Statistical Learning in Mediation Analysis

Chapter 4: Controlled and natural direct effects under exposure-induced confounding

David Benkeser
Emory University

Iván Díaz
Weill Cornell Medicine

MODULE 17

Summer Institute in Statistics for Clinical and Epidemiological Research

August 2022

Contents of this chapter

Controlled direct effects

Recall the definition of **controlled direct effects**, which compare:

- $E[Y(1, m^*)]$ = avg. outcome if given $A = 1$ and $M = m^*$ vs.
- $E[Y(0, m^*)]$ = avg. outcome if given $A = 0$ and $M = m^*$.

For example, we could define a **controlled direct effect** as

$$\text{CDE}(m^*) = E[Y(1, m^*)] - E[Y(0, m^*)] \ .$$

Notice that this effect parameter, depends on a **reference level** m^*.

- **Example**: If we enforced that all women were screened for breast cancer at age 50 (say, $m^* = 1$), what is the difference in breast cancer-related mortality comparing high SES ($A = 1$) vs low SES ($A = 0$).

Review: identification of controlled direct effects

We required **two randomization** assumptions to identify CDE(m^*).

The first is the **typical randomization assumption**, for $a = 0, 1$, $Y(a, m^*) \perp A \mid W$.

- No unmeasured confounders of Y and A.
- Enforced by randomizing A (possibly stratified on W).

The second is the often **more dubious** assumption that $M \perp Y(a, m^*) \mid A = a, W$.

- No unmeasured confounders of M and Y.
- Enforced by randomizing M (possibly stratified on W).

Review: identification of controlled direct effects}

Depending on the context, the second randomization condition **may not be enforceable** by design.

- How to randomly assign neutralizing antibody responses to a vaccine?

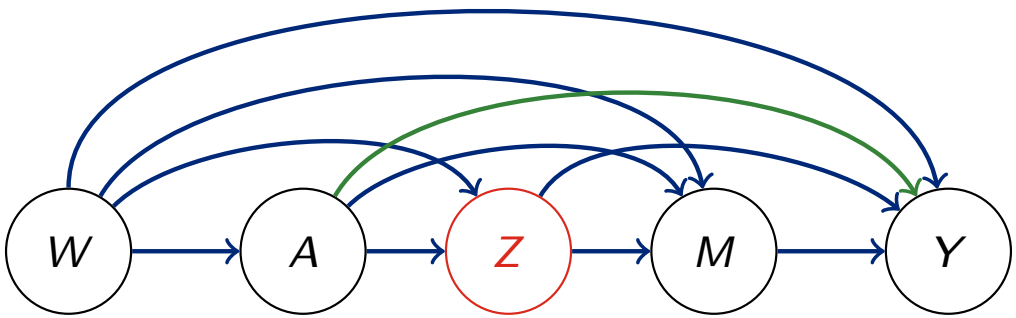
Under the above assumptions, we can identify controlled direct effects using a **G-computation formula**,

$$\text{CDE}(m^*) = E[E(Y \mid A = 1, M = m^*, W) - E(Y \mid A = 0, M = m^*, W)]$$

Time-varying mediator/outcome confounders

So far, our DAG **does not** contain confounders of the mediator-outcome relationship that are influenced by past treatment.

In other words, we have been assuming no such variable Z in the DAG below.



Fundamental question: Can we still estimate CDE(m^*) when such a Z is present?

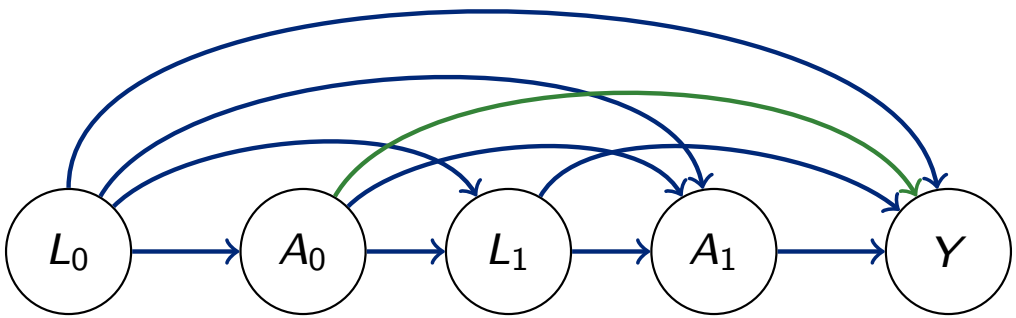
Time-varying mediator/outcome confounders

The previous identification formula **will not work**: $M \not\perp Y(a, m) \mid A = a, W$.

