Welcome to the Sentinel Innovation and Methods Seminar Series

The webinar will begin momentarily

Please visit www.sentinelinitiative.org for recordings of past sessions and details on upcoming webinars.

Note: closed-captioning for today's webinar will be available on the recording posted at the link above.

Sentinel

An Introduction to Negative Control and Proximal Causal Learning

Xu Shi

Department of Biostatistics
University of Michigan

January 28, 2022

Acknowledgment

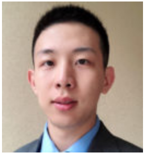
Yifan Cui
National U of Singapore

Oliver Dukes
U of Pennsylvania

Kendrick Li
U of Michigan

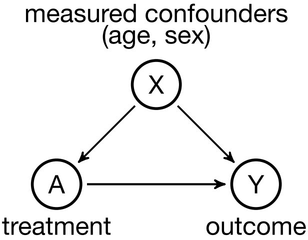
Wang Miao
Peking U

Eric Tchetgen Tchetgen
U of Pennsylvania



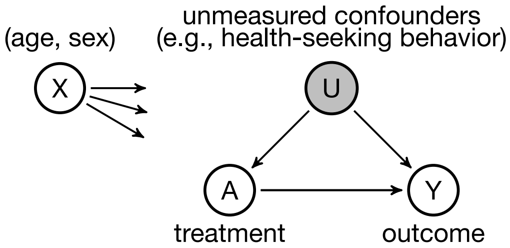
Andrew Ying
U of Pennsylvania

The "randomized" scenario in causal inference



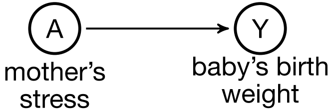
- Estimand: the average treatment effect ATE = $E[Y(1)] - E[Y(0)]$
- Key identification assumption: no unmeasured confounding
  - "Randomized" within each stratum of X
  - Not empirically verifiable

Unmeasured confounding: a threat to causal inference



- Hereafter all arguments are made implicitly conditional on X
- Unmeasured confounders *U*
  - At the center of much skepticism about observational studies
  - The instrumental variable (IV) methods require randomization
- A hidden treasure: negative control variable

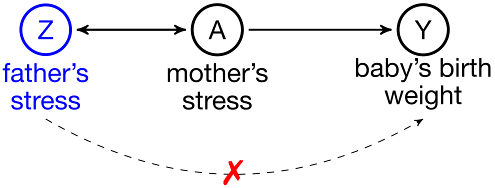
Does stress during pregnancy affect birth weight?



- Observational study on effect of mother’s stress on birth weight

Davey Smith 2008, 2012

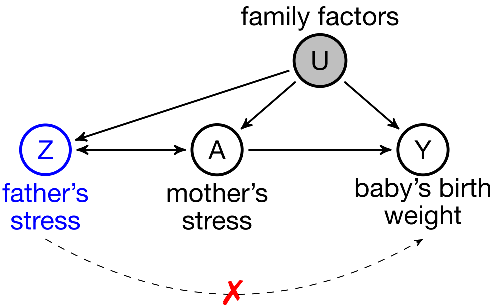
Does stress during pregnancy affect birth weight?



- Observational study on effect of mother's stress on birth weight
- No effect from father's stress after adjusting for mother's stress
    - Nonzero effect of father's stress indicates hidden bias

Davey Smith 2008, 2012

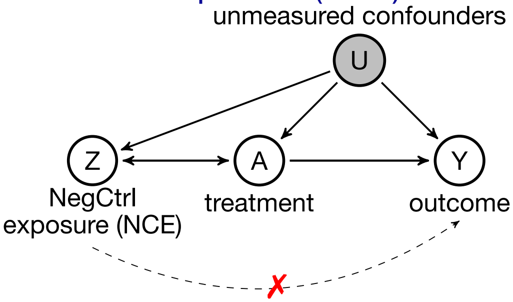
Does stress during pregnancy affect birth weight?



- Observational study on effect of mother's stress on birth weight
- No effect from father's stress after adjusting for mother's stress
    - Nonzero effect of father's stress indicates hidden bias
- Family factors could be an unmeasured confounder

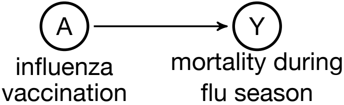
Davey Smith 2008, 2012

Negative control exposure (NCE)



- Z is an NCE if $\boxed{Y(a,z) = Y(a) \text{ and } Z \perp\!\!\!\perp Y(a) \mid U}$
  - (1) It does not causally affect Y
  - (2) It is associated with $Y(a)$ only through U

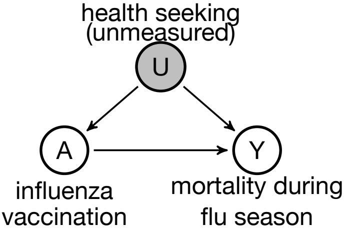
Does flu shot prevent 50% death in the elderly?



