Inverse probability of treatment weighting and marginal structural models

JA Labrecque

Erasmus Medical Center, Rotterdam, The Netherlands

Jan 30, 2020

Outline

All material available at https://github.com/jalabrecque/Berlin

1. What is causal inference?
2. Structural models
3. Inverse probability of treatment weights
4. Estimation of marginal structural model
5. MSM for time-varying exposures
6. Marginal vs conditional
7. Treatment models vs outcome models
8. Example
9. Exercise

1. What is causal inference?

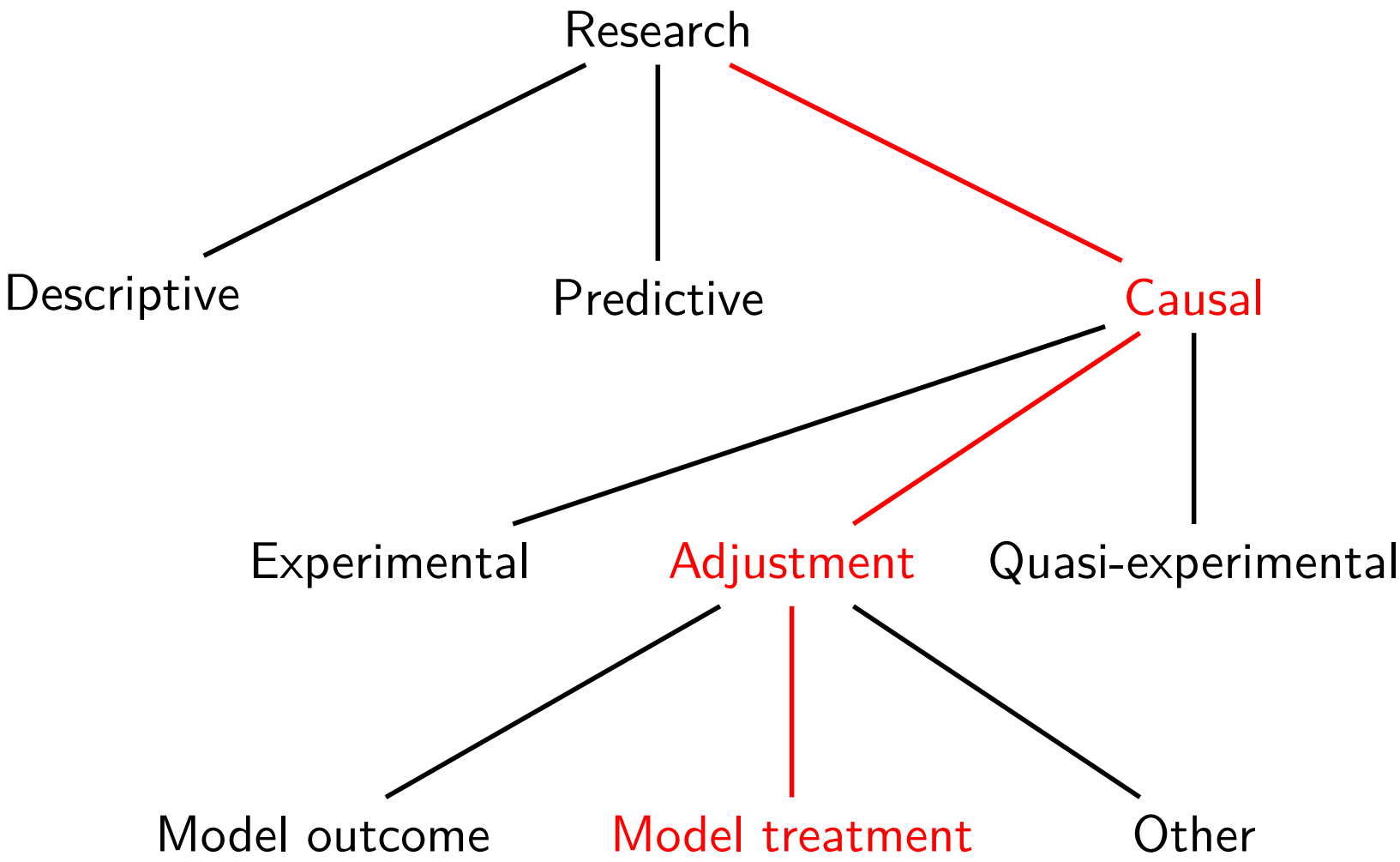
The research terrain

- The first question we should ask is what kind of study are we doing
  - Descriptive: what is the relationship between A and Y?
  - Predictive: what will Y be if I observed A=a?
  - Causal: how will Y change if I change A?

Causal inference

- Choices of ways to do causal inference:
  - Experimental
    - Randomized controlled trial
  - Confounder adjustment
    - Outcome regression
    - Propensity score
    - Doubly robust estimation
  - Quasi-experimenal
    - Instrumental variable
    - Regression discontinuity
    - Differences-in-differences

Where we are



The goal of causal inference

- What we want to know but can't observed: $E[Y^{a=1} - Y^{a=0}]$
- What we can observe: $E[Y|A = 1, L = l] - E[Y|A = 0, L = l]$
- Causal inference tells us how to model observed data to make an observable estimate equal to an unobservable causal estimate (and the assumptions required for them to be equal)

$$E[Y^{a=1} - Y^{a=0}] = E[Y|A = 1, L = l] - E[Y|A = 0, L = l]$$

2. Marginal structural models

Marginal structural models

- State the causal model (marginal)
- Find weights that balance covariates across levels of exposure
- Estimate the causal model using these weights

Marginal structural models

$$E[Y^a] = \beta_0 + \beta_A * a$$

Marginal structural models

- The left-hand side is a counterfactual
- There are no covariates

$$E[Y^a] = \beta_0 + \beta_A * a$$

Marginal structural models

Table 1.1

	$Y^{a=0}$	$Y^{a=1}$
Rheia	0	1
Kronos	1	0
Demeter	0	0
Hades	0	0
Hestia	0	0
Poseidon	1	0
Hera	0	0
Zeus	0	1
Artemis	1	1
Apollo	1	0
Leto	0	1
Ares	1	1
Athena	1	1
Hephaestus	0	1
Aphrodite	0	1
Cyclope	0	1
Persephone	1	1
Hermes	1	0
Hebe	1	0
Dionysus	1	0

Marginal structural models

	A	Y^a
Rheia	0	0
Rheia	1	1
Kronos	0	1
Kronos	1	0
Demeter	0	0
Demeter	1	0
Hades	0	0
Hades	1	0
Hestia	0	0
Hestia	1	0
Poseidon	0	1
Poseidon	1	0
Hera	0	0
Hera	1	0
Zeus	0	0
Zeus	1	1

Marginal structural models

- What we want is $E[Y^a]$ but what we have is $E[Y|A = a]$
- If we believed that $E[Y^a] = E[Y|A = a]$, we could substitute one for the other
- This is where our causal assumptions come into play

Marginal structural models

$$E[Y|A=a] = E[Y^a|A=a] = E[Y^a]$$

- On the left is what we observe, on the right is what we want
- The first equality is called the consistency assumption: $Y^a = Y$ for every person with $A = a$
- The second equality is called the exchangeability assumption: $Y^a \perp\!\!\!\perp A$
- You also need to assume positivity
- If we believe all these assumptions we can use observed data to estimate the MSM

How do we normally get exchangeability?

- We make a model where we put the outcome on one side of the equation and our exposure and potential confounders on the other side
- $$Y \sim A + \text{potential confounders} + \epsilon$$
- What are we doing here?
- We are making a statistical model of the outcome
- But our MSM does not condition on confounders so how can we get exchangeability without conditioning?

Marginal structural models

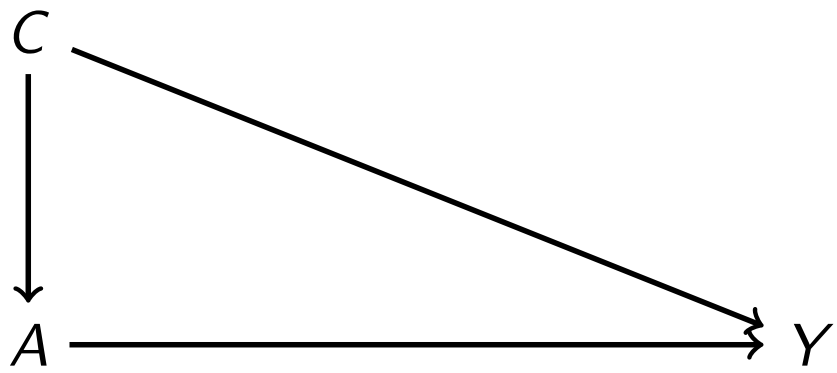
- So our goal, if we want to fit this model, is to make a dataset where we can achieve exchangeability without conditioning on confounders

$$E[Y^a] = \beta_0 + \beta_A * a$$

3. Inverse probability of treatment weights

Another way of dealing with confounding

- But, think back, what is the definition of a confounder?
- Loosely, it's a variable that is a cause of the outcome and associated with exposure
- We use regression to model the relationship between our confounders and the outcome
- But can we focus on the association with the exposure instead?



Artificial balancing

A very simple data set:

C	A	Y	n
0	0	0	6
1	0	1	4
0	1	1	4
1	1	2	6

C is a covariate, A is our exposure, Y is our outcome and n is the number of people in each stratum.

C	A	Y	n
0	0	0	6
1	0	1	4
0	1	1	4
1	1	2	6

C is a confounder of the effect of A on Y because p(C) among unexposed is $\frac{4}{10}$ and among the exposed $\frac{6}{10}$ AND because C causes Y (it increases it by 1).

C	A	Y	n
0	0	0	6
1	0	1	4
0	1	1	4
1	1	2	6

The true effect of A is 1. But C is a confounder so we get:

```
lm(Y ~ A, data = ds, weights = n)$coef["A"]
```

```
  A
1.2
```

```
lm(Y ~ A + C, data = ds, weights = n)$coef["A"]
```

```
A
1
```

Erasmus MC
Universitair Medisch Centrum Rotterdam

What if we simply added observations to the data set to balance C across exposures?

