Advanced Statistical Methods for Observational Studies

LECTURE 03

basic tools

MATCHING TO MORE THAN ONE CONTROL

Design of Observational Studies: chapter 8.5

matching to more than one control

- 1:k matching
  - One treated to exactly k control (and no control to multiple treatments)
  - Mostly implemented because of simplicity
- Variable control matching
  - One treated to any number of controls
  - And no control to multiple treatments

matching to more than one control

- Example: NICU
  - I’m going to use the absolute difference of pscore because it’s easy to see the point I’m making.
  - Obviously, this is typically a harder problem.

matching to more than one control

absolute difference: $|\hat{e}(\boldsymbol{x}_i) - \hat{e}(\boldsymbol{x}_j)|$

obs	e^(x)
1	0.54
2	0.43
3	0.57
4	0.53
5	0.57
6	0.54
7	0.53
8	0.57
9	0.43
10	0.57

$$
\begin{array}{c|cccccc}
 & 5 & 6 & 7 & 8 & 9 & 10 \\
\hline
1 & 0.03 & 0.00 & 0.01 & 0.03 & 0.11 & 0.03 \\
2 & 0.14 & 0.11 & 0.10 & 0.14 & 0.00 & 0.14 \\
3 & 0.00 & 0.04 & 0.04 & 0.00 & 0.14 & 0.00 \\
4 & 0.05 & 0.01 & 0.00 & 0.04 & 0.09 & 0.04
\end{array}
$$

This is matching with a variable number of controls.

matching to more than one control

- How to think about it:
  - Within a matched set the treated person is compared to the average of the controls in the set.
  - This means the control “stand in” is more carefully estimated.
  - Under strongly ignorable treatment assignment, this will tend to produce better results in the form of more precision and power of the tests.

matching to more than one control

- 1:k vs variable control matching
  - 1:k is more readily put into traditional methods of inference.
  - Variable controls will tend to be better matched to the treated.

implementing: more than one control

How to do these kinds of matches: (see DOS pages 178-179)

```
library(optmatch)
fullmatch()
```

Within `fullmatch()` set `min.controls=1`.

`max.controls` = maximum number of controls per treated

`omit.fraction` = determines number of controls to use.

In SAS, you can use `proc assign`.

matching to more than one control

- References for how to analyze: see citations DOS page 179

basic tools

FULL MATCHING

Design of Observational Studies: chapter 8.6

full matching

- Up until now, there's been exactly one treated within a set
- We've been thinking of estimating an effect for what could have happened if we switched the treated to controls.
  - Effect of treatment on the treated.
- Now we'll allow more than on treated within a set.
- Because everything up until now is a special case of this framework, it follows that full matching produces the best matched sets in terms of balance.

full matching

toy example distance matrix

$$\begin{array}{c|cccccc} & 5 & 6 & 7 & 8 & 9 & 10 \\ \hline 1 & 0.37 & 0.16 & 0.05 & 1.12 & 1.02 & 0.74 \\ 2 & 0.88 & 0.68 & 0.62 & 1.16 & 0.45 & 0.96 \\ 3 & 0.00 & 0.00 & 0.80 & 0.00 & 0.97 & 0.00 \\ 4 & 0.28 & 0.13 & 0.03 & 0.88 & 0.76 & 0.52 \end{array}$$

Sum of entries: 0.05+0.45+4*0.00+0.03 = **0.53**

There's one 1:1 set, one 1:4 set, and one 2:1 set.

full matching

toy example distance matrix

$$
\begin{array}{c|cccccc}
 & 5 & 6 & 7 & 8 & 9 & 10 \\
\hline
1 & 0.37 & 0.16 & 0.05 & 1.12 & 1.02 & 0.74 \\
2 & 0.88 & 0.68 & 0.62 & 1.16 & 0.45 & 0.96 \\
3 & 0.00 & 0.00 & 0.80 & 0.00 & 0.97 & 0.00 \\
4 & 0.28 & 0.13 & 0.03 & 0.88 & 0.76 & 0.52
\end{array}
$$

Sum of entries: 0.05+0.45+3*0.00+0.13 = **0.63**

There's three 1:1 set, one 1:3 se~~t, and one 2:1 se~~t.

implementing: full match

How to do these kinds of matches: (see DOS pages 183)

```
library(optmatch)
fullmatch()
```

Within `fullmatch()`, with no need to change settings.

In SAS, you can use `proc netflow` (but it is tough to do…).

efficiency

efficiency

- Our primary concern is bias.
- Bias is what the critics are going to hit us on.
- Bias doesn't go away as we get more and more data.
- Efficiency is good to pay attention to though.
- If we assume our naïve model and constant variance, and we standardize to infinite number of controls then

number of controls	1	2	4	6	10	∞
variance multiplier	2.00	1.50	1.25	1.17	1.10	1.00

- In the real world, going from 1:2 to 1:10 may actually not be as beneficial as it looks… this table assumes perfect matches are available.

practical issue

venturing out of the ivory tower.

assessing covariate balance

- Assessing covariate balance

unmatched

	High NICU	Low NICU	sd	Δ/sd
death	2.26%	1.25%	13.67%	0.07
birth weight (g)	2,454	2,693	739	-0.32
gestational age (months)	34.61	35.69	2.76	-0.39

matched

	High NICU	Low NICU	sd	Δ/sd
death	1.55%	1.94%	13.67%	-0.03
birth weight (g)	2,584	2,581	739	0.00
gestational age (months)	35.14	35.13	2.76	0.00

assessing covariate balance

- Standardize difference

(i) Create a weighted standard deviation using pre-match observations (i.e., use all observations).

$$s_{all,k} = \sqrt{\frac{s_{t,k}^2 + s_{c,k}^2}{2}}$$

where $s_{t,k}^2$ is the standard deviation of covariate $\boldsymbol{x}_k$ amongst the treated group prior to matching.

(i) Divide the difference of the observed means by the weighted standard deviation.

$$\frac{\overline{x_{t,k}} - \overline{x_{c,k}}}{s_{all,k}}$$

Links: 1 and 2

assessing covariate balance

- Assessing covariate balance

unmatched

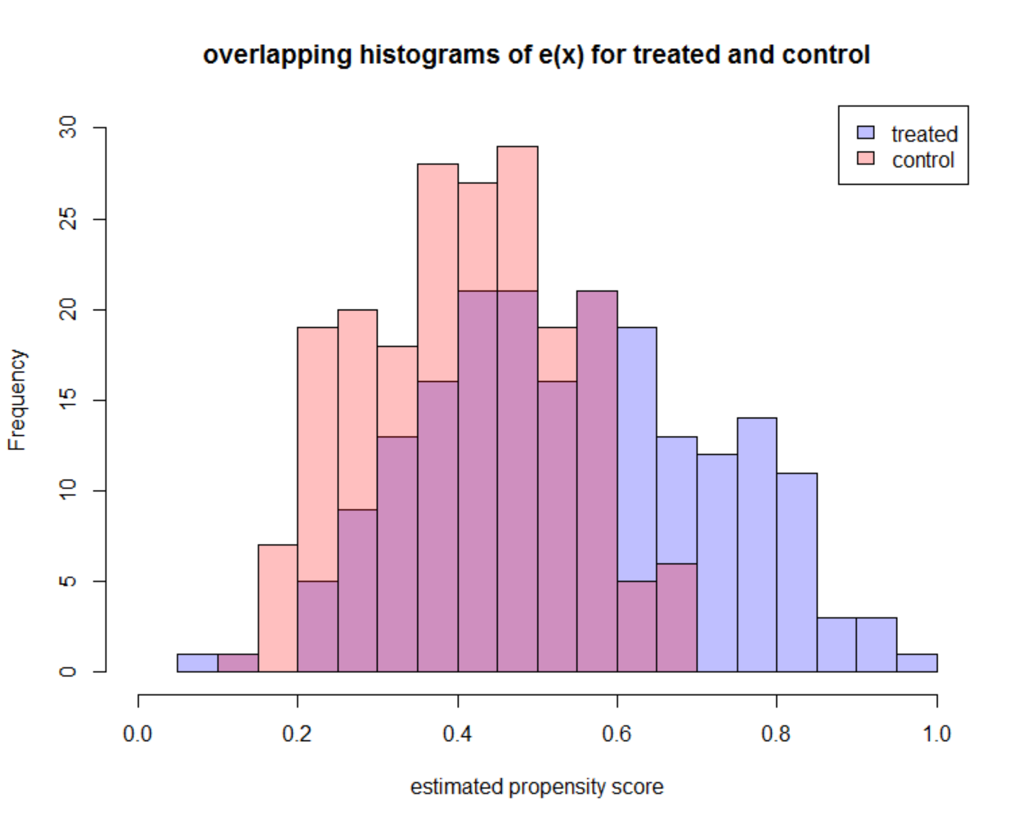
	High NICU	Low NICU	sd	Δ/sd
death	2.26%	1.25%	13.67%	0.07
birth weight (g)	2,454	2,693	739	-0.32
gestational age (months)	34.61	35.69	2.76	-0.39

