



# Advancing Pretrial Policy and Research (APPR)

Presented to the State of Washington  
Sentencing Guideline Commission

October 9, 2020

# APPR Project Tasks



## Data Management

Assessing Capacity  
Ensuring Quality  
Maintaining Security



## Validation

Testing Locally  
Assessing Bias  
Sharing Results



## Implementation

Assessing Fidelity  
Understanding Context  
Measuring Impacts



## Research

Studying PSA Use  
Understanding Decision Making  
Measuring Costs



## Dissemination

Reports  
Webinars  
Infographics

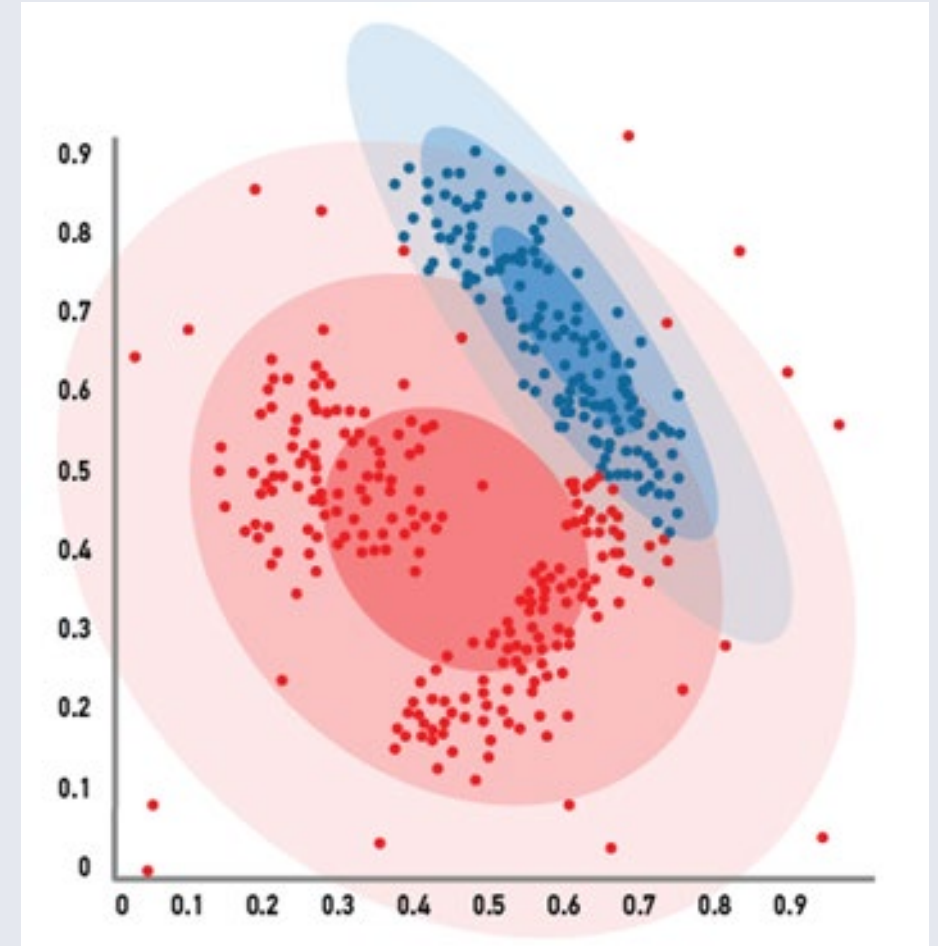
How well does the **current PSA** work in your jurisdiction?

Can a **localized PSA** improve upon the current PSA?

 **Validation**

# What is validation?

- A process to determine the accuracy of the PSA for a local population and context
- Does the assessment measure what it is intended to measure?
- Assess how well the PSA matches data from your county.