C	A	Y	n
0	0	0	6
1	0	1	9
0	1	1	4
1	1	2	6

What happens if we rerun our regression now that we've "artificially" balanced C?

```
lm(Y ~ A, data = ds, weights = n)$coef["A"]
```

```
A
1
```

- So we've found a way to achieve exchangeability (for measured covariates) without having to include it in our regression
- When you have a lot of confounders you can't simply add observations like this so we get a bit "mathy"
- Inverse probability of treatment weighting (IPTW) is just a way of balancing confounders using information from all confounders

C	A	Y	n	P(A=a|C=C)	IPTW	n*IPTW
0	0	0	6	0.6	1.67	10
1	0	1	4	0.4	2.50	10
0	1	1	4	0.4	2.50	10
1	1	2	6	0.6	1.67	10

- $P(A = 0|C = 0) = \frac{6}{10} = 0.6$
- $P(A = 0|C = 1) = \frac{4}{10} = 0.4$
- $P(A = 1|C = 0) = \frac{4}{10} = 0.4$
- $P(A = 1|C = 1) = \frac{6}{10} = 0.6$

Modeling the exposure

- In outcome regression, we model the outcome as a function of the exposure and confounders
- What if we, instead, model the exposure as a function of the confounders?
- $logit(P(A = a)) = \alpha + \bar{\beta}\bar{L}$
- A represents exposure and L represents a vector of confounders
- $P(A = a|L = l)$ is also called the propensity score (PS)
- It is the probability that you receive treatment A=a given covariates
- What would the propensity score be in a randomized trial?

What is the use in modeling the PS?

- The PS, it turns out, is what is known as a *balancing score*
- Balancing score is a score within which the covariates are balanced across levels of exposure for each value of the balancing score
- Let's look at the PS
  - We model the exposure A as a function of confounders L
  - We can calculate the PS that A=1 for each individual in our sample
  - If we choose all individuals with, for example, PS=0.3, this will include some people who have A=0 and some people who have A=1
  - If we compare the average covariate value among people with A=0 and people with A=1 among everyone with PS=0.3, we will find that the average value of L will be the same between these two groups
  - $A \perp\!\!\!\perp L | PS(L)$
  - What does this remind you of?

How do we calculate a propensity score?

- Methods to estimates PS
  - Most common: logistic regression
  - Machine learning
  - Covariate-balancing propensity score

What variables go in my propensity score

- When we're modeling the outcome, we have to include confounders in our model but including predictors of the outcome that aren't confounders will not hurt us
- This is not true of propensity score models. Even though we're modeling the exposure, including variables that are predictive of exposure but that are not confounders SHOULD NOT be included in the model
- What should be included in the propensity score model are variables which you would like to be balanced, *i.e.* confounders
- If variables are included that are correlated with exposure but are not risk factors for the outcome, this will cause bias amplification
- Bias amplification means that any bias in your study will be amplified by including this type of variable

A simple example

- We want to do a study of a drug that lowers blood pressure
- We think that age, sex and BMI are confounders. We make a table 1 of these confounders by treatment status:

	A0	A1
age	44.51	55.57
sex	0.55	0.47
BMI	26.20	27.80

- We also find that the treated group have 2.6mmHg higher blood pressure than the untreated group. These data are simulated so we know the true value: -5. Clearly the crude estimate is very confounded.
- We would like to us IPTW to estimate the effect of A on BP

A simple example

- We run a logistic regression of A on age, sex and BMI
- $logit(P(A)) \sim \alpha + age + sex + BMI$
- With that model we calculate the probability that each person receives the treatment A

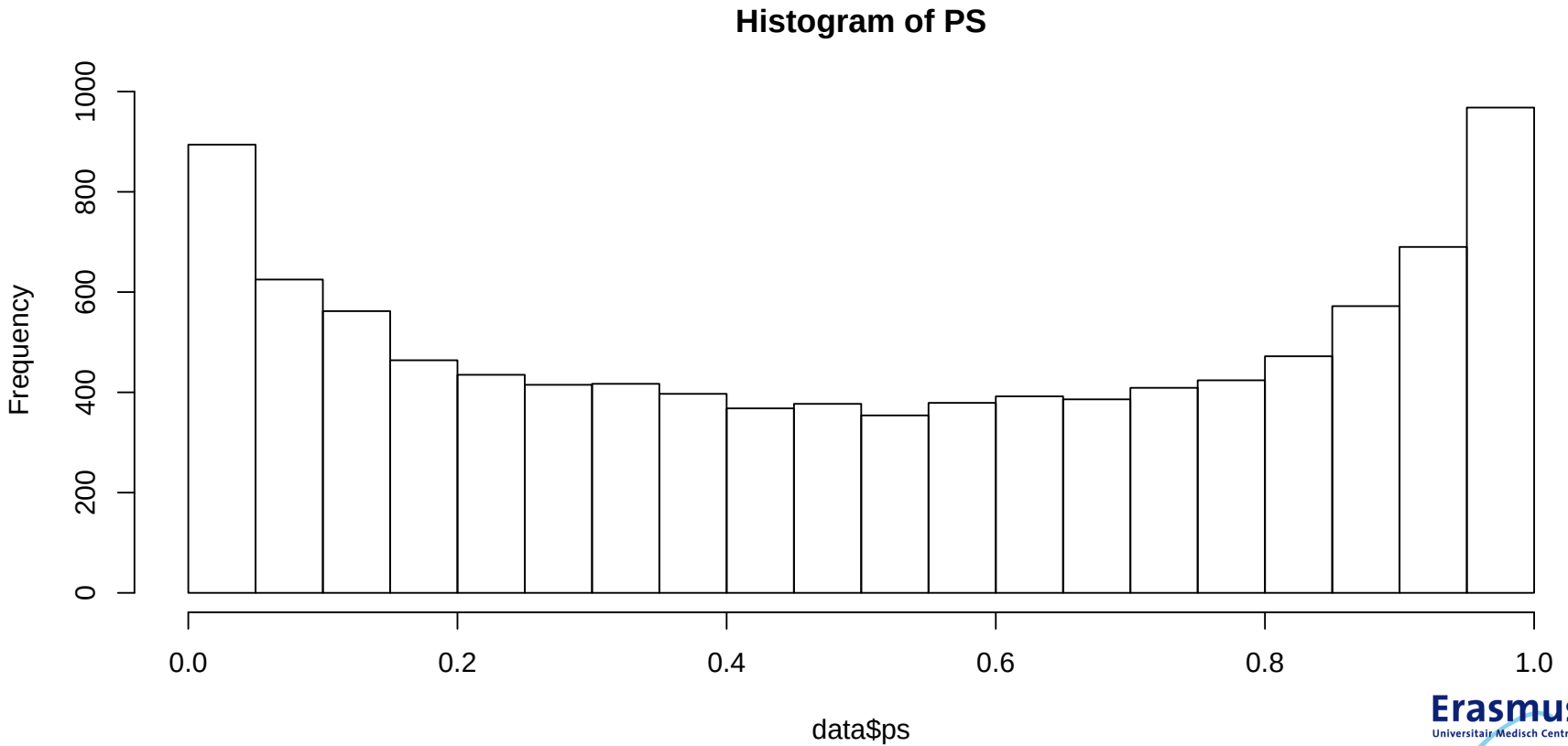
```
mod <- glm(A ~ age + sex + BMI, data=data, family = "binom

mod$coefficients
```

```
(Intercept)         age         sex         BMI
-18.0516298   0.2025868  -0.5905725   0.3051945
```

Erasmus MC
Universitair Medisch Centrum Rotterdam

A simple example

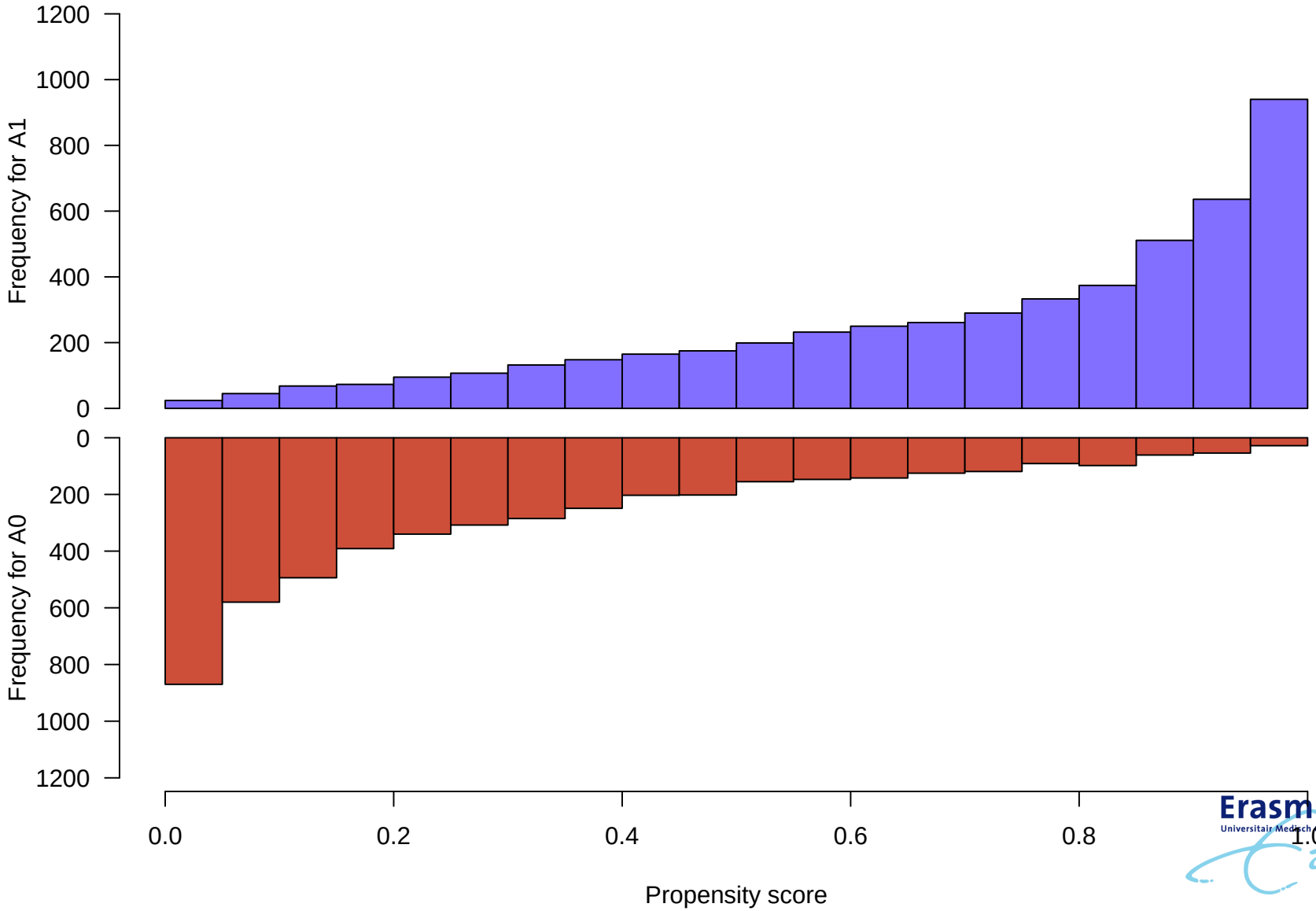
```
mod <- glm(A ~ age + sex + BMI, data=data, family = "binom
data$ps <- predict(mod, type = "response")
hist(data$ps,main="Histogram of PS")
```