matched

	High NICU	Low NICU	sd	Δ/sd
death	1.55%	1.94%	13.67%	-0.03
birth weight (g)	2,584	2,581	739	0.00
gestational age (months)	35.14	35.13	2.76	0.00

- The observed difference between the treated and control groups is judged by the typical variation in that covariate.

assessing covariate balance



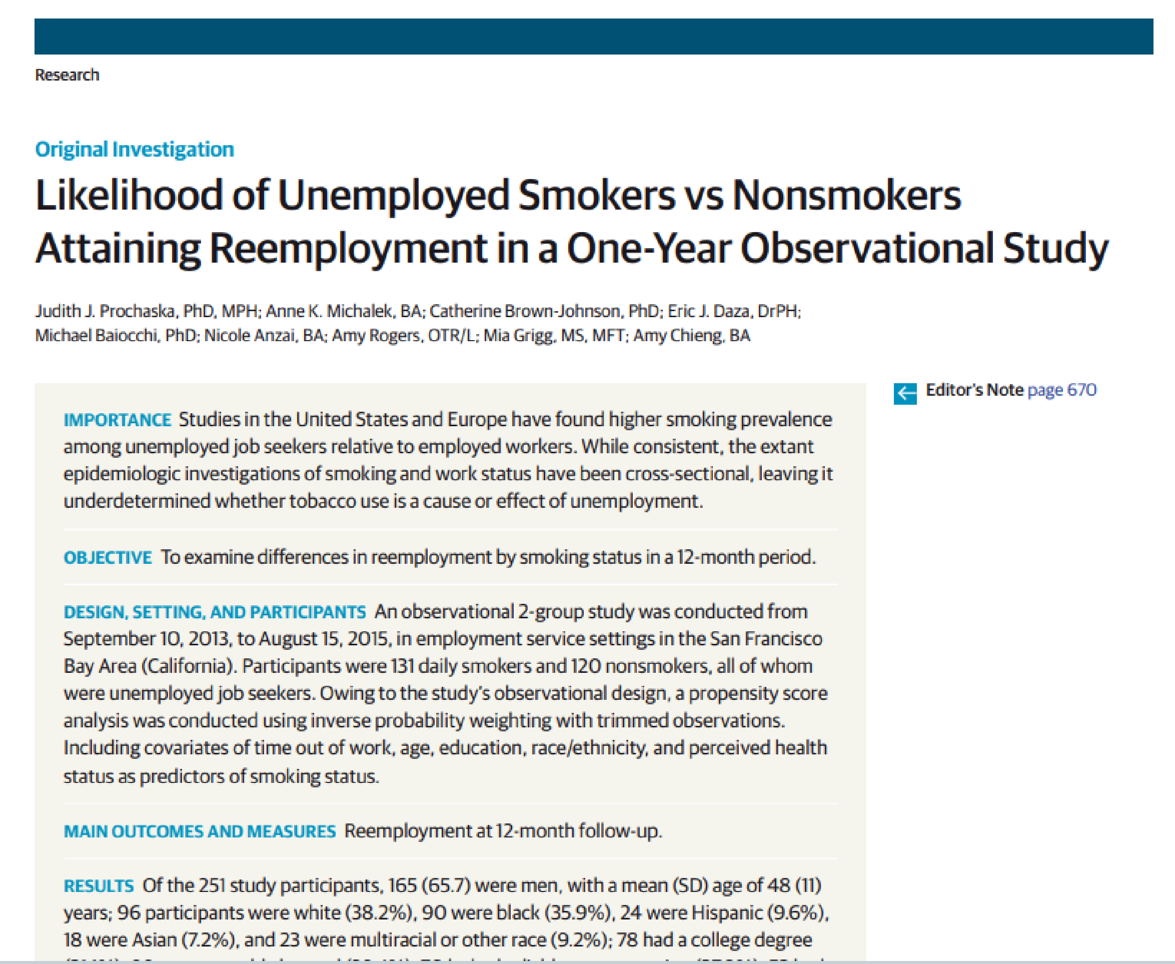
what are we estimating

TREATMENT EFFECT

treatment effect: smoking example

- Example: We collected data on people who reported for job training in the Bay Area. Roughly half smoked. We collected 20ish variables at baseline. We then looked at employment at 12 months.

treatment effect: smoking example



Research

Original Investigation

Likelihood of Unemployed Smokers vs Nonsmokers Attaining Reemployment in a One-Year Observational Study

Judith J. Prochaska, PhD, MPH; Anne K. Michalek, BA; Catherine Brown-Johnson, PhD; Eric J. Daza, DrPH; Michael Baiocchi, PhD; Nicole Anzai, BA; Amy Rogers, OTR/L; Mia Grigg, MS, MFT; Amy Chieng, BA

Editor's Note page 670

IMPORTANCE Studies in the United States and Europe have found higher smoking prevalence among unemployed job seekers relative to employed workers. While consistent, the extant epidemiologic investigations of smoking and work status have been cross-sectional, leaving it underdetermined whether tobacco use is a cause or effect of unemployment.

OBJECTIVE To examine differences in reemployment by smoking status in a 12-month period.

DESIGN, SETTING, AND PARTICIPANTS An observational 2-group study was conducted from September 10, 2013, to August 15, 2015, in employment service settings in the San Francisco Bay Area (California). Participants were 131 daily smokers and 120 nonsmokers, all of whom were unemployed job seekers. Owing to the study's observational design, a propensity score analysis was conducted using inverse probability weighting with trimmed observations. Including covariates of time out of work, age, education, race/ethnicity, and perceived health status as predictors of smoking status.