- Observational study on flu vaccine effectiveness
  - found 50% reduction in risk of all cause mortality during winter

Jackson et al. 2006; also considered using injury/trauma hospitalization to detect unmeasured confounding bias

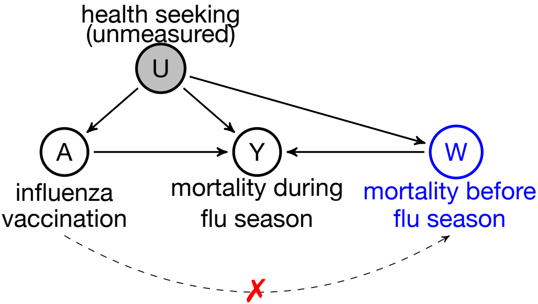
Does flu shot prevent 50% death in the elderly?



- Observational study on flu vaccine effectiveness
  - found 50% reduction in risk of all cause mortality during winter
- Potential unmeasured confounding by health seeking behavior

Jackson et al. 2006; also considered using injury/trauma hospitalization to detect unmeasured confounding bias

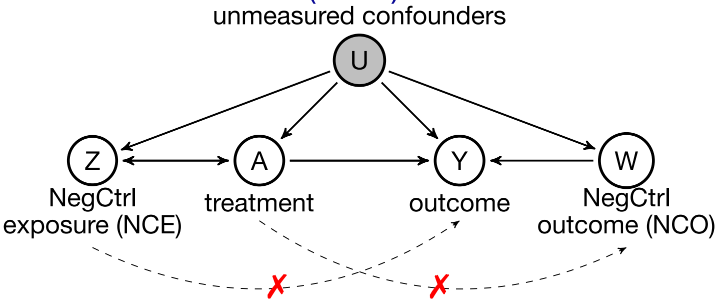
Does flu shot prevent 50% death in the elderly?



- Observational study on flu vaccine effectiveness
    - found 50% reduction in risk of all cause mortality during winter
- Potential unmeasured confounding by health seeking behavior
- Use mortality before flu season to detect confounding bias

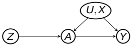
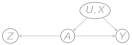
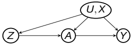
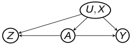
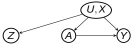
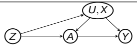
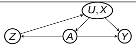
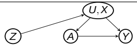
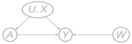
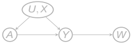
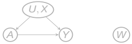
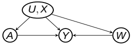
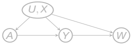
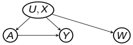
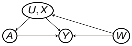
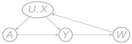
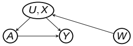
Jackson et al. 2006; also considered using injury/trauma hospitalization to detect unmeasured confounding bias

Negative control outcome (NCO)



- Z is an NCE if $\boxed{Y(a,z) = Y(a) \text{ and } Z \perp\!\!\!\perp Y(a) \mid U}$
  (1) It does not causally affect Y
  (2) It is associated with $Y(a)$ only through U
- W is an NCO if $\boxed{W(a,z) = W \text{ and } W \perp\!\!\!\perp (A,Z) \mid U}$
  (1) It is not causally affected by A
  (2) It is associated with (A,Z) only through U

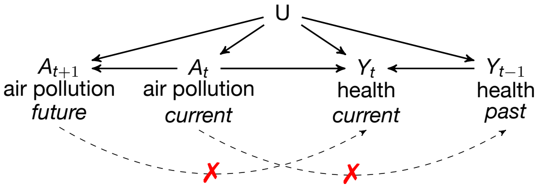
More examples that encode the NC assumptions

<table>
<tr><th colspan="4">Examples of NCE</th></tr>
<tr><td></td><td>$Z \to A$ (pre-treatment)</td><td>$A \to Z$ (post-treatment)</td><td>$Z \perp\!\!\!\perp A$</td></tr>
<tr><td>No arrow between U and Z (may violate U-comparability)</td><td>Instrumental variable (IV)
</td><td></td><td></td></tr>
<tr><td>$U \to Z$</td><td>Invalid IV
</td><td>Post-treatment proxy of U
</td><td>Proxy of U
</td></tr>
<tr><td rowspan="2">$Z \to U$</td><td colspan="3">May violate Assumptions if there is $W \to U$</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><th colspan="4">Examples of NCO</th></tr>
<tr><td></td><td>$W \to Y(a)$</td><td>$Y(a) \to W$</td><td>$Y(a) \perp\!\!\!\perp W \mid (U,X)$</td></tr>
<tr><td>No arrow between U and W (violate U-comparability)</td><td></td><td>Violate NCO definition
</td><td></td></tr>
<tr><td>$U \to W$</td><td></td><td>Violate NCO definition
</td><td></td></tr>
<tr><td rowspan="2">$W \to U$</td><td colspan="3">May violate Assumption if there is $Z \to U$</td></tr>
<tr><td></td><td>Violate NCO definition
</td><td></td></tr>
</table>

Examples of Z, A, U and W, Y, U relationships. Grey indicates violation of assumptions. (Shi, Miao, and Tchetgen 2020)

Negative controls are widely available

- Air pollution and health outcomes: the future $\not\Rightarrow$ the past [1]
  - NCE = future exposure; NCO = past outcome



- Genetics research and batch effect [2]
  - Use control genes to remove unwanted variation
- Drug/vaccine comparative effectiveness and safety [3]
  - Use secondary treatments or outcomes in electronic health records
  - Can combine multiple binary negative control variables