Erasmus MC
Universitair Medisch Centrum Rotterdam

A simple example

- We can also plot a histogram by treatment received



Erasmus MC
Universitair Medisch Centrum Rotterdam

Examples of going from PS to IPTW

- What is the IPTW of a treated participant with a PS=0.5?
  - They received treatment so it's simply $\frac{1}{PS} = \frac{1}{0.5} = 2$
  - So when we use the weighting, this person will count for 2 people

Examples of going from PS to IPTW

- What is the IPTW of a untreated participant with a PS=0.2?
  - They were not treated so the IPTW is $\frac{1}{1-PS} = \frac{1}{1-0.2} = 1.25$
  - So in the weighted sample, this person will count for 1.25 people

Examples of going from PS to IPTW

- What is the IPTW of a treated participant with a PS=0.2?
  - They were treated so the IPTW is $\frac{1}{PS} = \frac{1}{0.2} = 5$
  - So in the weighted sample, this person will count for 5 people
  - Think about it this way, this participant's PS was 0.2 so they had an 20% probability of being treated
  - Therefore, this person is kind of rare, they had a low probability of being treated but were treated anyway
  - Therefore we upweight this observation

Examples of going from PS to IPTW

- What is the IPTW of a treated participant with a PS=0.1?
  - They were treated so the IPTW is $\frac{1}{PS} = \frac{1}{0.1} = 10$

Examples of going from PS to IPTW

- What is the IPTW of a treated participant with a PS=0.01?
  - They were treated so the IPTW is $\frac{1}{PS} = \frac{1}{0.01} = 100$
  - This participant counts for 100 people. This means this is a very influential observation
  - We had better be very sure we've measured everything right about this person if they're going to count for so much
  - We'd like to be able to avoid observations that are so heavily weighted

DO NOT FORGET

$$\frac{1}{\mathrm{P}(\text{receiving the treatment they recevied})}$$

NOT

$$\frac{1}{\mathrm{P}(\text{being treated})}$$

Erasmus MC
Universitair Medisch Centrum Rotterdam

DO NOT FORGET

Statistics
in Medicine

P. C. AUSTIN AND E. A. STUART

While it was not the focus of the review, we noted that several sets of authors incorrectly defined the weights as the reciprocal of the propensity score, rather the reciprocal of the probability of receiving the treatment that was actually received.

Erasmus MC
Universitair Medisch Centrum Rotterdam

IPTW

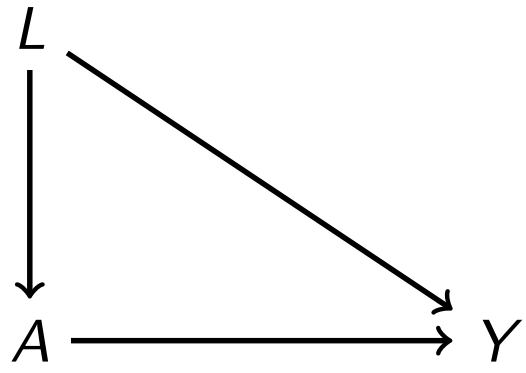
- Weighting by the inverse probability of sampling given group creates a pseudopopulation where there is no relationship between group and sampling status

- Weighting by the inverse probability of treatment received given confounders creates a new population where there is no relationship between confounders and treatment status

IPTW on a DAG

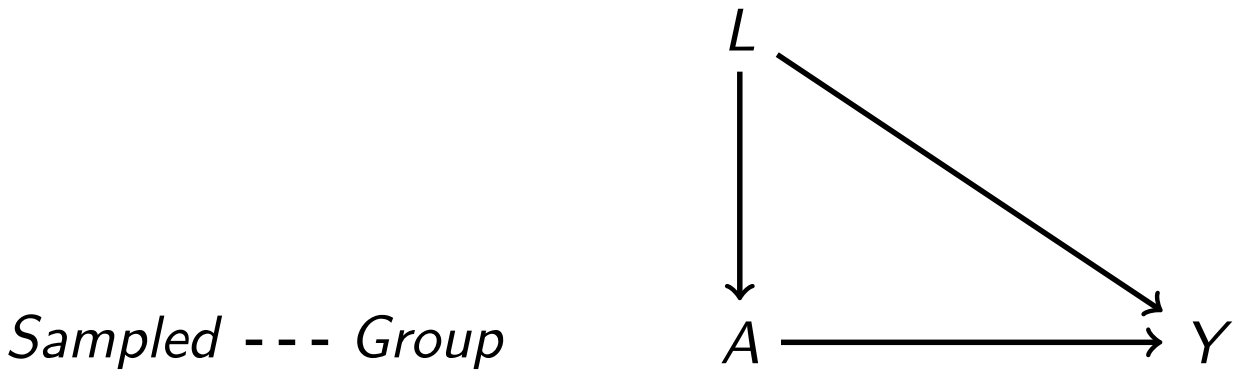
Unweighted

Sampled - - - *Group*

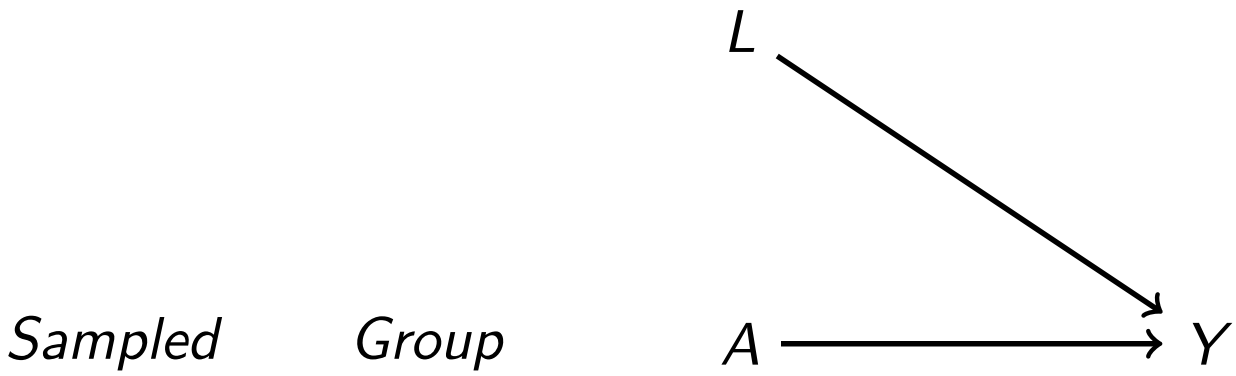


IPTW on a DAG

Unweighted



Weighted



Erasmus MC
Universitair Medisch Centrum Rotterdam

Causal inference with confounder adjustment

- Assumptions
  - Exchangeability (no bias)
    - The average outcome would be the same between the treated/untreated groups if they are set to have the same exposure
  - Consistency (well-defined interventions)
    - How are you going to intervene on the exposure
  - Positivity
    - Do you have both treated and control people in all strata of confounders

Exchangeability

- In more basic terms, this is the assumption that there is no confounding or selection bias
- In more complicated terms, it assumes the counterfactual Y^a is independent (not correlated with) the observed treatment

Consistency

- This assumption does the magic of tying real world observations to the world we could have observed if we had changed someone's exposure (counterfactual)
- For our purposes, you can think of this as the assumption of well-defined interventions
- What is the effect of reducing BMI by one unit on coronary heart disease?
  - This is not a well-defined intervention
  - There are many ways a person's BMI can be reduced not all of which will have the same effect

Positivity

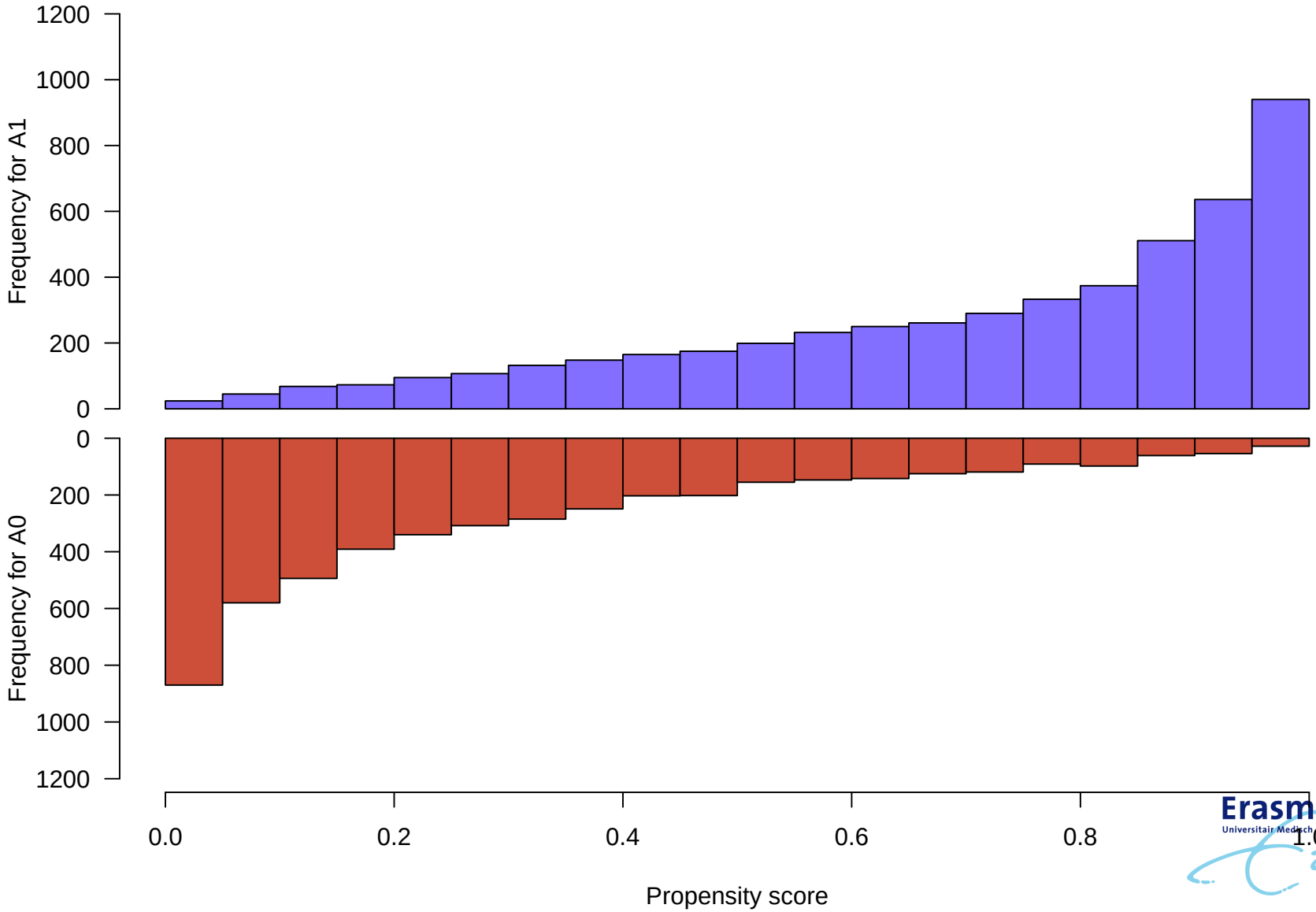
- There are both exposed and unexposed people for every possible combination of confounders
- Simple example, imagine sex is the only confounder in our study
  - Among men we have both exposed and unexposed people
  - But all the women in our study are exposed
  - Is it possible now to adjust for sex?
- Positivity ensures we don't have to extrapolate across levels of confounder
- One of the advantages of IPTW, you'll see, is that it's easier to check this assumption

