details on these tools, their development, and their predictive performance, see DeMichele et al. 2018 (PSA); Danner, VanNostrand, and Spruance 2015 (VPRAI); and Lowenkamp, Lemke, and Latessa 2008 (ORAS PAT).

TABLE 2

Comparison of Connecticut’s Point Scale and Commonly-Used Pretrial Risk Assessments

	Connecticut Point Scale	Public Safety Assessment	Virginia Pretrial Risk Assessment Instrument (VPRAI)	ORAS Pretrial Assessment Tool
Outcomes Predicted	FTA Rearrest	FTA Rearrest Rearrest for violent offense	FTA Rearrest	FTA Rearrest
Total items	14	9	8	7
Risk Factors Included				
Current charge	✓ 6 felony classifications, 5 misdemeanor classifications, and traffic	✓ Violent offense; violent offense and 20 years old or younger	✓ Felony drug, theft, or fraud	
Criminal/FTA history	✓ Prior FTA, number of convictions, number of pending charges, safety risk convictions or pending convictions, dangerous instrument	✓ Pending charge; prior misdemeanor, felony, or violent conviction; prior FTA in the past two years; prior FTA older than two years; prior sentence to incarceration	✓ Active supervision, pending charge(s), criminal history, two or more FTA, two or more violent convictions	✓ FTA warrants in the past 24 months, three or more prior jail incarcerations; age at first arrest
Employment	✓ Length at employer		✓ Employment at the time of arrest	✓ Employment at the time of arrest
Substance abuse	✓ History of substance abuse or mental health issues		✓ History of drug abuse (excluding alcohol)	✓ Illegal drug use during the past six months, severe drug use problem
Living situation	✓ Lives alone, with non-immediate family or roommate, with immediate family			✓ Lived at current residence or not lived at same residence for the past six months
Age		✓ At current arrest		
Education	✓ Total years of education			
Means of support	✓ None or incarcerated, reliance on others, self-reliance			
Marital status	✓			
Verifiable References	✓			

Sources: Hedlund, Jennifer. 2015. "Bail Revalidation: An Analysis of Weighted Release Criteria and Bail Points Final Report." Rep. Wethersfield, CT: Submitted to the Judicial Branch Court Support Services Division.

Virginia Department of Criminal Justice Services. 2018. Virginia Pretrial Risk Assessment Instrument – VPRAI. Instruction Manual - Version 4.3. Virginia Department of Criminal Justice Services.

"Public Safety Assessment: Risk Factors and Formula." n.d. Public Safety Assessment. Laura and John Arnold Foundation. Accessed August 14, 2019. <https://www.psapretrial.org/about/factors>.

The University of Toledo College of Law. n.d. "The Ohio Risk Assessment System (ORAS)." The University of Toledo Law Review. The University of Toledo. Accessed August 14, 2019. https://www.utoledo.edu/law/studentlife/lawreview/pdf/Trout_ORAS-Overview.pdf.

One way to think about the differences between the tools is that the Point Scale pays greater attention to protective factors, with positive points (reducing assessed risk) possible for marital status, living situation, means of support, length at current employer, educational attainment, and references. The weighting of all of these factors aside from marital status became greater in the 2015 revision, so this aspect of the Point Scale relative to other common pretrial risk assessment tools has only become more pronounced as it has evolved.

The most recent revision to the Point Scale proposed new point values and risk score groupings that differentiate groups based on likelihood of pretrial failure. As the risk level increases, each risk group experiences pretrial failure at a greater rate than the next lowest risk group. The analysis report provides results separately for individuals released by the court and non-court releases – it appears that the predictive discrimination (separation of the population into distinct risk groups) is stronger for non-court releases than for court releases. To date, none of the validation reports on the Point Scale have included calculation of an AUC statistic, which prevents a comparison of any kind across the tools on predictive discrimination. Including this in future validation studies would make it easier to both compare predictive performance across different release types and outcomes in Connecticut, and evaluate the Point Scale's performance relative to best practice standards for predictive accuracy. It is important to note, however, that it cannot be assumed that an AUC statistic observed for a pretrial risk assessment tool in another context would be the same if that tool were applied to a Connecticut defendant population; even if a tool performs well in a particular state or jurisdiction, it will not necessarily perform similarly in a different context.

