Evaluating Pre-trial Programs Using Interpretable Machine Learning Matching Algorithms for Causal Inference

Travis M. Seale-Carlisle*[1,2], Saksham Jain*[3], Courtney Lee[2], Caroline Levenson[2], Swathi Ramprasad[2], Brandon L. Garrett[2], Sudeepa Roy[3], Cynthia Rudin[3], and Alexander Volfovsky[3]

[1] University of Aberdeen

[2] Wilson Center for Science and Justice, Duke University

[3] Department of Statistical Science, Duke University

*contributed equally

Arrested Charged   **Pretrial Stage**

Case
Disposed

Time

During the pretrial stage, people may be:

- Detained in jail
- Released to the community without supervision
- Released to the community with supervision

Which option is best for *low-risk* individuals?

Option #1: Detaining low-risk people in jail

Compared to those who face similar charges, have similar case histories, and are released, people detained pretrial are more likely to

- miss crucial court appointments in the future (Lowenkamp et al., 2013; Stevenson, 2018)

- be convicted in the future (Dobbie et al., 2018; Leslie & Pope, 2017; Stevenson, 2018),

- receive harsh sentences in the future (Heaton et al., 2017; Leslie & Pope, 2017)

Which option is best for *low-risk* individuals?

- ~~Option #1: Detained in jail~~
- Option #2: Released to the community without supervision
- Option #3: Released to the community with supervision

What is a pretrial community supervision program?

Pre-trial programs may provide:
- Access to education
- Housing assistance
- Employment assistance
- Healthcare

In exchange for these services, pre-trial programs may require:
- Periodic check-ins with a case manager
- Drug testing and treatment
- Wearing an electronic monitor
- Taking part in therapy

These programs aim to _reduce new criminal activity_ and _failures to appear in court_ (e.g., Parmer & Merrit, 2019).

Some studies find that they do (e.g., Goldkamp & White, 2006; Lowenkamp & VanNostrand, 2013).
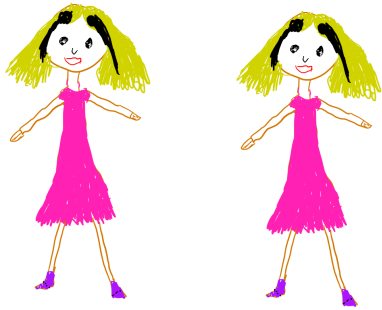
Others find that they do not (e.g., Cadigan & Lowenkamp, 2011; Robinson et al., 2011).

Bechtel et al., 2017 (review paper) says most studies on this topic are not rigorous (purely descriptive, not peer reviewed (!)).

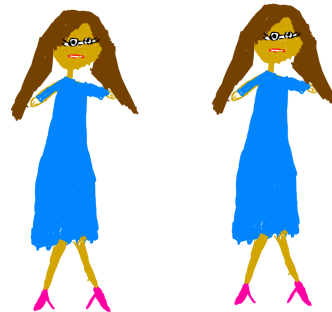Ideally, conduct RCT. Not possible here.
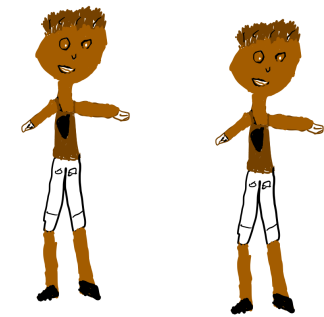
Instead do matching.

recidivated   didn't recidivate

control   treated

didn't recidivate   didn't recidivate

control   treated
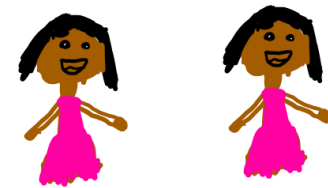
recidivated   didn't recidivate