Three things to check before estimation

1. Positivity
   - This is done by plotting a histogram of the propensity scores (not the weights) within each group
   - Positivity is satisfied when the histogram for the exposed and the unexposed overlap completely
   - This is known as looking for common support
   - If only a small portion of the histograms do not overlap, this might be *random violations* of positivity
   - If large portions of the histograms do not overlap, this might be a *structural violation* of positivity. If this occurs, it is likely to bias your effect estimates.

Three things to check before estimation

1. Positivity

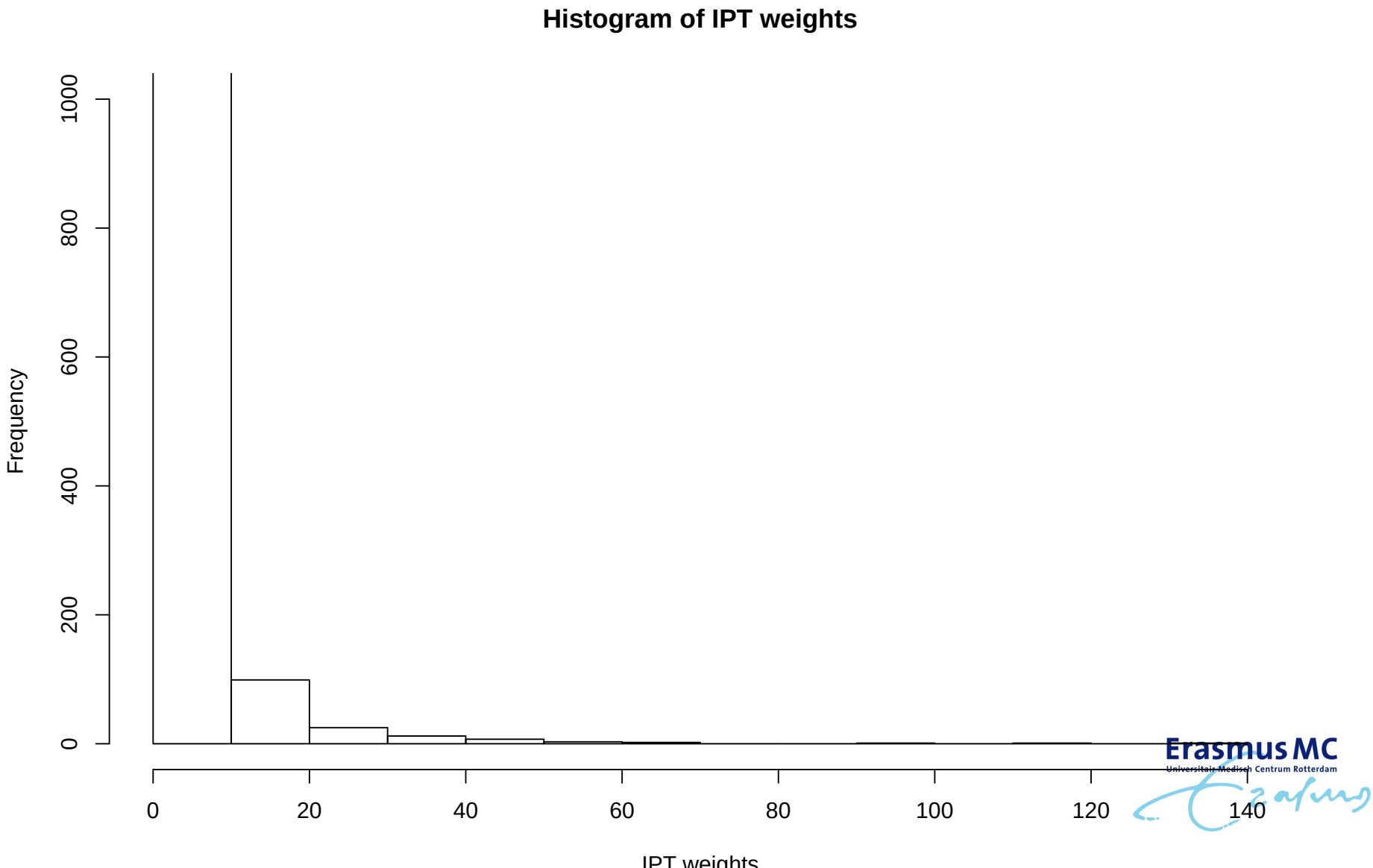


Three things to check before estimation

2. The mean and distribution of your weights
   - These weights should have a mean of 2
   - If you use different weights (*e.g.* stabilized weights) the mean might be something else
   - Check the range of your weights. Very large weights indicate that some observations are being given a lot of weight meaning they are very influential.

Three things to check before estimation

- Weights range: 1.0, 138.5
- Weights mean: 2.0



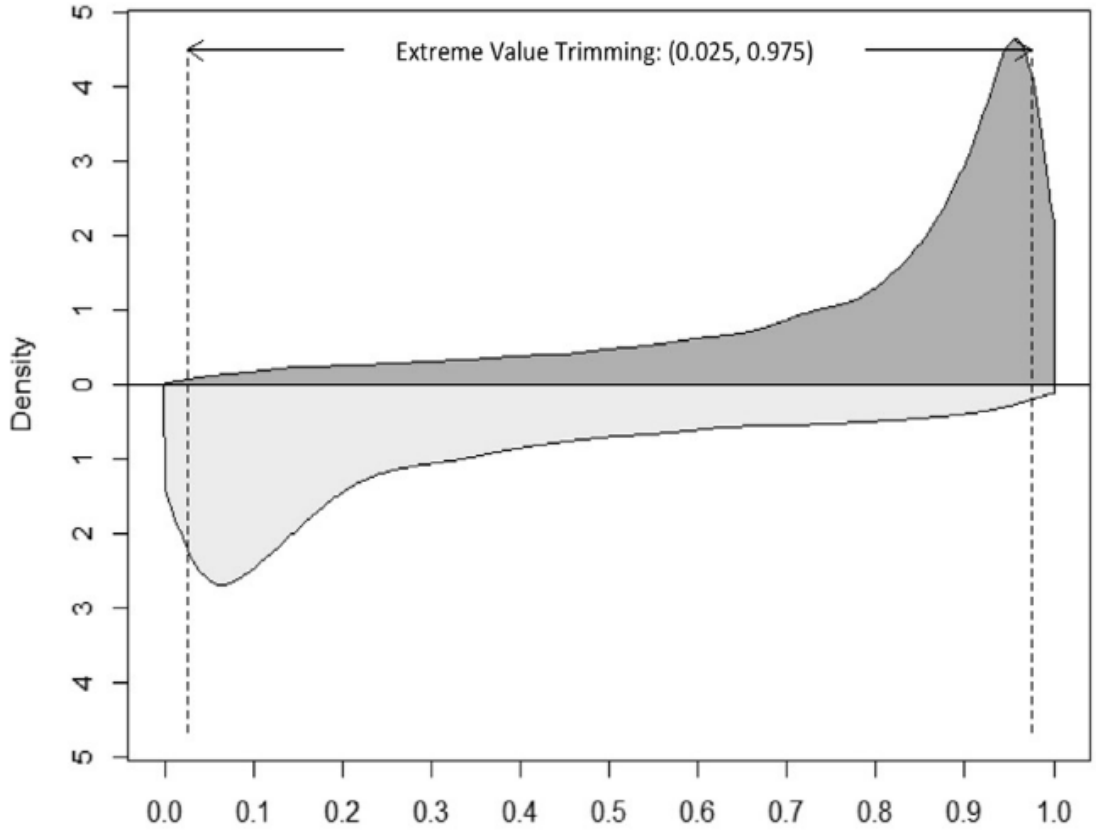
To check positivity and weight distribution

A

	Propensity Score Range*	Treated (n)	Untreated (n)
Before Trimming	0.0018, 0.9942	17,019	12,255
Trimmed		1,611	796
After Trimming**	0.025, 0.975	15,408	11,459

* Model includes 47 covariates,

** Extreme value trimming (0.025, 0.975)



(Shrier, Pang and Platt, 2017)

Erasmus MC
Universitair Medisch Centrum Rotterdam

Trimming

- Sometimes observations are trimmed (removed) if:
  - the PS is in a very high or very low percentile (influential data points)
  - in ranges of the PS where there is no overlap between the exposure and unexposed groups
- There is some debate around trimming
- Some think about this in terms of structural or random non-positivity
- Some prefer trimming as it can increase the precision of your estimates
- Ideally, trimming will increase your precision but not change your estimate very much
- If trimming changes, you have to make a decision about whether the observations being trimmed are representative or not