---

[1] Flanders et al. 2011 [2] Gagnon-Bartsch and Speed 2012 [3] Schuemie et al. 2014

Detection, reduction, and correction of bias

Limitation: application focused on bias detection; methods may require strong assumptions

Detect	[1]: Time-series study.
	[2]: invalid NCE.
Reduce	[3]: Time-series study.
	[4]: Standardized mortality ratio in occupational cohort study.
	[5]: Drug–outcome pairs with no plausible causal effect.
Correct	[6]: Time-to-event outcome.
	[7]: Generalized difference-in-differences using NCO.
	[8]: Calibration using NCO.
	[9]: Removing unwanted variation in gene-expression analysis.
	[10]: Nonparametric identification using double negative control.

[1] Flanders et al. 2011.

[2] Davey Smith 2012; Weisskopf, Tchetgen Tchetgen, and Raz 2016.

[3] Flanders, Strickland, and Klein 2017; Miao and Tchetgen Tchetgen 2017.

[4] Richardson et al. 2015.

[5] Schuemie et al. 2014, 2018.

[6] Richardson et al. 2014; Tchetgen Tchetgen, Sofer, and Richardson 2015.

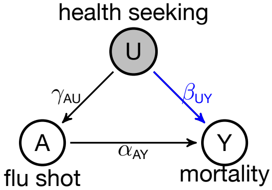
[7] Sofer et al. 2016; Glynn and Ichino 2019.

[8] Tchetgen Tchetgen 2014.

[9] Gagnon-Bartsch and Speed 2012; Jacob, Gagnon-Bartsch, and Speed 2016; Wang et al. 2017.

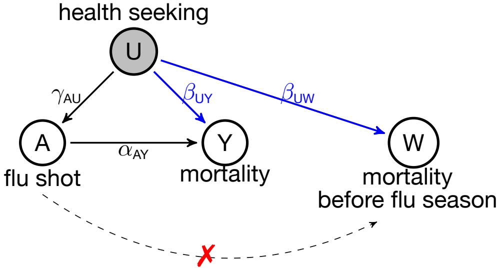
[10] Miao, Shi, and Tchetgen Tchetgen 2018; Miao, Geng, and Tchetgen Tchetgen 2018; Shi, Miao, and Tchetgen Tchetgen 2020.

Intuition behind identification



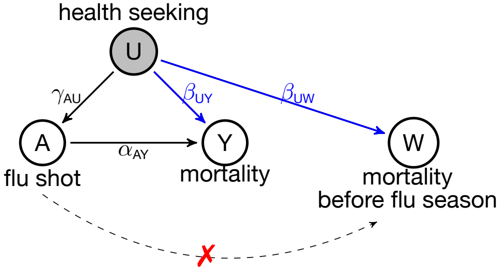
- Confounding bias is a product of *U*-*A* and *U*-*Y* association ($\gamma_{AU}\beta_{UY}$)

Intuition behind identification



- Confounding bias is a product of *U*-*A* and *U*-*Y* association ($\gamma_{AU}\beta_{UY}$)
  - *A*-*W* association is a product of *U*-*A* and *U*-*W* association ($\gamma_{AU}\beta_{UW}$)
  - Problem solved if *U*-*Y* association = *U*-*W* association

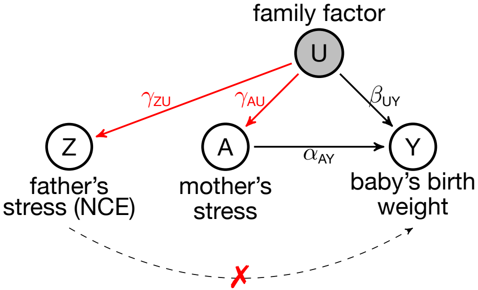
Intuition behind identification



- Confounding bias is a product of *U*-*A* and *U*-*Y* association ($\gamma_{AU}\beta_{UY}$)
  - *A*-*W* association is a product of *U*-*A* and *U*-*W* association ($\gamma_{AU}\beta_{UW}$)
  - Problem solved if *U*-*Y* association = *U*-*W* association
- Regress $\underbrace{Y \text{ on } A,}_{\alpha_{AY}+\gamma_{AU}\beta_{UY}}$ and $\underbrace{W \text{ on } A}_{\gamma_{AU}\beta_{UW}}$, then ATE = diff in coefs of *A*
- A special case: the difference-in-difference method[11]

[11] Richardson et al. 2014, 2015; Tchetgen Tchetgen, Sofer, and Richardson 2015; Sofer et al. 2016.