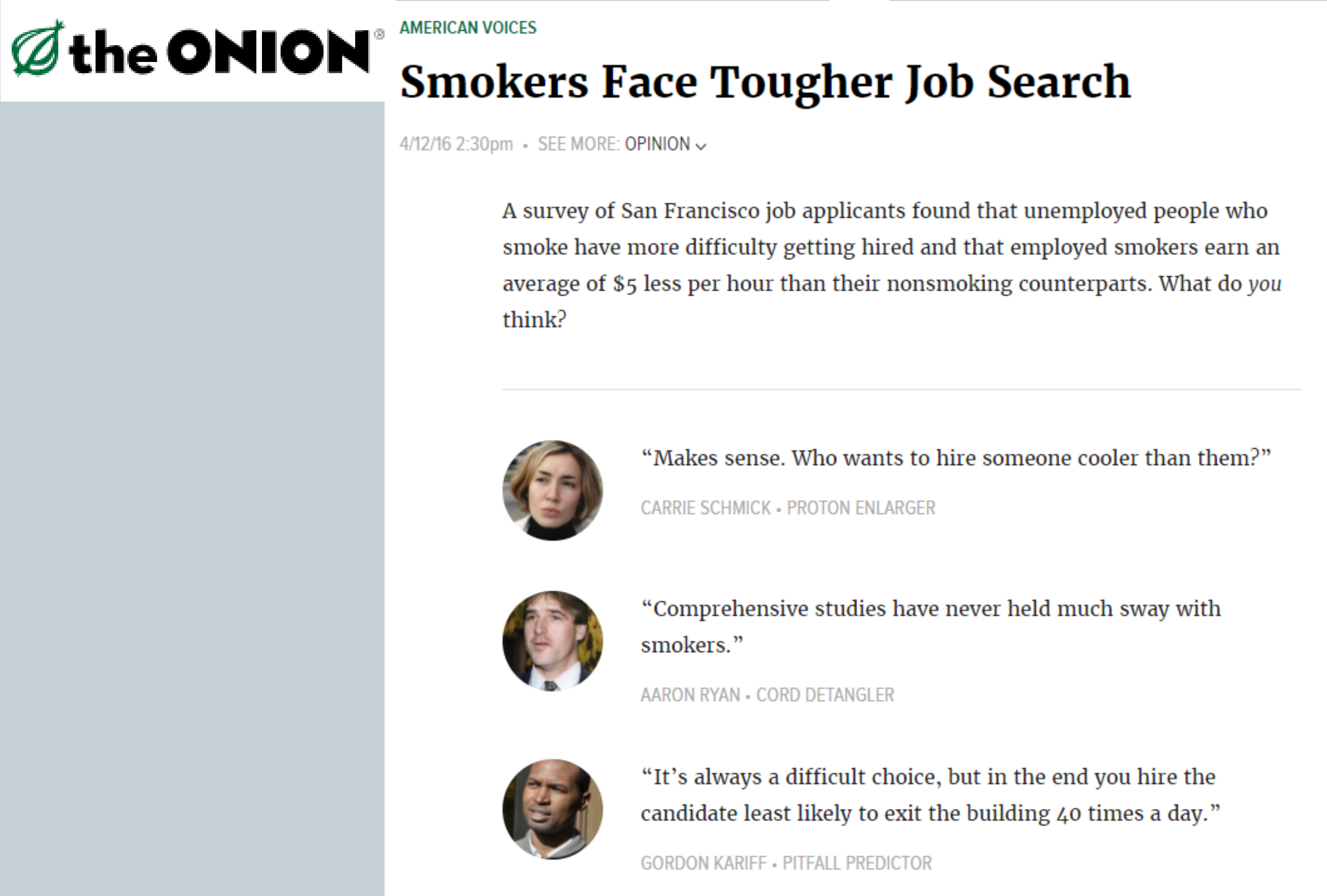
MAIN OUTCOMES AND MEASURES Reemployment at 12-month follow-up.

RESULTS Of the 251 study participants, 165 (65.7) were men, with a mean (SD) age of 48 (11) years; 96 participants were white (38.2%), 90 were black (35.9%), 24 were Hispanic (9.6%), 18 were Asian (7.2%), and 23 were multiracial or other race (9.2%); 78 had a college degree

treatment effect: smoking example

RESULTS Of the 251 study participants, 165 (65.7) were men, with a mean (SD) age of 48 (11) years; 96 participants were white (38.2%), 90 were black (35.9%), 24 were Hispanic (9.6%), 18 were Asian (7.2%), and 23 were multiracial or other race (9.2%); 78 had a college degree (31.1%), 99 were unstably housed (39.4%), 70 lacked reliable transportation (27.9%), 52 had a criminal history (20.7%), and 72 had received prior treatment for alcohol or drug use (28.7%). Smokers consumed a mean (SD) of 13.5 (8.2) cigarettes per day at baseline. At 12-month follow-up (217 participants retained [86.5%]), 60 of 108 nonsmokers (55.6%) were reemployed compared with 29 of 109 smokers (26.6%) (unadjusted risk difference, 0.29; 95% CI, 0.15-0.42). With 6% of analysis sample observations trimmed, **the estimated risk difference indicated that nonsmokers were 30% (95% CI, 12%-48%) more likely on average to be reemployed at 1 year relative to smokers**. Results of a sensitivity analysis with additional covariates of sex, stable housing, reliable transportation, criminal history, and prior treatment for alcohol or drug use (25.3% of observations trimmed) reduced the difference in employment attributed to smoking status to 24% (95% CI, 7%-39%), which was still a significant difference. Among those reemployed at 1 year, the average hourly wage for smokers was significantly lower (mean [SD], \$15.10 [\$4.68]) than for nonsmokers (mean [SD], \$20.27 [\$10.54]; F(1,86) = 6.50, P = .01).

treatment effect: smoking example



the ONION

AMERICAN VOICES

Smokers Face Tougher Job Search

4/12/16 2:30pm • SEE MORE: OPINION

A survey of San Francisco job applicants found that unemployed people who smoke have more difficulty getting hired and that employed smokers earn an average of $5 less per hour than their nonsmoking counterparts. What do *you* think?

"Makes sense. Who wants to hire someone cooler than them?"

CARRIE SCHMICK • PROTON ENLARGER

"Comprehensive studies have never held much sway with smokers."

AARON RYAN • CORD DETANGLER

"It's always a difficult choice, but in the end you hire the candidate least likely to exit the building 40 times a day."