- Z is a confounder of M and Y!

This DAG may look familiar to some of you...

- Define $L_0 = W_1$, $A_0 = A$, $Z = L_1$, $A_1 = M$.



The CDE(m^*) is exactly the longitudinal ATE defined by $E[Y(1, m^*)] - E[Y(0, m^*)]$!

Time-varying mediator/outcome confounders

The implication is that **identification and estimation** of CDE(m^*) is **exactly the same** as for a specific ATE in the three time-point longitudinal context.

All estimation techniques for causal inference from longitudinal studies immediately apply!

Identification of the CDE in the presence of intermediate confounders

Our goal is to contrast the mean outcome under an intervention that assigns $(A, M) = (a, m^*)$ for $a \in \{0, 1\}$.

From this we can get the contrast $E[Y(1, m^*)] - E[Y(0, m^*)]$, which is the CDE($m^*$).

As before, we require **two randomization** assumptions to identify CDE(m^*).

The first is the **typical randomization assumption**, for $a = 0, 1$, $Y(a, m^*) \perp A \mid W$.

- No unmeasured confounders of Y and A.
- Enforced by randomizing A (possibly stratified on W).

The second is the assumption that $M \perp Y(a, m^*) \mid (A = a, W, Z)$.

- No unmeasured confounders of M and Y.
- Enforced by randomizing M (possibly stratified on W and Z).

Identification of the CDE in the presence of intermediate confounders

Our goal is to infer what the mean outcome would be in the target population under the multi time-point intervention of interest.

We must be able to **observe the intervention of interest for each different "type" of individual** (as defined by recorded covariates) from this population:

- $P(A = a \mid W = w) > 0$ for each possible (w, a) and $a \in \{0, 1\}$;
- $P(M = m^* \mid W = w, A = a, Z = z) > 0$ for each possible (w, a, z);

As before, this is referred to as the **positivity condition**.

Identification of the CDE in the presence of intermediate confounders

$$E\left[Y \mid M = m^*, Z, A = a, W\right]$$

$$= E\left[Y(a, m^*) \mid M = m^*, Z, A = a, W\right] = E\left[Y(a, m^*) \mid Z, A = a, W\right]$$

$$=$$ mean counterfactual outcome among patients in treatment group $A = a$, with covariate values Z and W

$$E\left[E\left[Y \mid M = m^*, Z, A = a, W\right] \mid A = a, W\right]$$

$$= E\left[Y(a, m^*) \mid A = a, W\right] = E\left[Y(a, m^*) \mid W\right]$$

$$=$$ mean counterfactual outcome among patients with covariate value W

$$E\left[E\left[E\left[Y \mid M = m^*, Z, A = a, W\right] \mid A = a, W\right]\right]$$

$$= E\left[Y(a, m^*)\right]$$

$$=$$ mean counterfactual outcome

Identification of the CDE in the presence of intermediate confounders

The idea of inverse probability weighting naturally suggests a simple identification formula, as in the case of no intermediate confounders.

Individuals who received the entire regime of interest are not representative of the target population because of (time-varying) confounding.

What about upweighting their contribution to recover representativeness?

The generalized "propensity scores' ' are defined as

$$g_A(a \mid w) = P(A = a \mid W = w) > 0$$
$$g_M(m^* \mid w, a, z) = P(M = m^* \mid W = w, A = a, Z = z)$$

For a patient with variables $(W, Z) = (w, z)$, the composite probability of receiving intervention of interest is simply given by

$$\bar{g}(w, z) := g_A(a \mid w) g_M(m^* \mid w, a, z)$$

Identification of the CDE in the presence of intermediate confounders

If $\bar{g}(w, z)$ is small, a patient with values (w, z) and treatment-mediator profile $(a, m^\star)$ is an unlikely occurrence in the sampling population.

This patient needs to serve as stand-in for the many such patients not seen.

The **IPW identification formula** is given by

$$E\left[Y(a, m^\star)\right] = E\left[\left\{\frac{I(A = a, M = m^*)}{\bar{g}(W, Z)}\right\} Y\right]$$

for treatment-mediator profile (a, m^*) for $a \in \{0, 1\}$.