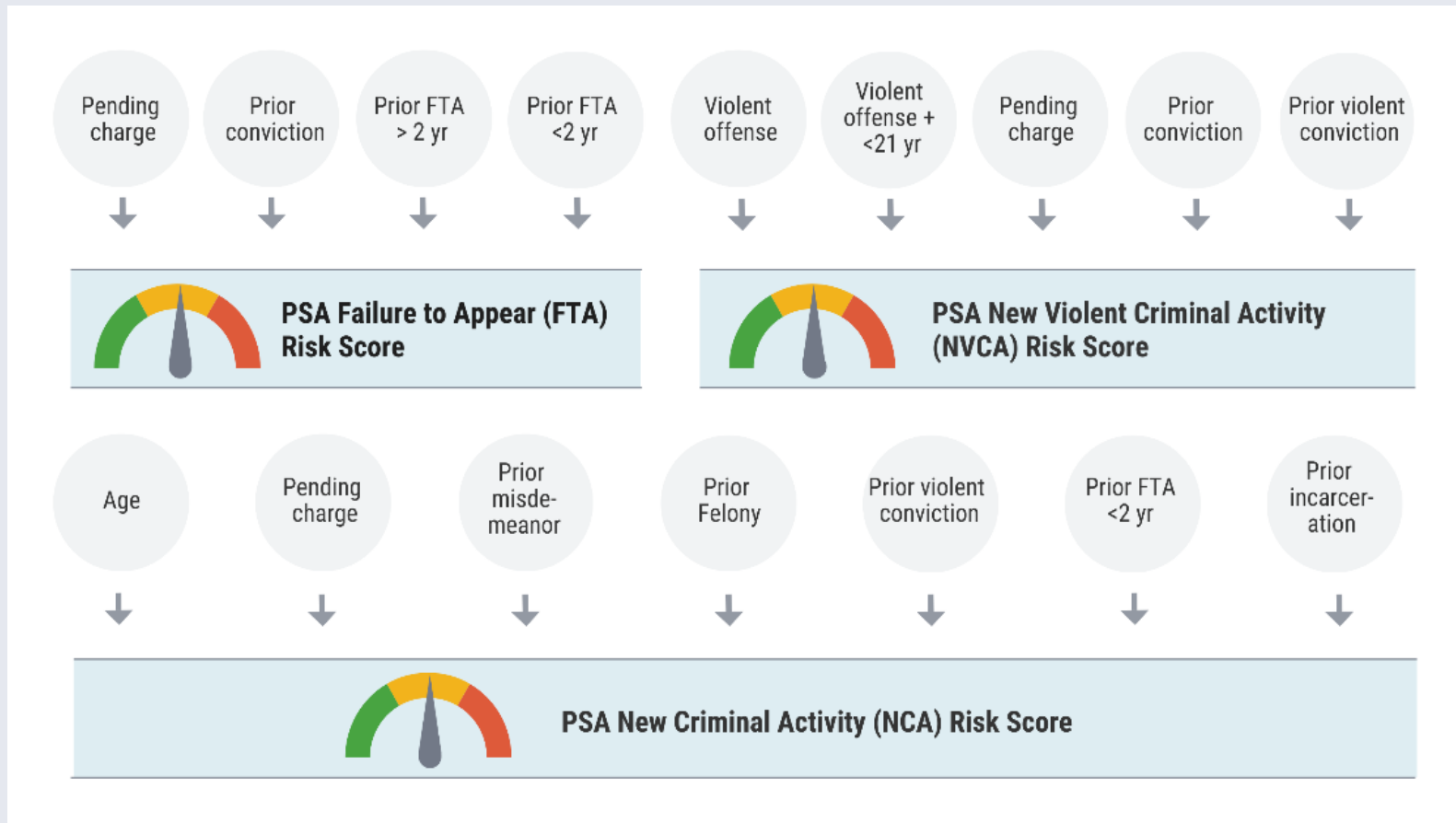






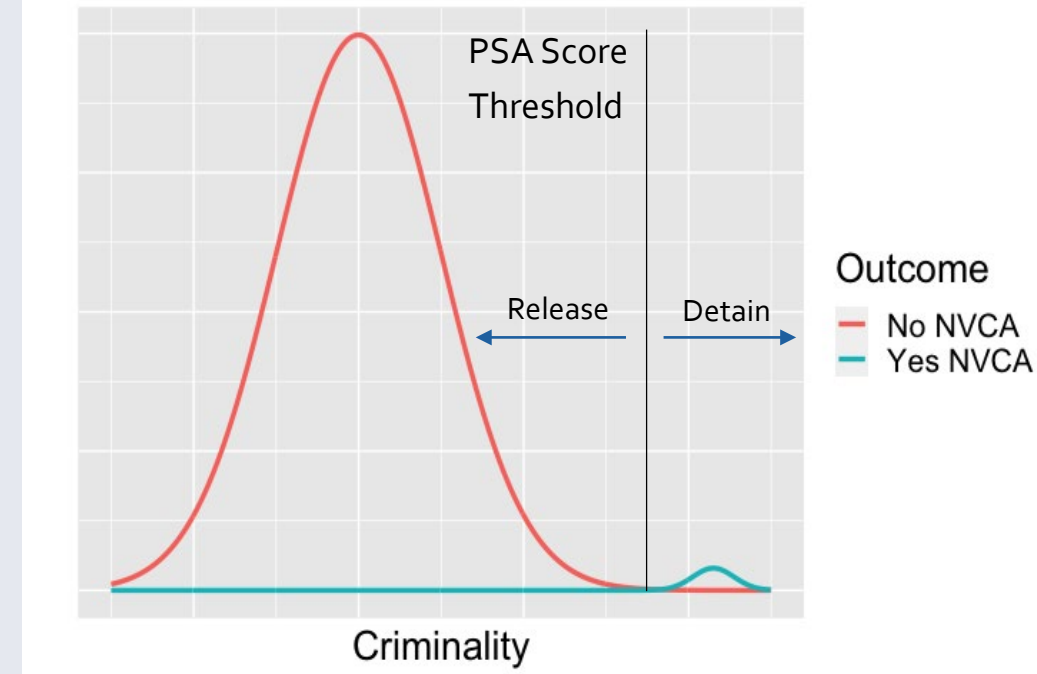
# Public Safety Assessment (PSA)