control   treated
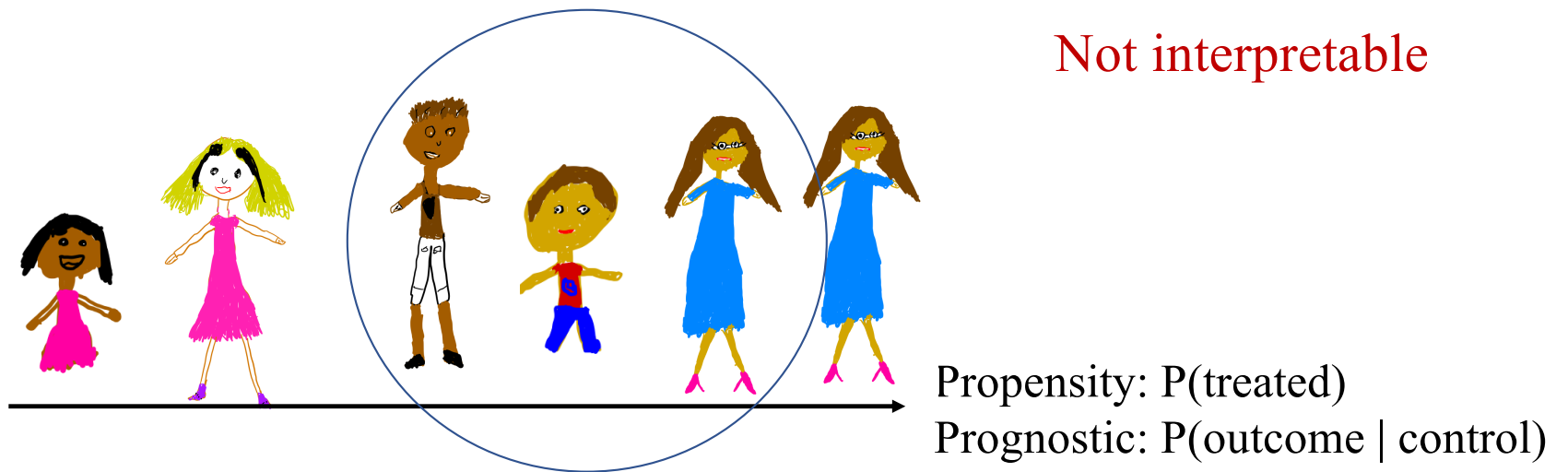
didn't recidivated   recidivated

control   treated

recidivated   recidivated

control   treated

Propensity score matching (Rosenbaum and Rubin, 1983) and
Prognostic score matching (Hansen, 2008) are not interpretable

Not interpretable

Propensity: P(treated)
Prognostic: P(outcome | control)

The Almost Matching Exactly Lab

Propensity Score Matching

- Creates *low quality* matched groups

- Requires analyst to pick variables for matching

- May provide inaccurate estimates of treatment effects
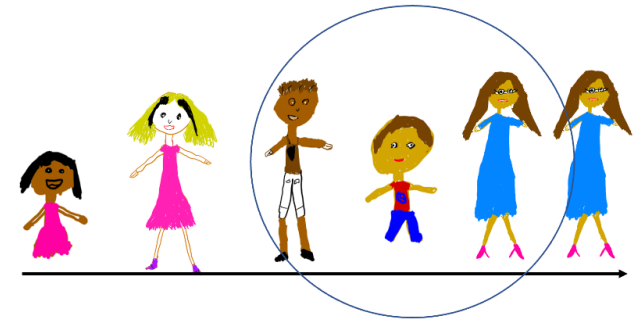
DAME-FLAME Matching

- Creates *high quality* matched groups

- Uses machine learning to find important variables for matching

- Provides accurate estimates of treatment effects
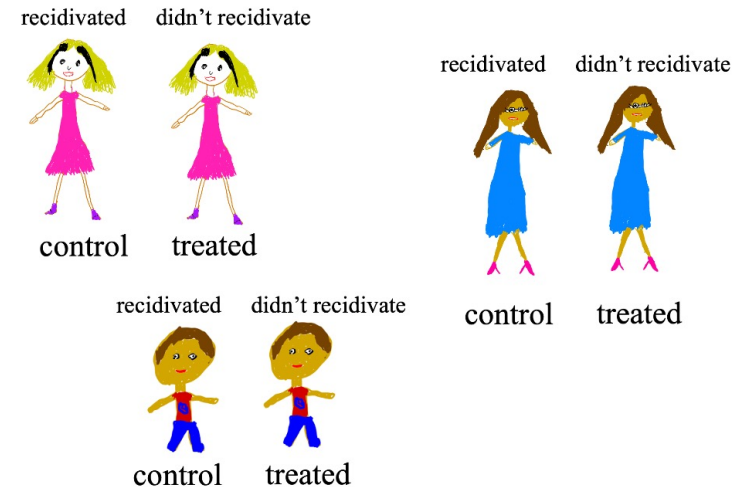
Our Study

Evaluate a pretrial program run by the Criminal Justice Resource Center (CJRC) in Durham, North Carolina.

Research question: Did the CJRC program reduce new criminal charges?

