

Sensitivity Analysis for Unmeasured Confounding: Formulation, Implementation, Interpretation

Joseph W Hogan

Department of Biostatistics
Brown University School of Public Health

CIMPOD, February 2016

Overview

- What is unmeasured confounding?
 - ▶ Ignorable treatment assignment: an untestable assumption
 - ▶ Representations of unmeasured confounding
- Example 1: binary treatment, binary outcome, no covariates
 - ▶ Defining average treatment effect
 - ▶ Bounds on treatment effect estimate
 - ▶ Sensitivity to unmeasured confounding
- Example 2: binary treatment, continuous outcome, covariates
 - ▶ Estimate ATE using G computation algorithm
 - ▶ Sensitivity to unmeasured confounding
- Summary & comparison of methods

Why is unmeasured confounding important?

- Important source of uncertainty in observational studies
- Sensitivity to assumptions is related to quality of evidence
- PCORI recommendations for reporting

Problems posed by unmeasured confounding

In observational studies, existence of unmeasured confounding can lead to biased estimates of causal effect

However it is not possible to test the 'no unmeasured confounding' null hypothesis.

Important questions about unmeasured confounding as they relate to drawing inference and reporting results about causal effects:

- How should it be represented?
- How to assess effects on bias and uncertainty?

Some (simple) notation

We use the **potential outcomes** framework

Y_0 = outcome if treatment not received

Y_1 = outcome if treatment received

The **observed data** for an individual are (A, Y, \mathbf{X})

$$A = \begin{cases} 1 & \text{if treatment received} \\ 0 & \text{if not} \end{cases}$$

$$Y = \begin{cases} Y_1 & \text{if } A = 1 \\ Y_0 & \text{if } A = 0 \end{cases}$$

\mathbf{X} = measured covariates

Average treatment effect (ATE)

$$E(Y_1 - Y_0)$$

What do we mean by unmeasured confounder?

First, need to define **ignorable treatment assignment**:

Treatment assignment is ignorable if there exists a subset $\mathbf{X}^* \subseteq \mathbf{X}$ such that

$$Y_0 \perp\!\!\!\perp A \mid \mathbf{X}^* \quad \text{and} \quad Y_1 \perp\!\!\!\perp A \mid \mathbf{X}^*$$

- For purposes of this talk, this is the same as 'no unmeasured confounders'.
- Means that treatment is randomized within levels of \mathbf{X}^*
- If this condition does not hold, there is unmeasured confounding

Representations of unmeasured confounding

Added variable representation

- There exists an unmeasured confounder U
- Formulate model of its relationship to outcome and treatment assignment

Potential outcomes representation

- The unmeasured confounder is the unobserved potential outcome
- Specify distribution of unobserved potential outcome, conditional on observed data

Potential outcome representation

- In potential outcome representation, the unmeasured confounder is the the unobserved potential outcome

$$Y_{1-A}$$

- Sensitivity analyses therefore based on comparing its distribution to that of the *observed* potential outcomes, e.g.,

$$P(Y_0 | A = 0) \quad \text{vs} \quad P(Y_0 | A = 1)$$

- Focus here: estimation of means

Breaking down ATE

Proportion receiving treatment: $p = \Pr(A = 1)$

The ATE is a difference of weighted averages

$$E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$$

$$E(Y_1) = pE(Y_1 | A = 1) + (1 - p)E(Y_1 | A = 0)$$

$$E(Y_0) = pE(Y_0 | A = 1) + (1 - p)E(Y_0 | A = 0)$$

What can be estimated from data?

Breaking down ATE

- First note that because $Y = Y_1$ when $A = 1$, can write

$$E(Y_1 | A = 1) = E(Y | A = 1)$$

- Likewise

$$E(Y_0 | A = 0) = E(Y | A = 0)$$

- What this means:
 - ▶ We can estimate $E(Y_0 | A = 0)$ using the sample mean of Y among $A = 0$
 - ▶ We can estimate $E(Y_1 | A = 1)$ using the sample mean of Y among $A = 1$

Example 1: Clofibrate trial

Coronary Drug Project Research Group, NEJM 1980

Table 1. Five-Year Mortality in Patients Given Clofibrate or Placebo, According to Cumulative Adherence to Protocol Prescription.