GORDON KARIFF • PITFALL PREDICTOR

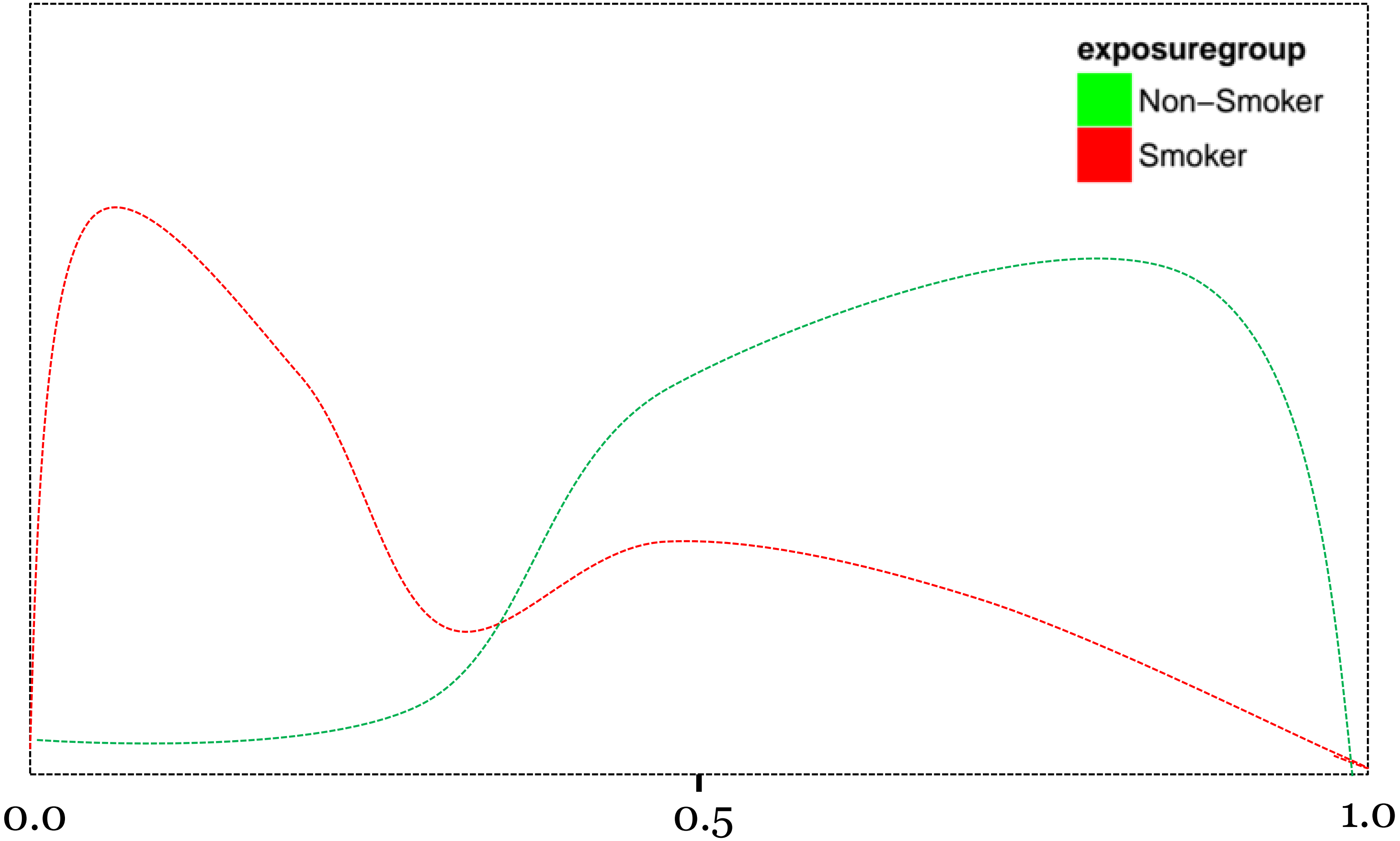
https://www.theonion.com/smokers-face-tougher-job-search-1819563077

treatment effect: smoking example

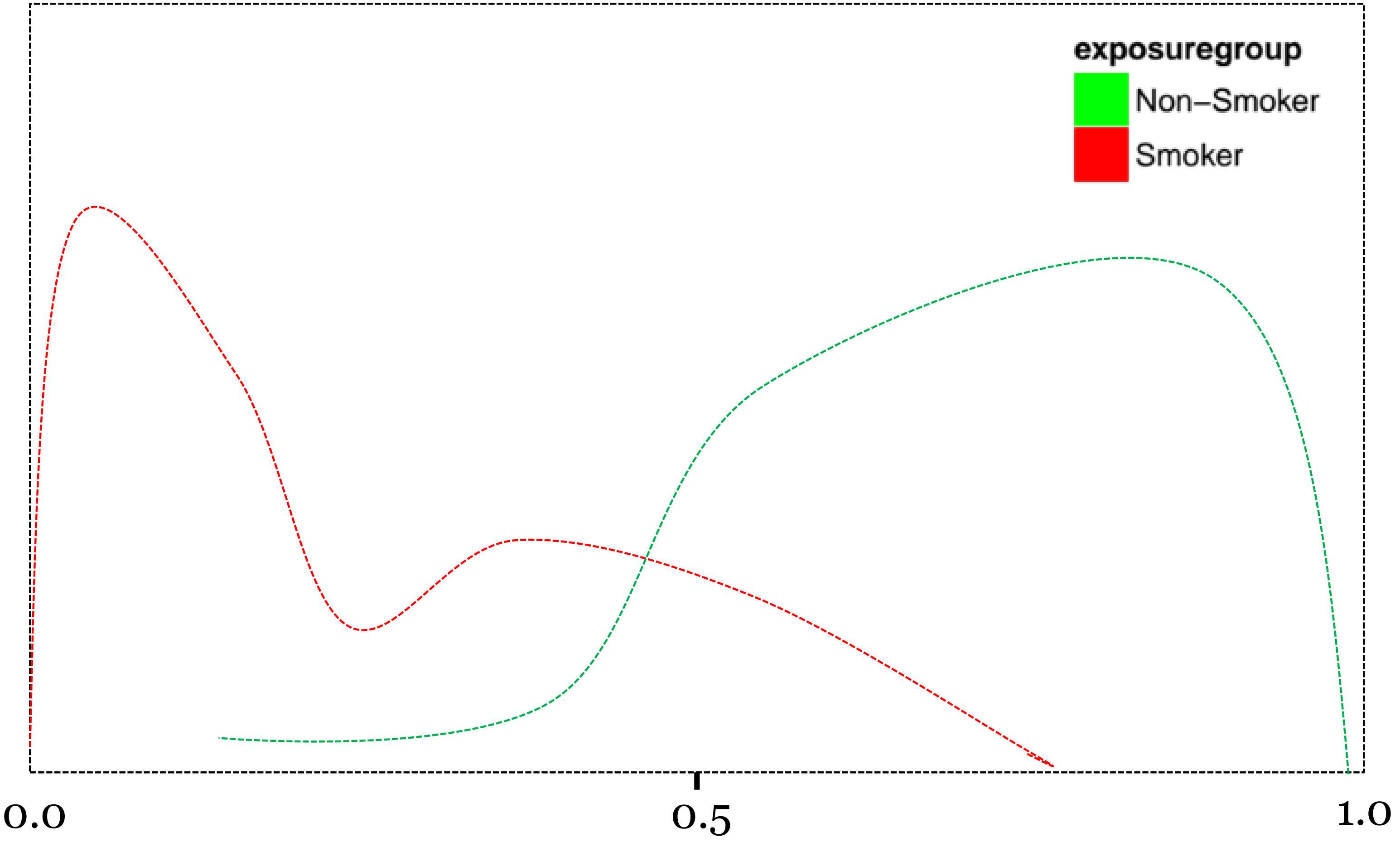
- Example: We collected data on people who reported for job training in the Bay Area. Roughly half smoked. We collected 20ish variables at baseline. We then looked at employment at 12 months.
- Let’s consider matching one treated to one control.