Assessments are made by scoring the PSA factors with a model



# Validation Trade-offs

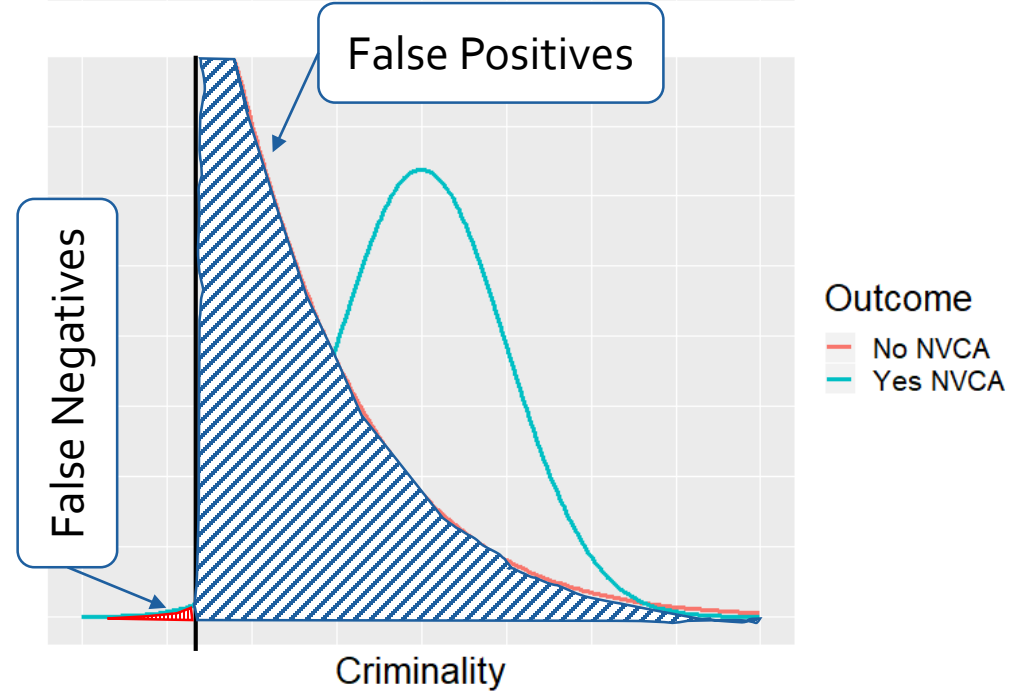
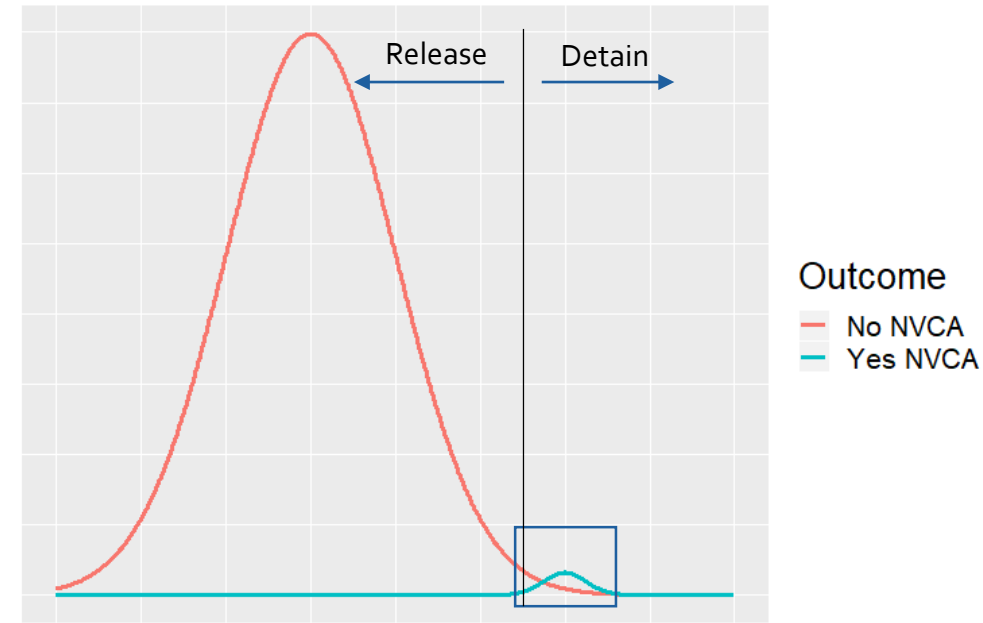
- There is a trade off in the types of errors.



# Validation Trade-offs

- There is a trade off in the types of errors.