ADHERENCE *	TREATMENT GROUP			
	CLOFIBRATE		PLACEBO	
	<i>no. of patients</i>	<i>% mortality †</i>	<i>no. of patients</i>	<i>% mortality †</i>
<80%	357	24.6±2.3 (22.5)	882	28.2±1.5 (25.8)
≥80%	708	15.0±1.3 (15.7)	1813	15.1±0.8 (16.4)
Total study group	1065	18.2±1.2 (18.0)	2695	19.4±0.8 (19.5)

Example 1: Clofibrate trial

Coronary Drug Project Research Group, NEJM 1980

Table 1. Five-Year Mortality in Patients Given Clofibrate or Placebo, According to Cumulative Adherence to Protocol Prescription.

ADHERENCE *	TREATMENT GROUP			
	CLOFIBRATE		PLACEBO	
	<i>no. of patients</i>	<i>% mortality †</i>	<i>no. of patients</i>	<i>% mortality †</i>
<80%	357	24.6±2.3 (22.5)	882	28.2±1.5 (25.8)
≥80%	708	15.0±1.3 (15.7)	1813	15.1±0.8 (16.4)
Total study group	1065	18.2±1.2 (18.0)	2695	19.4±0.8 (19.5)

Example 1: Clofibrate trial

Coronary Drug Project Research Group, NEJM 1980

Consider treatment arm only

$$\begin{aligned} Y &= 1 \text{ if died before 5 years} \\ &= 0 \text{ if not} \end{aligned}$$

$$\begin{aligned} A &= 1 \text{ if } \geq 80\% \text{ compliant with clofibrate} \\ &= 0 \text{ if not} \end{aligned}$$

Objective: estimate causal effect of complying with treatment

Example 1: Clofibrate trial

Proportion compliant: $\hat{p} = 708/1065 = .67$

Mortality proportion by compliance status

	$\hat{E}(Y A)$
$A = 1$	$106/708 = .15$
$A = 0$	$88/357 = .25$

Example 1: Clofibrate trial

- Observed data:

	$\hat{E}(Y_1 A)$	$\hat{E}(Y_0 A)$
$A = 1$.15	*
$A = 0$	**	.25

- Average treatment effect

$$\begin{aligned} \text{ATE} &= \hat{E}(Y_1) - \hat{E}(Y_0) \\ &= (.67)(.15) + (1 - .67)(**) - \{(.67)(*) + (1 - .67)(.25)\} \end{aligned}$$

- Point estimates

- ▶ Under ITA: $-.10$
- ▶ Lower bound: $-.65$
- ▶ Upper bound: $.35$

Example 1: Clofibrate trial

- Observed data:

	$\hat{E}(Y_1 A)$	$\hat{E}(Y_0 A)$
$A = 1$.15	.25
$A = 0$.15	.25

- Average treatment effect

$$\begin{aligned} \text{ATE} &= \hat{E}(Y_1) - \hat{E}(Y_0) \\ &= (.67)(.15) + (1 - .67)(.15) - \{(.67)(.25) + (1 - .67)(.25)\} \end{aligned}$$

- Point estimates

- ▶ **Under ITA: $-.10$**
- ▶ Lower bound: $-.65$
- ▶ Upper bound: $.35$

Example: Clofibrate trial

- Observed data:

	$\hat{E}(Y_1 A)$	$\hat{E}(Y_0 A)$
$A = 1$.15	1
$A = 0$	0	.25

- Average treatment effect

$$\begin{aligned} \text{ATE} &= \hat{E}(Y_1) - \hat{E}(Y_0) \\ &= (.67)(.15) + (1 - .67)(0) - \{(.67)(1) + (1 - .67)(.25)\} \end{aligned}$$