Question 1: Dual Outcomes Prediction

The first specific question posed for this review was whether separate pretrial tools that predict FTA and rearrest during the pretrial release period would be preferable to a single tool that predicts both outcomes. As the comparison above indicates, predicting both FTA and rearrest is a common feature of all of these tools as well as the Point Scale, and this is generally the case for locally-developed tools as well. However, the way that this is operationalized varies. Some tools combine the risk for both outcomes into a single score, while others generate separate scores for each outcome.

The Point Scale has been designed to generate a single score that incorporates the combined risk of FTA and rearrest. Both the VPRAI and ORAS Pretrial Assessment Tool are similarly designed. It is possible to design single-scale pretrial risk assessment tools that perform well predictively because the risk factors for FTA and rearrest tend to be similar.

The PSA also predicts both outcomes, but takes a different approach. The PSA generates separate scores for FTA risk, general rearrest risk, and violent rearrest risk. Some risk factors are used for all scores, others are used for only one score, and still others are combined or defined in different ways to predict different outcomes (see table 3). Using these factors, the PSA produces an FTA risk score of 1 to 6, a general rearrest risk score of 1 to 6, and a yes/no flag for elevated risk of violent rearrest. This provides decision-makers with a more nuanced picture of the likelihood of success for an individual, but also makes incorporating it into a decision framework more complicated. Preferred decision outcomes are simpler to define and articulate for a single unified risk score, and it is easier to track performance measures such as concurrence between risk score and decision outcome.

TABLE 3
Public Safety Assessment Items Used for Each Outcome Prediction

Item	FTA	New Arrest	New Violent Arrest
Age at current arrest		✓	
Current violent offense			✓
Current violent offense+ 20 years old or younger			✓
Pending charge at time of offense	✓	✓	✓
Prior misdemeanor conviction		✓	
Prior felony conviction		✓	
Prior conviction (misdemeanor or felony)	✓		✓
Prior violent conviction		✓	✓
Prior FTA past 2 years	✓	✓	
Prior FTA older than 2 years	✓		
Prior sentence to incarceration		✓	

Source: "Public Safety Assessment: Risk Factors and Formula." n.d. Public Safety Assessment. Laura and John Arnold Foundation. Accessed August 14, 2019. <https://www.psapretrial.org/about/factors>.

A note of caution on predicting violent rearrest or reoffending is needed. When policymakers are thinking about pretrial decisions, the risk they are generally most concerned with is the potential dangerousness of individuals if they are released pretrial. Accordingly, the predicted probability of rearrest for a violent offense is of great interest to criminal justice stakeholders and policymakers. The challenge is that violent rearrest rates during pretrial release tend to be very low, meaning that violent reoffending is a low probability outcome. For predictions of this nature, even the highest risk category will almost certainly contain many more individuals who do not experience

an arrest for a violent crime than those who do. For this reason, caution should be exercised in using any such scale or flag.

The Connecticut pretrial process and decision-making structure have incorporated a single score from the Point Scale for many years. It is worth examining whether predictive performance for FTA and rearrest would be improved by creating separate models predicting each from the factors available in the Point Scale, or otherwise available through administrative data. However, unless any improvement in predictive performance from generating separate scores proves to be very substantial, the benefits of creating separate scores for each outcome are unlikely to outweigh the difficulties of preparing CSSD staff and judges to utilize multiple risk scores in place of the single score to which they are accustomed.