treatment effect: smoking example

- Non-overlap
  - Look at a histogram



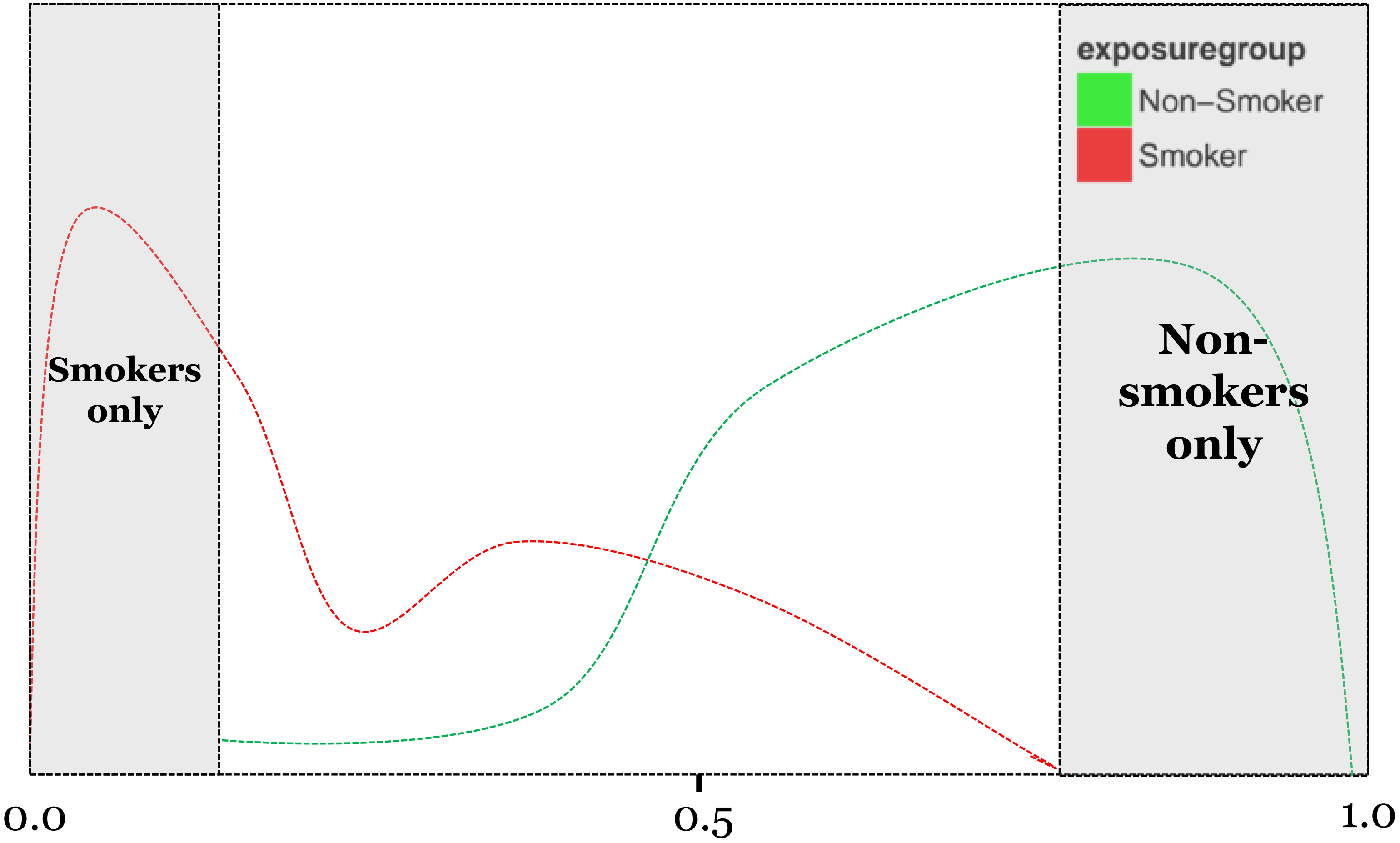
P(non-S|**X**) = Probability of being a non-smoker, given covariates



P(non-S|**X**) = Probability of being a non-smoker, given covariates

treatment effect: smoking example

- Non-overlap
  - Look at a histogram
  - Upper and lower



P(non-S|**X**) = Probability of being a non-smoker, given covariates

treatment effect: smoking example

RESULTS Of the 251 study participants, 165 (65.7) were men, with a mean (SD) age of 48 (11) years; 96 participants were white (38.2%), 90 were black (35.9%), 24 were Hispanic (9.6%), 18 were Asian (7.2%), and 23 were multiracial or other race (9.2%); 78 had a college degree (31.1%), 99 were unstably housed (39.4%), 70 lacked reliable transportation (27.9%), 52 had a criminal history (20.7%), and 72 had received prior treatment for alcohol or drug use (28.7%). Smokers consumed a mean (SD) of 13.5 (8.2) cigarettes per day at baseline. At 12-month follow-up (217 participants retained [86.5%]), 60 of 108 nonsmokers (55.6%) were reemployed compared with 29 of 109 smokers (26.6%) (unadjusted risk difference, 0.29; 95% CI, 0.15-0.42). With 6% of analysis sample observations trimmed, **the estimated risk difference indicated that nonsmokers were 30% (95% CI, 12%-48%) more likely on average to be reemployed at 1 year relative to smokers**. Results of a sensitivity analysis with additional covariates of sex, stable housing, reliable transportation, criminal history, and prior treatment for alcohol or drug use (25.3% of observations trimmed) reduced the difference in employment attributed to smoking status to 24% (95% CI, 7%-39%), which was still a significant difference. Among those reemployed at 1 year, the average hourly wage for smokers was significantly lower (mean [SD], \$15.10 [\$4.68]) than for nonsmokers (mean [SD], \$20.27 [\$10.54]; F(1,86) = 6.50, P = .01).

treatment effect: smoking example

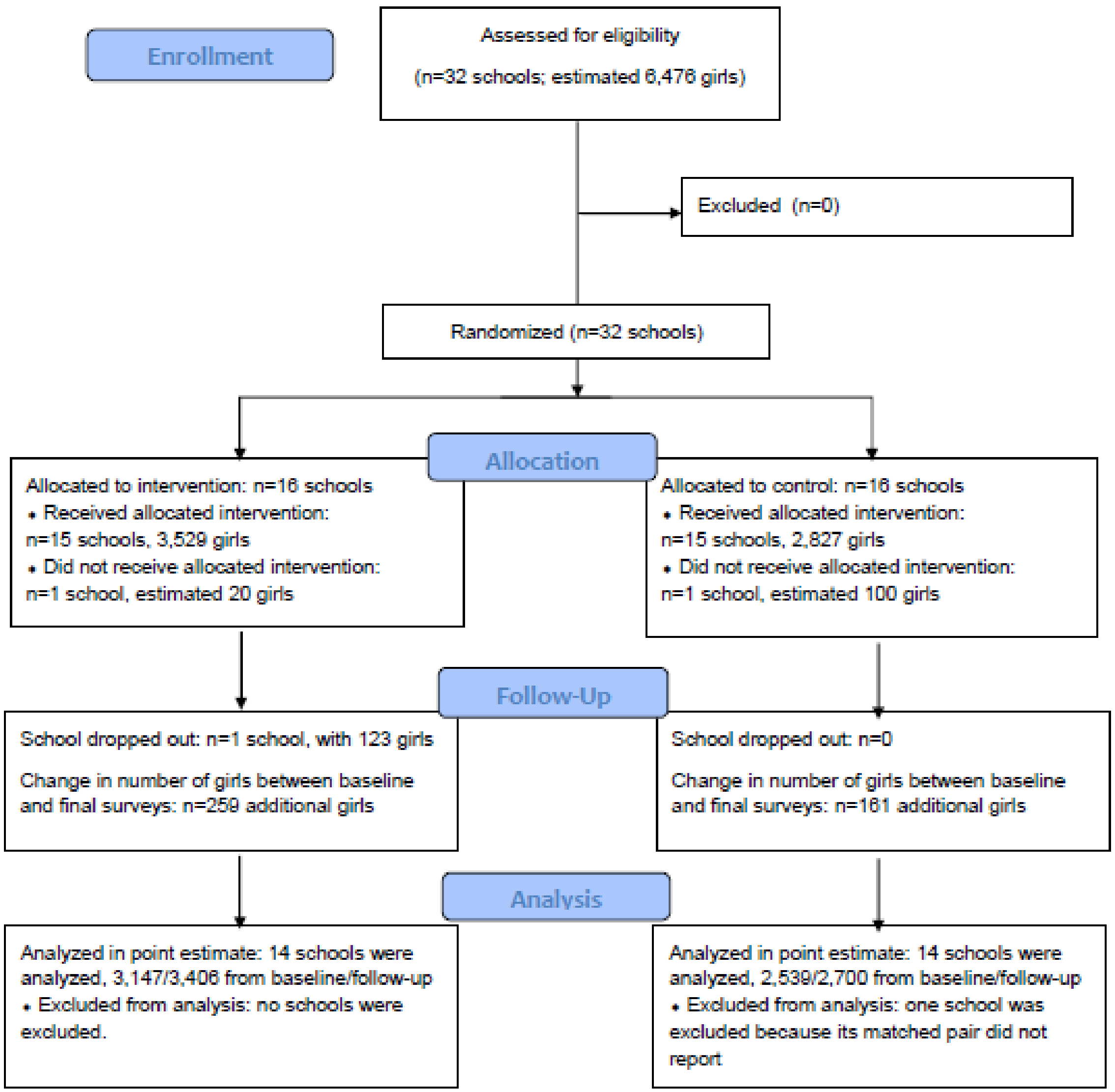
- Non-overlap
  - Look at a histogram
  - Upper and lower
  - Violation of strongly ignorable treatment assignment
  - Careful, need to consider what effect you’re estimating
    - What’s actually estimable and what isn’t
  - Focus on the 50% range because that’s actually where the debate is happening
  - Trim at the edges because that’s where you’re pretty sure the violation of SITA is going to happen
  - More detail here: *Crump et al*

