Sentinel

# Addressing Unmeasured Confounding in Observational Studies: Insights from Negative Control Methods and Proximal Causal Inference

**Xu Shi, PhD**
FDA Sentinel Innovation Center
Department of Biostatistics, University of Michigan

August 24, 2025

# Disclaimers

This work was supported by Task Order 75F40124F19011 under Master Agreement 75F40119D10037 from the U.S. Food and Drug Administration (FDA).

The views expressed in this presentation represent those of the presenter and do not necessarily represent the official views of the U.S. FDA.

# Acknowledgments

Yifan Cui
NUS

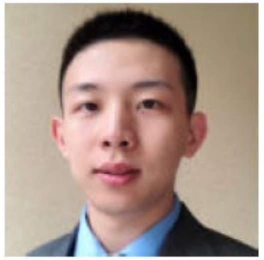
Oliver Dukes
Ghent U

Kendrick Li
St. Jude

Wang Miao
Peking U

Chan Park
UIUC

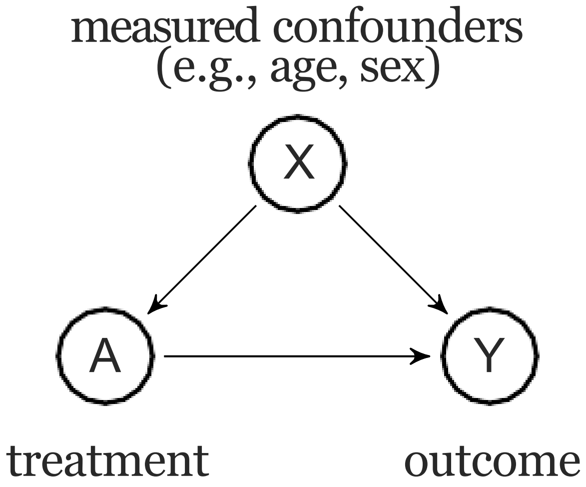
Eric Tchetgen Tchetgen
UPenn



Andrew Ying
UPenn

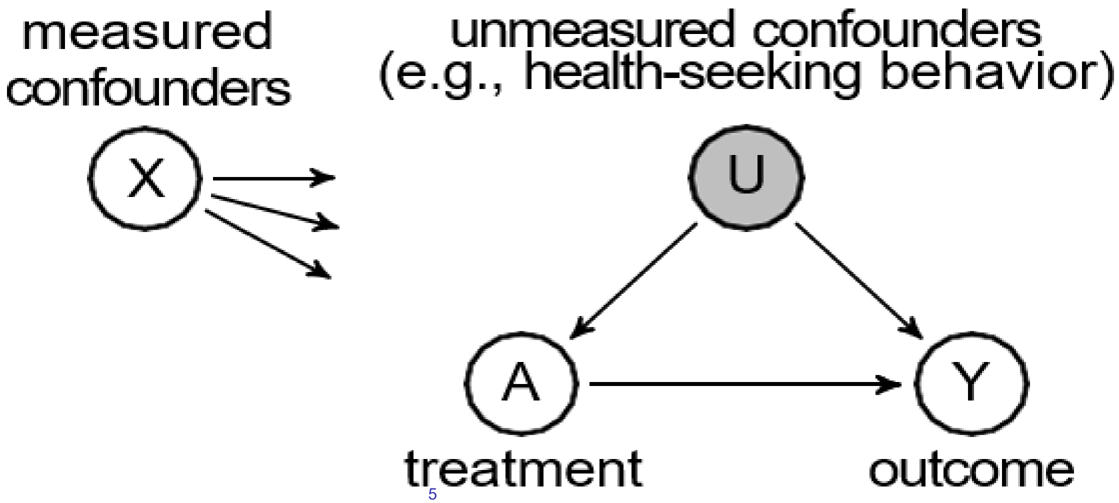
Prabrisha Rakshit
UPenn

# 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

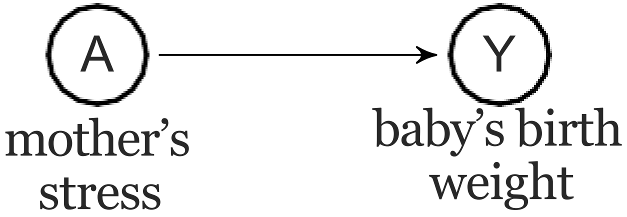



- Unmeasured confounders $U$
  - Often lead to skepticism about observational studies
  - The instrumental variable (IV) methods require randomization
- A hidden treasure: negative control variable
  - Examine associations where a causal link is not expected
  - Widely available, e.g., control genes, EHR/claims, air pollution study

---

Hereafter all arguments are made implicitly conditional on X

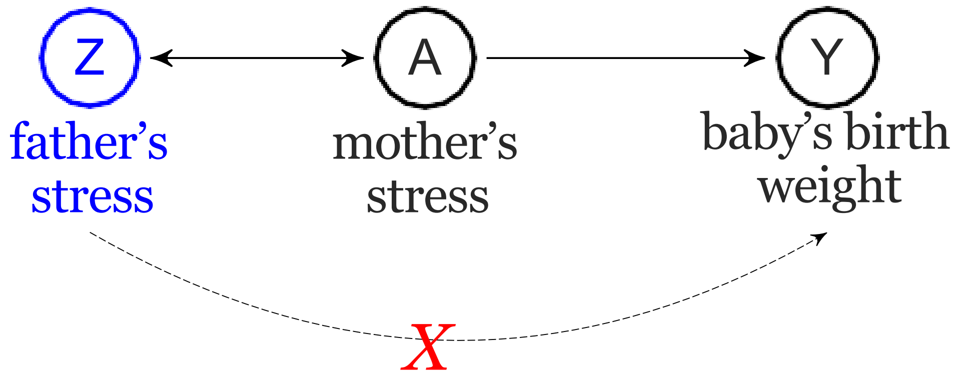
# Does stress during pregnancy affect birth weight?