In summary, prediction of both main pretrial failure/success measures is common practice in pretrial risk assessment. Within that overall conclusion, there is still the question of whether a tool should produce a single risk score that reflects the combined likelihood of those outcomes or separate scores speaking to them individually. There is not a clear consensus on this point, in large part because it requires assessing the tradeoff between ease of use for decision-making and (potential) enhancement in predictive performance. It should also be noted that, as a practical matter, the calculation of separate scores creates a new group of individuals – those that are high on one risk scale but low on the other. Therefore, the motivation for calculating two scores separately for FTA and rearrest should meet the two criteria that 1) there is an achievable improvement in predictive performance overall and 2) the agency is willing and capable of tailoring supervision and management practices to that new group of defendants.

Question 2: Pretrial Risk Assessment and Structured Decision-Making in a No-Bond Context

The second specific question posed for this review was how the tool and/or the decision-making framework building on the tool might need to change in a no-money bail system. With regard to the Point Scale itself, a transition to a pretrial decision-making environment without money bail does not change the pretrial outcomes that need to be predicted. Further, the most recent validation analysis of the Point Scale examined only the pretrial population released on a promise to appear, on the logic that the use of financial bond can impact appearance rates. As this validation population was not directly affected by financial bond, the validation results from that study are as relevant in a no-bond context as in the current Connecticut policy context.

Should money bail be eliminated and the Point Scale instrument retained, it will be important to conduct an updated validation study on the Point Scale, which may suggest other modifications to the risk prediction model. We would recommend including outcomes for all individuals released pretrial in that validation analysis. The logic for excluding them from the 2015 validation and Point Scale modification analysis was that the imposition of pretrial conditions or money bail affect the likelihood of misconduct and validation on a population not subject to either is a purer test of the predictive accuracy of the tool. While this makes sense, a move to a new policy structure would presumably involve broader use of pretrial conditions and supervision. Therefore, providing analysis and reporting results of pretrial outcomes by risk level for all individuals released pretrial will provide important information for policymakers and the public, even if revisions to the model are based solely on the population released on a promise to appear.

Moving away from money bail would require extensive changes to the structured pretrial decision-making tools and process. The first stage of decision-making could remain the same – those recommended and approved for release on a promise to appear based in part on Point Scale results under the status quo would continue to be so released. For those not released on promise to appear, there are currently two pretrial release accountability mechanisms: non-monetary conditions of release and imposition of bond. Eliminating money bail will substantially change the available options for pretrial release, and necessarily require a substitute for the decision aid judges use in matching bond amounts to pretrial risk and charge severity. Connecticut stakeholders planning for a no-bond context should consider repurposing the Financial Bond Guide, which already combines consideration of risk, charge severity, and mitigating/aggravating factors (Hedlund and Bantley 2009), substituting the preferred pretrial decision for the bond amounts currently in the tool. The exercise of populating such a matrix (see figure 1 for a sample matrix based in part on the Colorado Pretrial Assessment Tool) can guide Connecticut stakeholders in the reallocation of bond outcomes among:

- Promise to appear
- Release with conditions
- Detention

FIGURE 1

Example of Pretrial Release Matrix without Financial Bond Options (Mesa County, CO)

CPAT Category	1 Felony VRA Crimes (C.R.S 24-4.1-302)	2 Drug Distribution and Aggregation DARP	3 Domestic Violence DVSI 11 or Greater	4 Domestic Violence DVSI 10 or Less	5 Other Felony Crimes	6 Misdemeanor VRA Crimes (C.R.S. 24-4.1-302)	7 Misdemeanor and Traffic Does not include DUI
1	Enhanced	Enhanced	Basic	Court Reminder Calls Only	Court Reminder Calls Only	Court Reminder Calls Only	Court Reminder Calls Only
2	Enhanced	Enhanced	Enhanced	Basic	Court Reminder Calls Only	Court Reminder Calls Only	Court Reminder Calls Only
3	Intensive	Intensive	Intensive	Enhanced	Basic	Basic	Court Reminder Calls Only
4	Intensive	Intensive	Intensive	Intensive	Enhanced	Enhanced	Basic