- Point estimates

- ▶ Under ITA: $-.10$
- ▶ **Lower bound: $-.65$**
- ▶ Upper bound: $.35$

Example: Clofibrate trial

- Observed data:

	$\hat{E}(Y_1 A)$	$\hat{E}(Y_0 A)$
$A = 1$.15	0
$A = 0$	1	.25

- Average treatment effect

$$\begin{aligned} \text{ATE} &= \hat{E}(Y_1) - \hat{E}(Y_0) \\ &= (.67)(.15) + (1 - .67)(1) - \{(.67)(0) + (1 - .67)(.25)\} \end{aligned}$$

- Point estimates

- ▶ Under ITA: $-.10$
- ▶ Lower bound: $-.65$
- ▶ **Upper bound: $.35$**

Summary: Bounds

- No point estimates!
- Conveys lack of information in observed data
- Implicitly gives equal weight to all values within the interval
- Does not rely on any assumptions about unmeasured confounding
- Does not use covariate information
- In principle works for binary outcome, binary treatment
- Not plausible for continuous outcomes

Sensitivity Analysis

What does it mean to do sensitivity analysis?

- 'Unmeasured confounding' is a phenomenon that cannot be observed
- Sensitivity analysis
 - ▶ Make assumptions about things you can't see
 - ▶ Vary those assumptions to see how analysis changes

Sensitivity Analysis

Define two sensitivity parameters:

$$\delta_0 = E(Y_0 | A = 1) - E(Y_0 | A = 0)$$

$$\delta_1 = E(Y_1 | A = 1) - E(Y_1 | A = 0)$$

Interpretation: Compliance, captured by A , is a behavioral characteristic.

δ_0 = difference in mortality rate between compliers and non-compliers, under scenario that **none** received treatment

δ_1 = difference in mortality rate between compliers and non-compliers, under scenario that **all** received treatment

Example

If compliers have lower mortality, even in the absence of treatment, then $\delta_0 < 0$.

A Simple Sensitivity Analysis for the Clofibrate Trial

Rearrange to represent quantities that *cannot* be estimated in terms of those that *can* be estimated:

In terms of the Clofibrate trial:

$$\begin{aligned}\delta_0 &= E(Y_0 | A = 1) - .25 \\ \delta_1 &= .15 - E(Y_1 | A = 0)\end{aligned}$$

A Simple Sensitivity Analysis for the Clofibrate Trial

Rearrange terms:

$$E(Y_0 | A = 1) = .25 + \delta_0$$

$$E(Y_1 | A = 0) = .15 - \delta_1$$

Ignorable treatment assignment (no unmeasured confounding):

$$\delta_0 = \delta_1 = 0$$

Introduce unmeasured confounding:

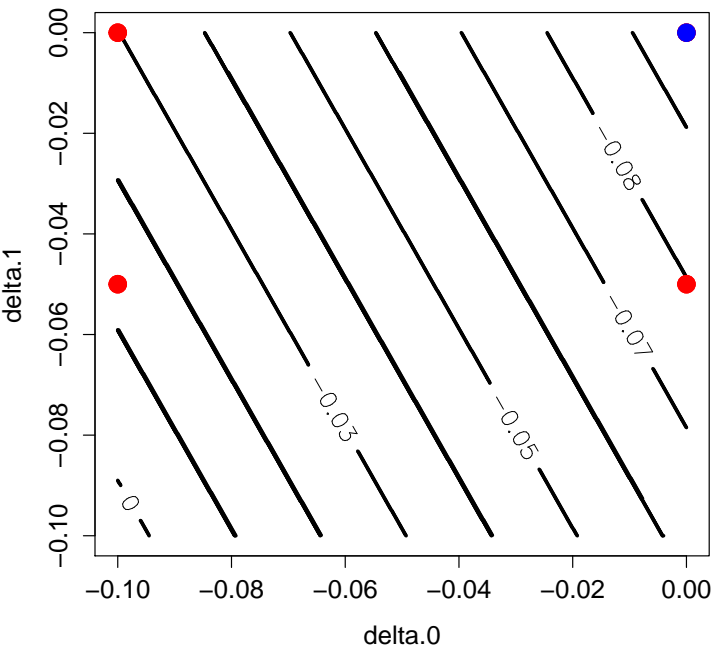
- Suppose those who take treatment ($A = 1$) tend to have lower mortality
- Then want to vary $\delta_1 < 0$ and $\delta_0 < 0$