treatment effect: smoking example

- Consider how to remove the observations that you can’t/don’t want to include in your study.
- This is roughly equivalent to the inclusion/exclusion criteria of a randomized controlled trial.
- Examine the pscore fitted model and see what parts of the covariate space are in the non-overlap
- Use a regression tree (or some other classifier) to make it intelligible. Citation: Traskin & Small (2011)

CONSORT Flow Diagram

Flow Diagram



Enrollment

Assessed for eligibility

(n=32 schools; estimated 6,476 girls)

Excluded (n=0)

Randomized (n=32 schools)

Allocation

Allocated to intervention: n=16 schools
- Received allocated intervention: n=15 schools, 3,529 girls
- Did not receive allocated intervention: n=1 school, estimated 20 girls

Allocated to control: n=16 schools
- Received allocated intervention: n=15 schools, 2,827 girls
- Did not receive allocated intervention: n=1 school, estimated 100 girls

Follow-Up

School dropped out: n=1 school, with 123 girls

Change in number of girls between baseline and final surveys: n=259 additional girls

School dropped out: n=0

Change in number of girls between baseline and final surveys: n=161 additional girls

Analysis

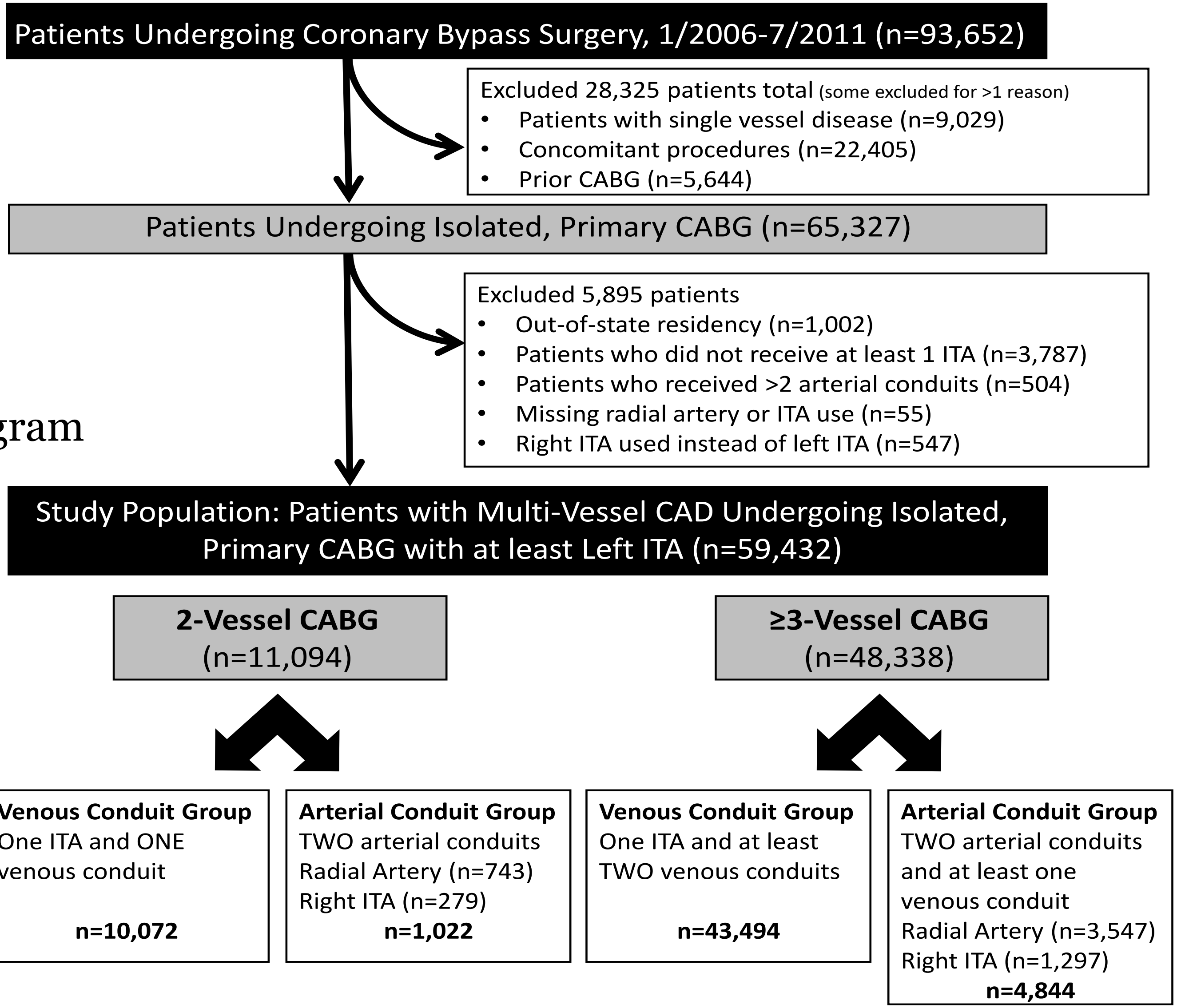
Analyzed in point estimate: 14 schools were analyzed, 3,147/3,406 from baseline/follow-up
- Excluded from analysis: no schools were excluded.

Analyzed in point estimate: 14 schools were analyzed, 2,539/2,700 from baseline/follow-up
- Excluded from analysis: one school was excluded because its matched pair did not report

Figure 1: Participant flow diagram for this study. See "Losses and Exclusions" for more discussion.

STROBE Flow Diagram

Figure 1: **Patient selection flow diagram**



Patients Undergoing Coronary Bypass Surgery, 1/2006-7/2011 (n=93,652)

Excluded 28,325 patients total (some excluded for >1 reason)

- Patients with single vessel disease (n=9,029)
- Concomitant procedures (n=22,405)
- Prior CABG (n=5,644)

Patients Undergoing Isolated, Primary CABG (n=65,327)

Excluded 5,895 patients

- Out-of-state residency (n=1,002)
- Patients who did not receive at least 1 ITA (n=3,787)
- Patients who received >2 arterial conduits (n=504)
- Missing radial artery or ITA use (n=55)
- Right ITA used instead of left ITA (n=547)

Study Population: Patients with Multi-Vessel CAD Undergoing Isolated, Primary CABG with at least Left ITA (n=59,432)

2-Vessel CABG
(n=11,094)

≥3-Vessel CABG
(n=48,338)

Venous Conduit Group
One ITA and ONE venous conduit
n=10,072

Arterial Conduit Group
TWO arterial conduits
Radial Artery (n=743)
Right ITA (n=279)
n=1,022