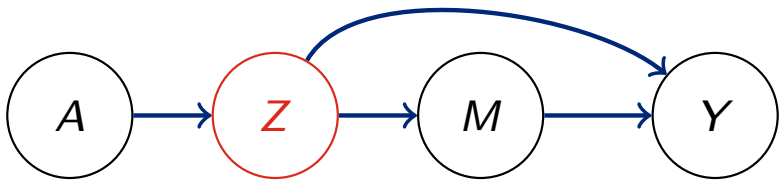
How does the risk of positivity violations compare to the single time-point setting?

Identification of the CDE in the presence of intermediate confounders

The equivalence between the IPW and G-computation identification formulas can be established through repeated uses of the law of total expectation.

$$
\begin{aligned}
E\left[\left\{\frac{I(A=a, M=m^*)}{\bar{g}(W,Z)}\right\} Y\right] &= E\left[E\left[\left\{\frac{I(A=a, M=m^*)}{\bar{g}(W,Z)}\right\} Y \,\middle|\, W, A, Z, M\right]\right] \\
&= E\left[\left\{\frac{I(A=a, M=m^*)}{\bar{g}(W,Z)}\right\} E\left[Y \,\middle|\, W, A=a, Z, M=m^*\right]\right] \\
&= E\left[\left\{\frac{I(A=a)}{g_A(a \mid W)\cancel{g_M(m^* \mid W, a, Z)}}\right\} E\left[Y \,\middle|\, W, A=a, Z, M=m^*\right] \cancel{E\left[I(M=m^*) \mid W, A, Z\right]}\right] \\
&= E\left[\left\{\frac{I(A=a)}{g_A(a \mid W)}\right\} E\left[E\left[Y \,\middle|\, W, A=a, Z, M=m^*\right] | A=a, W\right]\right] \\
&= E\left[E\left[E\left[Y \mid M=m^*, Z, A=a, W\right] \mid A=a, W\right]\right]
\end{aligned}
$$

Failure of naive approach



$$
\begin{aligned}
Y \mid M = m, Z = z, A = a &\sim \text{Normal}(1 + m + 2z, 1) \\
M \mid Z = z, A = a &\sim \text{Bernoulli}(\text{expit}(-1 + z)) \\
Z \mid A = a &\sim \text{Normal}(1 + a, 1) \\
A &\sim \text{Bernoulli}(0.5)
\end{aligned}
$$

Failure of naive approach

$$
\begin{aligned}
Y \mid M = m, Z = z, A = a &\sim \text{Normal}(1 + m + 2z, 1) \\
M \mid Z = z, A = a &\sim \text{Bernoulli}(\text{expit}(-1 + z)) \\
Z \mid A = a &\sim \text{Normal}(1 + a, 1) \\
A &\sim \text{Bernoulli}(0.5)
\end{aligned}
$$

Using the G-computation formula, we can compute mean counterfactual outcomes corresponding to different treatment profiles:

$$
\begin{aligned}
E[Y(1,1)] &= E[E[Y \mid M = 1, Z, A = 1] \mid A = 1] = E[2 + 2Z \mid A = 1] = 6 \\
E[Y(1,0)] &= E[E[Y \mid M = 0, Z, A = 1] \mid A = 1] = E[1 + 2Z \mid A = 1] = 5 \\
E[Y(0,1)] &= E[E[Y \mid M = 1, Z, A = 0] \mid A = 1] = E[2 + 2Z \mid A = 0] = 4 \\
E[Y(0,0)] &= E[E[Y \mid M = 0, Z, A = 0] \mid A = 1] = E[1 + 2Z \mid A = 0] = 3
\end{aligned}
$$

Failure of naive approach

Can controlled direct effects be read off the regression of Y on (M, Z, A)?

Controlled direct effects

$$E[Y(1,1) - Y(0,1)] = 2 \neq 0 = E[Y \mid M = 1, Z, A = 1] - E[Y \mid M = 1, Z, A = 0]$$
$$E[Y(1,0) - Y(0,0)] = 2 \neq 0 = E[Y \mid M = 0, Z, A = 1] - E[Y \mid M = 0, Z, A = 0]$$

Failure of the method of product of coefficients

$$
\begin{aligned}
Y \mid M = m, Z = z, A = a &\sim \text{Normal}(1 + m + 2z, 1) \\
M \mid Z = z, A = a &\sim \text{Normal}(-1 + z, 1) \\
Z \mid A = a &\sim \text{Normal}(1 + a, 1) \\
A &\sim \text{Bernoulli}(0.5)
\end{aligned}
$$

Method of product of coefficients:

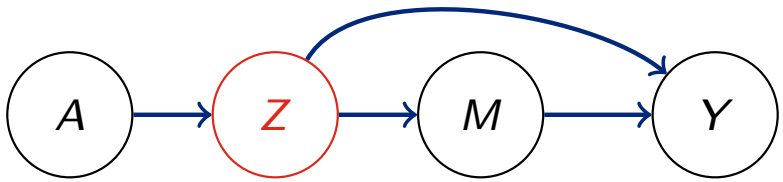
1. Run a linear regression of Y on all variables
2. Run a linear regression of M on all variables
3. The **indirect effect** is the product of the coefficient in step 1 associated to M and the coefficient in step 2 associated to A

Failure:

- The product of coefficients is 0
- Yet, there is an indirect effect $A \to Z \to M \to Y$

Failure of naive approaches to mediation analysis

Why does this happen?



The regression of Y on (A, Z, M) fixes Z: as such, some causal paths between A and Y are blocked.

This is a case of **treatment-confounder feedback**, whose presence often invalidates naive approaches to causal inference (see Chapter 20 of Hernán & Robins, 2020).

Simple estimation procedures

The IPW identification formula suggests a simple estimation strategy.

First, we may construct estimates $g_{A,n}$ and $g_{M,n}$ of propensity scores g_A and g_M using our favorite regression estimator for binary outcomes:

- $g_A(a \mid w)$: probability of $A = a$ given $W = \ell_0$;
- $g_M(m^* \mid w, a, z)$: probability of $M = m^*$ given $(W, Z) = (w, z)$ and $A = a$;

Then, we can compute the corresponding **IPW estimator** of $E[Y(1, 1, \ldots, 1)]$ as

$$\psi_{n,IPW} := \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{I(A_i = a, M_i = m^*)}{g_{A,n}(a \mid W_i) g_{M,n}(m^* \mid W_i, a, Z_i)} \right\} Y_i \ .$$

Illustrating IPW in R

Below we simulate data as described in the previous slides.

$$
\begin{aligned}
Y \mid M = m, Z = z, A = a &\sim \text{Normal}(1 + m + 2z, 1) \\
M \mid Z = z, A = a &\sim \text{Bernoulli}(\text{expit}(-1 + z + a)) \\
Z \mid A = a &\sim \text{Normal}(1 + a, 1) \\
A &\sim \text{Bernoulli}(0.5)
\end{aligned}
$$

```
# set a seed for reproducibility
set.seed(212)
n <- 5000
A <- rbinom(n, size = 1, p = 0.5)
Z <- rnorm(n, mean = A + 1, sd = 1)
M <- rbinom(n, size = 1, p = plogis(-1 + Z + A))
Y <- rnorm(n, mean = 1 + M + 2 * Z, 1)
```

Illustrating IPW in R

We can see that IPW yields approximately correct answers in each case.

```
# correct ps model for A
ps0 <- glm(A ~ 1)
g_0 <- predict(ps0, type = "response")
# correct ps model for M
ps1 <- glm(M ~ Z + A, family = binomial())
# estimates of P(M = 1 | A = 1, Z = Zi)
df_A_1 <- data.frame(A = 1, Z = Z)
g_1_A_1 <- predict(ps1, newdata = df_A_1, type = "response")
# estimates of P(M = 1 | A = 0, Z = Zi)
df_A_0 <- data.frame(A = 0, Z = Z)
g_1_A_0 <- predict(ps1, newdata = df_A_0, type = "response")
```

Illustrating IPW in R

```
# IPW estimate of E[Y(1,1)] ~ 6
mean( (A == 1 & M == 1) / (g_0 * g_1_A_1) * Y )
```

```
## [1] 5.96643
```

```
# IPW estimate of E[Y(1,0)] ~ 5
mean( (A == 1 & M == 0) / (g_0 * (1 - g_1_A_1)) * Y )
```

```
## [1] 5.186773
```

```
# IPW estimate of E[Y(0,1)] ~ 4
mean( (A == 0 & M == 1) / ((1 - g_0) * g_1_A_0) * Y )
```

```
## [1] 4.016724
```

```
# IPW estimate of E[Y(0,0)] ~ 3
mean( (A == 0 & M == 0) / ((1 - g_0) * (1 - g_1_A_0)) * Y )
```

```
## [1] 3.001915
```

Simple estimation procedures

The G-computation formula suggests another natural estimation strategy.