		Predicted	
		NVCA	No NVCA
Actual	NVCA	40	<b>10</b>
	No NVCA	<b>1000</b>	9000

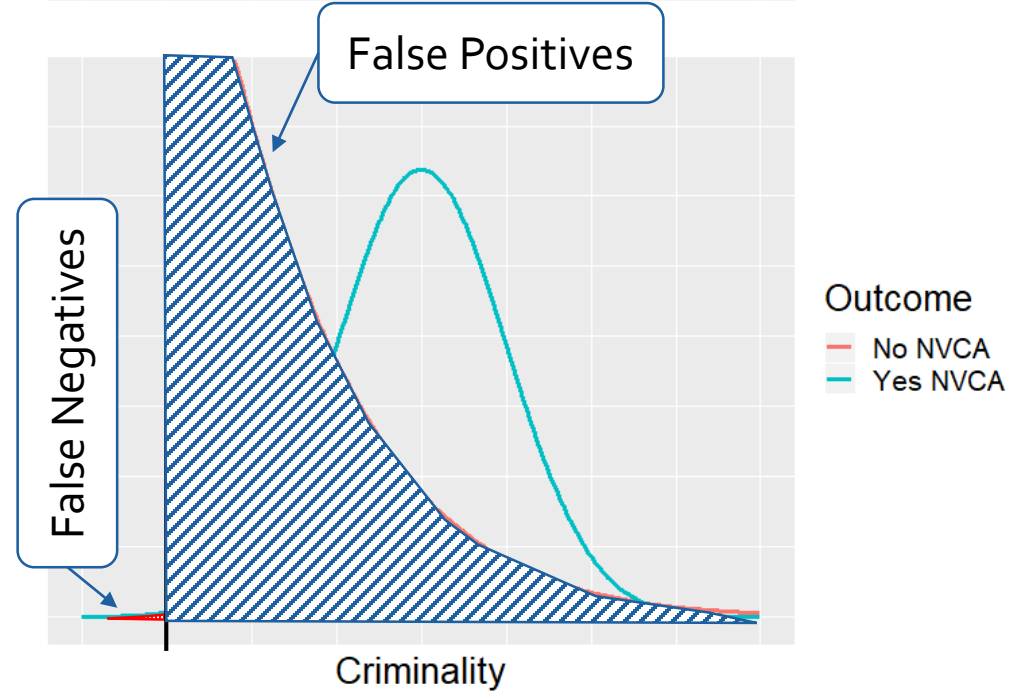
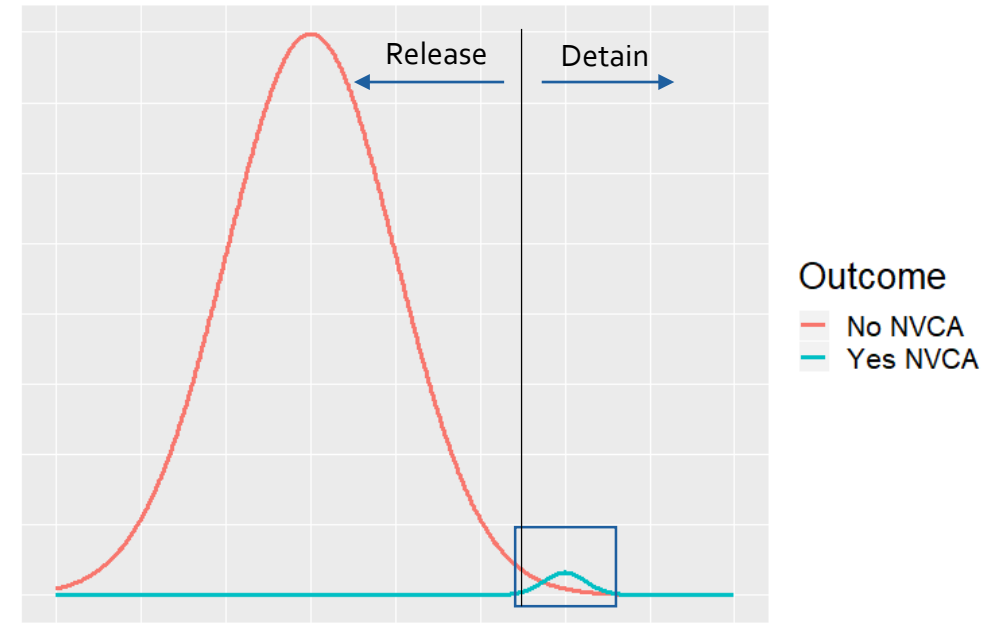




# Validation Trade-offs

- There is a trade off in the types of errors.
  - Move decision left to catch more new violent crimes?