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

Davey Smith, 2008, 2012

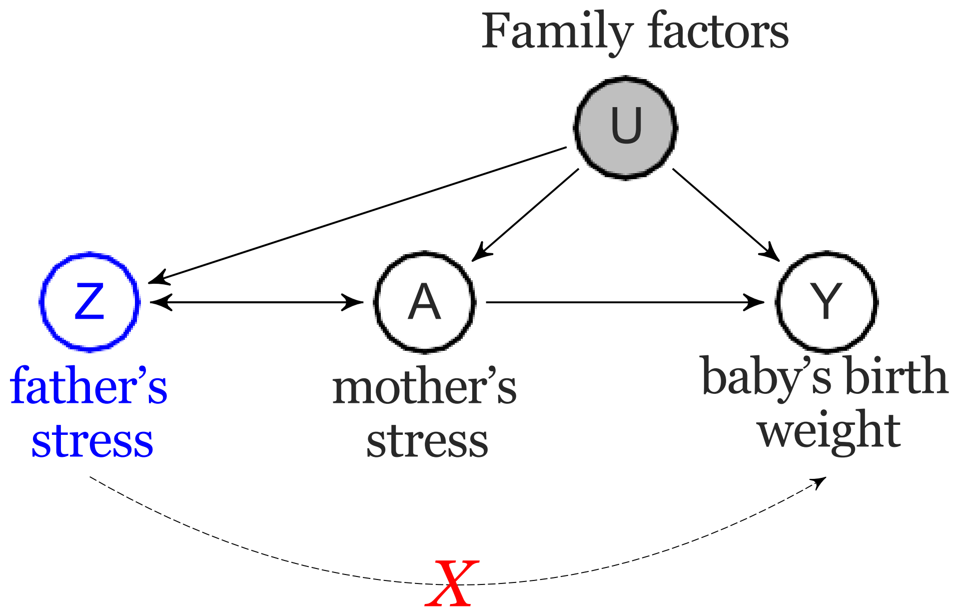
# Does stress during pregnancy affect birth weight?




- 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

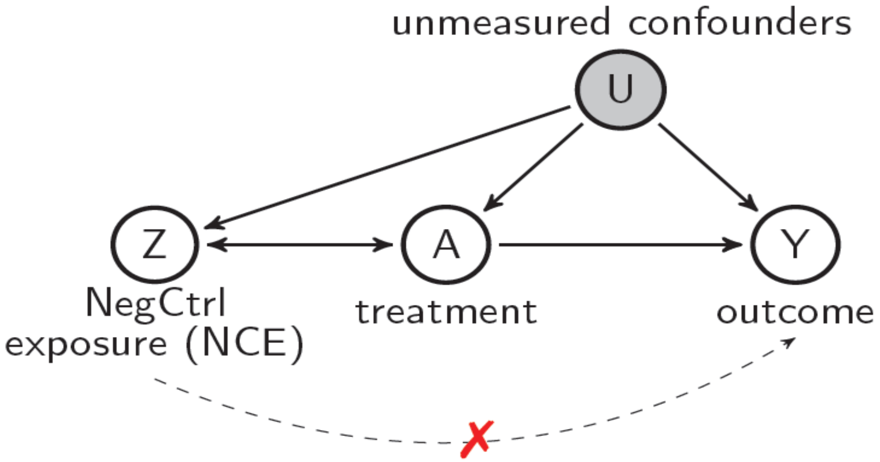
# Does stress during pregnancy affect birth weight?




- 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?

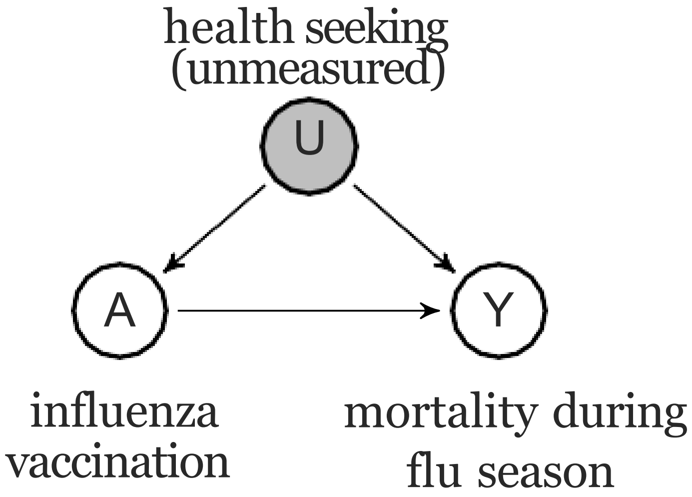

$A \longrightarrow Y$

influenza vaccination → mortality during flu season

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

Jackson et al., [2006]