Example sensitivity analysis (partial)

δ_0	δ_1	Interpretation	Tx effect
0	0	ITA	-0.10 (-0.14, -0.05)
-0.10	0	mortality 10% lower among compliers, in absence of treatment	-0.03 (-0.08, .02)
0	-0.05	mortality 5% lower among compliers, in presence of treatment	-0.08 (-0.13, -0.03)
-0.10	-0.05		-0.01 (-0.06, .04)

Summary of sensitivity analysis

- Possible unmeasured confounder: compliers engage in other healthy behaviors
- This unmeasured confounder may explain observed treatment effect
- Tipping point that changes point estimate to zero: see graph



Next example: Continuous outcome with observed confounders

Example 2: HER Study

- Epidemiologic study of HIV in women, 1993-99
- Want to examine effect of antiviral therapy initiation on CD4 count six months later
- Potential confounders: observed covariates at time of treatment decision

		Antiviral Therapy	
		A = 0	A = 1
		(n = 246)	(n = 111)
Outcome	CD4 at 6 months	278	289
Covariates	Baseline log VL	3.4	3.3
	Baseline symptoms	.59	.68
	Baseline CD4	271	250

	A	X1	X2	X3	Y	Y0	Y1
[92,]	0	0	181	3.13	145	145	.
[93,]	0	2	385	3.14	236	236	.
[94,]	0	0	396	3.15	610	610	.
[119,]	0	0	247	3.49	252	252	.
[120,]	0	0	455	3.53	495	495	.
[121,]	0	0	268	3.54	328	328	.
[122,]	0	1	93	3.55	63	63	.
[162,]	0	0	40	3.91	40	40	.
[163,]	0	0	191	3.92	209	209	.
[164,]	0	0	337	3.93	173	173	.
[165,]	0	1	6	3.94	6	6	.
[7,]	1	2	176	1.70	236	.	236
[8,]	1	0	484	1.70	504	.	504
[34,]	1	1	156	2.78	162	.	162
[35,]	1	2	130	2.84	44	.	44
[105,]	1	3	67	4.87	70	.	70
[106,]	1	0	174	4.96	288	.	288
[107,]	1	0	117	5.13	212	.	212

Causal inference via G-computation algorithm

- 1 Fit regression $[Y_1 | X, A = 1]$
- 2 Fit regression $[Y_0 | X, A = 0]$
- 3 Use these to generate prediction of Y_1, Y_0 for whole sample
- 4 Estimated ATE is difference of averages

$$\widehat{ATE} = (1/n) \sum_{i=1}^n \hat{Y}_{1i} - \hat{Y}_{0i}$$

	A	Y0.hat	Y1.hat	diff
[92,]	0	201.78	213.47	11.69
[93,]	0	383.58	432.41	48.83
[94,]	0	408.14	431.82	23.68
[119,]	0	257.47	282.24	24.77
[120,]	0	458.12	494.80	36.68
[121,]	0	276.66	303.86	27.20
[122,]	0	100.62	131.78	31.16
[162,]	0	49.07	73.93	24.86
[163,]	0	195.22	228.35	33.13
[164,]	0	335.97	377.06	41.09
[165,]	0	8.72	45.44	36.72
[7,]	1	210.20	209.91	-0.29
[8,]	1	522.32	512.37	-9.95
[34,]	1	176.73	190.85	14.12
[35,]	1	143.26	170.58	27.32
[105,]	1	34.21	124.55	90.34
[106,]	1	157.70	217.14	59.44
[107,]	1	99.40	160.50	61.10