Three things to check before estimation

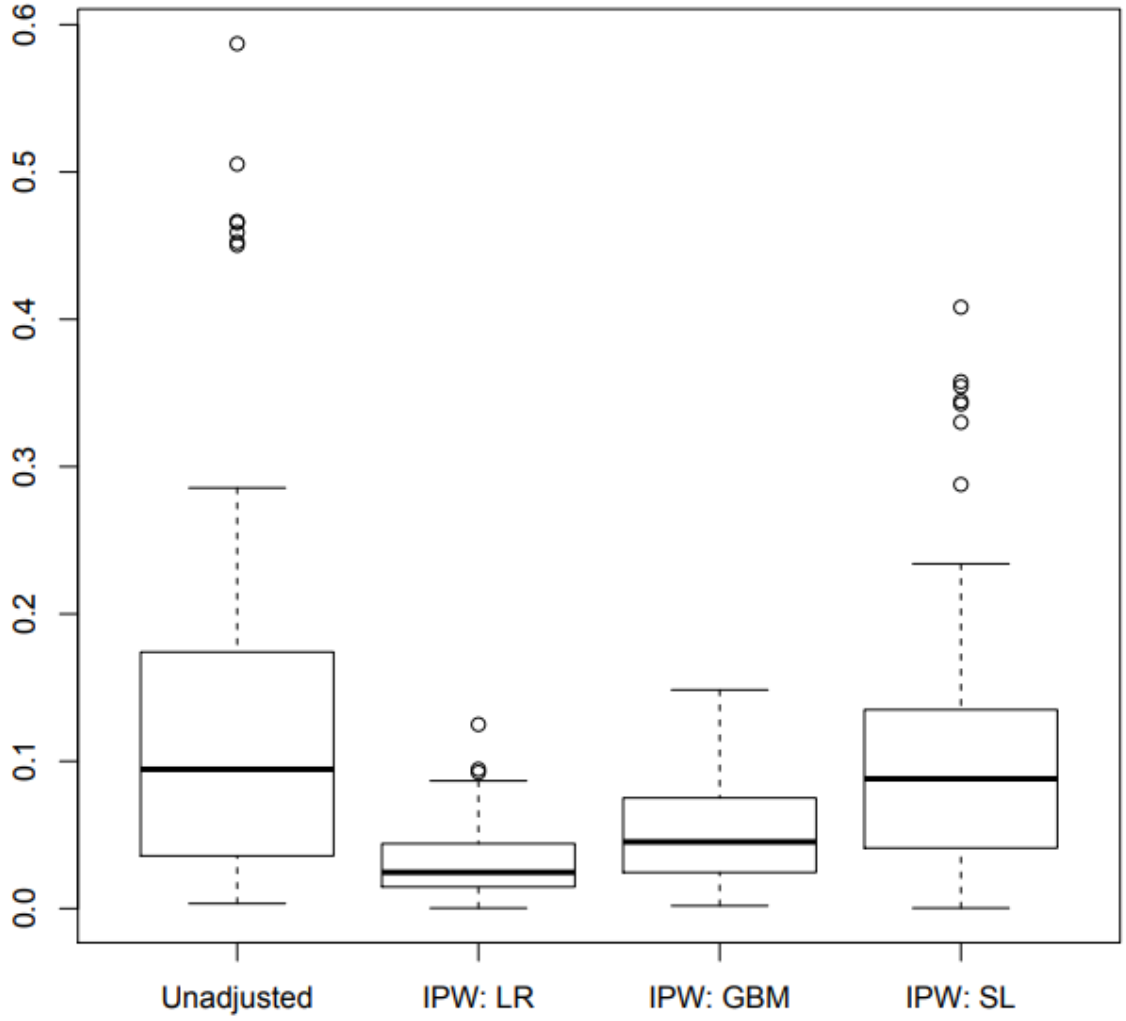
3. Balance in covariates
   - This can be done by using the 'w=' option in R (or the WEIGHT function in SAS)
   - You can reweight any descriptive command using the 'w=' option to check the balance across exposure groups in the weighted pseudopopulations
   - Some suggest checking for balance in higher order variables as well (squared terms or interactions between variables)
   - Again, you can go back and change your treatment model if you're not happy with the balance
   - If we reweight the table one from our example before. On the left the original table one and on the right the weighted table one:

	A0	A1
age	44.51	55.57
sex	0.55	0.47
BMI	26.20	27.80

	A0	A1
age	50.42	50.49
sex	0.51	0.50
BMI	27.06	27.06

Balance

- Don't assume your model is going to do the balancing for you
- A more complex model will not necessarily give you better balance



Erasmus MC
Universitair Medisch Centrum Rotterdam

Balance

- Many programs will show balance statistics based on standardized differences
- If you have external knowledge about which variables are stronger confounders (*i.e.* more strongly related to the outcome), you could prioritize balance among those variables

TABLE. Balance Diagnostics: Standardized Mean Differences in National Health and Nutrition Examination Survey

		Propensity Score Quintiles						
Var.	**Orig.**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Match**	**IPW**
Logistic regression[a] (PA = 0.70)								
Gender	0.138	0.102	0.104	0.029	0.200	0.031	0.006	0.023
Age	0.592	0.257	0.171	0.099	0.311	0.164	0.002	0.014
Race	0.315	0.317	0.112	0.344	0.415	0.287	0.120	0.052
Educ.	0.512	0.538	0.417	0.280	0.238	0.302	0.133	0.029
Marital	0.488	0.432	0.239	0.272	0.233	0.261	0.094	0.023
Household income	0.297	0.628	0.512	0.551	0.644	0.591	0.147	0.066
Poverty	0.453	0.087	0.126	0.114	0.004	0.146	0.049	0.000

(Moodie and Stephens, 2017)

4. Estimation of marginal structural model

Ok. I estimated my PS. Now what?

- The PS is continuous, so you can't just analyze people with the same value
- There are a number of things you can do with the PS once you've estimated it
  - Stratify
  - Match
  - IPTW

Estimation

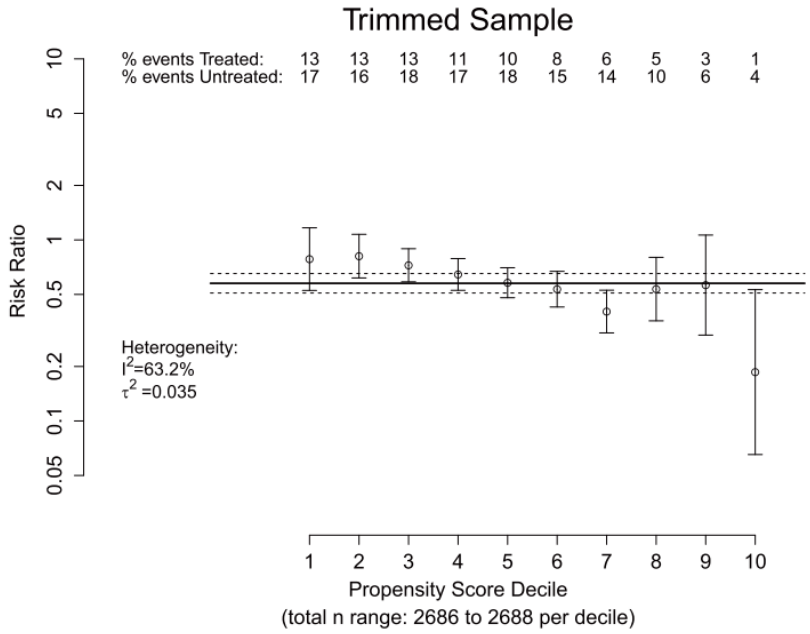
- Once you are satisfied with the balance you can estimate your effect
- Now, after all this, we finally come back to our MSM: $E[Y^a] = \beta_0 + \beta_1 * a$
- Because or weights have removed the assocation between A and L, we can estimate this model directly without having to include any covariates
- Estimate a bivariate regression of your outcome on your exposure (do not include any covariates) and weight this model by your IPTW weights
- $Y\ \alpha + A$, weights=IPTW
- Voila, you have now used MSM estimated with IPTW to estimate a causal effect
- In our simple example from before, our crude estimate was 2.6. If we reweight that model using our IPTW weights we get -4.8 which is much closer to the true value of -5.

Effect modification with IPTW

- What if you want to estimate effects in different subgroups
- You can include interaction terms in your outcome model
- $Y\ \alpha + A + V + A * V$, weights=IPTW
  - If V is binary, the coefficient for A will be the effect of A when V=0
  - The coefficient for A plus the coefficient for A*V will be the effect of A when V=1
  - The coefficient for V has no causal interpretation because we are only making causal assumptions for A (positivity, consistency and exchangeability)

What if the PS is an effect modifier?

- You can estimate the IPTW effect among quantiles of the PS



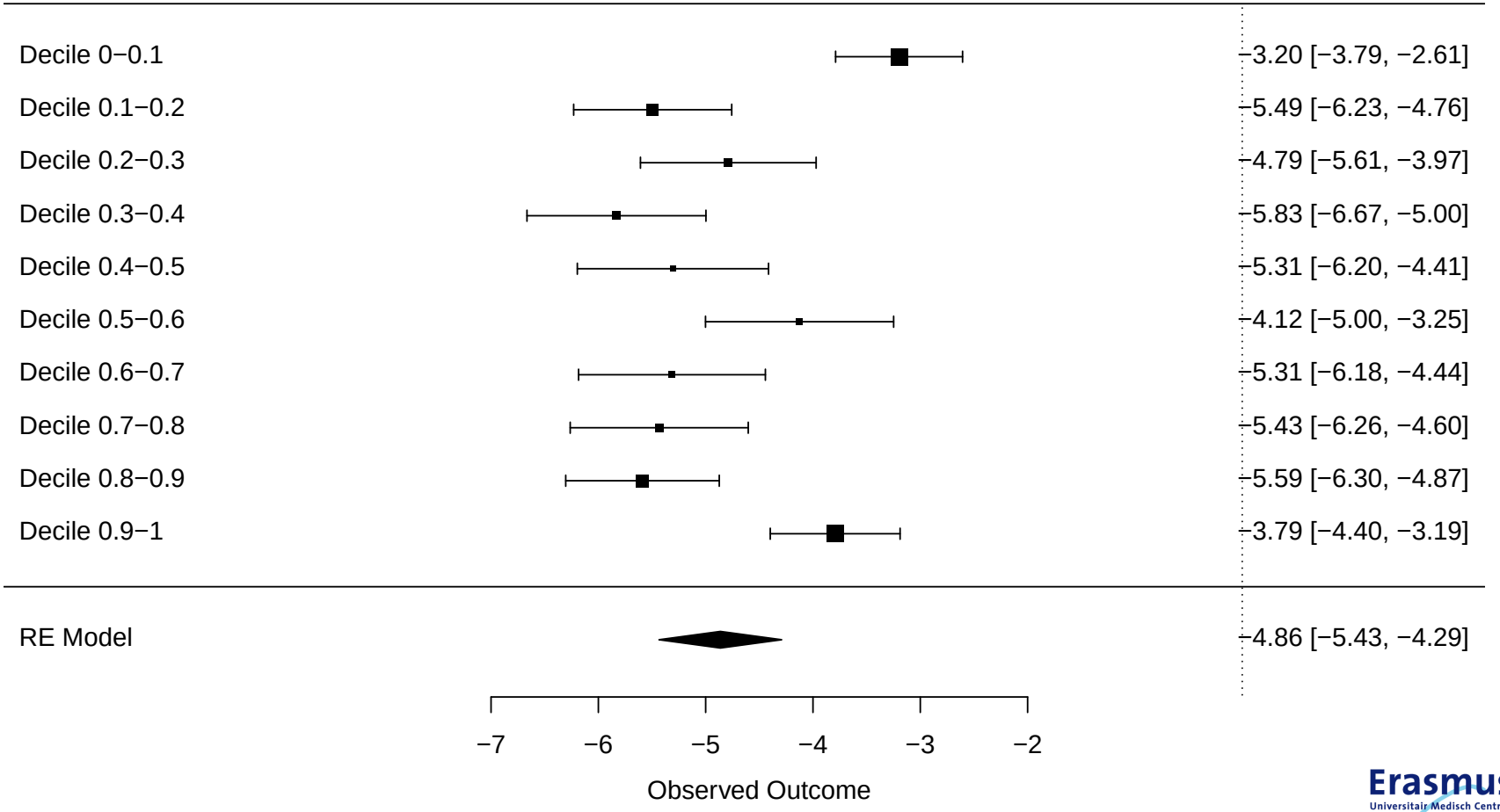
showing that the between-study variance equals 0.035 and that 63.2% of the variability in effect estimates is due to between-study heterogeneity rather than sampling error. Although the I^2 suggests moderate relative heterogeneity, all of the effect estimates across the propensity score range would lead to the same decision and are clinically homogeneous. Therefore, reporting one overall effect estimate might be considered appropriate in our example. In other data, reporting only a single estimate might lead to the loss of important information for medical decision making. Finally, the top of Fig. 2 indicates the absolute proportion of treated and untreated participants who had an event, which is also required for informed decision making