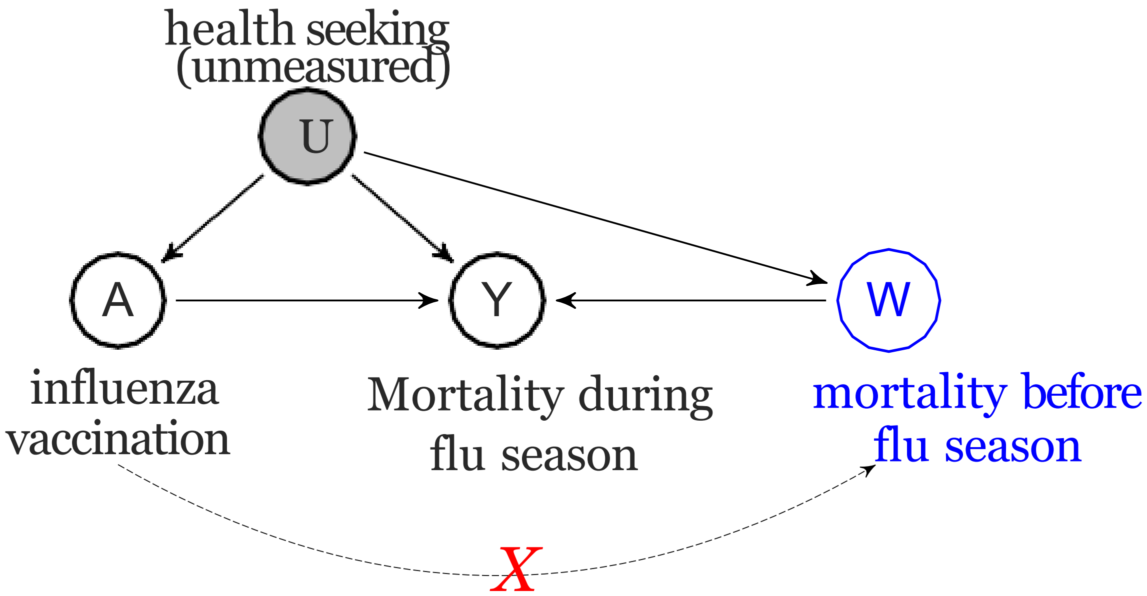
# Does flu shot prevent 50% death in the elderly?


- 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

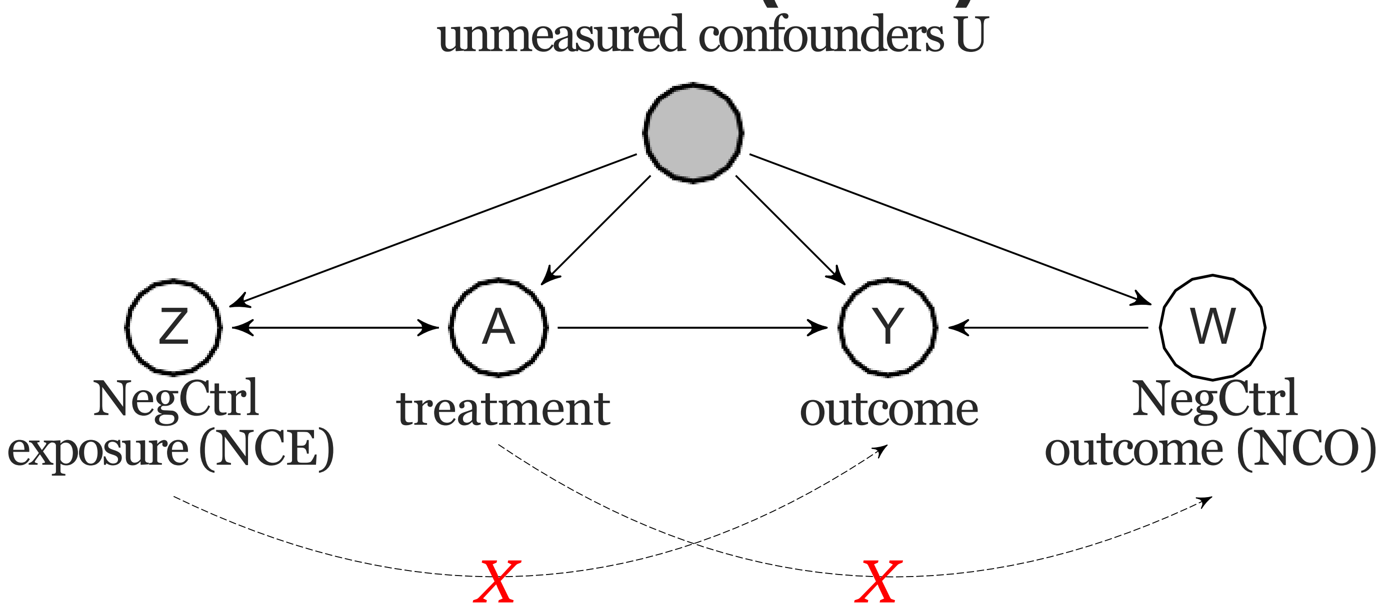
# Does flu shot prevent 50% death in the elderly?