Source: Mesa County SMART Praxis, <https://higherlogicdownload.s3-external-1.amazonaws.com/PRETRIAL/Mesa%20County%20Pretrial%20SMART%20Praxis%20-%20Pretrial%20Stakeholder%20Group%202013.pdf?AWSAccessKeyId=AKIAVRDO7IERBJP4KSQZ&Expires=1569880605&Signature=1%2B3z4C8OCxxcxhE%2BEQuZbxbo%2BC0%3D>.

Notes: CPAT stands for Colorado Pretrial Assessment Tool.

Perhaps the most critical question on pretrial risk and decision-making is what role preventive detention would play, if any. For all of the problems associated with money bond, it does allow for at least the possibility of release in all Connecticut cases.³ Careful calibration of the criteria for

³ Preventive detention is unavailable to individuals unless charged with a capital offense in Connecticut (Connecticut Sentencing Commission 2017), and Connecticut abolished the death penalty in 2012.

pretrial detention will be necessary to avoid its overly broad application. At a minimum, it seems the prevalence of such detention should be no more than the 7.5 percent of total arrests and 13.6 percent of custodial arrests that result in defendants having a financial bond ordered and being detained until the verdict under the status quo (Connecticut Sentencing Commission 2017, p. 32). As a point of comparison, after New Jersey's nearly complete elimination of money bond, 6.4 percent of defendants charged by complaint-summons or complaint warrants and 19.5 percent of those charged by complaint warrants were detained in 2018 (New Jersey Courts 2019).

It is important to note that pretrial misconduct rates were quite low for even the highest risk portion of the defendant population as identified by the Point Scale. After the revision of the Point Scale in the most recent update by Connecticut State, FTA rates in the highest category (-12 or fewer points) were less than 25 percent for non-court releases and just over 18 percent for court releases (Hedlund 2015). New arrest rates in the highest risk category were about 30 percent for non-court releases and less than 20 percent for court releases. While it would be helpful to examine the combined rates for any misconduct (the proportion who either failed to appear, or were rearrested, or both), this indicates that even "high risk" individuals as assessed by the Point Scale avoid pretrial misconduct more often than not. This suggests that however preventive detention might be employed in Connecticut, it should require more than even the highest risk Point Scale score to result in this outcome. And, more fundamentally, the definition of "high risk" as defined in the Point Scale should also be reconsidered and adjusted.

In summary, the Point Scale tool itself would not need to be changed for a no-bond context, although if retained it should be revalidated and recalibrated (i.e., adjusting cutpoints) as needed. The structured decision-making tool for bond-setting would need to be replaced, and can be repurposed as a way for Connecticut stakeholders to determine the preferred pretrial release decision and pathway for defendants based on their pretrial risk level and current charge severity.

Question 3: Racial and Ethnic Disparity

The final specific question posed for this review was whether elements in the current tool are potentially discriminatory and should be modified or removed to avoid exacerbating racial and ethnic disparities in pretrial detention. Concerns about the potential for risk assessments to reproduce or exacerbate the racial and ethnic disparities found throughout the justice system have gained extensive public attention, and have focused largely on pretrial risk assessment.

Several specific critiques have been advanced regarding risk assessment tools and equity that are relevant to consideration of the Point Scale. First, there is concern that the data used by these tools are biased. Sometimes this critique is focused on dynamic factors such as educational attainment or employment (see DOJ 2014), but the strongest predictors of misconduct in risk assessment tools are criminal history factors such as prior arrests, which are also correlated with race (Skeem and Lowenkamp 2016). In fact, some recent research suggests that inclusion of sociodemographic factors mitigates disparities, which is largely driven by disparities in criminal history across racial and ethnic groups (Hamilton et al. 2019a; Hamilton et al. 2019b). Past criminal justice outcomes reflect both individual offending behavior and system behavior, including differential treatment of individuals in the justice system based on their race, ethnicity, or other factors. This is not only true of outcomes included as risk factors, but of the outcomes tools are designed to predict, such as new arrests.