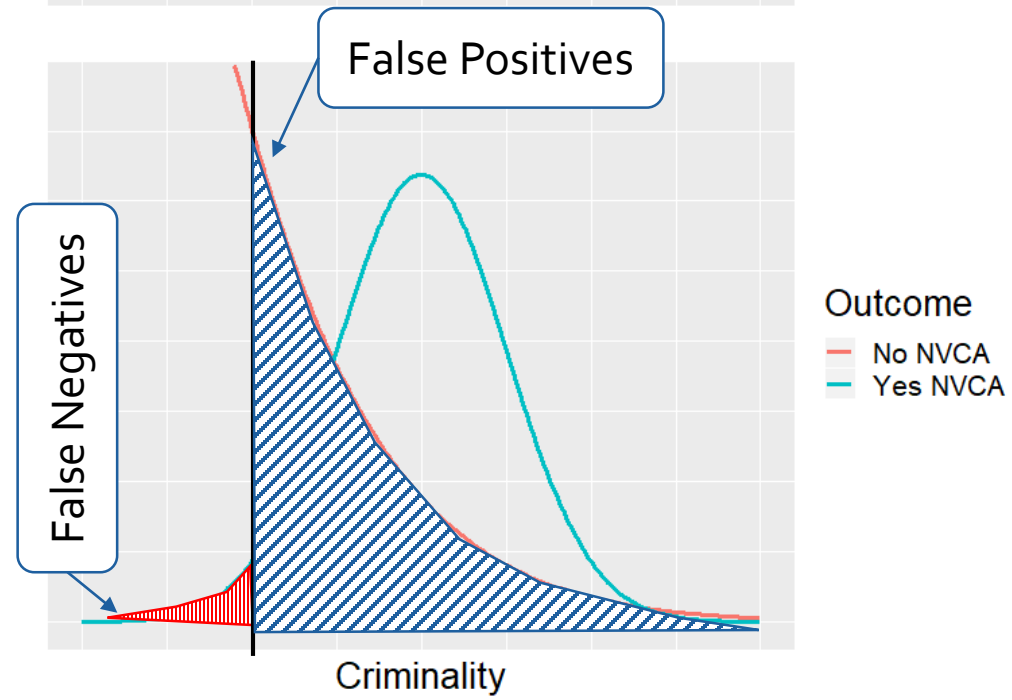
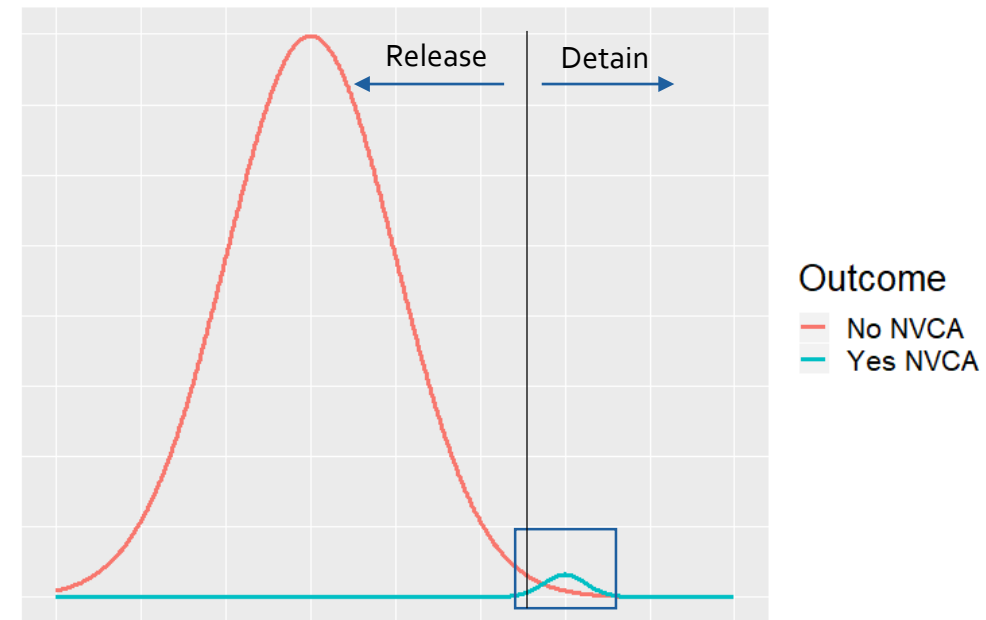
		Predicted	
		NVCA	No NVCA
Actual	NVCA	45	<b>5</b>
	No NVCA	<b>1500</b>	8500



# Validation Trade-offs

- There is a trade off in the types of errors.
  - Move decision left to catch more new violent crimes?
  - Move decision right to prevent unjust detention?

