Targeted register analyses

PhD short course

Case study: State-Level Masking Mandates and COVID-19 Outcomes in the United States

A Demonstration of the Causal Roadmap

Andrew Mertens

University of California, Berkeley, Division of Biostatistics

EPIDEMIOLOGY

Articles

Articles & Issues Collections Multimedia For Authors Journal Info

INFECTIOUS DISEASES

State-Level Masking Mandates and COVID-19 Outcomes in the United States
A Demonstration of the Causal Roadmap

Wong, Angus K.; Balzer, Laura B.

Author Information

Epidemiology: March 2022 - Volume 33 - Issue 2 - p 228-236
doi: 10.1097/EDE.0000000000001453

Motivation

- United States had a wide variability in state-level recommendations for mask wearing to reduce COVID-19 transmission
- Prior research:
  - The majority of studies aiming to evaluate the impact relied on epidemic modeling rather than causal inference techniques
  - A survey found strong correlation between mask wearing and COVID-19 case rates
    - No control for confounding
  - An event study design (similar to difference-in-differences) found that mandates for public masking were strongly associated with reductions in county-level COVID-19 growth rates, but mandates for employee-only masking were not (Lyu and Wehbly 2020).
  - In contrast, using an econometric approach based on linear structural equations, a study found that early implementation of a national mandate for employee-only masking could have saved upwards of 47,000 US lives by the end of May 2020 (Chernozhukov et al. 2021).
- This paper applied the Causal Roadmap with the goal of evaluating the effect of delays in state-level public masking mandates on the relative growth of COVID-19 cases and deaths.

Roadmap- Overview

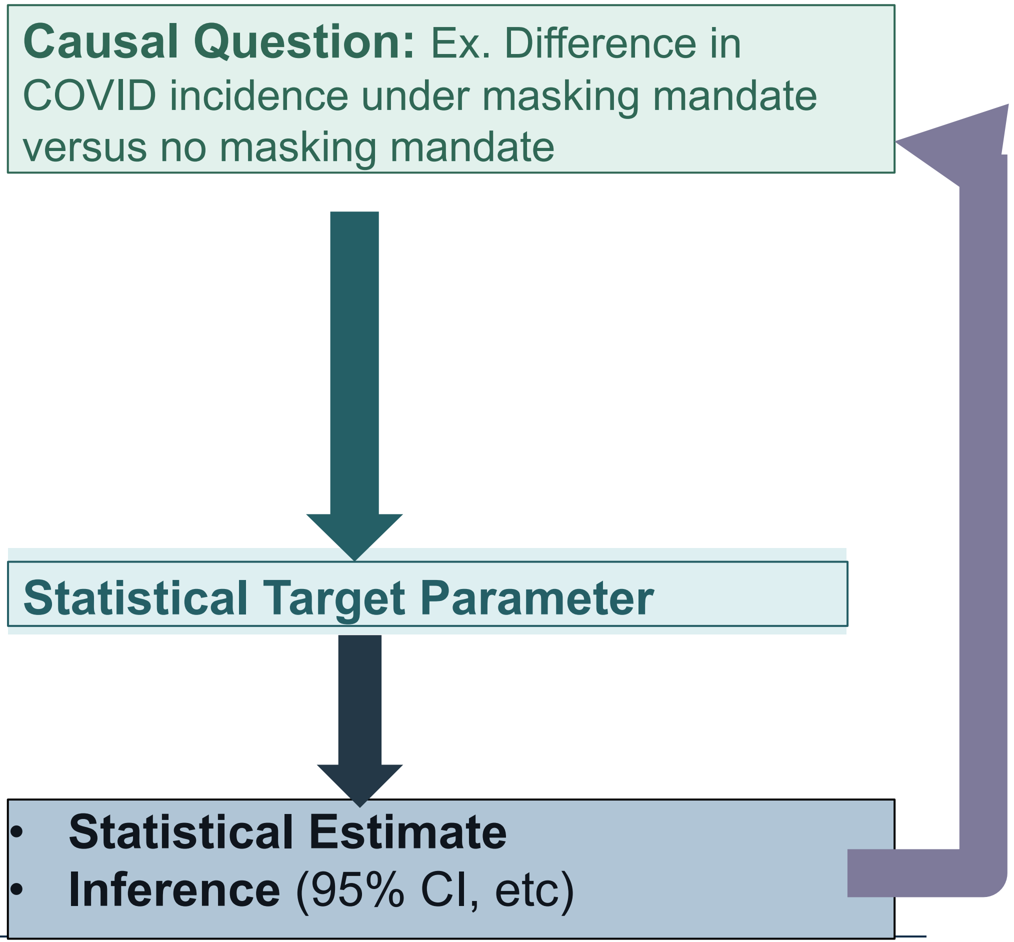
1. **Causal question**
   - Translate scientific question into causal parameter (defined in terms of counterfactual outcomes)