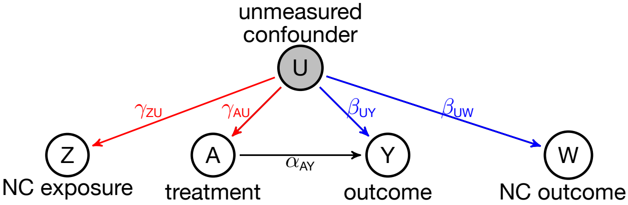
Intuition behind identification



- Confounding bias is a product of *U*-*A* and *U*-*Y* association ($\gamma_{AU}\beta_{UY}$)
  - *Z*-*Y* association is a product of *U*-*Z* and *U*-*Y* association ($\gamma_{ZU}\beta_{UY}$)
  - Problem solved if *U*-*A* association = *U*-*Z* association
- Regress $\underbrace{Y \text{ on } A}_{\alpha_{AY} + \gamma_{AU}\beta_{UY}}$ and $\underbrace{Z}_{\gamma_{ZU}\beta_{UY}}$, then ATE = diff in coefs of *A* and *Z*
- A special case: air pollution studies[12]

[12] Flanders et al. 2011; Flanders, Strickland, and Klein 2017; Miao and Tchetgen Tchetgen 2017.

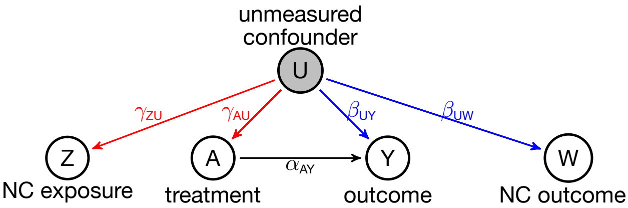
Intuition behind identification



- What if $\beta_{UY} \neq \beta_{UW}$ and $\gamma_{AU} \neq \gamma_{ZU}$?
  - *A*-*W* association ($\gamma_{AU}\beta_{UW}$) recovers the confounding bias ($\gamma_{AU}\beta_{UY}$) up to a scale $\frac{\beta_{UY}}{\beta_{UW}}$
  - We cannot identify either β_{UY} or β_{UW}, but we can identify the ratio

[13] Miao, Shi, and Tchetgen Tchetgen 2018; Miao, Geng, and Tchetgen Tchetgen 2018; Shi, Miao, and Tchetgen 2020; Shi, Miao, and Tchetgen Tchetgen 2020; Tchetgen et al. 2020.

Intuition behind identification



- What if $\beta_{UY} \neq \beta_{UW}$ and $\gamma_{AU} \neq \gamma_{ZU}$?
  - *A*-*W* association ($\gamma_{AU}\beta_{UW}$) recovers the confounding bias ($\gamma_{AU}\beta_{UY}$) up to a scale $\frac{\beta_{UY}}{\beta_{UW}}$
  - We cannot identify either β_{UY} or β_{UW}, but we can identify the ratio
- Double negative control: use both an NCE and an NCO[13]
  - Identify the ratio using the NCE: $\frac{\beta_{UY}}{\beta_{UW}} = \frac{Z\text{-}Y \text{ association}}{Z\text{-}W \text{ association}} = \frac{\gamma_{ZU}\beta_{UY}}{\gamma_{ZU}\beta_{UW}}$
  - W recovers bias up to a scale; Z recovers that scale

[13] Miao, Shi, and Tchetgen Tchetgen 2018; Miao, Geng, and Tchetgen Tchetgen 2018; Shi, Miao, and Tchetgen 2020; Shi, Miao, and Tchetgen Tchetgen 2020; Tchetgen et al. 2020.

From linear additive model to nonparametric identification

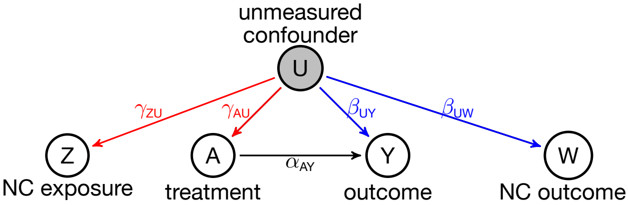
- For simplicity, consider the following two linear additive models
  - $E[Y \mid A, Z, U] = \beta_0 + \beta_A A + \beta_U U$
    $E[W \mid A, Z, U] = \gamma_0 + \gamma_U U$
  - The causal effect is $\beta_A = E[Y(1) - Y(0)] = E[Y(1) - Y(0) \mid U]$
- One can show that $\boxed{E[Y \mid A, Z] = \beta_0^* + \beta_A A + \beta_U^* E[W \mid A, Z]}$
  - Regress Y on A and $\widehat{W}$, where $\widehat{W}$ is predicted from $E[W \mid A, Z]$
- The ATE can be identified nonparametrically[14]
  - $E[Y(a)] = E[h(a, W)]$, $h()$ satisfies $\boxed{E[Y \mid A, Z] = E[h(A, W) \mid A, Z]}$
  - e.g., in the linear model above, $h(A, W) = \beta_0^* + \beta_A A + \beta_U^* W$
  - Requires Z and W to be sufficiently informative about U

[14] Miao, Shi, and Tchetgen Tchetgen 2018; Miao, Geng, and Tchetgen Tchetgen 2018; Shi, Miao, and Tchetgen Tchetgen 2020; Shi, Miao, and Tchetgen Tchetgen 2020; Tchetgen et al. 2020.