Causal inference via G-computation algorithm

Inference about ATE

	Est.	s.e.
Unadjusted	9.7	18.7
Adjusted (GCA)	30.4	11.4

Representing unmeasured confounding

With no unmeasured confounding, potential outcome means equal across treatment groups

$$\begin{aligned}E(Y_1 | A = 1, \mathbf{x}) - E(Y_1 | A = 0, \mathbf{x}) &= 0 \\E(Y_0 | A = 1, \mathbf{x}) - E(Y_0 | A = 0, \mathbf{x}) &= 0\end{aligned}$$

Can represent unmeasured confounding as differences in potential outcome means

$$\begin{aligned}\eta_1 &= E(Y_1 | A = 1, \mathbf{x}) - E(Y_1 | A = 0, \mathbf{x}) \\ \eta_0 &= E(Y_0 | A = 1, \mathbf{x}) - E(Y_0 | A = 0, \mathbf{x})\end{aligned}$$

These can also depend on \mathbf{X}

Representing unmeasured confounding

Also relates to treatment effect

$$\eta_1 - \eta_0 = E(Y_1 - Y_0 | A = 1, \mathbf{x}) - E(Y_1 - Y_0 | A = 0, \mathbf{x})$$

Examples:

- Confounding by indication: those receiving treatment are less healthy

$$\eta_0 < 0 \quad \text{and} \quad \eta_1 < 0$$

- Treatment prescribed preferentially to those who will benefit more

$$\eta_1 > \eta_0$$

Implementation

Can show that this amounts to adjusting imputed values as follows

- For those with $A = 0$

$$\widehat{Y}_{1i}(\eta_1) = \widehat{Y}_{1i} - \eta_1$$

- For those with $A = 1$

$$\widehat{Y}_{0i}(\eta_0) = \widehat{Y}_{0i} + \eta_0$$

When the sensitivity parameters do not depend on \mathbf{x} ,

$$\begin{aligned} \text{ATE}(\eta_0, \eta_1) &= \text{ATE}(0, 0) - \{\eta_1 P(A = 0) + \eta_0 P(A = 1)\} \\ &= \text{ATE}(0, 0) - \text{unmeasured confounding bias} \end{aligned}$$

How to select values for η_0, η_1

- Recall that η 's are differences in conditional means

$$\eta_1 = E(Y_1 | A = 1, \mathbf{x}) - E(Y_1 | A = 0, \mathbf{x})$$

$$\eta_0 = E(Y_0 | A = 1, \mathbf{x}) - E(Y_0 | A = 0, \mathbf{x})$$

- Simple measurement scale: residual SD from observed-data regressions

$$\eta_1 = \lambda_1 \sigma_1$$

$$\eta_0 = \lambda_0 \sigma_0$$

where the σ 's are residual SD

- Our approach: Use single value of λ

$$\eta_1 = \lambda \sigma_1$$

$$\eta_0 = \lambda \sigma_0$$

Illustration using HERS Data

Objectives

- Show how methods implemented
- Compare representations of
 - ▶ Robustness of findings
 - ▶ Changes in degree of sensitivity when new variable are added
- Interpret results in context

Illustration using HERS Data

Analysis 1: Adjust for these confounders

- baseline log viral load
- baseline HIV symptom level (1 to 10)

