

Risk assessment and racial bias

Alexandra Chouldechova

Assistant Professor, Statistics and Public Policy
Heinz College, Carnegie Mellon University

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

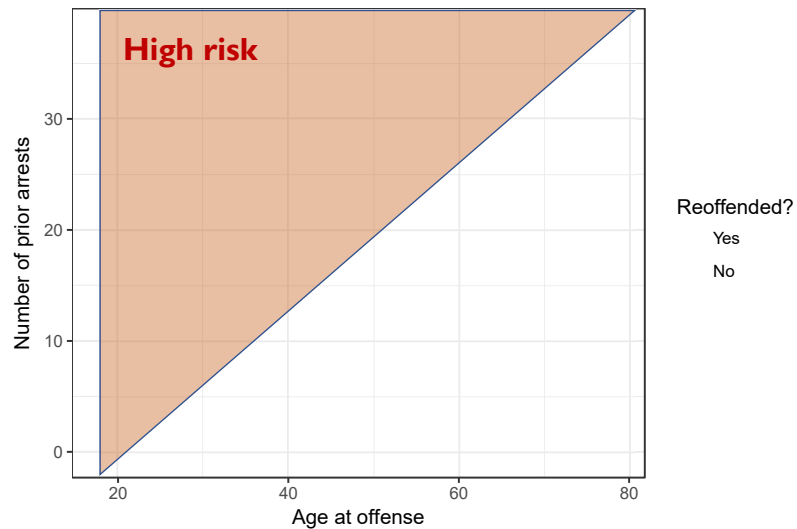


Risk assessment

A brief history and taxonomy

- Dates back to the **1920s** in the work of Ernest Burgess, who developed a tool to predict recidivism risk for offenders released in Illinois
- Early versions were based on **clinical judgment**
 - Take the factors that experts believe to be associated with reoffense risk
 - Develop a scoring system that tallies up risk factors
- **Actuarial instruments** (based on statistical models or “machine learning”) produce scores that have higher predictive accuracy

Actuarial instruments



First generation

- Unstructured
- Based on “clinical” judgment

Second generation (STATIC-99)

- Actuarial tools (statistical models)
- Based on static factors (E.g., age, criminal history)

Third generation (LSI-R)

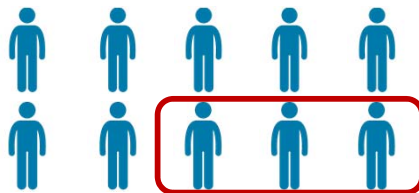
- Mix of Actuarial and clinical judgment
- Consider static and dynamic items (e.g., attitudes, behavioral health)
- Risk-Needs assessment

Fourth generation (COMPAS, ORAS)

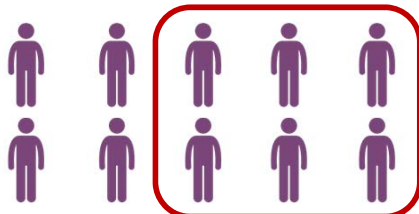
- Integrates case planning and risk management
- Helps guide decisions about interventions and supervision

Predictive (racial) bias



- Experts can be biased in their assessments
 - E.g., overestimate risk for some groups relative to others
- Risk assessment models can have the same issue
- An **unbiased** tool would predict reoffense likelihood with **equal accuracy across groups**
- **Immediate problems:**
 - There are many different ways of measuring **accuracy**
 - We observe **rearrest**, not **reoffense**
 - Are we happy with an unbiased prediction of **who gets caught**?
- Let's look at an example.



More individuals from the **Purple** group get flagged as **high risk**.

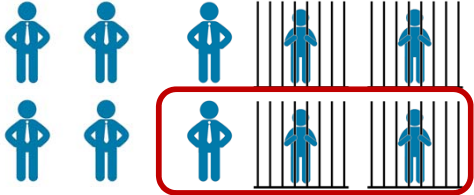


Does this mean that the risk assessment tool is **biased**?

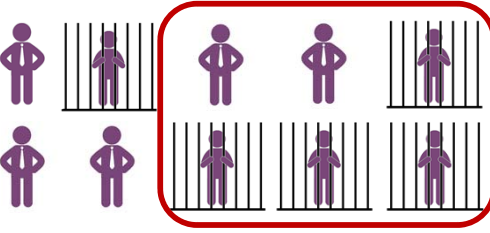
Reoffends
Doesn't reoffend

Let's look at outcomes.