Double negative control in practice

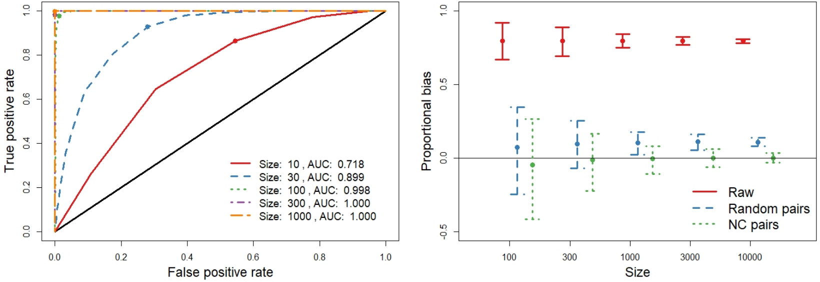
- Two stage least squares (TSLS) under linear models
  - Stage I: regress W on A and Z, and obtain fitted values $\widehat{W}$
  - Stage II: regress Y on A, adjusting for $\widehat{W}$ (as if it is U)
- Can use existing instrumental variable software packages
  - `SYSLIN` in SAS; `ivregress`, `ivreg`, `ivreg2` in Stata; `gmm`, `sem`, `ivpack`, `AER` in R
  - e.g., `gmm::gmm(g=Y~A+W+X,x=~A+Z+X)` in R

How to find a candidate negative control variable?



- Data-driven Automated Negative Control Estimation (DANCE)
  - Identifies triplets of negative control variables
  - Aggregates ATEs obtained from all pairs of negative controls
  - Limitation: can only detect a special type of negative control
- Rationale: all paths from $\{W, Z\}$ to $\{Y, A\}$ pass through U
  - Therefore $\Sigma_{\{W,Z\},\{Y,A\}} = \begin{pmatrix} cov(W,Y) & cov(W,A) \\ cov(Z,Y) & cov(Z,A) \end{pmatrix}$ is rank deficient
  - Such a rank constraint can be determined using statistical tests

Performance of the DANCE algorithm



- High AUC in negative control detection
- Low bias in effect estimation using the detected negative controls

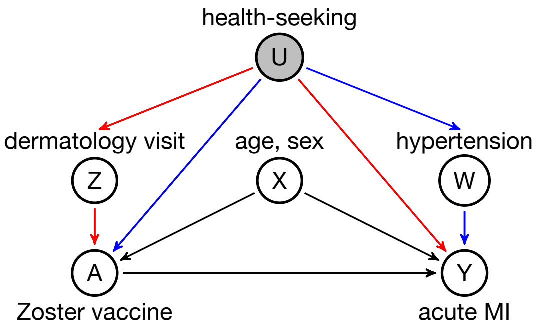
Proximal Causal Learning

Are two cheap, noisy measures better than one expensive, accurate measure?[15]

- Hard to eliminate measurement error
  - The "no unmeasured confounding" assumption depends on investigator's ability to accurately measure covariates capturing all potential sources of confounding
  - The most one can hope for is that covariate measurements are at best proxies of the true underlying confounding mechanism
- Easier to get to the right kind of measurement error
  - Acknowledge that covariates are imperfect proxies of confounders
  - Find proxies that satisfy certain assumptions
  - Allow the "no unmeasured confounding" to be violated

[15] Browning and Crossley 2009.

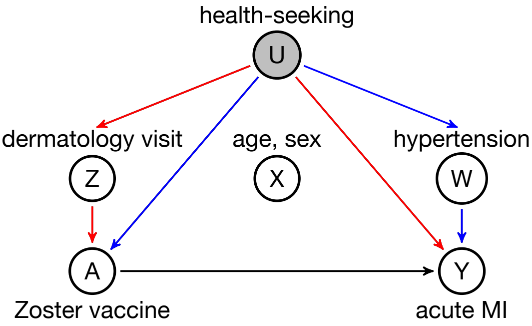
An example in vaccine safety study



- Adverse effect of a new Zoster vaccine on acute MI
- Plan to adjust for the following confounders:
  - age, sex (*X*)
  - dermatology visit (*Z*), hypertension (*W*)

Tchetgen et al. 2020

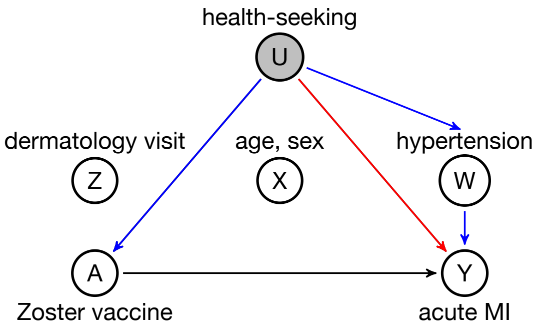
An example in vaccine safety study



- Adverse effect of a new Zoster vaccine on acute MI
- Plan to adjust for the following confounders:
  - age, sex (X)
  - dermatology visit (Z), hypertension (W)