2. **Observed data & statistical model**
   - Model should reflect uncertainty

3. **Identify**
   - Translate causal parameter to statistical parameter under explicit causal assumptions

4. **Estimate**

5. **Interpret**



Causal Roadmap Steps

1. The Causal Question
   - The Research Question
   - The Causal Model
   - The Causal Parameter
2. Observed Data and Statistical Model
3. Identification
4. Statistical Estimation and Inference
5. Interpretation of Results

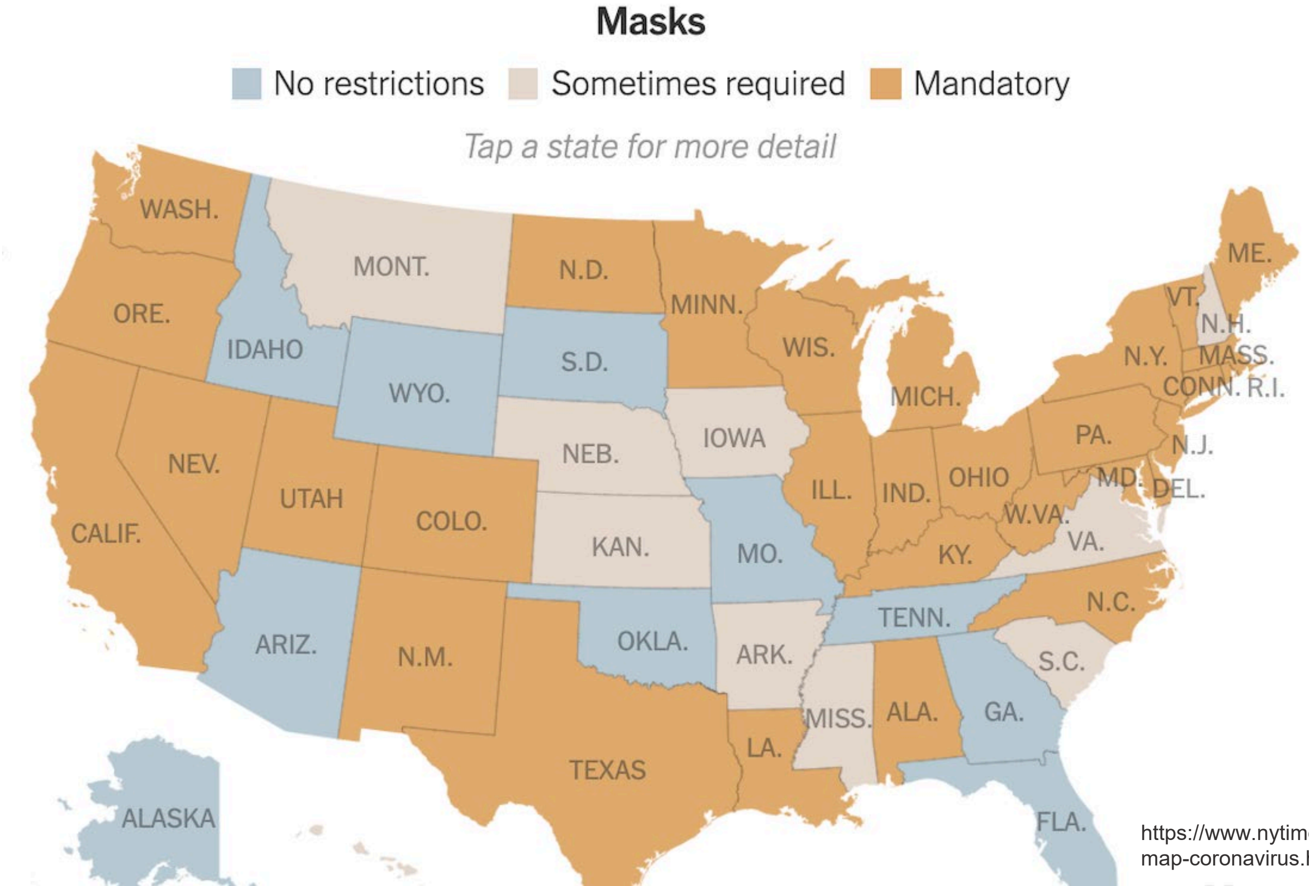
Causal Roadmap Steps

1. The Causal Question
   - The Research Question
   - The Causal Model
   - The Causal Parameter
2. Observed Data and Statistical Model
3. Identification
4. Statistical Estimation and Inference
5. Interpretation of Results

The Research Question

"We sought to investigate the effect of public masking mandates in US states on COVID-19 at the national level in Fall 2020. Specifically, we aimed to evaluate how the relative growth of COVID-19 cases and deaths would have differed if all states had issued a mandate to mask in public by 1 September 2020 versus if all states had delayed issuing such a mandate."

The Research Question

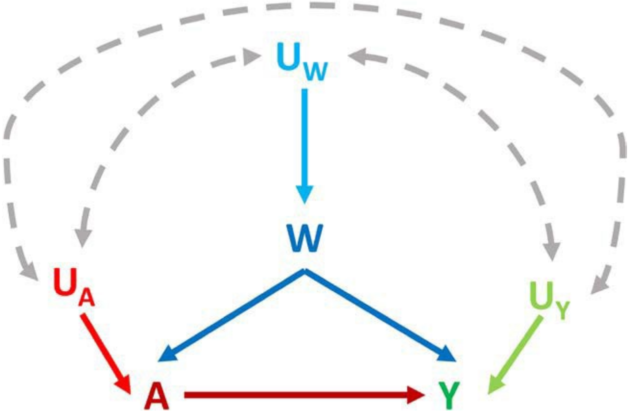


https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html?smtyp=cur&smid=tw-nytimes

The Causal Model

Structural causal model:

$$W = fW\,(UW\,)$$
$$A = f\,A\,(W, UA)$$
$$Y = fY\,(W,\,A, UY\,)$$



- Did not specify the functional form of the structural equations.
  - No parametric assumptions
- Did not specify any independence assumptions between the unmeasured factors
  - AKA, unmeasured common causes of W, A, Y

The Causal Parameter

Causal rate ratio

$$CRR = \frac{E[Y_1]}{E[Y_0]}$$

The ratio in the expected relative growth in COVID-19 cases (deaths) if all 50 states had early versus delayed implementation of the masking policy.

Causal Roadmap Steps

1. The Causal Question
   - The Research Question
   - The Causal Model
   - The Causal Parameter
2. Observed Data and Statistical Model
3. Identification
4. Statistical Estimation and Inference
5. Interpretation of Results

The Observed Data

- The observed data O were at the state level and consisted of the measured confounders W, the masking mandate indicator A, and the relative growth outcome Y.

- We assumed the observed data were generated according to a process compatible with the causal model, which did not place any restrictions on the set of possible distributions of the observed data.
  - Thereby, the statistical model was nonparametric.

- Assumption of no interference between states (i.e., the outcome for one state was not impacted by another's policy).