Venous Conduit Group
One ITA and at least TWO venous conduits
n=43,494

Arterial Conduit Group
TWO arterial conduits and at least one venous conduit
Radial Artery (n=3,547)
Right ITA (n=1,297)
n=4,844

CAD, coronary artery disease; CABG, coronary artery bypass grafting; ITA, internal thoracic artery

treatment effect on [group]

CAREFUL CONSIDERATIONS

treatment effect on [insert group]

- Effect estimates: ATE, TonT, TonC and CACE
- Causal effect of the treatment

$$Y_i(D_i = 1) - Y_i(D_i = 0) = \Delta_i$$

- Average Treatment Effect

$$E_i[Y_i(D_i = 1) - Y_i(D_i = 0)] = \bar{\Delta}_i$$

- Treatment effect on the Treated

$$E_i[Y_i(D_i = 1) - Y_i(D_i = 0)|d_i = 1] = \bar{\Delta}_i{}^T$$

- Treatment effect on the Control

$$E_i[Y_i(D_i = 1) - Y_i(D_i = 0)|d_i = 0] = \bar{\Delta}_i{}^C$$

- Complier average causal effect

$$E_i[Y_i(D_i = 1) - Y_i(D_i = 0)|i \in complier] = \bar{\Delta}_i{}^{IV}$$

treatment effect on [insert group]

observation	Y(d=1)	Y(d=0)	delta	dose	included in effect
1	8	9	-1	0	-1
2	5	3	2	1	2
3	4	5	-1	0	-1
4	6	7	-1	0	-1
5	10	11	-1	0	-1
6	3	4	-1	0	-1
7	3	1	2	1	2
8	-1	0	-1	0	-1
9	5	6	-1	0	-1
10	2	0	2	1	2
				average:	0

Average Treatment Effect

treatment effect on [insert group]

observation	Y(d=1)	Y(d=0)	delta	dose	included in effect
1	8	9	-1	0	
2	5	3	2	**1**	2
3	4	5	-1	0	
4	6	7	-1	0	
5	10	11	-1	0	
6	3	4	-1	0	
7	3	1	2	**1**	2
8	-1	0	-1	0	
9	5	6	-1	0	
10	2	0	2	**1**	2
				average:	2

Treatment Effect on the Treated

treatment effect on [insert group]

observation	Y(d=1)	Y(d=0)	delta	dose	included in effect
1	8	9	-1	**0**	-1
2	5	3	2	1	
3	4	5	-1	**0**	-1
4	6	7	-1	**0**	-1
5	10	11	-1	**0**	-1
6	3	4	-1	**0**	-1
7	3	1	2	1	
8	-1	0	-1	**0**	-1
9	5	6	-1	**0**	-1
10	2	0	2	1	
				average:	-1

Treatment Effect on the Control

subsetting

TOO MUCH DATA

dealing with lots of observations

- If you get lots of observations then you should be happy.
- If you try to put them all into a matching algorithm then you will be sad.
- The complexity of matching algorithms grows really fast so cutting down the problem into smaller chunks helps a lot.
- Look at your covariates:
  - Is there one or two that are binary or categorical?
  - Break your data set into separate data sets and match within a given level of a variable (or variables).
  - Choose variables that are prognostically important.
  - It's nice if these variables are close to uniformly distributed (e.g., p=0.5, or p=<1/3, 1/3, 1/3>).

dealing with lots of observations

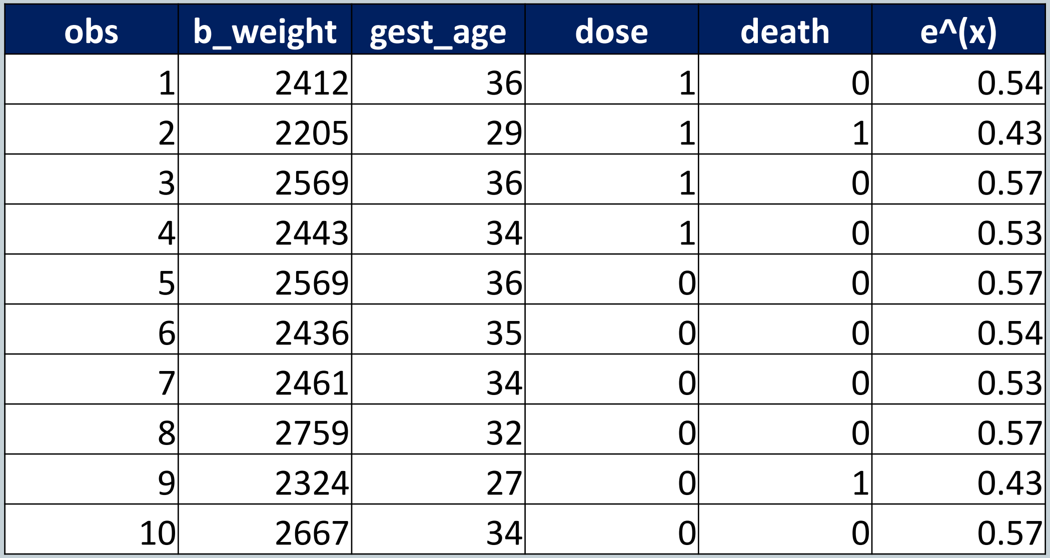
- In the NICU example, we had millions of babies.
- I subsetted the data on gestational age (i.e., 26 weeks only matched to 26 weeks).
- For larger gestational age groups, I further subsetted on birth weight.
  - This was much less satisfactory because it's more continuous.
  - I picked arbitrary boundaries and didn't look back…
- You can fret about the matching method, but do not confuse that for the quality of the match which is assessed by looking at the covariates.

practical issue

venturing out of the ivory tower.

missing covariates

- Missing covariates



obs	b_weight	gest_age	dose	death	e^(x)
1	2412	36	1	0	0.54
2	2205	29	1	1	0.43
3	2569	36	1	0	0.57
4	2443	34	1	0	0.53
5	2569	36	0	0	0.57
6	2436	35	0	0	0.54
7	2461	34	0	0	0.53
8	2759	32	0	0	0.57
9	2324	27	0	1	0.43
10	2667	34	0	0	0.57

missing covariates

- Missing covariates

obs	b_weight	gest_age	dose	death	e^(x)
1	2412	36	1	0	
2	NA	29	1	1	
3	2569	36	1	0	
4	2443	34	1	0	
5	2569	36	0	0	
6	2436	NA	0	0	
7	2461	34	0	0	
8	2759	32	0	0	
9	2324	27	0	1	
10	2667	34	0	0	

missing covariates

- Missing covariates

obs	b_weight	bw_mis	gest_age	ga_mis	dose	death
1	2412	0	36	0	1	0
2	NA	1	29	0	1	1
3	2569	0	36	0	1	0
4	2443	0	34	0	1	0
5	2569	0	36	0	0	0
6	2436	0	NA	1	0	0
7	2461	0	34	0	0	0
8	2759	0	32	0	0	0
9	2324	0	27	0	0	1
10	2667	0	34	0	0	0