Tchetgen et al. 2020

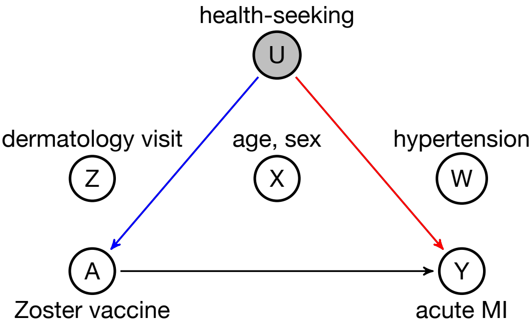
An example in vaccine safety study



- Adverse effect of a new Zoster vaccine on acute MI
- Plan to adjust for the following confounders:
  - age, sex (*X*)
  - dermatology visit (*Z*), hypertension (*W*)

Tchetgen et al. 2020

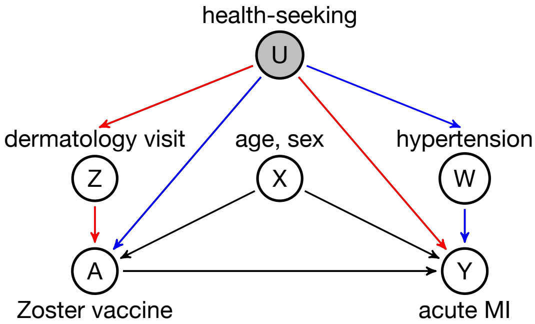
An example in vaccine safety study



- Adverse effect of a new Zoster vaccine on acute MI
- Plan to adjust for the following confounders:
  - age, sex (X)
  - dermatology visit (Z), hypertension (W)

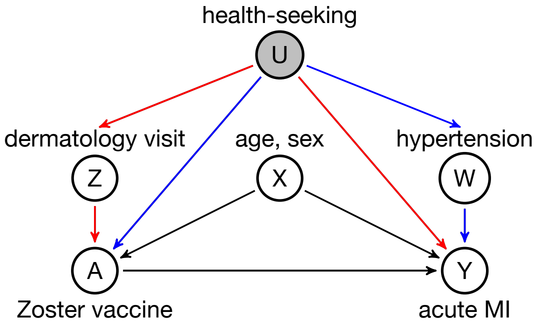
Tchetgen et al. 2020

An example in vaccine safety study



Tchetgen et al. 2020

An example in vaccine safety study



- Three types of confounding variables
  - Common causes of the treatment and outcome: age, sex (X)
  - Treatment-inducing confounding proxy: dermatology visit (Z)
  - Outcome-inducing confounding proxy: hypertension (W)

Tchetgen et al. 2020

Classical vs Proximal causal inference

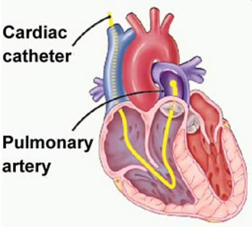
- **Classical causal inference: fails when U exists**
  - Confounders $\{X, W, Z\}$: age, sex, dermatology visit, hypertension
  - Standard g-formula: $E[Y(a)] = E[m(a, X, W, Z)]$
  - $m(a, x, w, z)$ is the outcome model
    $m(a, x, w, z) = E[Y \mid A = a, X = x, W = w, Z = z]$
  - Estimation via g-computation

- **Proximal causal inference**
  - $X =$ age, sex; $W =$ hypertension; $Z =$ dermatology visit
  - Proximal g-formula: $E[Y(a)] = E[h(a, W, X)]$
  - $h(a, w, x)$ is the outcome bridge function
    $E[h(a, w, x) \mid A = a, Z = z, X = x] = E[Y \mid A = a, Z = z, X = x]$
  - Estimation via proximal g-computation or two-stage least squares

---

Greenland and Robins 1986; In slight abuse of notation let $\sum$ denote an integral in the case of a continuous variable

Application to the SUPPORT study



- Right heart catheterization (RHC) procedure
  - Performed to measure blood flow and pressures in the heart
  - Many physicians believed that measurements from the RHC can guide therapy and lead to better outcomes for critically ill patients
  - Due to the popularity and strong belief of the procedure, conducting a clinical trial was unethical

- The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT)[16]
  - Evaluate the effectiveness of RHC among adults admitted to the intensive care unit (ICU)
  - 2184 patient managed with RHC, 3551 without RHC

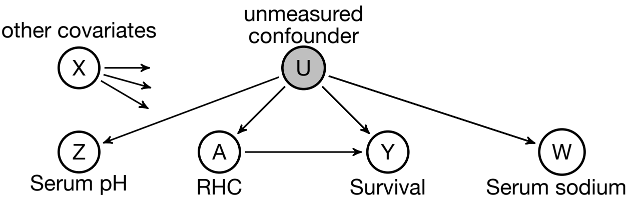
[16] Connors et al. 1996.

A controversial result

- The SUPPORT study found that RHC was harmful
- Potential confounding
  - Confounding bias might show harmful effect of RHC
  - Patients for whom RHC was performed might have been a lot sicker
- This data set has been analyzed by many researchers[17]
  - Majority relying on the no unmeasured confounding assumption

[17] Lin, Psaty, and Kronmal 1998; Tan 2006; Li, Morgan, and Zaslavsky 2018; Mao and Li 2020.