(Shrier, Pang and Platt, 2017)

Erasmus MC
Universitair Medisch Centrum Rotterdam

What if the PS is an effect modifier?

- $I^2 = 82.2$
- $\tau^2 = 0.69$



Erasmus MC
Universitair Medisch Centrum Rotterdam

MSM, step by step

1. Model your treatment as a function of your confounders
2. Check balance of the covariates between treatment groups
3. Check positivity
4. Check mean and distribution of weights
5. Go back to step 1 if you aren't satisfied with the results in steps 2, 3, and 4
6. Estimate the marginal structural model

Confidence intervals

- Two ways of estimating confidence intervals:
  - Non-parametric bootstrap
  - Robust variance estimator (conservative)
- Can also estimate p-values (but don't!)

Advantages of PS for adjustment

- Directly checks positivity
- Dimension reduction
- Outcome model non-parametric
- Can play with treatment model to check balance
- Looks like a trial

Disadvantages of PS for adjustment

- Will be biased with non-positivity
- Will be biased with misspecification of the treatment model

Other ways you can use the PS

- Stratification
  - Stratify the PS, estimate the effect within each stratum and pool
- Targeted maximum likelihood estimation
  - Doubly robust
  - Uses the propensity score to create the "clever covariate"
- Matching
  - Match people with similar PSes and conduct a matched analysis

Additional points

- Propensity score methods are not quasi-experimental. It only adjusts for the confounders that you put into the propensity score.
- Think about whether it is easier to model the exposure or model the outcome

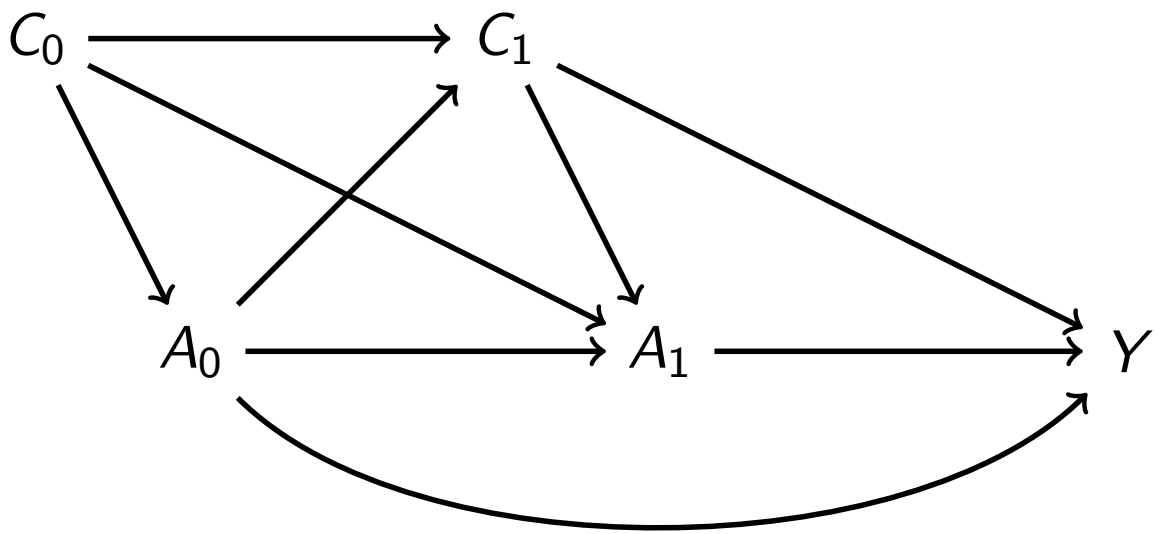
Weights for censoring

- Weights can be used in a very similar way to correct for losses to follow up
- Estimate the probability of NOT being lost to follow-up given baseline covariates
- Weight the observed observations using the inverse of these weights
- This creates a pseudopopulation that resembles the population if no one was lost to follow-up

5. MSM with time-varying exposures

MSM with time-varying exposures: why we need them

- If you want to know the effect of A_0 and A_1 in the graph below, what do you adjust for?
- Pre-1986 there would have been no way to answer this



Weights for MSM with time-varying exposure

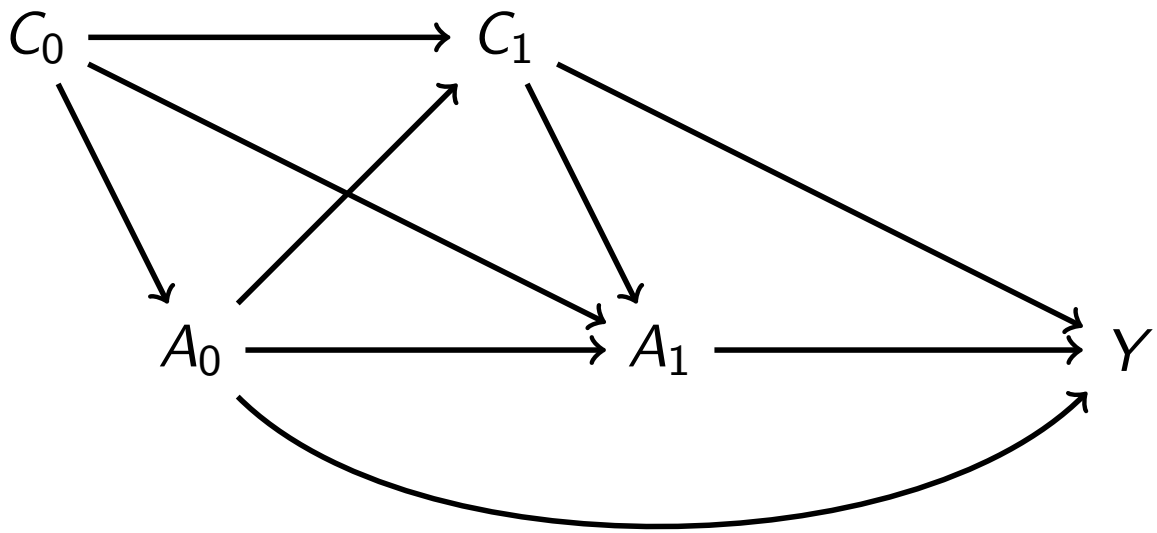
- Unstabilized:

$$W^{\bar{A}} = \prod_{k=0}^{K} \frac{1}{f(A_k|A_{k-1}, \bar{L}_k)}$$

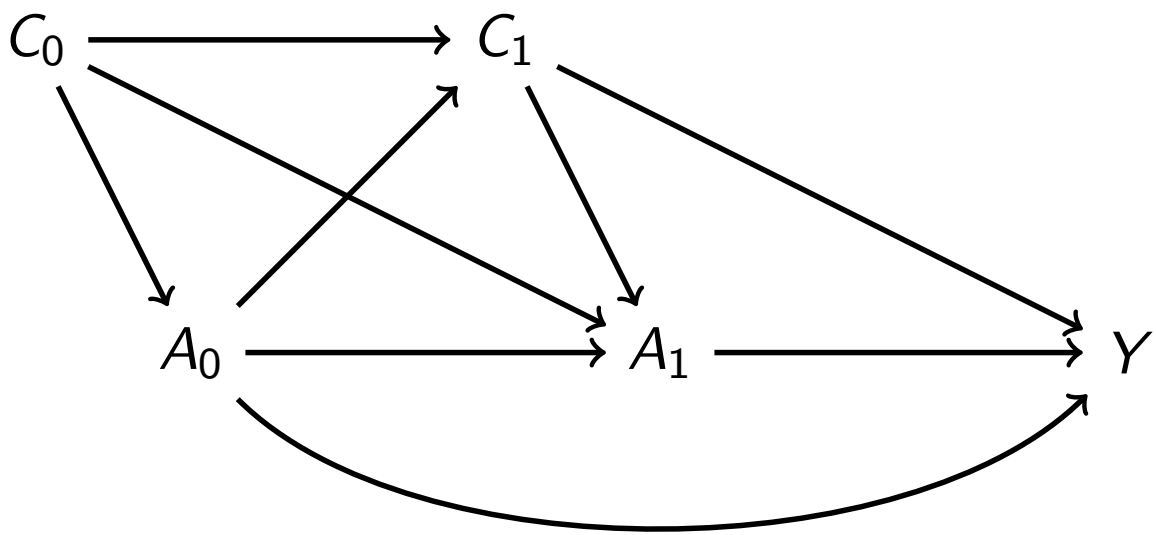
- Stabilized:

$$SW^{\bar{A}} = \prod_{k=0}^{K} \frac{f(A_k|A_{k-1})}{f(A_k|A_{k-1}, \bar{L}_k)}$$