Treatment: CJRC pretrial supervision

Control: Pretrial release without any supervision or restriction

Depending on the individual, the CJRC offers:

- Access to education
- Housing assistance
- Employment assistance
- Healthcare

Depending on the individual, the CJRC requires:

- Periodic check-ins with a case manager
- Drug testing and treatment
- Wearing an electronic monitor
- Taking part in therapy

https://www.dconc.gov/county-departments/departments-a-e/criminal-justice-resource-center

Data

- CJRC (January 1st, 2016 – July 31st, 2019)
  - Contains demographic, employment, and criminal history information
- Durham Jail
  - Contains bond information
- ACIS (January 1st, 2016 – July 31st, 2020)
  - Contains case outcome, sentencing, and new criminal activity information

- 1,415 people recommended for CJRC programs
  - 687 of them took part (**Treatment**)
  - 728 people were recommended but did not take part (**Control**)
- Most people were Black (76%), male (67%), and 17-30 years old (60%)
- The most common charge with an assaultive misdemeanor (44%)
- Most people had their charges dismissed (82%); few were convicted (17%)

Treatment: Of 687 treated people, 295 had a new criminal charge (43%).

Control: Of 728 controls, 303 had a new criminal charge (42%).

Propensity score says no treatment effect: effect size -0.0375 [-0.517, 0.44].

Propensity Score Matched Group: Unit 924

Unit ID	treated	GENDER	RACE	CLASS CHARGE	YEARS AT RESIDENCE	VETERAN	PRIOR ADULT CONVICTIONS	AGE	VPRAI SCORE
Query: 924	1	Male	Black	H or I	≥ 1 year	no	0 - 5	18 - 20	3
1397	0	Female	Black	2 or 3	none	no	0 - 5	27 - 29	2
904	0	Female	Black	H or I	<1 year	no	0 - 5	18 - 20	2
242	0	Female	White	A1 or 1	≥ 1 year	no	0 - 5	35 - 39	0
275	0	Female	Black	2 or 3	<1 year	no	6 - 10	40 - 44	7
1006	0	Male	Black	H or I	<1 year	no	6 - 10	35 - 39	6

Annotations: Treated · Male · Black · Class H or I Felony · ≥1 year at residence · not veteran · 0-5 prior adult convictions · 18-20 yrs · VPRAI 3

DAME-FLAME Matched Group: Unit 924

Unit ID	treated	GENDER	RACE	CLASS CHARGE	YEARS AT RESIDENCE	VETERAN	PRIOR ADULT CONVICTIONS	AGE	VPRAI SCORE
Query: 924	1	Male	Black	H or I	≥ 1 year	no	0 - 5	18 - 20	3
25	0	Male	Black	H or I	≥ 1 year	no	0 - 5	27 - 29	3
170	0	Male	Black	H or I	≥ 1 year	no	0 - 5	24 - 27	3
182	0	Male	Black	H or I	≥ 1 year	no	0 - 5	24 - 27	3
289	1	Male	Black	H or I	≥ 1 year	no	0 - 5	45 - 49	3
322	1	Male	Black	H or I	≥ 1 year	no	0 - 5	21 - 23	3
373	0	Male	Black	H or I	≥ 1 year	no	0 - 5	27 - 29	3
408	1	Male	Black	H or I	≥ 1 year	no	0 - 5	24 - 27	3
411	1	Male	Black	H or I	≥ 1 year	no	0 - 5	21 - 23	3
471	0	Male	Black	H or I	≥ 1 year	no	0 - 5	35 - 39	3
512	0	Male	Black	H or I	≥ 1 year	no	0 - 5	30 - 34	3
591	1	Male	Black	H or I	≥ 1 year	no	0 - 5	27 - 29	3
744	0	Male	Black	H or I	≥ 1 year	no	0 - 5	24 - 27	3
748	0	Male	Black	H or I	≥ 1 year	no	0 - 5	21 - 23	3
749	0	Male	Black	H or I	≥ 1 year	no	0 - 5	27 - 29	3
914	1	Male	Black	H or I	≥ 1 year	no	0 - 5	21 - 23	3
985	1	Male	Black	H or I	≥ 1 year	no	0 - 5	21 - 23	3
1015	1	Male	Black	H or I	≥ 1 year	no	0 - 5	24 - 27	3
1039	1	Male	Black	H or I	≥ 1 year	no	0 - 5	21 - 23	3
1059	1	Male	Black	H or I	≥ 1 year	no	0 - 5	40 - 44	3
1130	0	Male	Black	H or I	≥ 1 year	no	0 - 5	24 - 27	3
1255	1	Male	Black	H or I	≥ 1 year	no	0 - 5	18 - 20	3
1323	1	Male	Black	H or I	≥ 1 year	no	0 - 5	24 - 27	3