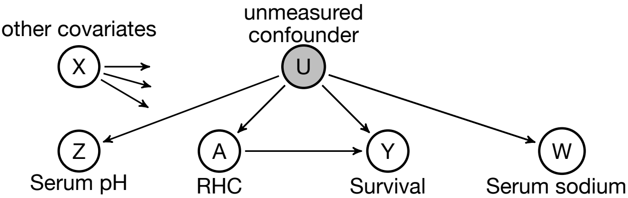
Candidate proxies in the SUPPORT study



- The SUPPORT study collected 72 covariates including
  - demographics, comorbidity, vital signs, functional status
  - physiological status measured from a blood test during the initial 24 hours in the ICU $\Rightarrow$ 10 candidate proxies

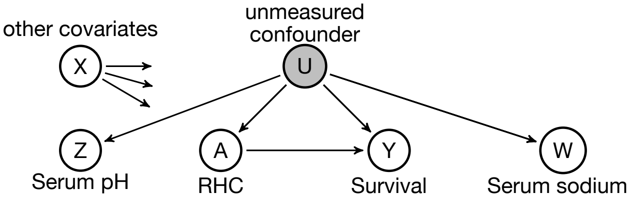
- We applied our DANCE algorithm to find valid proxies
  - Most frequently selected pair: ph and sod
  - ph = Serum pH; sod = Serum sodium

Methods



- We evaluate effect of RHC on survival time in days
    - Assumed a linear additive model
- Estimation
    - Proximal two stage least squares
    - Inverse probability weighting to adjust for the other covariates X

Results



Proxy variables	RHC effect (95% CI)
W = ph, Z = sod	-0.44 (-1.00, 0.11)
W = sod, Z = ph	-0.40 (-1.09, 0.30)
Average over all detected (W,Z) pairs	-0.71 (-1.50, 0.08)
Naive adjustment	-1.29 (-1.83, -0.75)

- RHC was not significantly associated with survival time
- Note that the role of Z and W are exchangeable
  - Our results remained invariant to the choice of *W* and *Z*
  - This verifies that the graph is correctly specified

Summary

- Negative controls and proxies can adjust for confounding bias
- Can directly use off-the-shelf software packages
- A data-driven pipeline of negative control detection and adjustment
- Current work by the proximal causal inference group

Review on negative controls	arXiv:2009.05641
Nonparametric identification	arXiv:1609.08816
Binary variable setting	arXiv:1808.04906
Outcome confounding bridge	arXiv:1808.04945
Proximal causal inference	arXiv:2009.10982
Outcome and treatment confounding bridge	arXiv:2011.08411
Longitudinal data setting	arXiv:2109.07030
Panel data setting	arXiv:2108.13935
Proximal mediation analysis	arXiv:2109.11904

Bibliography I

[1] Sander Greenland and James M Robins. "Identifiability, exchangeability, and epidemiological confounding". In: International journal of epidemiology 15.3 (1986), pp. 413–419.

[2] Alfred F Connors, Theodore Speroff, Neal V Dawson, Charles Thomas, Frank E Harrell, Douglas Wagner, Norman Desbiens, Lee Goldman, Albert W Wu, Robert M Califf, et al. "The effectiveness of right heart catheterization in the initial care of critically Ill patients". In: Jama 276.11 (1996), pp. 889–897.

[3] Danyu Y Lin, Bruce M Psaty, and Richard A Kronmal. "Assessing the sensitivity of regression results to unmeasured confounders in observational studies". In: Biometrics (1998), pp. 948–963.

[4] Lisa A Jackson, Michael L Jackson, Jennifer C Nelson, Kathleen M Neuzil, and Noel S Weiss. "Evidence of bias in estimates of influenza vaccine effectiveness in seniors". In: International Journal of Epidemiology 35.2 (2006), pp. 337–344.

[5] Zhiqiang Tan. "A distributional approach for causal inference using propensity scores". In: Journal of the American Statistical Association 101.476 (2006), pp. 1619–1637.

[6] George Davey Smith. "Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings?" In: Basic & Clinical Pharmacology & Toxicology 102.2 (2008), pp. 245–256.

[7] Martin Browning and Thomas Crossley. "Are two cheap, noisy measures better than one expensive, accurate one?" In: American Economic Review 99.2 (2009), pp. 99–103.

[8] W Dana Flanders, Mitchel Klein, Lyndsey A Darrow, Matthew J Strickland, Stefanie E Sarnat, Jeremy A Sarnat, Lance A Waller, Andrea Winquist, and Paige E Tolbert. "A method for detection of residual confounding in time-series and other observational studies". In: Epidemiology 22.1 (2011), p. 59.

[9] George Davey Smith. "Negative control exposures in epidemiologic studies. Comments on "Negative controls: a tool for detecting confounding and bias in observational studies"". In: Epidemiology 23.2 (2012), pp. 350–351.

[10] Johann A Gagnon-Bartsch and Terence P Speed. "Using control genes to correct for unwanted variation in microarray data". In: Biostatistics 13.3 (2012), pp. 539–552.

Bibliography II

[11] David B Richardson, Dominique Laurier, Mary K Schubauer-Berigan, Eric J Tchetgen Tchetgen, and Stephen R Cole. "Assessment and indirect adjustment for confounding by smoking in cohort studies using relative hazards models". In: American Journal of Epidemiology 180.9 (2014), pp. 933–940.