The Observed Data

- **A** = indicator that a state issued a mandate requiring masks or cloth face coverings in public indoor and outdoor spaces, when it was not possible to maintain at least 6 feet distance, by the target date of 1 Sept. 2020
  - roughly with the start of the school year for K-12 schools and higher education.
  - Hypothesis that masking mandate in place before this target date would help limit subsequent spread of COVID-19 given the anticipated shift in behavior from summertime to classroom-based activities.
- **Y**= state-specific COVID-19 relative growth, defined as the cumulative incidence ratio between confirmed cases (or deaths) a set number of days after the target date divided by confirmed cases (deaths) on the target date.

The Observed Data

- **W:** covariate set included:
  - population demographics (e.g., distributions of age and ethnicity)
  - socio-economic measures
  - prevalence of co-morbidities
  - measures of population density
  - commuting patterns
  - political leaning
  - prior public health policies (e.g., gathering restrictions, stay-at-home orders)
  - COVID-19 outcomes per-capita (e.g., COVID-19 tests, cases, and deaths)
  - changes in Google's residential mobility indicators 7- and 14-days before the target date.

The Unobserved Data

Unmeasured confounders

- **Key unmeasured confounder:** the state's COVID-19 the epidemic trajectory near the target date (e.g., upslope, apex, downslope, low plateau)
- Perceived or actual compliance with previous public health policies
- Strength of the state's public health department.

Causal Roadmap Steps

1. The Causal Question
   - The Research Question
   - The Causal Model
   - The Causal Parameter
2. Observed Data and Statistical Model
3. Identification
4. Statistical Estimation and Inference
5. Interpretation of Results

Step 3: Identification

Assumptions needed to express the causal parameter as some function of the observed data distribution

- Temporality
  - Probably mostly true: may be errors in reporting case or death dates
- No interference
  - Not true: Dependence between states
- Consistency
  - True: An individual's potential outcome under her observed exposure history is precisely her observed outcome
  - Aka the exposure is defined with enough specificity that different variants of the exposure do not have different effects on the outcome
- No unmeasured common causes of the exposure A and the outcome Y
  - Not true!
  - **The causal effect of interest was not identifiable despite the natural experiment occurring between states.**

Step 3: Identification

Statistical estimand: adjusted cumulative incidence ratio

$$aCIR = \frac{E[E(Y|A = 1, W)]}{E[E(Y|A = 0, W)]}$$

Step 3: Identification

Data support conditions

- Positivity assumption
  - AKA data overlap or “experimental treatment assignment assumption”
- Theoretical (structural) positivity violation: if it were impossible for some state to have mask mandates
  - Not violated
- Practical positivity violation: no variation in A in some strata of W
  - I.E. if no variation in pollical affiliation and implementing mandates
  - More likely the higher dimension of W and the smaller the N
  - Reduced the adjustment set W through screening based on univariate associations with the outcome

Causal Roadmap Steps

1. The Causal Question
   - The Research Question
   - The Causal Model
   - The Causal Parameter
2. Observed Data and Statistical Model
3. Identification
4. Statistical Estimation and Inference
5. Interpretation of Results

Statistical Estimation and Inference

- TMLE with Super Learner
- Learners in SL library:
  - empirical mean
  - generalized additive models (gam)
  - recursive partitioning and regression trees (rpart)
  - extreme gradient boosting (xgboost)
  - multivariate adaptive regression splines (earth)
  - all leaners paired with covariate screening algorithm.
- Statistical inference via the estimated influence curve.
- All analyses were conducted in with the ltmle package

Causal Roadmap Steps

1. The Causal Question
   - The Research Question
   - The Causal Model
   - The Causal Parameter
2. Observed Data and Statistical Model
3. Identification
4. Statistical Estimation and Inference
5. Interpretation of Results

Step 5: Interpretation of Results

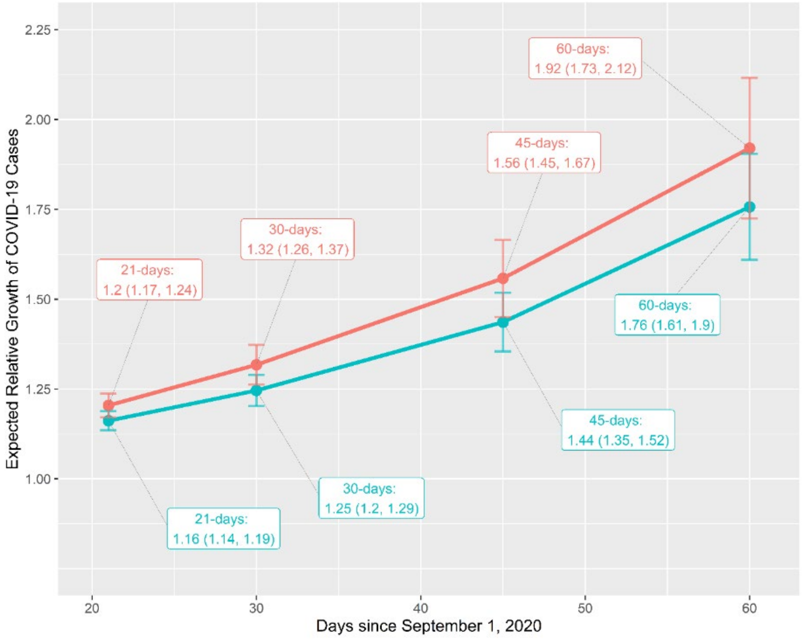
- 25 states had the mask mandate in place by 1 Sept. 2020, 25 did not.
- Early acting states had:
  - higher percentage of Black and Hispanic residents
  - higher population density
  - fewer Republican voters
  - more per-capita COVID-19 tests, cases, and deaths prior to 1 Sept
  - more likely to have previously issued orders to stay at home and for school masking

	Early Masking (N=25)	Delayed Masking (N=25)
Caucasian, median % (IQR)	68.5 (55.6, 75.9)	78.3 (63.1, 82)
Age, median years (IQR)	38.8 (37.7, 39.9)	38.2 (36.7, 39.1)
Smoker, median % (IQR)	17 (14.1, 19.3)	17.2 (15.6, 19.3)
Population density, median people per km^2 (IQR)	67.9 (24.6, 160.7)	26.8 (9.6, 62.3)
Urbanicity in 2010, median % (IQR)	81 (73.2, 88)	66.4 (64, 75.1)
Public transportation usage, median % (IQR)	1.8 (0.9, 5.8)	1.2 (0.8, 2)
Prior COVID-19 policies		
Implemented prior lockdown policy, n (%)	24 (96%)	19 (76%)
Implemented gathering restrictions, n (%)	25 (100%)	24 (96%)
Prior COVID-19 outcomes		
Confirmed cases 30 days prior, median per 100,000 residents (IQR)	1390.8 (888, 1717.7)	1036.1 (851.2, 1452.3)
Mobility change 7 days prior, median % (IQR)	9 (8,11)	8 (6,10)

Step 5: Interpretation of Results

- Early implementation of the state-level public masking mandate was associated with a **9% relative reduction** in the relative growth of COVID-19 cases after 2 months.

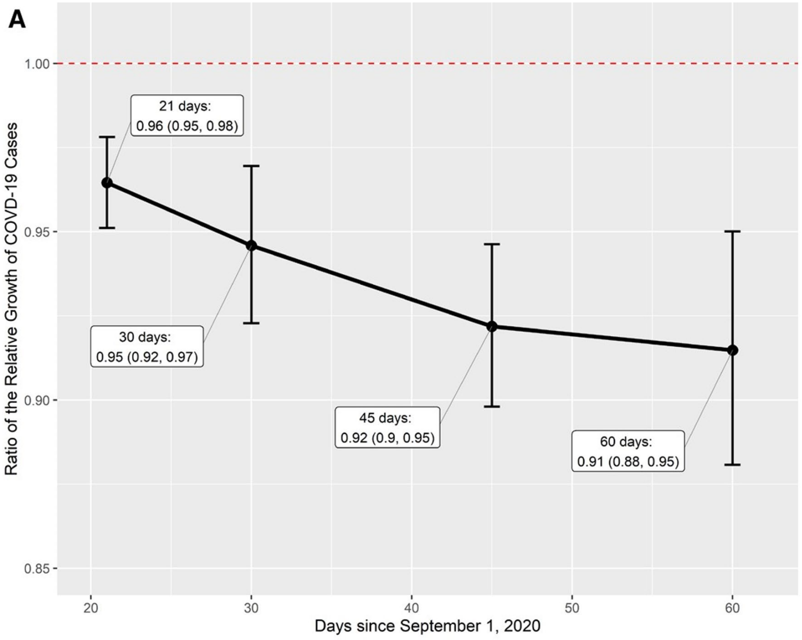
Expected relative growth of COVID-19 cases under early implementation of the public masking mandate