A second line of critique has focused on concerns that risk-guided decision-making leads to more punitive or restrictive justice outcomes for people of color. The concern relates directly to the previous discussion of how Point Scale results might relate to outcomes like detention. A recent analysis of data in New York City by the Center for Court Innovation (CCI) (Picard et al. 2019)⁴ examined this question, finding that administrative data there could be used to predict pretrial failure with roughly equal accuracy among white, black, and Hispanic defendants, but that the “false positives” (people in the high-risk category who did not experience pretrial failure) would disproportionately be black. CCI’s analysis concluded that a hybrid strategy combining risk score with charge severity would produce the most equitable detention outcomes (though not necessarily the lowest overall detention rate). While the results might look different for Connecticut populations, their approach and discussion of the implications of different modes of using assessment results to guide pretrial decisions could serve as a model for further analysis in Connecticut.

The third line of critique relates to lack of transparency of risk assessment tools. This critique is strongest for proprietary tools, which are a “black box” and often do not provide any opportunity for the public to understand how the risk score for an individual is produced. At the same time, when a risk assessment tool has the factors included and their weights made publicly available, as is the case in the Point Scale, there is greater transparency regarding the basis for decisions than if an individual actor such as a judge is making discretionary decisions. The Connecticut State University research team also supported transparency in regard to disparity by analyzing score outcomes by race and ethnicity (Hedlund 2015). Specifically, that analysis looked at the distribution of white, black, and Hispanic defendants who scored below zero (recommended for bond), or at zero or above (recommended for release on promise to appear). On this metric, Point Scale results after the recalibration are largely equal across the groups. However, as discussed in the previous section, the distribution of scores across the Point Scale continuum is incorporated into the Financial Bond Guide. Therefore, assessing predictive equity for the full range of Point Scale risk scores is important.

In order to better understand any potential issues with the Point Scale and racial and ethnic disparities, and in line with the transparency necessary on this critical issue, further analysis of the Point Scale should at a minimum provide detail on the distribution of the individual risk factors in the tool by racial and ethnic groups and the distribution of risk scores by racial and ethnic groups. The former is a first step in determining whether there are factors in the tool that have disparity-enhancing effects on scores. A second analytical step related to any such factors could be validation of a model that does not include them, so that the tradeoff between equity and predictive performance can be clearly considered. Finally, we have already recommended calculation of the AUC statistic for the Point Scale, and further recommend that this be done separately by race and ethnicity to assess predictive equity, or that the tool performs equally well predicting outcomes for members of each group. In so doing, the Commission can also examine the extent to which racial and ethnic groups are subject to different prediction errors. It is not desirable to overestimate risk for one racial or ethnic group while underestimating risk for another. This second analytic step recommended above can help assess that.

⁴ The CCI researchers developed a predictive model for pretrial outcomes based on available data as an exploratory analysis: they were not testing a tool in use or intended to be put in use in New York.

Going beyond the tool itself to consider the pretrial decision-making structure, it is important to collect and publish data on release outcomes by risk level and race and ethnicity, so that internal and public stakeholders can assess how the tool specifically and larger system reforms generally are performing on disparity-reduction goals.

In summary, further data analysis is necessary in order to fully understand how the Point Scale, in whole and in its specific items, interacts with racial and ethnic disparities in Connecticut. This is the case under the status quo use of the tool, and it is important to develop and publicize performance measures on equity in prediction and equity in pretrial outcomes going forward as Connecticut considers and adopts pretrial reforms generally.

Summary Recommendations

We have made recommendations relative to the scope of this report throughout. In this section, we summarize them for ease of use.