[12] Martijn J Schuemie, Patrick B Ryan, William DuMouchel, Marc A Suchard, and David Madigan. "Interpreting observational studies: why empirical calibration is needed to correct p-values". In: Statistics in Medicine 33.2 (2014), pp. 209–218.

[13] Eric Tchetgen Tchetgen. "The control outcome calibration approach for causal inference with unobserved confounding". In: American Journal of Epidemiology 179.5 (2014), pp. 633–640.

[14] David B Richardson, Alexander Keil, Eric J Tchetgen Tchetgen, and Glinda S Cooper. "Negative control outcomes and the analysis of standardized mortality ratios". In: Epidemiology 26.5 (2015), pp. 727–732.

[15] Eric J Tchetgen Tchetgen, Tamar Sofer, and David Richardson. "Negative outcome control for unobserved confounding under a Cox proportional hazards model". In: (2015). Available at `https://biostats.bepress.com/harvardbiostat/paper192/`.

[16] Laurent Jacob, Johann A Gagnon-Bartsch, and Terence P Speed. "Correcting gene expression data when neither the unwanted variation nor the factor of interest are observed". In: Biostatistics 17.1 (2016), pp. 16–28.

[17] Tamar Sofer, David B Richardson, Elena Colicino, Joel Schwartz, and Eric J Tchetgen Tchetgen. "On negative outcome control of unobserved confounding as a generalization of difference-in-differences". In: Statistical Science 31.3 (2016), pp. 348–361.

[18] Marc G Weisskopf, Eric J Tchetgen Tchetgen, and Raanan Raz. "Commentary: on the use of imperfect negative control exposures in epidemiologic studies". In: Epidemiology 27.3 (2016), pp. 365–367.

[19] W Dana Flanders, Matthew J Strickland, and Mitchel Klein. "A new method for partial correction of residual confounding in time-series and other observational studies". In: American Journal of Epidemiology 185.10 (2017), pp. 941–949. **This paper develops a regression-based method taking future air pollution as a negative control exposure to reduce residual confounding bias in a time-series study on air pollution effects.**

Bibliography III

[20] Wang Miao and Eric J Tchetgen Tchetgen. "Invited commentary: bias attenuation and identification of causal effects with multiple negative controls". In: American Journal of Epidemiology 185.10 (2017), pp. 950–953.

[21] Jingshu Wang, Qingyuan Zhao, Trevor Hastie, and Art B Owen. "Confounder adjustment in multiple hypothesis testing". In: Annals of Statistics 45.5 (2017), pp. 1863–1894. **This paper unifies unmeasured confounding adjustment methods in multiple hypothesis testing and provides theoretical guarantees for these methods.**

[22] Fan Li, Kari Lock Morgan, and Alan M Zaslavsky. "Balancing covariates via propensity score weighting". In: Journal of the American Statistical Association 113.521 (2018), pp. 390–400.

[23] W Miao, X Shi, and EJ Tchetgen Tchetgen. "A confounding bridge approach for double negative control inference on causal effects". In: arXiv preprint arXiv:1808.04945 (2018). **This paper introduces the confounding bridge function that links primary and negative control outcome distributions for identification of the average treatment effect leveraging a negative control exposure via a simple two-stage least squares procedure.**

[24] Wang Miao, Zhi Geng, and Eric J Tchetgen Tchetgen. "Identifying causal effects with proxy variables of an unmeasured confounder". In: Biometrika 105.4 (2018), pp. 987–993. **This paper establishes sufficient conditions for nonparametric identification of the average treatment effect using double negative control.**

[25] Martijn J Schuemie, George Hripcsak, Patrick B Ryan, David Madigan, and Marc A Suchard. "Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data". In: Proceedings of the National Academy of Sciences 115.11 (2018), pp. 2571–2577.

[26] Adam Glynn and Nahomi Ichino. "Generalized Nonlinear Difference-in-Difference-in-Differences". In: V-Dem Working Paper 90 (2019). Available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3410888`.

[27] Huzhang Mao and Liang Li. "Flexible regression approach to propensity score analysis and its relationship with matching and weighting". In: Statistics in Medicine (2020).

[28] Xu Shi, Wang Miao, and Eric Tchetgen Tchetgen. "A Selective Review of Negative Control Methods in Epidemiology". In: arXiv preprint arXiv:2009.05641 (2020).

Bibliography IV

[29] Xu Shi, Wang Miao, and Eric J Tchetgen Tchetgen. "Multiply robust causal inference with double negative control adjustment for categorical unmeasured confounding". In: Journal of the Royal Statistical Society: Series B (Statistical Methodology) 82.2 (2020), pp. 521–540. **This paper provides a general semiparametric framework for obtaining inferences about the average treatment effect under categorical unmeasured confounding and negative controls.**

[30] Eric J Tchetgen Tchetgen, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao. "An Introduction to Proximal Causal Learning". In: arXiv preprint arXiv:2009.10982 (2020). **This paper provides a new view of confounding adjustment: do we really want to adjust for variables that are purely "propensity" vs "prognostic" as confounders, or are they in fact proxies of true underlying confounding mechanism operating in a given observational study.**

Questions?