- 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

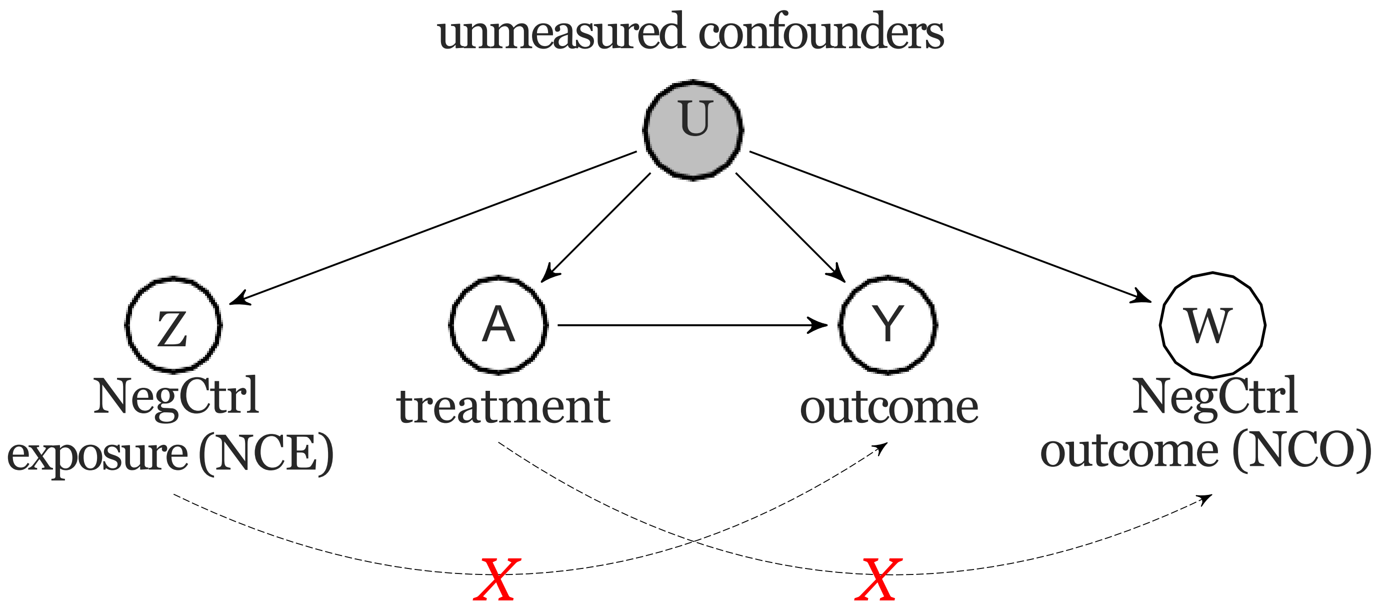
Jackson et al., 2006; also considered using injury/trauma hospitalization to detect unmeasured confounding bias

# Negative control outcome (NCO)



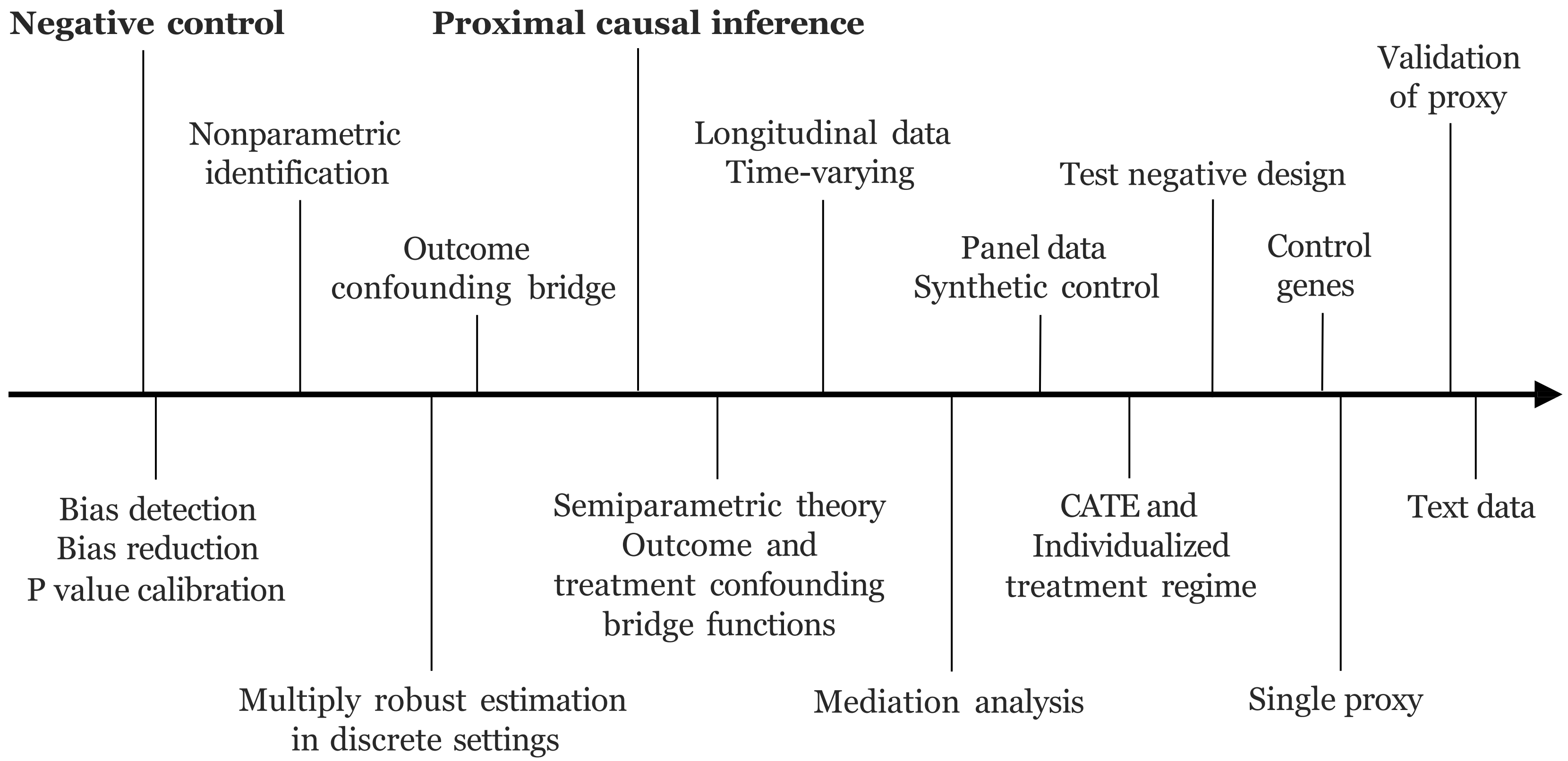

- $Z$ is an NCE if $\boxed{Y(a,z) = Y(a) \text{ and } Z \perp\!\!\!\perp Y(a) \mid U}$
  (1) It does not causally affect $Y$
  (2) It is associated with $Y(a)$ only through $U$
- $W$ is an NCO if $\boxed{W(a,z) = W \text{ and } W \perp\!\!\!\perp (A, Z) \mid U}$
  (1) It is not causally affected by $A$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

# 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 $U$

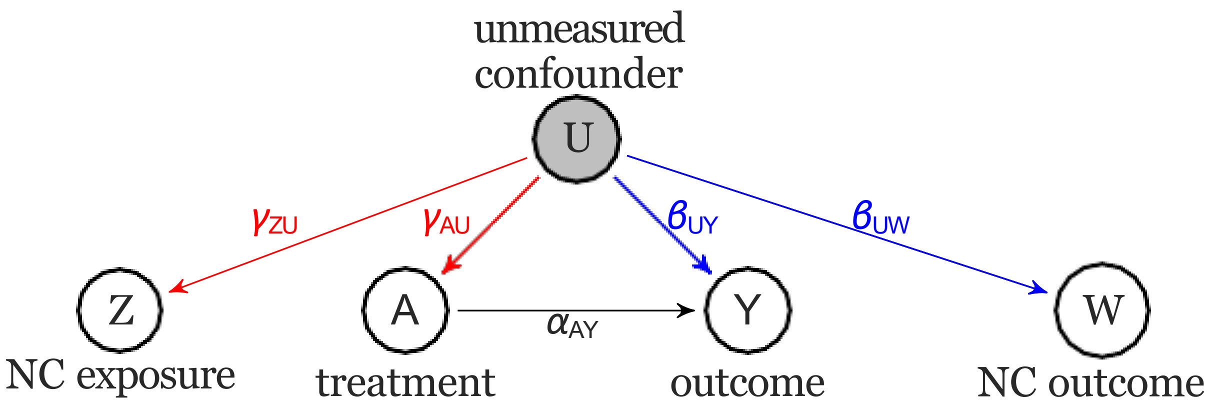
# Double negative control methods and proximal causal inference




The double negative control method has been written into the US FDA's Prescription Drug User Fee Act (PDUFA) reauthorization for fiscal years 2023-2027, in response to the growing stakeholder interest to understand how methodological advances with negative controls can improve real-world studies.

Chen, Bhattacharya, and Keith, 2024; Cui et al., 2020; Dukes, Shpitser, and Tchetgen Tchetgen, 2021; Kummerfeld, Lim, and Shi, 2024; Lipsitch, Tchetgen Tchetgen, and Cohen, 2010; Miao, Geng, and Tchetgen Tchetgen, 2018; Miao, Shi, and Tchetgen Tchetgen, 2018; Qi, Miao, and Zhang, 2024; Shen and Cui, 2023; Shi et al., 2021, 2020; Shi, Miao, and Tchetgen, 2020; Sverdrup and Cui, 2023; Tchetgen, Park, and Richardson, 2023; Tchetgen Tchetgen et al., 2020; Ying et al., 2021

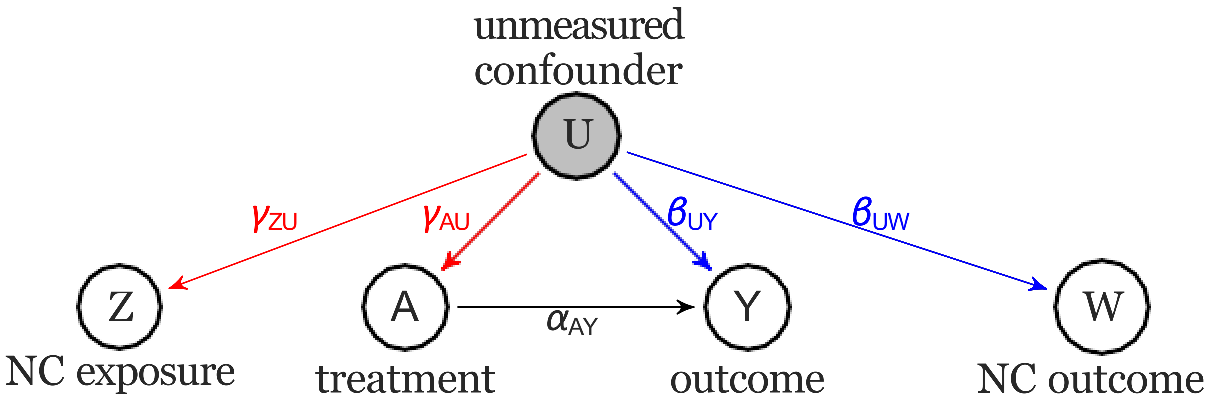
# How to find a candidate negative control variable?




- 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

Kummerfeld, Lim, and Shi, 2022

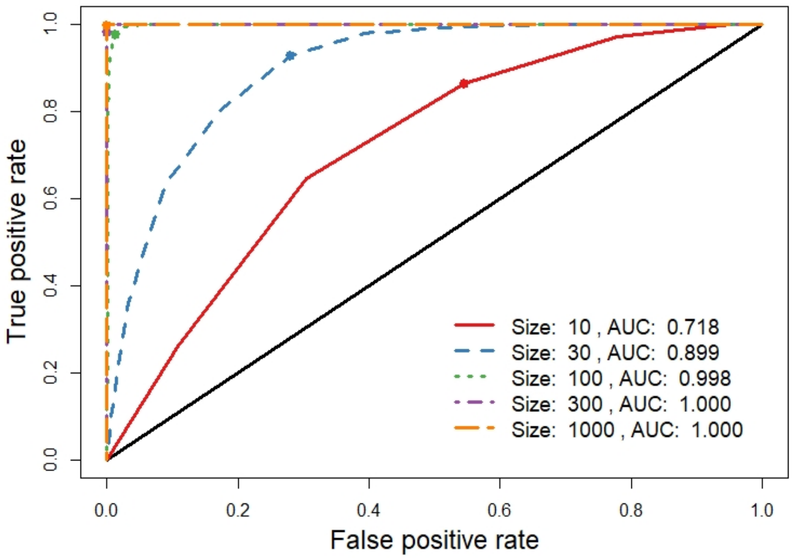
# How to find a candidate negative control variable?

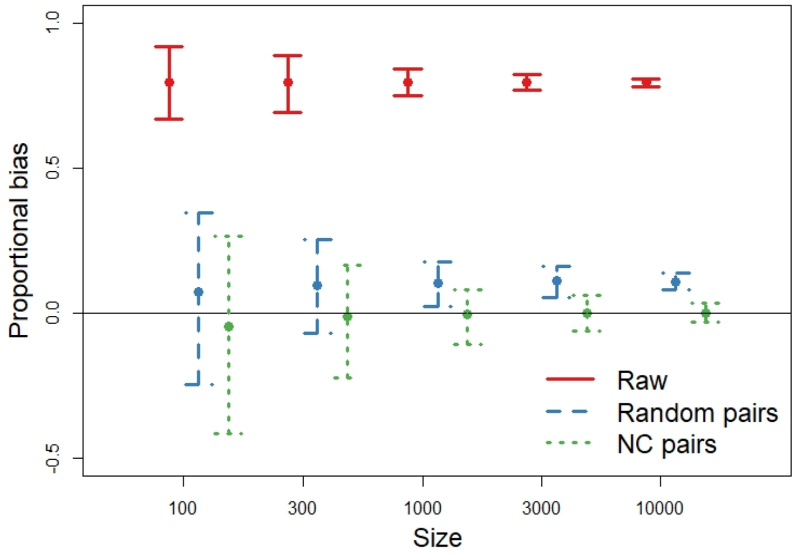



- Rationale: all paths from $\{W, Z\}$ to $\{Y, A\}$ pass through $U$
    - Therefore $\Sigma_{\{W,Z\},\{Y,A\}} = \begin{bmatrix} cov(W,Y) & cov(W,A) \\ cov(Z,Y) & cov(Z,A) \end{bmatrix}$ is rank deficient
    - Such a rank constraint can be determined using statistical tests

Kummerfeld, Lim, and Shi, [2022]

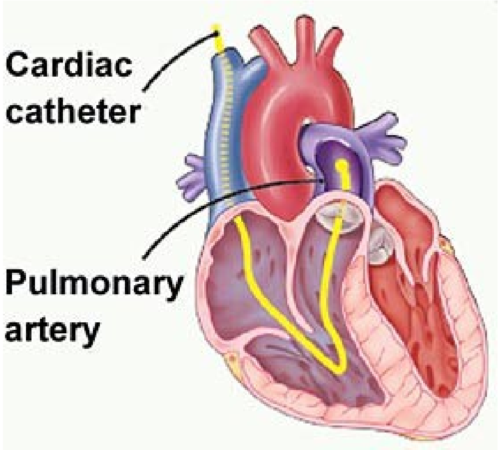
# Performance of the DANCE algorithm




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

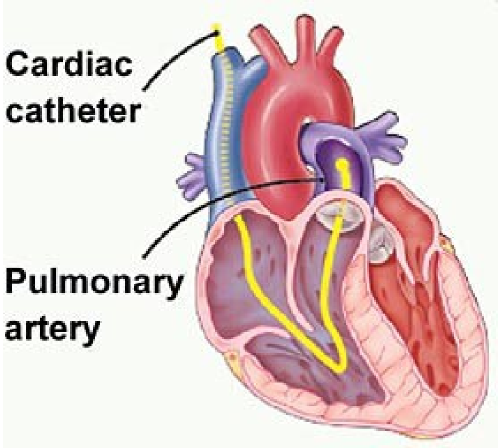
# Application to the SUPPORT study




- 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

Connors et al., 1996

# 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)
  - Evaluate the effectiveness of RHC among adults admitted to the intensive care unit (ICU)
  - 2184 patients managed with RHC, 3551 without RHC

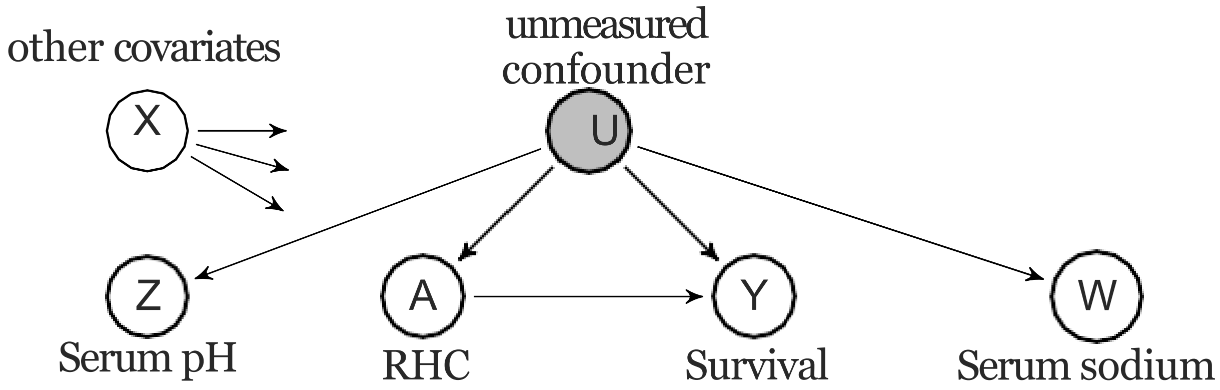
Connors et al., 1996

# A controversial result in the literature

- 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
  - Majority relying on the no unmeasured confounding assumption

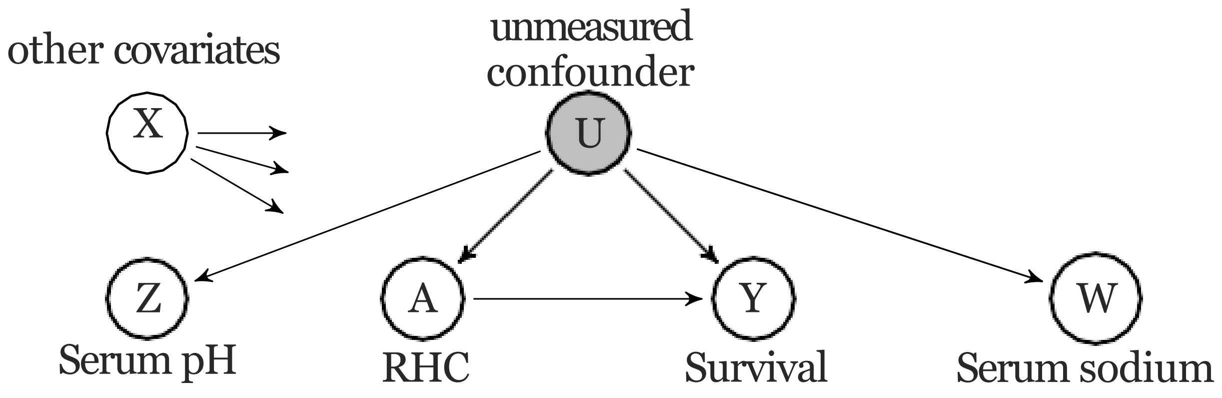
Li, Morgan, and Zaslavsky, 2018; Lin, Psaty, and Kronmal, 1998; Mao and Li, 2020; Tan, 2006

# Candidate proxies in the SUPPORT study



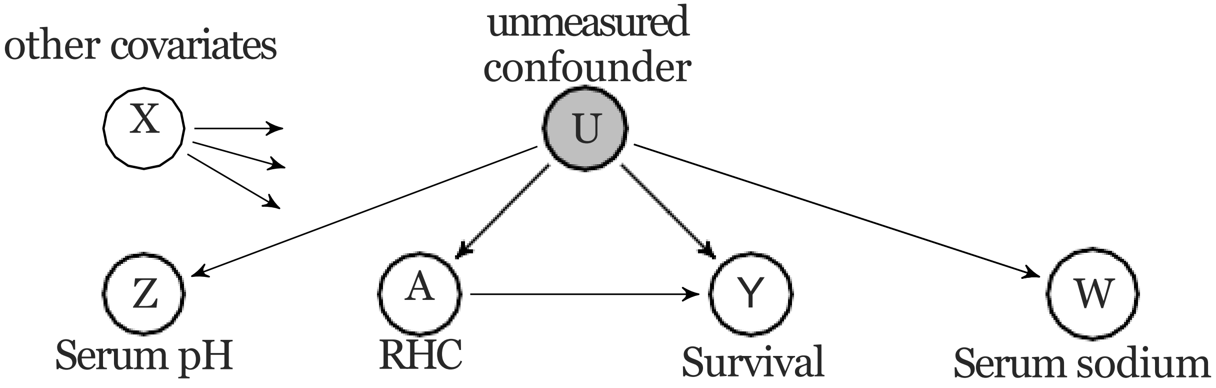

- The SUPPORT study collected 72 covariates including
  - demographics, comorbidity, vital signs, functional status
  - physiological status measured from a blood test during the initial 24 hours in the ICU ⇒ 10 candidate proxies
- We applied our DANCE algorithm to find valid proxies
  - Most frequently 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

# FDA PDUFA (Prescription Drug User Fee Act) VII commitment

| FY 23 | FY 24 | FY 25 | FY 26 | FY 27 |
|---|---|---|---|---|
| Oct 1, 2022 Sep 30, 2023 | Oct 1, 2023 Sep 30, 2024 | Oct 1, 2024 Sep 30, 2025 | Oct 1, 2025 Sep 30, 2026 | Oct 1, 2026 Sep 30, 2027 |
| By September 30, 2023, FDA will hold a public workshop on post-market safety studies in pregnant women | By September 30, 2024, FDA will publish a pregnancy workshop report describing the proposed framework | By September 30, 2025, FDA will publish on its website an update on facilitation of public and sponsor access to Sentinel's distributed data network | | By September 30, 2027, FDA will publish a report on the results of the negative control and pregnancy development projects |
| By September 30, 2023, FDA will hold a public workshop on the use of negative controls | By September 30, 2024, FDA will initiate 5 pregnancy demonstration projects to assess performance of:<br>1. Pregnancy registries vs. electronic healthcare databases studies for signal detection (relatively common exposure to medication in pregnancy)<br>2. Single arm safety studies vs. signal detection methods using electronic healthcare data (anticipated low exposure to medication in pregnancy)<br>3. Pregnancy registries vs. electronic healthcare database studies for signal evaluation (relatively common exposure to medication in pregnancy)<br>4. MCM as a composite outcome in signal detection and evaluation when true risk for some but not all specific malformations<br>5. EHR and claims-linked healthcare data algorithm for a pregnancy-related outcome, or composite of outcomes, after vaccine use in pregnant patients. | By September 30, 2025, FDA will analyze, and report on the use of Sentinel for regulatory purposes (e.g., labeling changes, PMRs, PMCs) | | |
| For FY23-27, FDA will report its obligations for updated PDUFA VI commitments in PDUFA Financial Report with detail for spending categories (e.g., data infrastructure, analytical capabilities, safety issue analyses, etc.)<br>*Recurring annually* | By September 30, 2024, FDA will initiate methods projects:<br>1. Sentinel tools: negative control automation<br>2. Double negative control adjustment | | | |

- Evaluate the use of the DANCE algorithm in Sentinel settings
  - Plasmode simulation study & empirical study with a safety endpoint
  - Convert the tool into the Sentinel system (led by SOC)

# Bibliography (1)


- Chen, Jacob, Rohit Bhattacharya, and Katherine Keith (2024). “Proximal causal inference with text data”. In: *Advances in Neural Information Processing Systems* 37, pp. 135983–136017.
- Connors, Alfred F, Theodore Speroff, Neal V Dawson, Charles Thomas, Frank E Harrell, Douglas Wagner, Norman Desbiens, Lee Goldman, Albert W Wu, Robert M Califf, et al. (1996). “The effectiveness of right heart catheterization in the initial care of critically III patients”. In: *Jama* 276.11, pp. 889–897.
- Cui, Yifan, Hongming Pu, Xu Shi, Wang Miao, and Eric Tchetgen Tchetgen (2020). “Semiparametric proximal causal inference”. In: *arXiv preprint arXiv:2011.08411.*
- Davey Smith, George (2008). “Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings?” In: *Basic & Clinical Pharmacology & Toxicology* 102.2, pp. 245–256.
- Lipsitch, Marc et al. (2010) “Negative controls: a tool for detecting confounding and bias in observational studies.” Epidemiology (Cambridge, Mass.) vol. 21,3 (2010): 383-8. doi:10.1097/EDE.0b013e3181d61eeb
- Dukes, Oliver, Ilya Shpitser, and Eric J Tchetgen Tchetgen (2021). “Proximal mediation analysis”. In: *arXiv preprint arXiv:2109.11904.*
- Jackson, Lisa A, Michael L Jackson, Jennifer C Nelson, Kathleen M Neuzil, and Noel S Weiss (2006). “Evidence of bias in estimates of influenza vaccine effectiveness in seniors”. In: *International journal of epidemiology* 35.2, pp. 337–344.

# Bibliography (2)


Kummerfeld, Erich, Jaewon Lim, and Xu Shi (2022). "Data-driven Automated Negative Control Estimation (DANCE): Automated Search for and Validation of Negative Controls". In: *arXiv preprint arXiv:2210.00528.*

Kummerfeld, Erich, Jaewon Lim, and Xu Shi (2024) "Data-driven Automated Negative Control Estimation (DANCE): search for, validation of, and causal inference with negative controls". In: *Journal of Machine Learning Research* 25.229, pp. 1–35.

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

Document

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

Lipsitch, Marc, Eric Tchetgen Tchetgen, and Ted Cohen (2010). "Negative controls: a tool for detecting confounding and bias in observational studies". In: *Epidemiology (Cambridge, Mass.)* 21.3, p. 383.

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

Miao, Wang, Zhi Geng, and Eric J Tchetgen Tchetgen (2018). "Identifying causal effects with proxy variables of an unmeasured confounder". In: *Biometrika* 105.4, pp. 987–993.

# Bibliography (3)


Miao, Wang, Xu Shi, and Eric Tchetgen Tchetgen (2018). "A confounding bridge approach for double negative control inference on causal effects". In: *arXiv e-prints*, arXiv–1808.

Qi, Zhengling, Rui Miao, and Xiaoke Zhang (2024). "Proximal learning for individualized treatment regimes under unmeasured confounding". In: *Journal of the American Statistical Association* 119.546, pp. 915–928.

Shen, Tao and Yifan Cui (2023). "Optimal treatment regimes for proximal causal learning". In: *Advances in Neural Information Processing Systems* 36, pp. 47735–47748.

Shi, Xu, Wang Miao, Mengtong Hu, and Eric Tchetgen Tchetgen (2021). "Theory for identification and Inference with Synthetic Controls: A Proximal Causal Inference Framework". In: *arXiv preprint arXiv:2108.13935*.

Shi, Xu, Wang Miao, Jennifer C Nelson, and Eric J Tchetgen Tchetgen (2020). "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, pp. 521–540.

Shi, Xu, Wang Miao, and Eric Tchetgen Tchetgen (2020). "A selective review of negative control methods in epidemiology". In: *Current epidemiology reports*, pp. 1–13.

Sverdrup, Erik and Yifan Cui (2023). "Proximal causal learning of conditional average treatment effects". In: *International Conference on Machine Learning*. PMLR, pp. 33285–33298.

# Bibliography (4)


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

Tchetgen, Eric Tchetgen, Chan Park, and David Richardson (2023). "Single proxy control". In: *arXiv preprint arXiv:2302.06054*.

Tchetgen Tchetgen, Eric J, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao (2020). "An Introduction to Proximal Causal Learning". In: *arXiv preprint arXiv:2009.10982*.

Ying, Andrew, Wang Miao, Xu Shi, and Eric J Tchetgen Tchetgen (2021). "Proximal Causal Inference for Complex Longitudinal Studies". In: *arXiv preprint arXiv:2109.07030*.

# Thank You