Recidivism rate: 40%

Blue group has lower recidivism rate (4/10 vs. 5/10) .




Recidivism rate: 50%

Reoffense rate among those flagged as high risk is the same across groups:
 $2 / 3 = 67\%$ vs.
 $4 / 6 = 67\%$

This is called **predictive parity**

PPV	67%	67%
------------	-----	-----

PRO PUBLICA

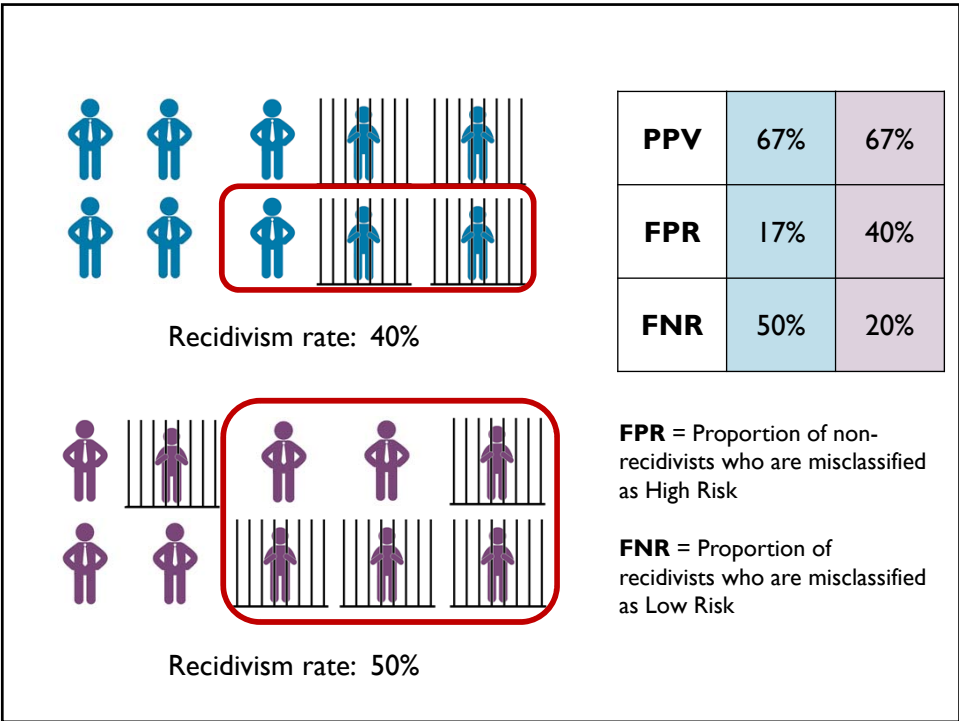


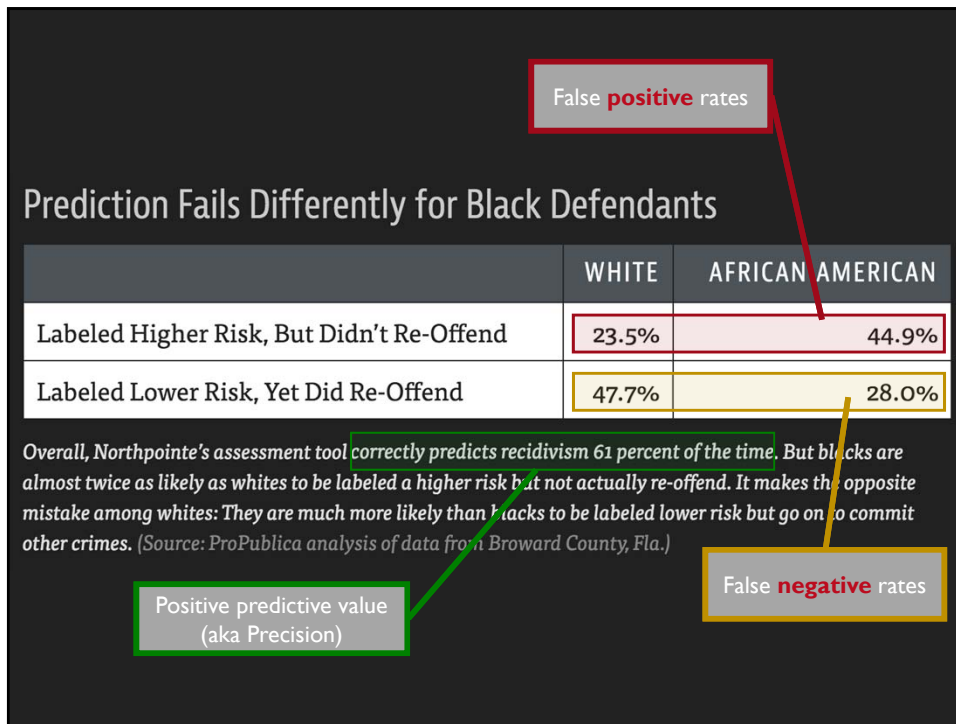
Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