Step 5: Interpretation of Results

- Early implementation of the state-level public masking mandate was associated with a **9% relative reduction** in the relative growth of COVID-19 cases after 2 months.

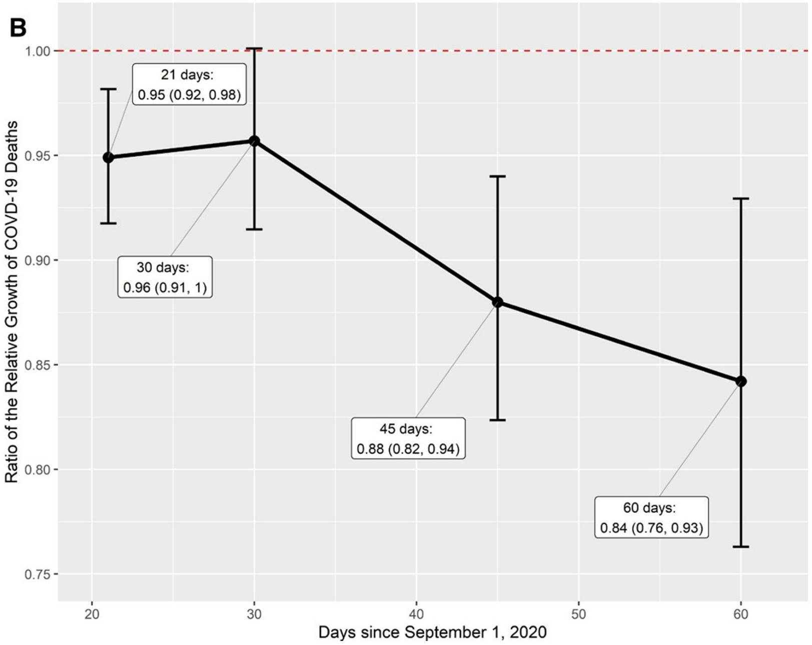
Expected relative growth of COVID-19 cases under early implementation of the public masking mandate



Step 5: Interpretation of Results

- Early implementation of the state-level public masking mandate was associated with a **16% relative reduction** in the change of COVID-19 deaths after 2 months

Expected relative growth of COVID-19 deaths under early implementation of the public masking mandate