Analysis 2: Adjust for the same confounders, plus

- baseline CD4 count

Analysis via potential outcomes method

Table of residual SD

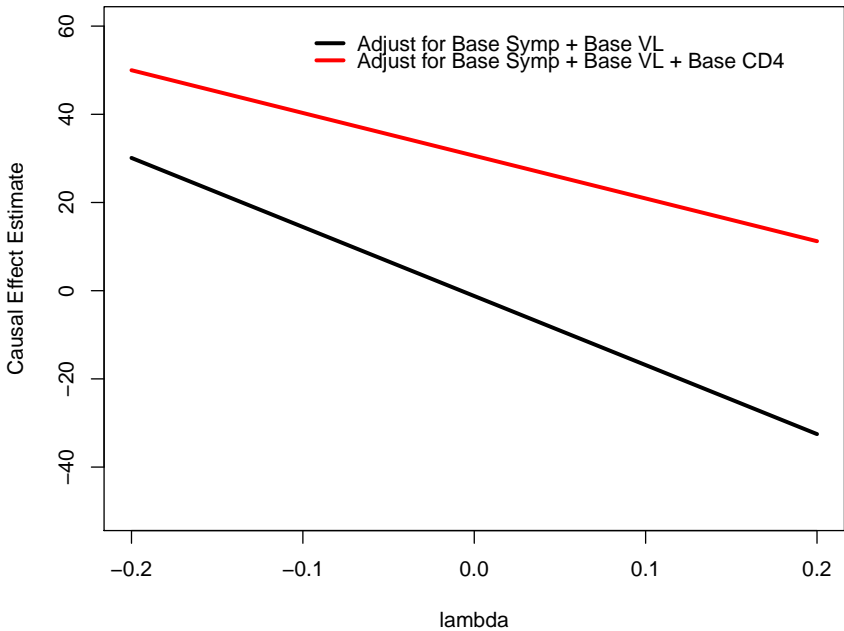
Confounders adjusted for	σ_0	σ_1
Base Sympt, Base VL	152	161
Base Sympt, Base VL, Base CD4	90	102

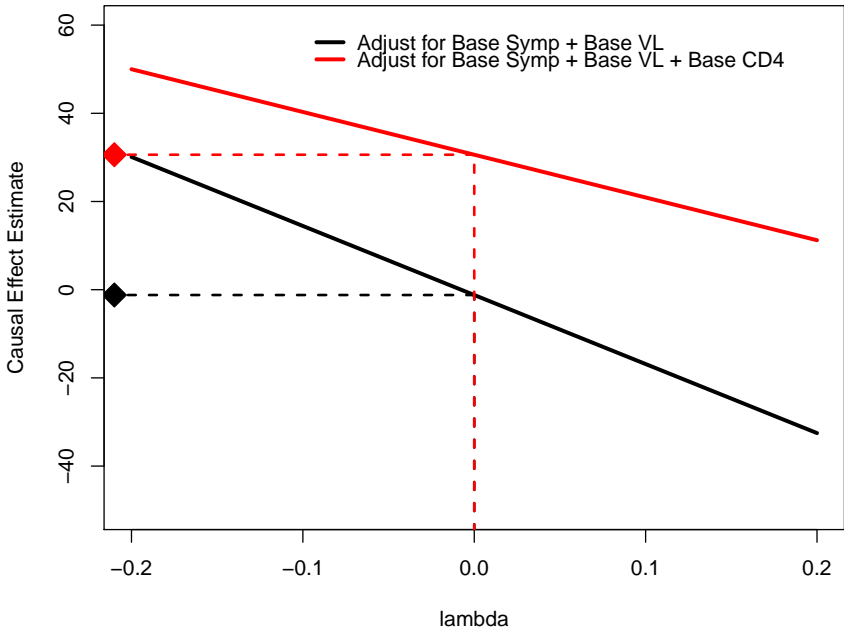
Example: If $\lambda = -.1$, then

$$\eta_0 = E(Y_0 | A = 1, \mathbf{x}) - E(Y_0 | A = 0, \mathbf{x}) = (-.1)(90) = -9$$

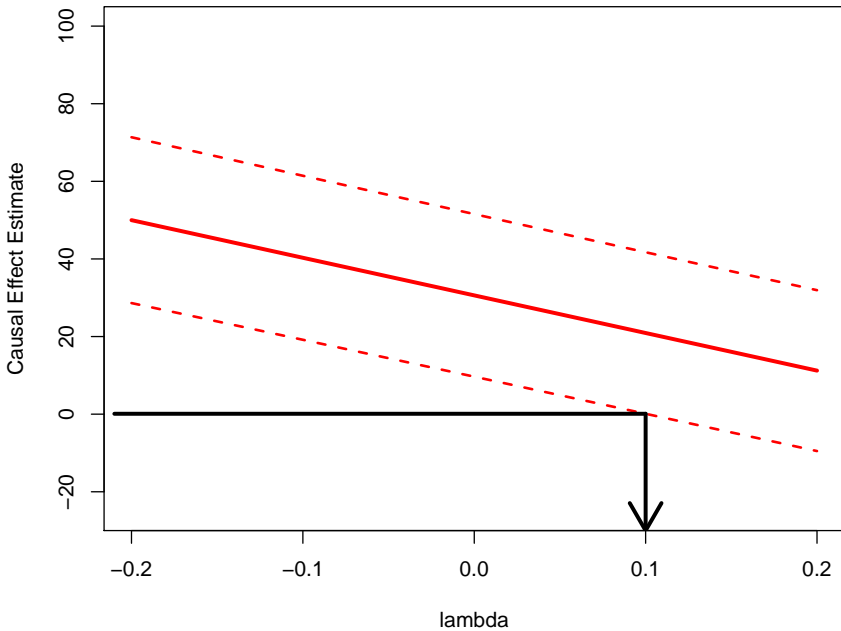
Implies 'confounding by indication'

- If left untreated, CD4 would be lower for those who actually received treatment
- This difference applies within groups having same values of measured confounders





Tipping Point



Potential outcomes analysis: Robustness

Estimated ATE under no unmeasured confounding

$$\widehat{\text{ATE}}(0, 0) = 30.4 \text{ (8.2, 52.8)}$$

Confidence interval will include 0 when $\lambda \geq 0.1$:

$$\eta_1 = E(Y_1 | A = 1, \mathbf{X}) - E(Y_1 | A = 0, \mathbf{X}) = (.1)(102) = 10.2$$

$$\eta_0 = E(Y_0 | A = 1, \mathbf{X}) - E(Y_0 | A = 0, \mathbf{X}) = (.1)(90) = 9.0$$

i.e., when unmeasured confounding implies those selected to receive treatment would, on average, have better outcomes than those not selected, within groups having the same \mathbf{X} values.

Summary and conclusions

Compared two methods for assessing effect of unmeasured confounding

- Bounds
 - ▶ Convey lack of information
 - ▶ No assumptions about unmeasured confounding
 - ▶ Gives ranges that are usually too large to be helpful
 - ▶ Cannot use with continuous outcomes

Summary and conclusions

- Sensitivity analysis based on differences in potential outcomes
 - ▶ Unmeasured confounding = differences in potential outcome means
 - ▶ Allows use of covariates
 - ▶ We illustrated with GCA, but can use with other methods
 - ▶ Allows transparent assessment of robustness

How to do inference using G-computation algorithm

Step 1: Fit a model for $E(Y_1 | X_1, X_2, X_3)$

- Can do this with regression of Y on \mathbf{X} among $A = 1$

$$E(Y | X_1, X_2, X_3, A = 1) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

Call:

```
glm(formula = Y ~ V, subset = (A == 1))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.73174	44.00156	0.198	0.843
V1	5.51940	9.18138	0.601	0.549
V2	1.01917	0.07724	13.195	<2e-16 ***
V3	6.33678	9.75998	0.649	0.518

```
> sigma.1
```

```
[1] 102.1384
```

How to do inference using G-computation algorithm

Step 2: Use this model to generate predicted values of Y_1 , including for those with $A = 0$

$$\hat{Y}_{1i} = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i}$$

	A	X1	X2	X3	Y	Y0	Y1	Y0.hat	Y1.hat
[92,]	0	0	181	3.13	145	145	.		213.47
[93,]	0	2	385	3.14	236	236	.		432.41
[94,]	0	0	396	3.15	610	610	.		431.82
[119,]	0	0	247	3.49	252	252	.		282.24
[120,]	0	0	455	3.53	495	495	.		494.80
[121,]	0	0	268	3.54	328	328	.		303.86
[122,]	0	1	93	3.55	63	63	.		131.78
[162,]	0	0	40	3.91	40	40	.		73.93
[163,]	0	0	191	3.92	209	209	.		228.35
[164,]	0	0	337	3.93	173	173	.		377.06
[165,]	0	1	6	3.94	6	6	.		45.44

A X1 X2 X3 Y Y0 Y1 Y1.hat

[92,]	0	0	181	3.13	145	145	.	213.47
[93,]	0	2	385	3.14	236	236	.	432.41
[94,]	0	0	396	3.15	610	610	.	431.82
[119,]	0	0	247	3.49	252	252	.	282.24
[120,]	0	0	455	3.53	495	495	.	494.80
[121,]	0	0	268	3.54	328	328	.	303.86
[122,]	0	1	93	3.55	63	63	.	131.78
[162,]	0	0	40	3.91	40	40	.	73.93
[163,]	0	0	191	3.92	209	209	.	228.35
[164,]	0	0	337	3.93	173	173	.	377.06
[165,]	0	1	6	3.94	6	6	.	45.44
[7,]	1	2	176	1.70	236	.	236	209.91
[8,]	1	0	484	1.70	504	.	504	512.37
[34,]	1	1	156	2.78	162	.	162	190.85
[35,]	1	2	130	2.84	44	.	44	170.58
[105,]	1	3	67	4.87	70	.	70	124.55
[106,]	1	0	174	4.96	288	.	288	217.14
[107,]	1	0	117	5.13	212	.	212	160.50

How to do inference using G-computation algorithm

Step 3: Repeat this process for Y_0

Call:

```
glm(formula = Y ~ V, subset = (A == 0))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	89.0177	27.4414	3.244	0.00135	**
V1	-7.4478	6.1926	-1.203	0.23027	
V2	0.9663	0.0456	21.190	< 2e-16	***
V3	-20.0045	6.0580	-3.302	0.00110	**

```
> sigma.0
```

```
[1] 90.25616
```

	A	Y0	Y1	Y0.hat	Y1.hat

[92,]	0	145	.	201.78	213.47
[93,]	0	236	.	383.58	432.41
[94,]	0	610	.	408.14	431.82
[119,]	0	252	.	257.47	282.24
[120,]	0	495	.	458.12	494.80
[121,]	0	328	.	276.66	303.86
[122,]	0	63	.	100.62	131.78
[162,]	0	40	.	49.07	73.93
[163,]	0	209	.	195.22	228.35
[164,]	0	173	.	335.97	377.06
[165,]	0	6	.	8.72	45.44
[7,]	1	.	236	210.20	209.91
[8,]	1	.	504	522.32	512.37
[34,]	1	.	162	176.73	190.85
[35,]	1	.	44	143.26	170.58
[105,]	1	.	70	34.21	124.55
[106,]	1	.	288	157.70	217.14
[107,]	1	.	212	99.40	160.50

A Y0.hat Y1.hat

[92,]	0	201.78	213.47
[93,]	0	383.58	432.41
[94,]	0	408.14	431.82
[119,]	0	257.47	282.24
[120,]	0	458.12	494.80
[121,]	0	276.66	303.86
[122,]	0	100.62	131.78
[162,]	0	49.07	73.93
[163,]	0	195.22	228.35
[164,]	0	335.97	377.06
[165,]	0	8.72	45.44
[7,]	1	210.20	209.91
[8,]	1	522.32	512.37
[34,]	1	176.73	190.85
[35,]	1	143.26	170.58
[105,]	1	34.21	124.55
[106,]	1	157.70	217.14
[107,]	1	99.40	160.50