DAME-FLAME vs. Propensity Score Matching
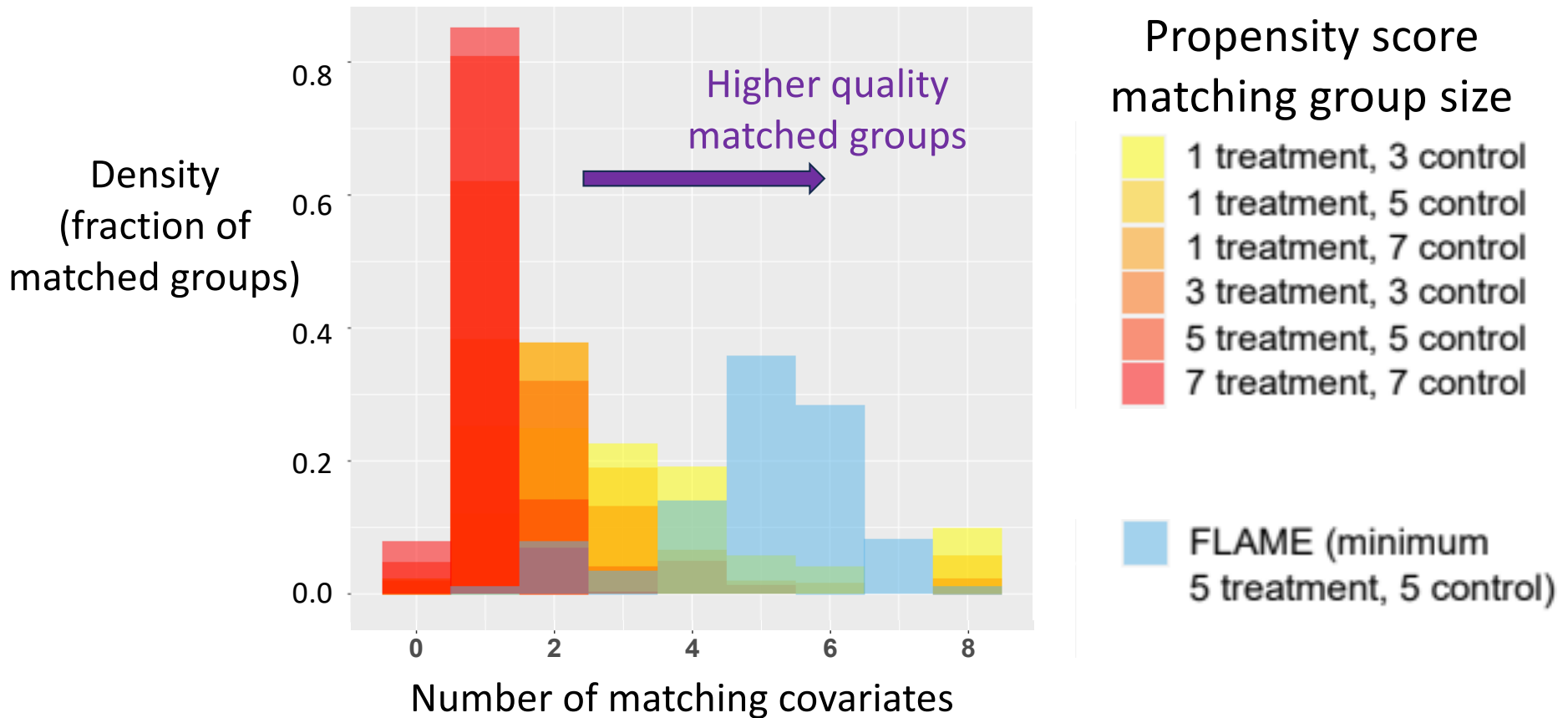
Result

- DAME-FLAME estimated average treatment effect of .018, with 95% CI overlapping zero. Pretrial program had no significant effect on new criminal charges.

- This result replicates Bechtel et al. (2017).

- Pretrial programs may need to consider alternative strategies to reduce new criminal activity. (Do more studies though.)

- DAME-FLAME is quite useful in evaluating pre-trial programs and can be useful in other areas in criminology where random assignment is not possible, ethical, or feasible.

Thanks