Set $m^* = 1$. The estimand is simply

$$E\left[E\left[E\left(Y \mid M = 1, Z, A = 1, W\right) \mid A = 1, W\right]\right].$$

Y	M	Z	A	W
0	1	0.7	1	2.1
1	1	-0.2	0	-1.6
1	1	2.0	1	0.3
0	0	6.9	1	1.4
1	0	3.1	0	0.9
0	1	-5.2	1	-3.1
1	0	5.2	1	2.5
0	1	-1.1	1	-0.1

Simple estimation procedures

The G-computation formula suggests another natural estimation strategy.

Set $m^* = 1$. The estimand is simply

$$E\left[E\left[E\left(Y \mid M = 1, Z, A = 1, W\right) \mid A = 1, W\right]\right].$$

Y	M	Z	A	W
0	1	0.7	1	2.1
1	1	-0.2	0	-1.6
1	1	2.0	1	0.3
0	0	6.9	1	1.4
1	0	3.1	0	0.9
0	1	-5.2	1	-3.1
1	0	5.2	1	2.5
0	1	-1.1	1	-0.1

STEP 1: Regress Y on Z and W among those with $M = 1$, $A = 1$. $\longrightarrow$ $\bar{Q}_{Y,n}$

Simple estimation procedures

The G-computation formula suggests another natural estimation strategy.

Set $m^* = 1$. The estimand is simply

$$E\left[E\left[E\left(Y \mid M = 1, Z, A = 1, W\right) \mid A = 1, W\right]\right].$$

Y	M	Z	A	W	$\bar{Q}_{Y,n}(Z, W)$
0	1	0.7	1	2.1	0.34
1	1	-0.2	0	-1.6	0.19
1	1	2.0	1	0.3	0.26
0	0	6.9	1	1.4	0.31
1	0	3.1	0	0.9	0.29
0	1	-5.2	1	-3.1	0.15
1	0	5.2	1	2.5	0.36
0	1	-1.1	1	-0.1	0.25

STEP 1: Regress Y on Z and W among those with $M = A = 1$. $\longrightarrow$ $\bar{Q}_{Y,n}$

Compute $\bar{Q}_{Y,n}(Z, W)$ for every patient.

Simple estimation procedures

The G-computation formula suggests another natural estimation strategy.

Set $m^* = 1$. The estimand is simply

$$E\left[E\left[E\left(Y \mid M = 1, Z, A = 1, W\right) \mid A = 1, W\right]\right].$$

Y	M	Z	A	W	$\bar{Q}_{Y,n}(Z, W)$
0	1	0.7	1	2.1	0.34
1	1	-0.2	0	-1.6	0.19
1	1	2.0	1	0.3	0.26
0	0	6.9	1	1.4	0.31
1	0	3.1	0	0.9	0.29
0	1	-5.2	1	-3.1	0.15
1	0	5.2	1	2.5	0.36
0	1	-1.1	1	-0.1	0.25

STEP 2: Regress $\bar{Q}_{Y,n}(Z, W)$ on W among those with $A = 1$. $\longrightarrow$ $\bar{Q}_{Z,n}$

Simple estimation procedures

The G-computation formula suggests another natural estimation strategy.

Set $m^* = 1$. The estimand is simply

$$E\left[E\left[E\left(Y \mid M=1, Z, A=1, W\right) \mid A=1, W\right]\right].$$

Y	M	Z	A	W	$\bar{Q}_{Y,n}(Z,W)$	$\bar{Q}_{Z,n}(W)$
0	1	0.7	1	2.1	0.34	0.40
1	1	-0.2	0	-1.6	0.19	0.22
1	1	2.0	1	0.3	0.26	0.20
0	0	6.9	1	1.4	0.31	0.36
1	0	3.1	0	0.9	0.29	0.21
0	1	-5.2	1	-3.1	0.15	0.34
1	0	5.2	1	2.5	0.36	0.20
0	1	-1.1	1	-0.1	0.25	0.29

STEP 2: Regress $\bar{Q}_{Y,n}(Z,W)$ on W among those with $A=1$. $\longrightarrow$ $\bar{Q}_{Z,n}$

Compute $\bar{Q}_{Z,n}(W)$ for every patient.

Simple estimation procedures

The G-computation formula suggests another natural estimation strategy.