Step 5: Interpretation of Results

Associations were considerably larger when using an unadjusted approach

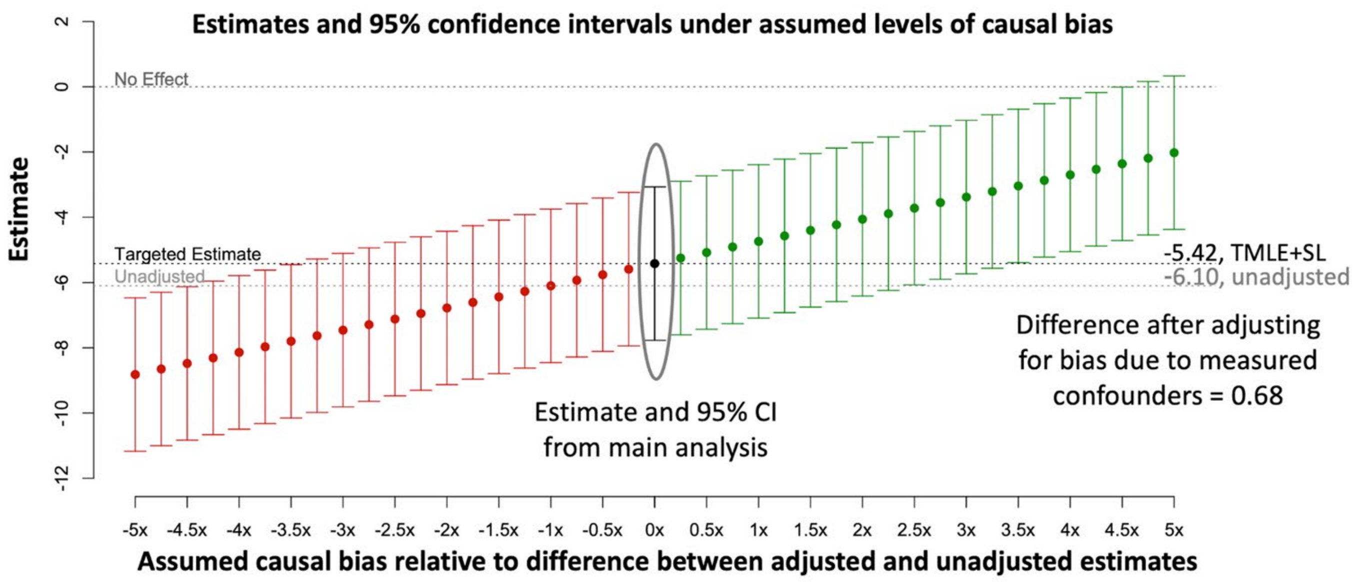
Not adjusting for confounders increased the estimated 2-month associations:

- from 9% to 28% for cases
- from 16% to 29% for death

	Early (95% CI)	Delayed (95% CI)	RR (95% CI)	RD (95% CI)
TMLE- 60 days	1.76 (1.61, 1.9)	1.92 (1.73, 2.12)	0.91 (0.88, 0.95)	-0.16 (-0.24, -0.09)
Unadjusted-60 days	1.54 (1.44, 1.65)	2.16 (1.83, 2.49)	0.72 (0.61, 0.85)	-0.61 (-0.96, -0.27)

Sensitivity analysis option

Causal Gap Adjusted Confidence Intervals



Courtesy of "Targeted-Learning Based Statistical Analysis Plan" Webinar by Susan Gruber on 28 April 2021

Step 5: Interpretation of Results

Limitations. Bias due to:

1. Measurement error
2. Independence assumption between states likely violated
    1. Confidence intervals overly precise
3. Incomplete control for measured confounders and complex dependence
    1. High dimensional covariate set compared to number of observations
4. Unmeasured confounding.
    1. Examples:
        1. perceived epidemic trajectory
        2. perceived compliance with prior public health policies
        3. strength of the state's public health department
        4. complex infectious disease dynamics
    2. Likely adjusted results were closer to true effect than the unadjusted and possible adjusting for unmeasured confounders would adjust effects towards the null

Step 5: Interpretation of Results

Conclusion

- Inference qualitatively agrees with prior research
- The causal effect of interest was not identifiable
  - Unmeasured confounding, positivity violations, and interference
  - Still specified a statistical estimand that best answered our research question.
  - Estimated avoiding parametric assumptions
- The causal roadmap allowed elaborating and critically evaluating the assumptions
  - Specified statistical parameters that would equal causal parameters if the identifiability assumptions did, in fact, hold.
- Estimated with TMLE and SuperLearner to minimize statistical bias
- Causal Roadmap provided framework to interpret the resulting point estimates and inferences appropriately:
  - Associations providing best answers to critical policy decisions.

Discussion - What do you think?

1. Identifiability:
   - Can you ever truly believe you have this in observational data?
2. Causal roadmap
   - Helpful, or just new terminology and structure to what good research already does?
   - Sufficient in limitations of the analysis?
   - Should the roadmap push to not do flawed observational analyses?
3. Targeted learning
   - When is juice worth the squeeze?