- **Revalidation in a post-bond context.** Should such a change occur in Connecticut and the Point Scale instrument be retained, it will be important to conduct an updated validation study on the Point Scale, which may suggest other modifications to the risk prediction model. We further recommend including outcomes for all people released pretrial in that validation analysis, not just those released on a promise to appear.
- **Validation to allow for comparison with performance of other tools.** The AUC statistic is the common metric for comparatively assessing the predictive validity of risk assessment tools. Including calculation of the AUC statistic would allow Connecticut to compare the predictive performance of the Point Scale in a way that was not possible in our review. We further recommend sub-analyses be done by race and ethnicity to assess predictive equity, or that the tool performs equally well predicting outcomes for members of each group.
- **Predicting multiple pretrial outcomes with a single tool.** It is worth examining whether predictive performance for FTA and rearrest would be improved by creating separate models predicting each from the factors available in the Point Scale, or otherwise available through administrative data. However, even if improvement is realized by doing so, a second decision must be made about whether the added complication of providing multiple scores would have a net negative effect on pretrial decision-making in Connecticut.
- **Revision to structured decision-making.** Connecticut stakeholders planning for a no-bond context should consider repurposing the Financial Bond Guide that already combines consideration of risk, charge severity, and mitigating/aggravating factors, substituting the preferred pretrial decision for the bond amounts currently in the tool. If this repurposing is not pursued, a new pretrial decision-making matrix incorporating risk level and other factors should be developed and tested.
- **Risk and preventive detention.** Even the high-risk individuals as assessed by the Point Scale avoid pretrial misconduct more often than not. This suggests that if preventive detention becomes an option in Connecticut, it should require more than even the highest risk Point Scale score to result in this outcome.
- **Analysis of the Point Scale relative to equity.** Further analysis of the Point Scale should at a minimum provide detail on the distribution of the individual risk factors in the tool by racial and ethnic groups, and on the distribution of risk scores by racial and ethnic groups.

- **Performance measurement on equity and disparity.** Going beyond the tool itself to consider the pretrial decision-making structure, it is important to collect and publish data on release outcomes by risk level and race and ethnicity, so that internal and public stakeholders can assess how the tool specifically and larger system reforms generally are performing on disparity-reduction goals.

References

- Baird, Chris, Theresa Healy, Kristen Johnson, Andrea Bogie, Erin Wicke Dankert, Chris Scharenbroch. 2013. *A Comparison of Risk Assessment Instruments in Juvenile Justice*. Oakland, CA: National Council on Crime and Delinquency.
- Berk, Richard, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. 2017. *Fairness in Criminal Justice Risk Assessments: The State of the Art*. arXiv Preprint.
- Connecticut Sentencing Commission. 2017. *Report to the Governor and the General Assembly on Pretrial Release and Detention in Connecticut*. New Britain, CT: Connecticut Sentencing Commission. http://ctsentencingcommission.org/wp-content/uploads/2018/04/Pretrial_Release_and_Detention_in_CT_2.14.2017.pdf.
- Corbett-Davies, Sam, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. 2017. *Algorithmic Decision Making and the Cost of Fairness*. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM.
- Danner, Mona J.E., Marie VanNostrand, and Lisa M. Spruance. 2015. *Risk-Based Pretrial Release Recommendation and Supervision Guidelines: Exploring the Effect of Officer Recommendations, Judicial Decision-Making, and Pretrial Outcome*. St. Petersburg, FL: Luminosity, Inc.
- Dawes, Robyn M., David Faust, and Paul E. Meehl. 1989. *Clinical Versus Actuarial Judgment*. *Science* (243): 1668-1674.
- DeMichele, Matthew, Peter Baumgartner, Michael Wenger, Kelle Barrick, Megan Comfort, and Shilpi Misra. 2018. *The Public Safety Assessment: A Re-Validation and Assessment of Predictive Utility and Differential Prediction by Race and Gender in Kentucky*. Research Triangle Park, NC: RTI International.
- Department of Justice. 2014. *Attorney General Eric Holder Speaks at the National Association of Criminal Defense Lawyers 57th Annual Meeting and 13th State Criminal Justice Network Conference*. <https://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-national-association-criminal-defense-lawyers-57th>.
- Desmarais, Sarah L., and Evan M. Lowder. 2019. *Pretrial Risk Assessment Tools: A Primer for Judges, Prosecutors, and Defense Attorneys*. Chicago: John D. and Catherine T. MacArthur Foundation.
- Gottfredson, Stephen D., and Laura J. Moriarty. 2006. *Clinical Versus Actuarial Judgments in Criminal Justice Decisions: Should One Replace the Other?* *Federal Probation*, 70(15).
- Hamilton, Zachary, Melissa A. Kowalski, Alex Kigerl, and Douglas Routh. 2019a. *Optimizing Youth Risk Assessment Performance: Development of the Modified Positive Achievement Change Tool in Washington State*. *Criminal Justice and Behavior*.
- Hamilton, Zachary, Melissa A. Kowalski, Roger Schaefer, and Alex Kigerl. 2019b. *Recrafting Youth Risk Assessment: Developing the Modified Positive Achievement Change Tool for Iowa*. *Deviant Behavior*.