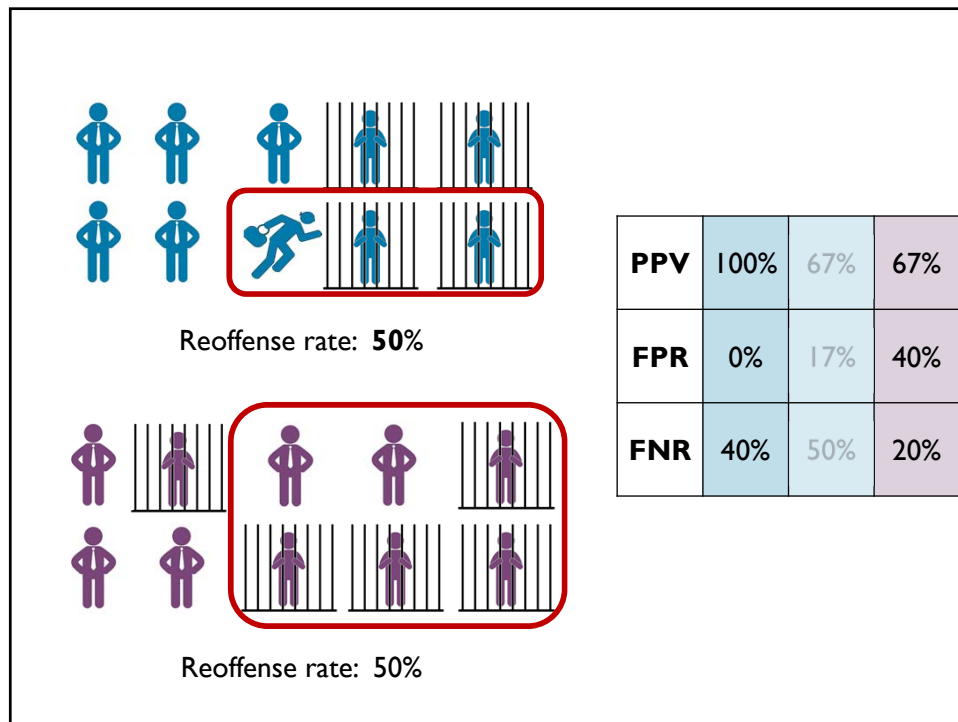
Source:
Julia Angwin,
Jeff Larson,
Surya Mattu and
Lauren Kirchner, ProPublica

Machine Bias

There's software used across the country to predict future criminals.
And it's biased against blacks.







Disentangling concerns

- Risk assessment tools make mistakes
 - So do human assessors
- When thinking about whether the tool could be useful, it can help to step away from issues of model inaccuracy
- Try the “Oracle test”

What questions remain?



Risk assessment tool



Oracle



Omitted objective bias

Any questions that remain may help clarify concerns about:

- **Choice of target variable**
 - Is rearrest the right target?
- **Disconnect between prediction target and decision criteria**
 - Should future dangerousness or failure to appear risk factor into bail decisions?
- **Explainability**
 - Is it enough to know that the individual is high risk, or do you also need to know why?
- **Effects of interventions**
 - What interventions are at your disposal to reduce risk, and are any of them likely to be effective for the given individual?



Oracle



**Fairness is
a process
property**

An unbiased risk assessment tool may lead to biased outcomes and may have disparate impact depending on how it is used.

Thank you.

Carnegie Mellon University
HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT
achould@cmu.edu

Some additional resources

<https://csgjusticecenter.org/wp-content/uploads/2014/07/Risk-Assessment-Instruments-Validated-and-Implemented-in-Correctional-Settings-in-the-United-States.pdf>

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/?utm_term=.0a8059546911

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2826600