missing covariates

- Missing covariates

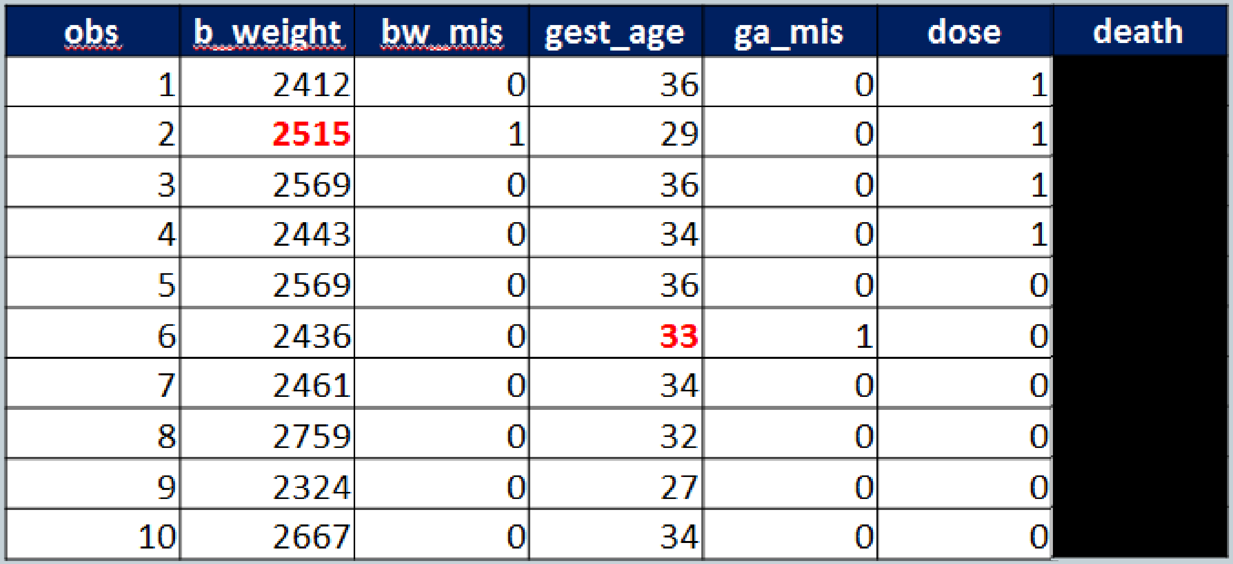
obs	b_weight	bw_mis	gest_age	ga_mis	dose	death
1	2412	0	36	0	1	0
2	**2515**	1	29	0	1	1
3	2569	0	36	0	1	0
4	2443	0	34	0	1	0
5	2569	0	36	0	0	0
6	2436	0	**33**	1	0	0
7	2461	0	34	0	0	0
8	2759	0	32	0	0	0
9	2324	0	27	0	0	1
10	2667	0	34	0	0	0

(i) Build pscores using the imputed value and the missing indicators.
(ii) Use imputed values and missing indicators in calculating the Mahalanobis distance.

a small but important point

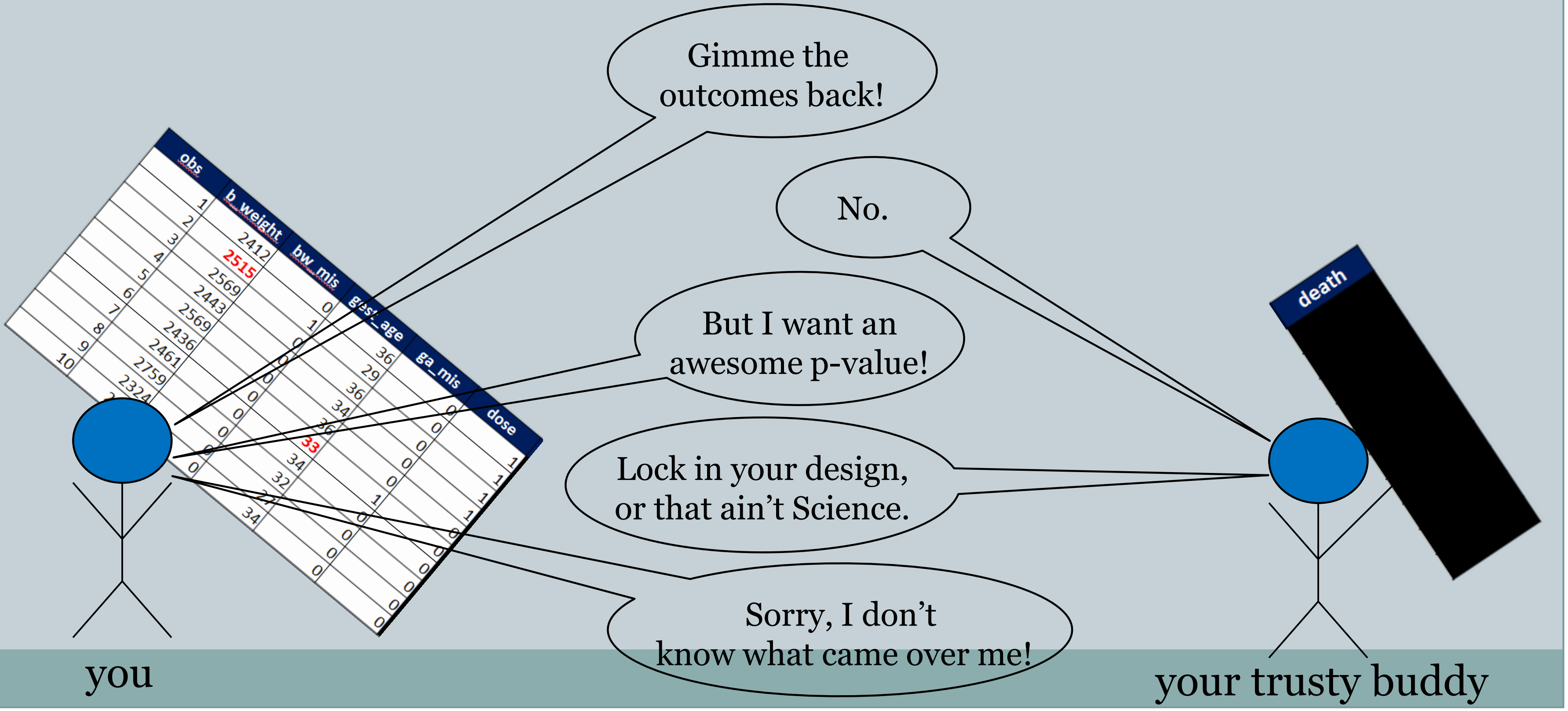
obs	b_weight	bw_mis	gest_age	ga_mis	dose	death
1	2412	0	36	0	1	0
2	**2515**	1	29	0	1	1
3	2569	0	36	0	1	0
4	2443	0	34	0	1	0
5	2569	0	36	0	0	0
6	2436	0	**33**	1	0	0
7	2461	0	34	0	0	0
8	2759	0	32	0	0	0
9	2324	0	27	0	0	1
10	2667	0	34	0	0	0

a small but important point



obs	b_weight	bw_mis	gest_age	ga_mis	dose	death
1	2412	0	36	0	1	[redacted]
2	2515	1	29	0	1	[redacted]
3	2569	0	36	0	1	[redacted]
4	2443	0	34	0	1	[redacted]
5	2569	0	36	0	0	[redacted]
6	2436	0	33	1	0	[redacted]
7	2461	0	34	0	0	[redacted]
8	2759	0	32	0	0	[redacted]
9	2324	0	27	0	0	[redacted]
10	2667	0	34	0	0	[redacted]

a small but important point



obs	b_weight	bw_mis	gest_age	ga_mis	dose
1	2412	0	36	0	1
2	2515	1	29	0	1
3	2569	0	36	0	1
4	2443	0	34	0	1
5	2569	0	36	0	0
6	2436	0	33	1	0
7	2461	0	34	0	0
8	2759	0	32	0	0
9	2324	0	27	0	0
10	[illegible]	0	34	0	0

death

[redacted]

Gimme the outcomes back!
No.
But I want an awesome p-value!
Lock in your design, or that ain't Science.
Sorry, I don't know what came over me!
you
your trusty buddy

fin.