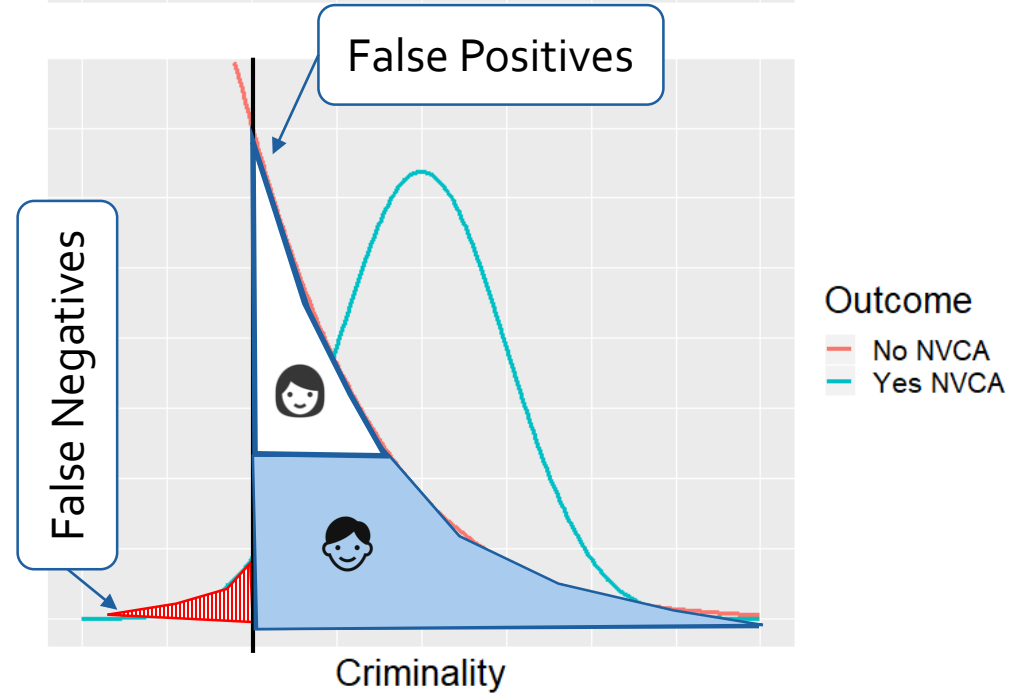
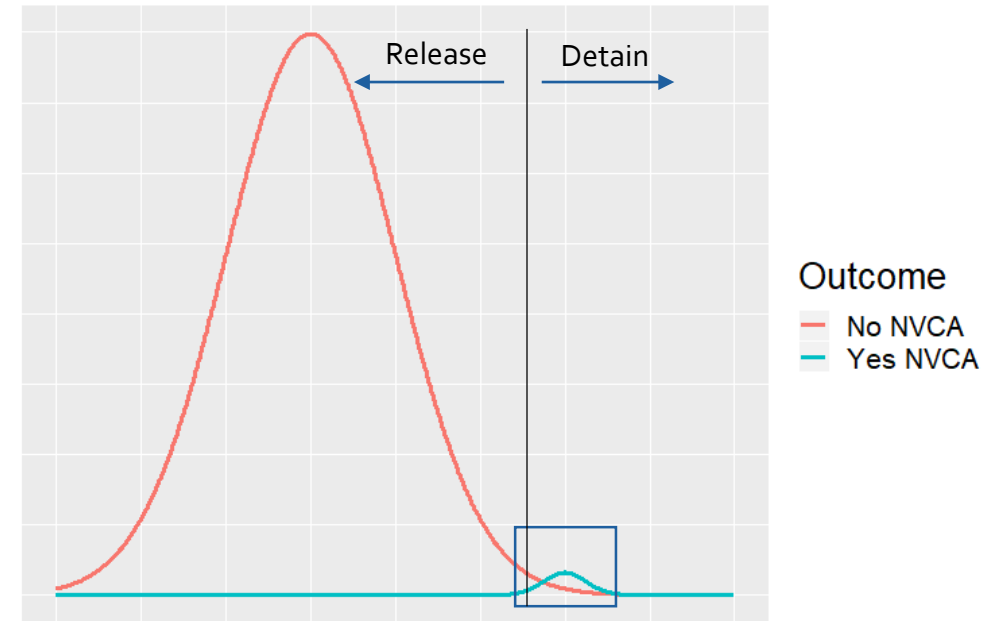
		Predicted	
		NVCA	No NVCA
Actual	NVCA	35	<b>15</b>
	No NVCA	<b>500</b>	9500



# Validation Trade-offs

- There is a trade off in the types of errors.
  - Move decision left to catch more new violent crimes?
  - Move decision right to prevent unjust detention?
- Questions of bias:
  - Does the PSA give similarly accurate predictions for different groups?
  - Or does the PSA systematically underpredict one group over another?

	False Positives	Total	Percent
Male	400	9500	4.2%
Female	100	550	18.2%
	<b>500</b>	<b>10500</b>	<b>4.8%</b>



# Assessing Predictive Bias

- Predictive bias
  - › "A situation in which an examination is used to predict a specific criterion for a particular population, and is **found to give systematically different predictions for subgroups of the population** who are identical on that specific criterion"
  - › Switon (1981). *Predictive Bias in Graduate Admissions Tests*. ETS
- There are several ways to assess predictive bias
- Difficult to satisfy multiple measures of predictive bias simultaneously
- Does the assessment provide equal probabilities of FTA, NCA, and NVCA for PSA scores regardless of race or sex?

*Figure 1. Predicted probabilities of failure to appear by Public Safety Assessment (PSA) FTA score between whites and blacks*

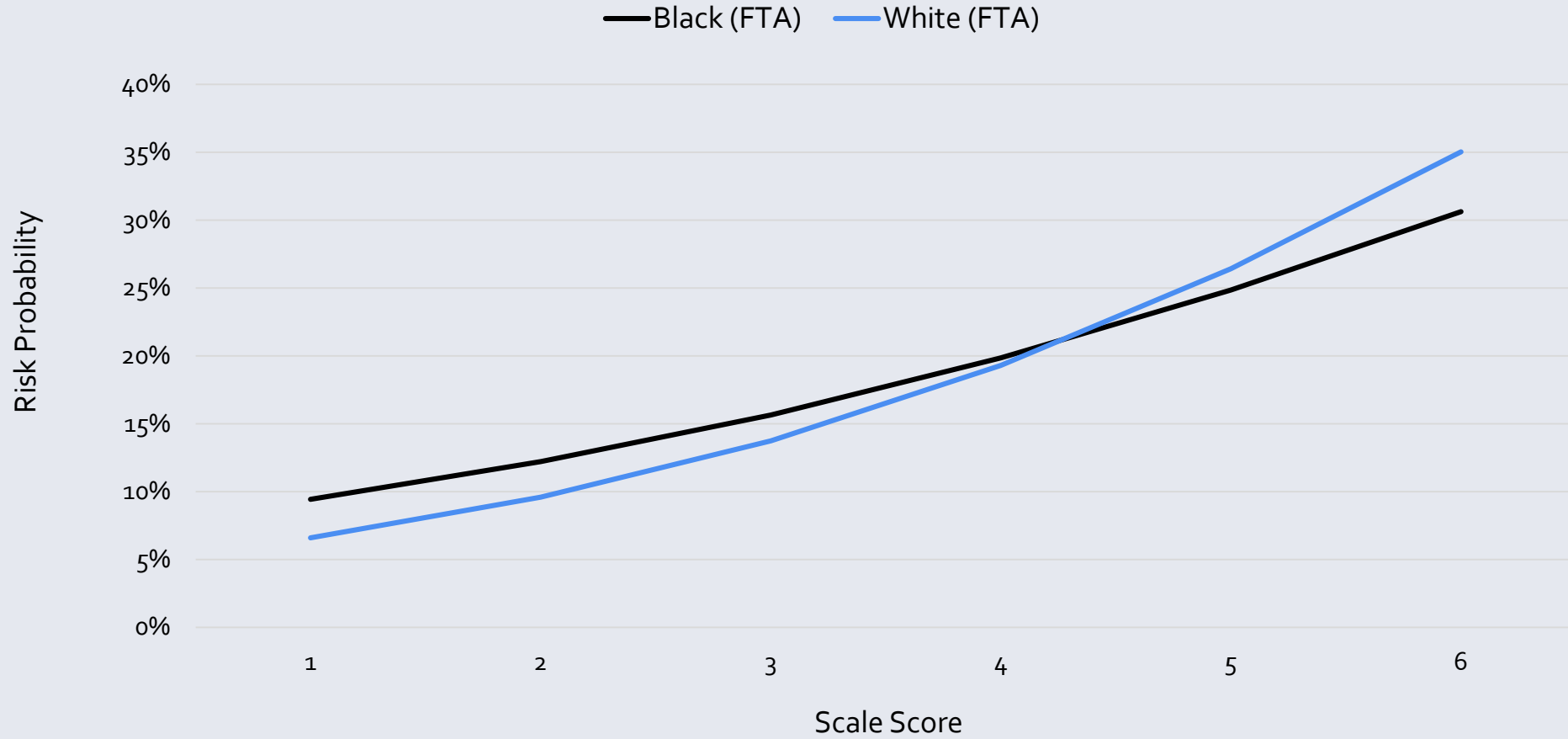
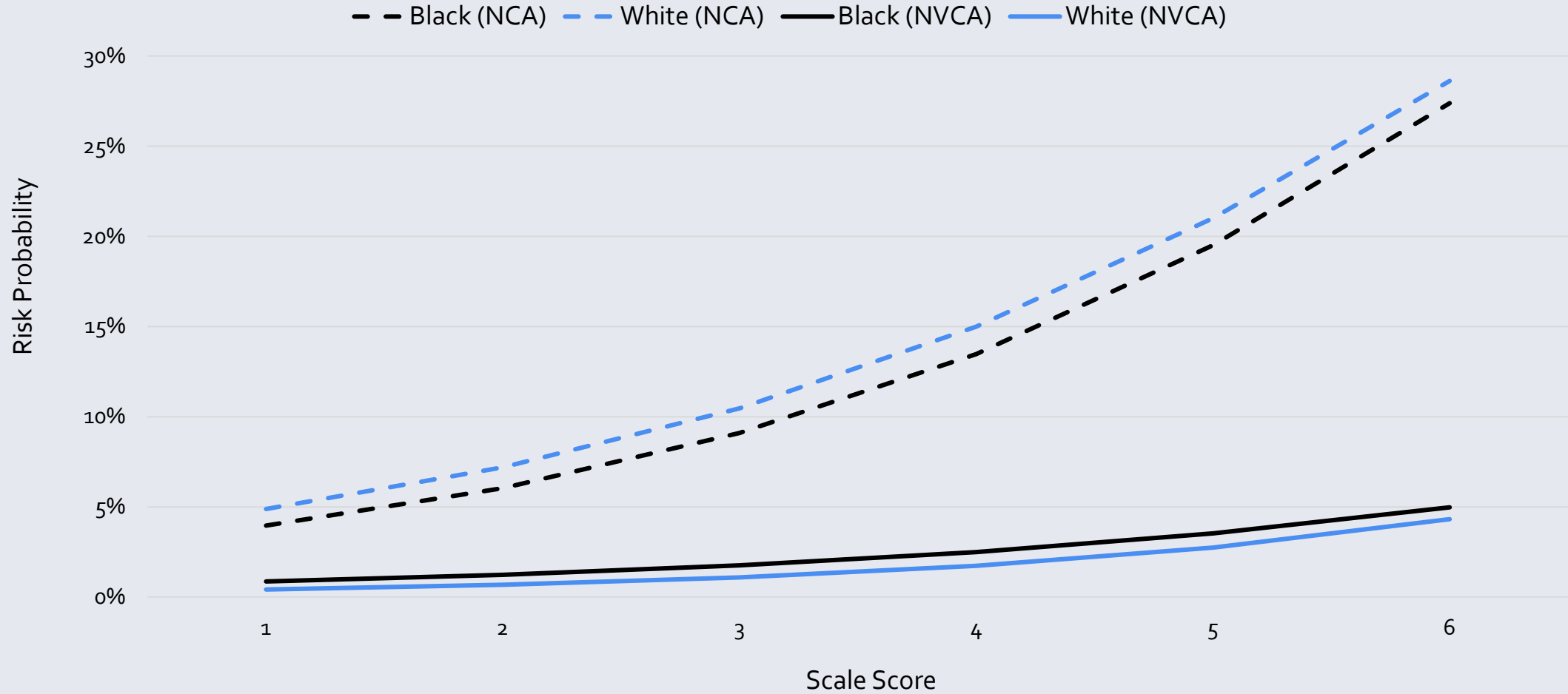


Figure 2. Predicted probabilities of pretrial NCA and NVCA by Public Safety Assessment (PSA) NCA and NVCA scores between whites and blacks





# RTI's PSA Validation Plan



Validation master plan and site-specific plans



Validation of PSA scoring algorithm



Validation of site-specific PSA scoring algorithms



Scheduled and Selective Revalidation



Reproducibility and Archiving

# Validation Plans



- Validation master plan uploaded to the Open Science Foundation (OSF) website
  - <https://osf.io/smz5v>
- Site-specific addendums will be made if there are data idiosyncrasies that need to be accommodated

## Validation Master Plan: *Table of Contents*

1	Introduction .....
1.1	The PSA Decision Process.....
1.2	Opportunities for Improving the Current PSA Scoring Formula.....
1.3	Primary Goals: PSA Validation and Localization.....
1.4	Societal Costs of the False Positive/False Negative Tradeoff.....
2	Validation, Fairness and Bias .....
2.1	Validation.....
2.2	Fairness, Bias, and Classification Errors .....
3	Version Control and Workflow .....
4	Analysis Plan Preregistration .....
5	Data and the Data Management Plan.....
6	Descriptive Statistics.....
7	PSA Scoring .....
8	PSA Validation Analysis Plan .....
8.1	Validity .....
8.2	Fairness and Bias.....
9	Missing Outcome for Detained Defendants .....
10	Localized PSA Analysis Plan.....
10.1	Local Factor Weight Estimation .....
10.2	Fitting Machine Learning Models .....
10.3	Comparing Localized PSA to Standard PSA Scoring .....
10.4	Local Assessment of Validity, Fairness and Bias.....
10.5	Revalidation Frequency .....
10.6	Localized PSA Scoring Formula and the Release Conditions Matrix .....

# Validation of Current PSA Scoring Algorithm



- Detailed output summarizing how the current PSA scores recent cases in your jurisdiction
- Includes:
  - › Descriptive statistics
    - Prior to analysis, what does my jurisdiction's data look like?
  - › Model validity metrics
    - How do the predictions of the outcomes (FTA, NCA, NVCA) differ from the actual outcomes?
  - › Measures of statistical bias
    - Does my model impact subgroups differently?

# Localized PSA: Site-Specific Scoring Algorithms



- Use your jurisdiction's data to *develop* models
- Assess different model types, with the goal of reducing predictive bias
- Determine site PSA classification goals and trade-offs
  - Work with sites to determine how they value false positives vs. false negatives
- Communicate and educate about the final scoring algorithm to ensure understand and transparency

# Revalidation



## Scheduled revalidation of localized PSA

- › Currently recommending biannual revalidation
- › Compensate for changes in implementation over time
- › Final timing recommendations will account for site capacity and expected rate of system improvement



## Selective revalidation of localized PSA

- › Ensure robustness in response to system changes/shocks
- › Policy change
- › COVID19
- › Establish guidelines for when to engage in select revalidation

# Reproducibility and Archiving

- An R package is being developed to run the validation analyses
  - › R is an open source language and environment for statistical computing
  - › Using the R package ensures reproducibility
- Data, validation report, and R package will be archived on OSF and shared with AV
- The R package could also be used by jurisdictions for revalidation local PSA data