- PS for A_0: $P(A_0 = 1|C_0)$
- PS for A_1: $P(A_1 = 1|A_0, C_0, C_1)$

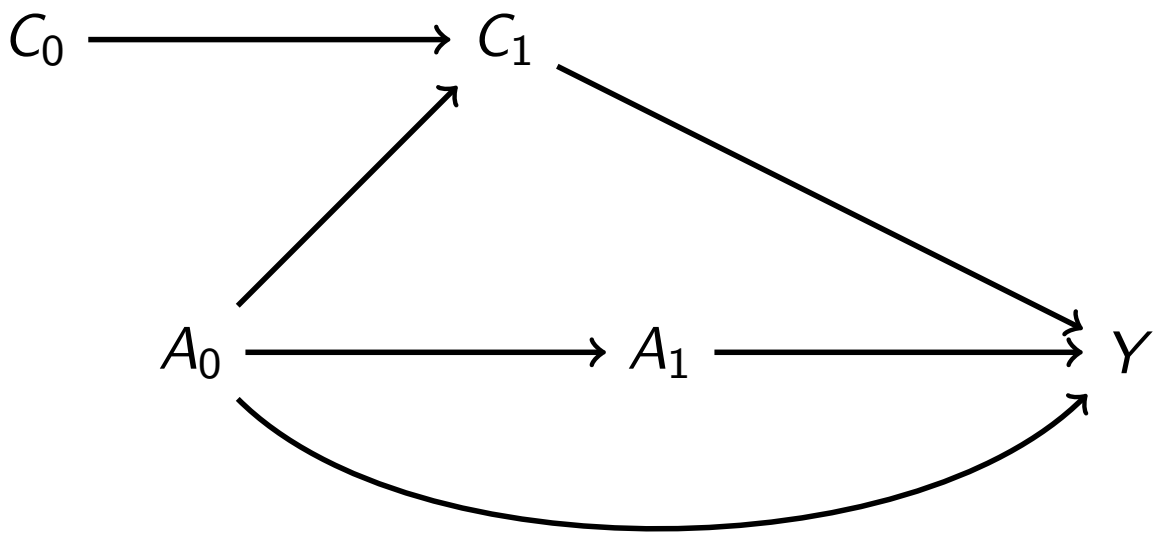


- $IPTW_{A_0}$: $\frac{1}{P(A_0=a_0|C_0)}$
- $IPTW_{A_1}$: $\frac{1}{P(A_1=a_1|A_0,C_0,C_1)}$
- $IPTW_{A_0,A_1}$: $\frac{1}{P(A_0=a_0|C_0)} * \frac{1}{P(A_1=a_1|A_0,C_0,C_1)}$
- Nice paper and software by Jackson 2016 (PMID:27479649) to assess balance for such weigths



Marginal structural models for time-varying exposures

- You can estimate a MSM:
  $E_{IPTW_{A_0,A_1}}[Y^a] = \beta_0 + \beta_{A_0} * a_0 + \beta_{A_1} * a_1$
- Creates a pseudopopulation with the following characteristic:



Erasmus MC
Universitair Medisch Centrum Rotterdam

6. Marginal versus conditional

Marginal versus conditional estimates

- These terms are a mess. Different people can understand TOTALLY different things
- Marginal estimates are the average causal effect in the population if you expose and unexpose everyone
  - IPTW
  - GEE
- Conditional effects are the average effect if you keep all covariates constant
  - Outcome regression
  - Many others
- This means that the IPTW and outcome regression estimates can be different when the outcome is binary or when there is effect estimation

```r
n <- 10000
c <- rbinom(n,1,0.5)
a <- rbinom(n,1,0.1 + 0.5*c)
y <- a + a*c + c + rnorm(n)
```

- What parameters might we be interested in here?
  - $\psi_{A|C=0} = 1$
  - $\psi_{A|C=1} = 2$
  - $\psi_{marginal} = 1.5$

```
n <- 10000
c <- rbinom(n,1,0.5)
a <- rbinom(n,1,0.1 + 0.5*c)
y <- a + a*c + c + rnorm(n)
```

- What estimate do you expect for ‘a’ here?

```
n <- 10000
c <- rbinom(n,1,0.5)
a <- rbinom(n,1,0.1 + 0.5*c)
y <- a + a*c + c + rnorm(n)

lm(y ~ a + c)
```

- What estimate do you expect for 'a' here?

```
n <- 10000
c <- rbinom(n,1,0.5)
a <- rbinom(n,1,0.1 + 0.5*c)
y <- a + a*c + c + rnorm(n)

lm(y ~ a + c) %>% summary %$% coefficients %>% round(2)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    -0.08       0.01   -5.68        0
a               1.71       0.03   67.20        0
c               1.25       0.02   51.62        0
```

- When using IPTW or MSM, we get the right answer without having to specify the interaction between A and C

```
ps <- predict(glm(a ~ c, family=binomial),
              type = "response")
iptw <- (1/ps)*a + (1/(1-ps)*(1-a))

lm(y ~ a, w=iptw) %>% summary %$% coefficients %>% round(2)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.49       0.02   27.00        0
a               1.48       0.03   57.34        0
```

Erasmus MC
Universitair Medisch Centrum Rotterdam

Challenges in interpreting results from 'multiple regression' when there is interaction between covariates

Ian Shrier,[1,2] Annabelle Redelmeier,[3] Mireille E Schnitzer,[4] Russell J Steele[3]

Abstract

Properly interpreting research results is the foundation of evidence-based medicine. Most observational studies use multiple regression

2. This is an observational study and the researchers expect that participants who were doing balance exercises were also more likely to be doing general warm-up exercises (cat-

Figure 1:

Erasmus MC
Universitair Medisch Centrum Rotterdam

7. Treatment models vs outcome models

Treatment models vs outcome models

- Both treatment and outcome models can be causal
- Treatment models have the advantage that you can play around with the model to get the right balance without looking at the outcome
- Ideally, do both and the estimates agree
- Some people say, "think about which you can model better, exposure or outcome." I'm not sure this is good advice.

8. Example

Richardson et al

The effect of free health care on polypharmacy: a comparison of propensity score methods and multivariable regression to account for confounding

Kathryn Richardson[1,2*], Rose Anne Kenny[1,2,3] and Kathleen Bennett[4]

[1] *The Irish Longitudinal Study on Ageing (TILDA), Trinity College Dublin, Dublin, Ireland*
[2] *Department of Medical Gerontology, Trinity College Dublin, Dublin, Ireland*
[3] *Trinity College Institute of Neuroscience, St James's Hospital, Dublin, Ireland*
[4] *Department of Pharmacology and Therapeutics, Trinity Centre for Health Sciences, St James' Hospital, Dublin, Ireland*

ABSTRACT

Purpose Differing healthcare access has implications for public health. In Ireland, eligibility for free public health care is means tested. Here, we examine the association between healthcare access and polypharmacy while accounting for underlying socio-economic and health status differences.

Methods Self-reported regular medication use, history of diagnosed health conditions, disability, socio-demographics, and objective measures of depression and anxiety for adults aged 50–69 years ($n = 5796$) were ascertained from the population-representative Irish Longitudinal Study on Ageing. Objective measures of frailty, cognition, hypertension, and body mass index were also assessed for 4241 participants. The associations between free healthcare access and polypharmacy and use of 15 medication classes were estimated using multivariable modified Poisson regression, adjustment for the propensity score, and inverse probability of treatment weighting by the propensity score.

Results Polypharmacy was reported by 22% and 7% of the 1932 and 3864 participants with and without public healthcare coverage. Public patients had a 21–38% greater risk of polypharmacy depending on the method used to account for confounding. Results were less robust using propensity score weighting. There was evidence that classes of cardiovascular drugs, drugs for acid-related disorders, and analgesics were used more commonly in public patients. Associations were mostly unaffected after also accounting for objective health measures but were significantly attenuated after accounting for frequency of healthcare visits.

Conclusions Publically funded health care in Ireland leads to greater medication use in people aged 50–69 years. This may reflect over-prescribing to public patients or restricted use among those who pay out of pocket.

Richardson et al

- They use outcome regression, outcome regression adjusted for PS and IPTW
- Correctly defined the IPT weights

Propensity scores. Logistic regression was used to estimate the PS representing the probability of 'exposure' to public healthcare coverage given an individual's covariates. For IPTW regression, individuals with public coverage were then weighted by the reciprocal of their PS, and individuals without public coverage were weighted by one minus the reciprocal of their PS. Weights were then stabilized by multiplying by the prevalence of public healthcare coverage

Richardson et al

- Checked balance, Checked positivity, trimmed extreme PS values, checked effect modification by PS

minus the prevalence for those without coverage.
Once the weights were applied, balance was assessed by examining both standardized differences and the degree of confounding remaining as above.[26]

Following usual recommendations for PS analyses, we excluded participants whose PS falls outside the overlapping range between the two sets of patients.[15] The IPTW method can be sensitive to individuals with large weights (i.e. PS very close to 0 or 1), and there may be uncertainty over the coding of such individual's covariates and treatment at the extremes of the PS distribution. We therefore, as a sensitivity analysis, applied asymmetrical trimming and excluded those with PS lying outside the 1st and 99th percentiles of the PS distribution in those with and without public health care, respectively.[19] To examine whether the association between public healthcare coverage and polypharmacy was modified by the PS, we estimated polypharmacy rate ratios in strata defined by the quintiles of the PS. We tested for heterogeneity by performing a four-degree-of-freedom likelihood ratio testing for interaction between public healthcare coverage and indicators for each quintile in a modified