Set $m^* = 1$. The estimand is simply

$$E\left[E\left[E\left(Y \mid M = 1, Z, A = 1, W\right) \mid A = 1, W\right]\right].$$

Y	M	Z	A	W	$\bar{Q}_{Y,n}(Z,W)$	$\bar{Q}_{Z,n}(W)$
0	1	0.7	1	2.1	0.34	0.40
1	1	-0.2	0	-1.6	0.19	0.22
1	1	2.0	1	0.3	0.26	0.20
0	0	6.9	1	1.4	0.31	0.36
1	0	3.1	0	0.9	0.29	0.21
0	1	-5.2	1	-3.1	0.15	0.34
1	0	5.2	1	2.5	0.36	0.20
0	1	-1.1	1	-0.1	0.25	0.29

STEP 3: Average out values of $\bar{Q}_{Z,n}(W)$ over all patients to get $\hat{E}[Y(1,1)] = \bar{Q}_{W,n}$.

Illustration of G-computation

We will now demonstrate that the G-computation formula gives correct answers for our simulated data.

Goal: compute $E[E[Y \mid M = m^*, Z, A = a] \mid A = a]$ for different values of (a, m^*).

A helpful way to think about regression quantities

$$E[\underbrace{Z}_{\text{outcome}} \mid \underbrace{S = s}_{\text{stratification}}, \underbrace{C}_{\text{covariates}}]$$

Considering the inner expectation, we have

$$E[\underbrace{Y}_{\text{outcome}} \mid \underbrace{M = m^*, A = a}_{\text{stratification}}, \underbrace{Z}_{\text{covariate}}] .$$

Illustration of G-computation

For example, if $a = 1, m^* = 1$,

```
# full data.frame
full_data <- data.frame(A = A, Z = Z, M = M, Y = Y)
# subset data to observations with A = 1 & M = 1
data_11 <- subset(full_data, A == 1 & M == 1)
# fit regression of Y ~ Z
fit_11 <- glm(Y ~ Z, data = data_11)
fit_11
```

```
##
## Call:  glm(formula = Y ~ Z, data = data_11)
##
## Coefficients:
## (Intercept)            Z
##       1.899        2.031
##
## Degrees of Freedom: 2140 Total (i.e. Null);  2139 Residual
## Null Deviance:       9740
## Residual Deviance: 1959  AIC: 5892
```

Illustration of G-computation

The fitted regression gives us the estimate

$$\widehat{E}[Y \mid M = 1, Z, A = 1] = 1.9 + 2.03Z \ .$$

Now, we need to estimate the outer expectation,

$$E[\underbrace{E[\ Y \mid M = 1, Z, A = 1]}_{\text{outcome}} \mid \underbrace{A = 1}_{\text{stratification}} \ ]$$

I.e., regression with outcome $1.9 + 2.03Z$ in observations with $A = 1$ and no covariates.

Illustration of G-computation

```
# get predicted value for everyone
full_data$Q2n_11 <- predict(fit_11, newdata = full_data)
# subset data to observations with A = 1
data_1 <- subset(full_data, A == 1)
# fit regression
fit_1 <- glm(Q2n_11 ~ 1, data = data_1)
# intercept is estimate of E[Y(1,1)]
(gcomp_1 <- fit_1$coefficients[1])
```

```
## (Intercept)
##     5.96347
```

Improved estimation procedures

Much like in the single time-point case, a hybrid between the G-computation and IPW estimators can be constructed, and enjoys improved properties.

The **augmented IPW (AIPW) estimator** is given by

$$\hat{E}[Y(1,1)] := \bar{Q}_{W,n} + \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{I(A_i = 1, M_i = 1)}{\bar{g}_n(W_i, Z_i)}\right\}\left[Y_i - \bar{Q}_{Y,n}(Z_i, W_i)\right]$$
$$+ \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{I(A_i = 1)}{g_{A,n}(1 \mid W_i)}\right\}\left[\bar{Q}_{Y,n}(Z_i, W_i) - \bar{Q}_{Z,n}(W_i)\right]$$

where $\bar{Q}_{W,n}$ is simply the G-computation estimator.

Since it builds upon estimates of all outcome regressions and propensity scores, the construction of this estimator requires more effort than for estimators seen so far.

However, this estimator enjoys double-robustness, and can be used to construct valid confidence intervals, even when flexible learning strategies (e.g., Super Learner) are used to estimate the outcome regressions and propensity scores.

AIPW in R

```
gcomp_1 +
  mean(
    (A == 1 & M == 1) / (g_0 * g_1_A_1) * (Y - full_data$Q2n_11) +
    (A == 1) / g_0 * (full_data$Q2n_11 - gcomp_1)
  )
```

```
## (Intercept)
##    5.964714
```

Improved estimation procedures

What does **double robustness** refer to in the context of multi time-point interventions?