Hedlund, Jennifer. 2015. *Bail Revalidation: An Analysis of Weighted Release Criteria and Bail Points*. Wethersfield, CT: Judicial Branch, Court Support Services Division.

Hedlund, Jennifer, and Kathleen Bantley. 2009. *The Development of Guidelines for Financial Bond Recommendations*. Wethersfield, CT: Judicial Branch, Court Support Services Division.

Knight, Raymond A., and David Thornton. 2007. *Evaluating and Improving Risk Assessment Schemes for Sexual Recidivism: A Long-Term Follow-Up of Convicted Sexual Offenders*. Washington, DC: U.S. Department of Justice.

Leadership Conference on Civil and Human Rights. 2018. *The Use of Pretrial "Risk Assessment" Instruments: A Shared Statement of Civil Rights Concerns*. Washington, DC: The Leadership Conference on Civil and Human Rights.

Lowenkamp, Christopher T., Richard Lemke, and Edward Latessa. 2008. "The development and validation of a pretrial screening tool." *Federal Probation*, 72(3), 2-9.

New Jersey Courts. 2019. *Annual Report to the Governor and the Legislature on Criminal Justice Reform, 2018*. Trenton, NJ: Administrative Office of the Courts.

Picard, Sarah, Matt Watkins, Michael Rempel, and Ashmini G. Kerodal. *Beyond the Algorithm: Pretrial Reform, Risk Assessment, and Racial Fairness*. New York: Center for Court Innovation.

"Public Safety Assessment: Risk Factors and Formula." n.d. Public Safety Assessment. Laura and John Arnold Foundation. <https://www.psapretrial.org/about/factors>.

Skeem, Jennifer L., and Christopher Lowenkamp. 2015. *Risk, Race, & Recidivism: Predictive Bias and Disparate Impact*. *Criminology* (54).

Stevenson, Megan. 2018. *Assessing Risk Assessment in Action*. Minneapolis: Minnesota Law Review.

The University of Toledo College of Law. n.d. "The Ohio Risk Assessment System (ORAS)." The University of Toledo Law Review. The University of Toledo. https://www.utoledo.edu/law/studentlife/lawreview/pdf/Trout_ORAS-Overview.pdf.

Thornton, David, and D. Richard Laws. 2009. *Cognitive approaches to the assessment of sexual interest in sexual offenders*. Hoboken, NJ: Wiley-Blackwell.

Virginia Department of Criminal Justice Services. 2018. *Virginia Pretrial Risk Assessment Instrument – VPRAI. Instruction Manual - Version 4.3*. Virginia Department of Criminal Justice Services.