Erasmus MC
Universitair Medisch Centrum Rotterdam

Richardson et al

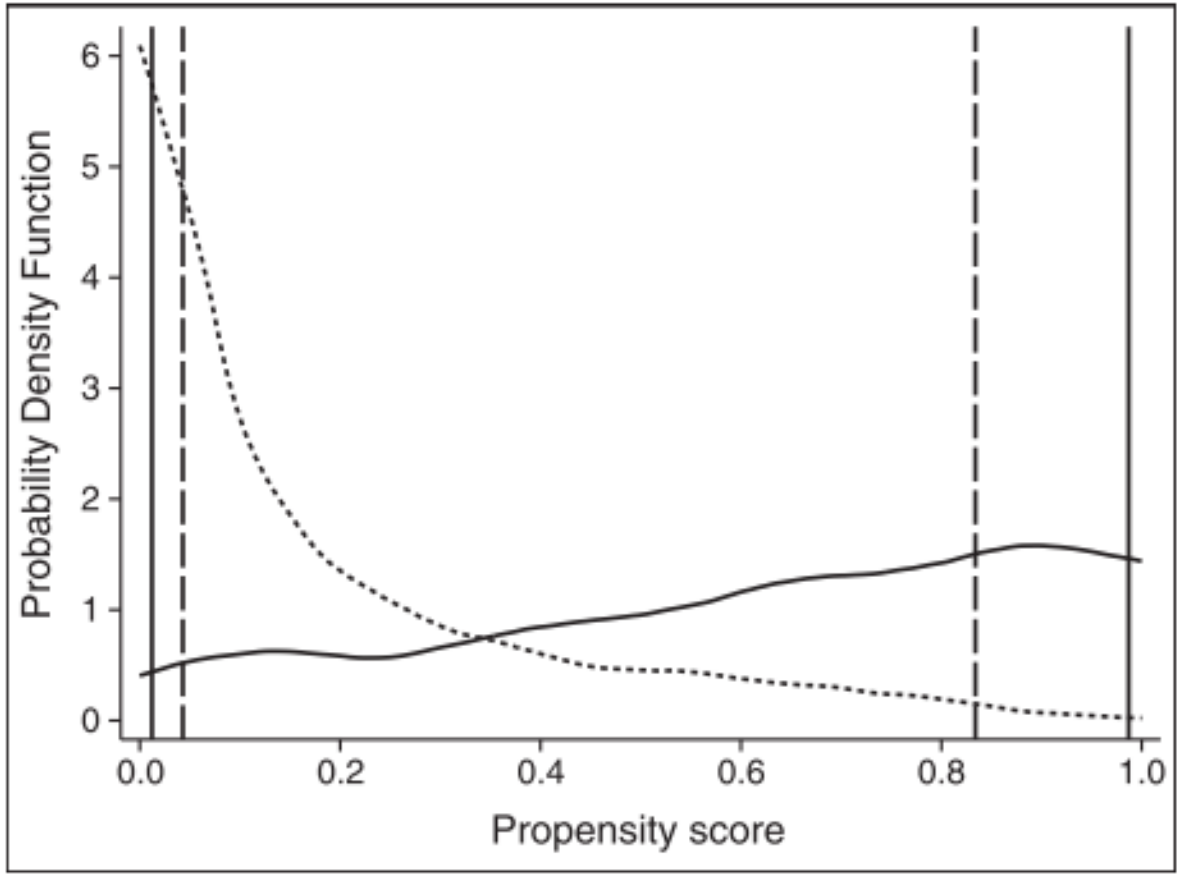
- Covariate balance

Variable	Unweighted ($n=5796$)			Inverse probability of treatment weighted ($n=5494$)		
	Public health care		Change in RR† (%)	Public health care		Change in RR† (%)
	No	Yes		No	Yes	
Women	52.3	59.3	1	54.9	56.9	0
Age (years)	58 (53–62)	60 (55–65)	−13	59 (54–64)	59 (54–64)	0
Married	78.9	60.3	0	73.2	73.1	0
Third/higher education	41.1	15.9	−8	29.8	30.9	2
Urban	53.3	48.5	1	49.1	48.1	0
Employed full time	35.4	9.4	−33	25.2	28.3	2
Household income < €20 000§	13.7	48.0	−10	25.7	25.9	−1
Renting§	3.1	21.0	1	8.1	8.7	0
Current smoker	16.0	29.3	−1	20.2	20.3	0
Health conditions						
Angina	2.1	6.2	−12	3.1	3.0	1
Arrhythmia	5.0	6.5	−3	5.4	4.5	1
Diabetes	4.9	9.6	−14	6.2	5.6	2
Heart attack	2.0	4.9	−8	3.0	2.5	2
High cholesterol	37.7	39.6	−2	37.4	36.6	1
Hypertension	28.1	39.4	−15	32.5	32.3	0
Other cardiovascular diseases	2.5	3.5	−2	2.9	2.1	2
Arthritis	19.3	29.0	−8	23.2	22.8	1
Asthma	8.2	11.9	−4	9.0	8.4	1
Chronic lung disease	2.4	6.4	−4	3.7	3.6	0
Osteoporosis	7.5	10.4	−2	8.8	9.5	0
Stomach ulcer	6.2	9.5	−2	6.8	6.8	1
Cataracts	3.9	6.1	−3	4.9	4.5	1
Glaucoma	1.3	2.4	−1	2.1	1.8	0
Anxiety	4.5	7.3	−3	5.3	5.8	0
Depression	4.3	9.2	−4	6.3	5.9	1
Moderate/severe chronic pain	19.8	34.1	−18	24.2	23.9	0
Disability (ADL or IADL)	5.5	14.2	−14	8.7	8.6	0
Impairments	1 (0–2)	2 (0–4)	−30	1 (0–2)	1 (0–3)	0
Depressive symptoms§	3 (0–7)	5 (1–11)	−10	3 (1–8)	3 (1–8)	0
Anxiety symptoms§	5 (3–7)	6 (3–9)	−2	5 (3–8)	5 (3–8)	1

Erasmus MC
Universitair Medisch Centrum Rotterdam

Richardson et al

- Density of PS by exposure



Erasmus MC
Universitair Medisch Centrum Rotterdam

Richardson et al

- Checked weights

selected covariates, there remained imbalances in the excluded education and income variables. The mean (standard deviation) and range of the stabilized IPTW weights using *a priori* and backwards stepwise selected covariates were 1.00 (1.21) and 0.4–28.5 and 1.01 (0.82) and 0.4–12.6, respectively. Applying asymmetrical trimming at the 1st and 99th percentiles

Richardson et al

- Results

	Full sample ($n = 5796$)		
Model	RR	95%CI	SE
Crude	3.05	2.65, 3.50	0.22
Multivariable model			
A priori covariates†	1.38	1.14, 1.67	0.13
Backwards stepwise‡	1.36	1.13, 1.63	0.13
Regression adjusted for propensity score			
A priori covariates†	1.33	1.13, 1.56	0.11
Trimmed§	1.36	1.14, 1.61	0.12
Backwards stepwise‡	1.32	1.12, 1.55	0.11
Trimmed§	1.36	1.14, 1.63	0.12
IPTW model			
A priori covariates†	1.25	1.04, 1.50	0.12
Trimmed§	1.26	1.04, 1.53	0.12
Backwards stepwise‡	1.21	0.99, 1.49	0.13
Trimmed§	1.30	1.06, 1.59	0.13

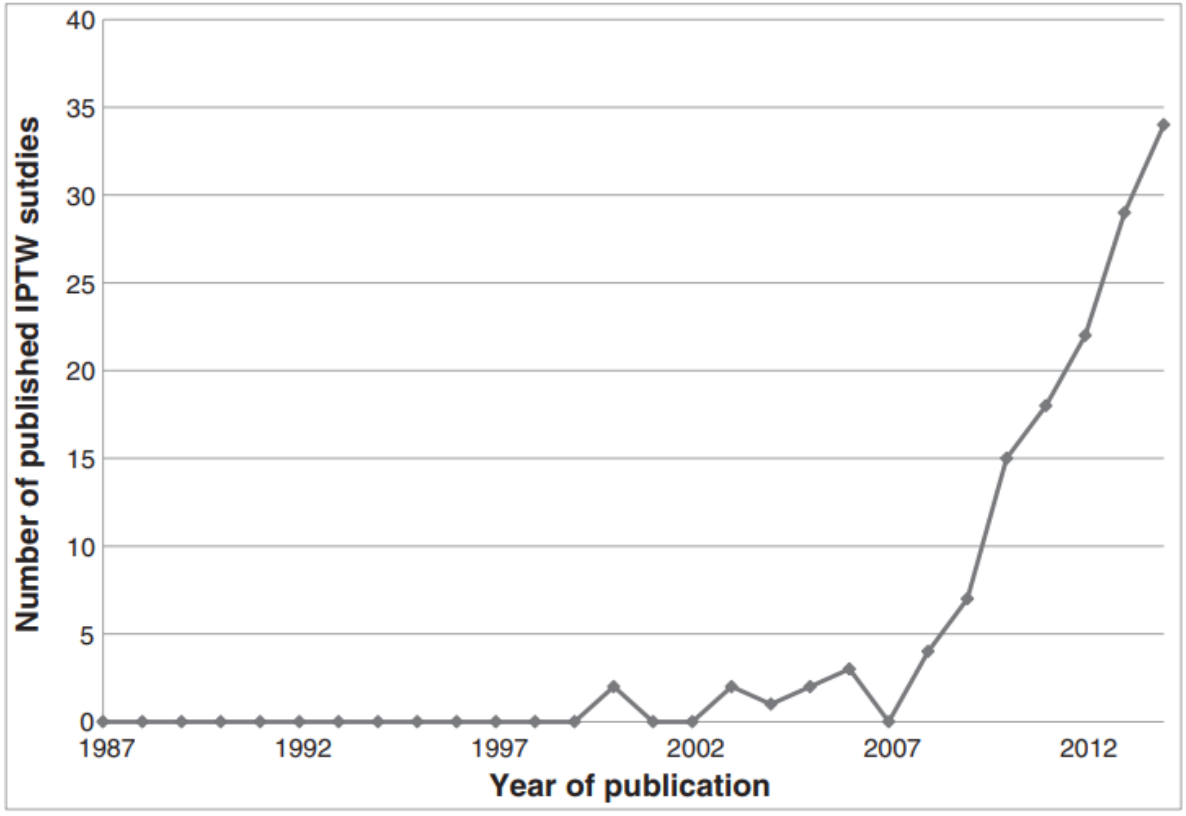
Erasmus MC
Universitair Medisch Centrum Rotterdam

Richardson et al

- This is a nice IPTW paper
  - Correct definition of IPT weights
  - Check balance
  - Check distribution of PS
  - Check weights
  - Check positivity
  - Check how trimming affects results
  - Check heterogeneity by PS

Exercise

- Now we will do a short exercise where you will estimate a PS, calculate the IPT weights and estimate a causal effect.
- Feel free to email me: j.labrecque@erasmusmc.nl



Erasmus MC
Universitair Medisch Centrum Rotterdam

References

- Causal Inference, Hernan and Robins
- Modern Epidemiology, Rothman, Greenland and Lash
- Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Austin and Stuart, 2015. Statistics in Medicine 34(28):3661-3679.
- Propensity Score Methods for Confounding Control in Non-Experimental Research. Brookhart et al, 2013. Circ Cardiovasc Qual Outcomes 6(5): 604-611.
- Moodie and Stephens, 2017. Treatment prediction, balance and propensity score adjustment. *Epidemiology* 28(5):e51-53.
- Shrier, Pang and Platt, 2017. Graphic report of the results from propensity score method analyses. *J Clin Epi* 88:154-159.