The estimator $\hat{E}[Y(1,1)]$ built upon estimators

$$\bar{Q}_n := (\bar{Q}_{Z,n}, \bar{Q}_{Y,n}) \quad \text{and} \quad g_n := (g_{A,n}, g_{M,n})$$

is doubly-robust, in the sense that it is consistent (i.e., hits the target) provided either $\bar{Q}_n$ hits the target $\bar{Q}_0$ or g_n hits the target g_0.

	Scen. 1	Scen. 2	Scen. 3
$\bar{Q}_{Y,n}$		✓	✓
$\bar{Q}_{Z,n}$		✓	
g_M	✓		
g_A	✓		✓

Improved estimation procedures

The targeted maximum likelihood estimation (TMLE) framework provides a recipe for constructing a G-computation estimator that is also doubly-robust. (Bang & Robins 2005; van der Laan & Gruber, 2012)

This estimator is constructed like the 'sequential regression' form of the G-computation estimator, but includes a refinement step after each $\bar{Q}$ estimator is obtained.

Unlike the AIPW, TMLE has the advantage that the estimate is always in the parameter space.

Improved estimation procedures

Properties of estimation procedures outlined

		$\bar{Q} + \bar{g}$		$\bar{Q} + \bar{g}$		$\bar{Q} + \bar{g}$	
	difficulty	**target**	**ci**	**target**	**ci**	**target**	**ci**
IPW	+			✓		✓	
G-COMP	++	✓				✓	
AIPW	+++	✓		✓		✓	✓
TMLE	++++	✓	✓	✓	✓	✓	✓

$\bar{Q} + g$: outcome regressions estimated well but not propensity scores

$\bar{Q} + g$: propensity scores estimated well but not outcome regressions

$\bar{Q} + g$: outcome regressions and propensity scores estimated well

target : does the estimator hit the right target?

ci : is valid inference possible and readily available, even when flexible learning strategies (such as Super Leaner) are used?

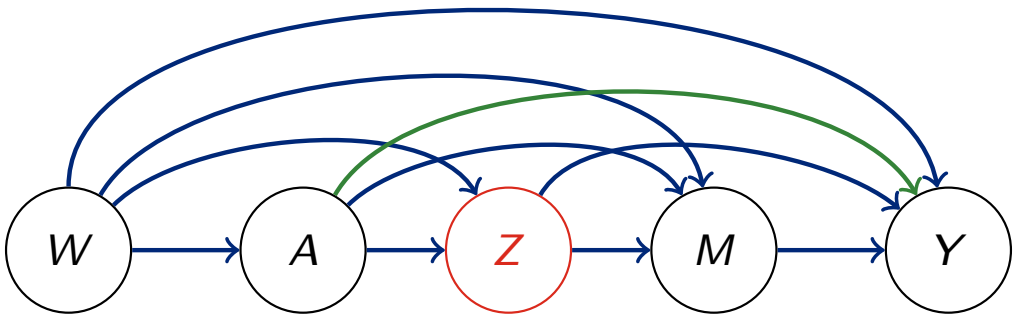
Improved estimation procedures

There is substantial work underway to produce novel estimators with even better properties. All of these innovations are based on the idea of TMLE.

- **Enhanced robustness**
  - Consistent estimation is possible under a wider range of scenarios than depicted on previous slides.
  - For this, more complicated procedures are needed and are being developed. (Luedtke et al. 2017; Rotnitzky et al., 2017)
- **Robust inference**
  - Typical double-robustness only refers to consistency.
  - However, constructing doubly-robust CI and tests is a much more important task. It is also very difficult in multi time-point settings. (Benkeser et al., 2017)
- **Targeted estimation of propensity scores**
  - With many confounders, good performance may be difficult to achieve in smaller samples.
  - Collaborative TMLE allows a smarter, data-driven selection of propensity score estimators to improve performance in such cases.

Non-identifiability of the natural direct and indirect effect

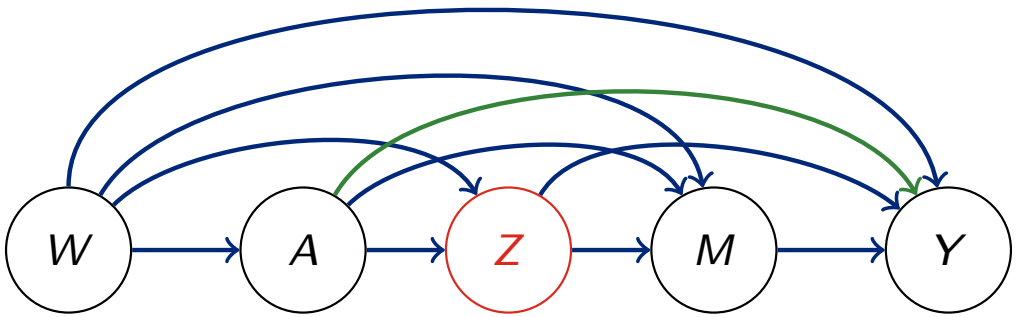
Recall the directed acyclic graph:



And recall the definition of the natural direct and indirect effects

$$ATE = \underbrace{E[Y(1, M(1))] - E[Y(1, M(0))]}_{\text{natural indirect effect}} + \underbrace{E[Y(1, M(0))] - E[Y(0, M(0))]}_{\text{natural direct effect}} .$$

Non-identifiability of the natural direct and indirect effect



And let us also recall one of the fundamental assumptions required for identification of the natural direct and indirect effect:

$$Y(1, m) \perp M(0) \mid W \text{ for all } m$$

Fundamental problem: **This assumption does not hold in the presence of a variable Z as in the above DAG**

Non-identifiability of the natural direct and indirect effect

$$Y(1, m) \perp M(0) \mid W \text{ for all } m$$

Fundamental problem: **This assumption does not hold in the presence of a variable Z as in the above DAG**

To see why, consider an intervention that sets $(A = 1, M = m)$. This intervention generates several counterfactuals:

- $Z(1)$ The mediator-outcome confounder that would have been observed under $A = 1$
- $Y(1, m)$ The outcome that would have been observed under $(A = 1, M = m)$

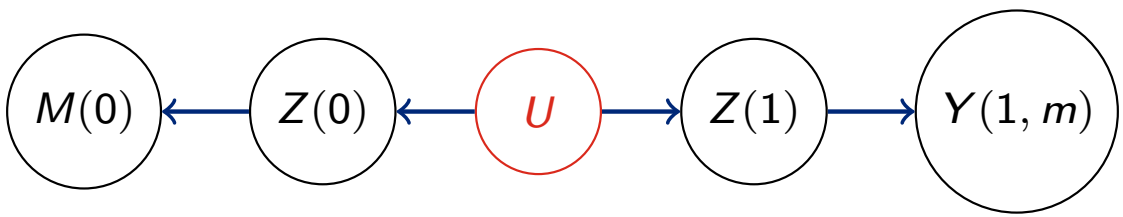
Consider also an intervention that sets $A = 0$. This generates a counterfactual

- $Z(0)$ The mediator-outcome confounder that would have been observed under $A = 0$
- $M(0)$ The mediator that would have been observed under $A = 0$

Non-identifiability of the natural direct and indirect effect

Key observation: $Z(1)$ **and** $Z(0)$ **are associated through unmeasured mechanisms** U **(e.g., genes).**

This implies $Y(1, m)$ and $M(0)$ are associated (after adjustment for W) through the following path:



This path cannot be blocked by any measured variable.

Non-identifiability of the natural direct and indirect effect

The assumption $Y(1, m) \perp M(0) \mid W$ for all m is referred to as a **cross-world counterfactual assumption.**

This assumption cannot be confirmed or disproved to hold, even in studies that randomize both the treatment and mediator

- The presence of intermediate confounders implies the cross-world assumption is violated
- In other words, the cross-world assumption **is stronger** than the assumption of no intermediate confounders
- In chapter 6 we will study natural mediation effects that do not require cross-world assumptions but do require no intermediate confounders (stochastic interventions)

References and additional reading

References:

Luedtke, Alexander R., et al. "Sequential double robustness in right-censored longitudinal models." arXiv preprint arXiv:1705.02459 (2017).

Rotnitzky, Andrea, James Robins, and Lucia Babino. "On the multiply robust estimation of the mean of the g-functional." arXiv preprint arXiv:1705.08582 (2017).

Bang, Heejung, and James M. Robins. "Doubly robust estimation in missing data and causal inference models." Biometrics 61.4 (2005): 962-973.

van der Laan, Mark J., and Susan Gruber. "Targeted minimum loss based estimation of causal effects of multiple time point interventions." The international journal of biostatistics 8.1 (2012).

Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC