Understanding the Use of Negative Controls to Assess the Validity of Non-Interventional Studies of Treatment Using Real-World Evidence

Virtual Public Workshop
March 8th, 2023 | 10:00 a.m. – 3:00 p.m. ET

FDA U.S. FOOD & DRUG ADMINISTRATION | Duke MARGOLIS CENTER for Health Policy

Welcome & Introduction

Rachele Hendricks-Sturrup, DHSc, MSc, MA

Research Director of Real World Evidence, Duke-Margolis Center for Health Policy

Agenda

Session	Title
1	Introduction to Negative Controls
2	Overview of Analytic Techniques: A Review of Negative Controls
Break 12:15 a.m. - 12:45 p.m.	
3	Utilizing Negative Controls in Safety and Effectiveness: Methods development and key considerations
4	Key Stakeholder Perspectives

Statement of Independence

The Robert J. Margolis, MD, Center for Health Policy is part of Duke University, and as such it honors the tradition of academic independence on the part of its faculty and scholars. Neither Duke nor the Margolis Center take partisan positions, but the individual members are free to speak their minds and express their opinions regarding important issues.

For more details on relevant institutional policies, please refer to the Duke Faculty Handbook, including the Code of Conduct and other policies and procedures. In addition, regarding positions on legislation and advocacy, Duke University policies are available at http://publicaffairs.duke.edu/government.

Virtual Meeting Reminders

- Attendees are encouraged to contribute throughout the meeting with questions in the Zoom Q&A function.
- This meeting is being recorded, and the recording and slide deck will be posted on the Duke-Margolis event page in the weeks following the meeting.

Introduction to Negative Controls

Negative Controls in Drug Safety and Effectiveness Studies

Background, Challenges, and Considerations

Fang Tian, PhD, MPH, MHS

Division of Epidemiology-I, Office of Surveillance and Epidemiology,
Center for Drug Evaluation and Research, FDA
March 8, 2023

Disclosure and Disclaimer

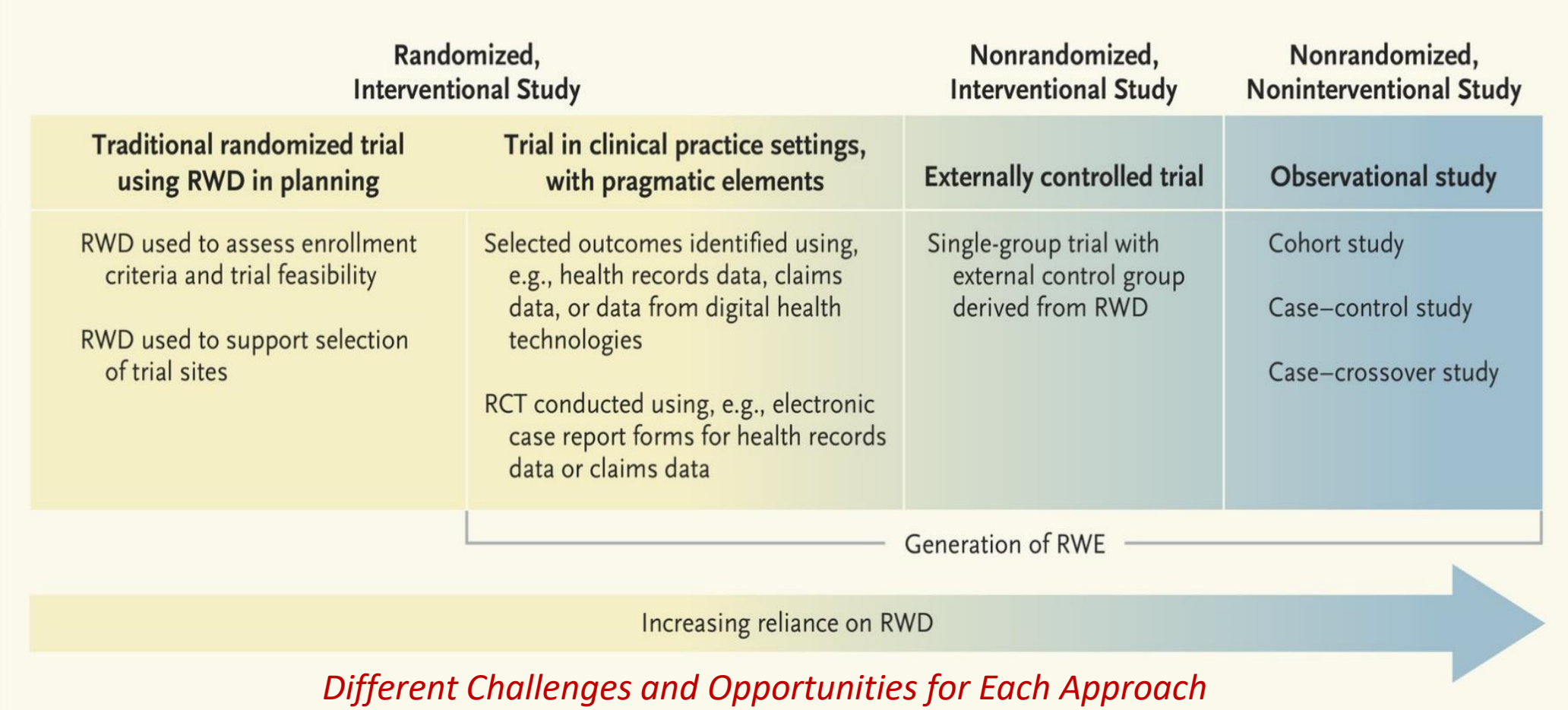
- The views and opinions expressed in the following slides are those of the individual presenter and should not be construed to represent FDA's views or policies.

- No conflicts of interest to disclose

Outline

- Background: Real-World Data and Real-World Evidence
- Negative Controls
  - Basic Concept
  - Why Negative Controls?
- Challenges and Considerations
- CDER Experience and Perspectives
- PDUFA VII Commitment on Negative Controls (Sentinel Initiative)

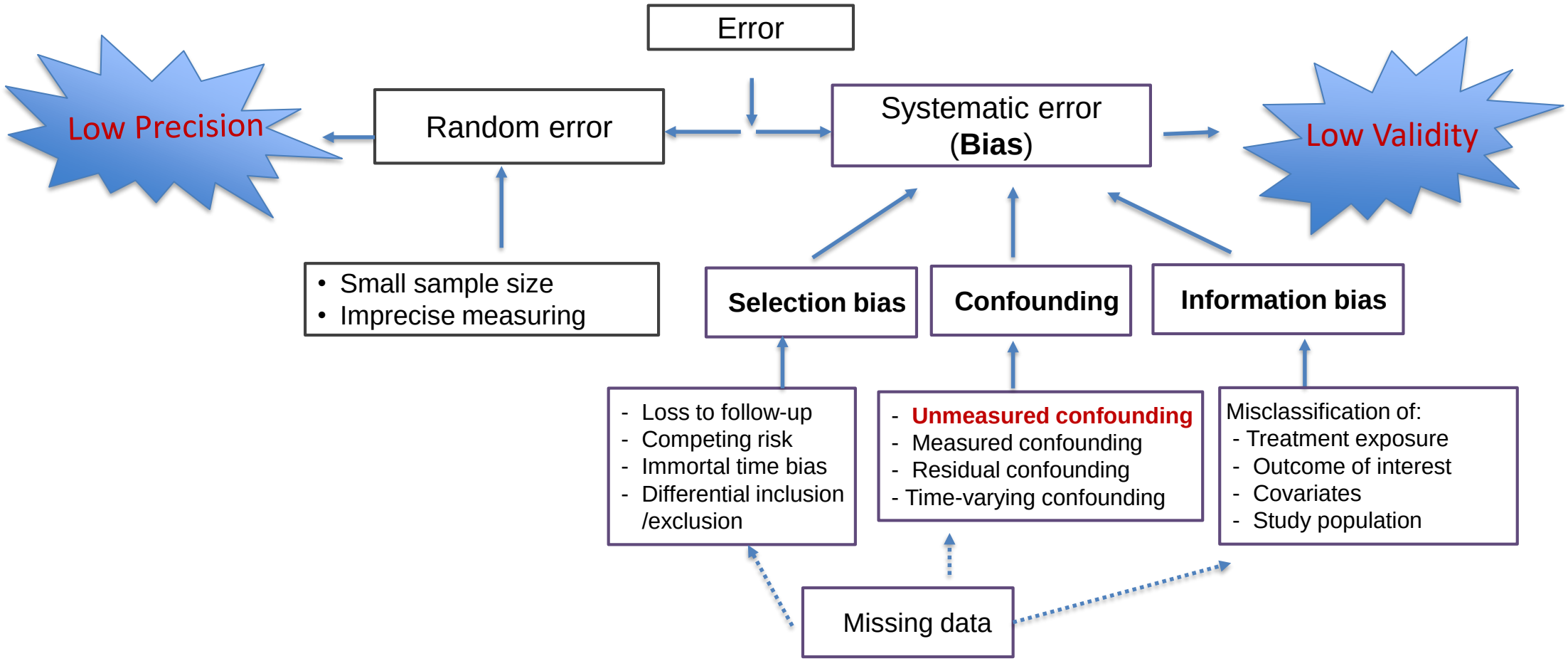
Real-World Evidence – Where Are We Now?



Different Challenges and Opportunities for Each Approach

Reference: Concato J and Corrigan-Curay J. Real-World Evidence – Where Are We Now? NEJM 2022 May 5;386(18):1680-1682.

Threats to the Validity of Non-Randomized Studies Using RWD



Addressing Unmeasured Confounding in Observational Studies

Design Phase

- Case only designs:
  - Case-crossover
  - Case-time control
  - Self-controlled case series
- Instrumental variable
- Difference-in-difference
- Trend-in-trend
- Negative controls

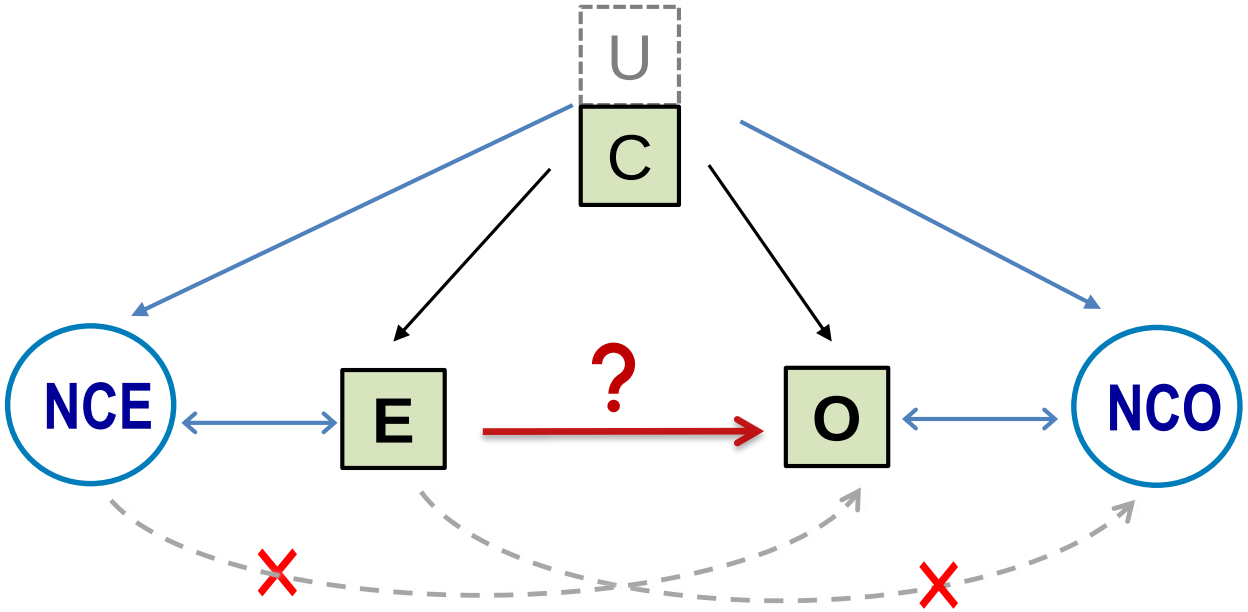
Analysis Phase

- Quantitative bias analysis
  - (e.g., E-value)
- Sensitivity analysis
  - (e.g., stratified analysis)
- Regression discontinuity
- Propensity score calibration
- Prior event rate ratio (PERR)
- Others:
  - Bayesian twin regression
  - Perturbation variable
  - Missing cause approach, etc.

Negative Controls: the Basic Concept

- **Negative Control Exposure (NCE):**
  - A variable that shares the same potential source of bias with the exposure of interest but is not causally related to the outcome of interest.
    - ✓ If an effect is observed between the NCE and outcome of interest, this may suggest that unmeasured confounding or an unmeasurable source of bias is influencing the results
- **Negative Control Outcome (NCO):**
  - A variable that shares the same potential source of bias with the outcome of interest but is not causally related to the exposure of interest.
    - ✓ If an effect is observed between the exposure and NCO, it may indicate that unmeasured confounding or an unmeasurable source of bias is influencing the results.

Causal DAG for Negative Controls: Addressing Unmeasured Confounding



E: Exposure

O: Outcome

C: Measured confounder

U: Unmeasured confounder

NCE: Negative control exposure

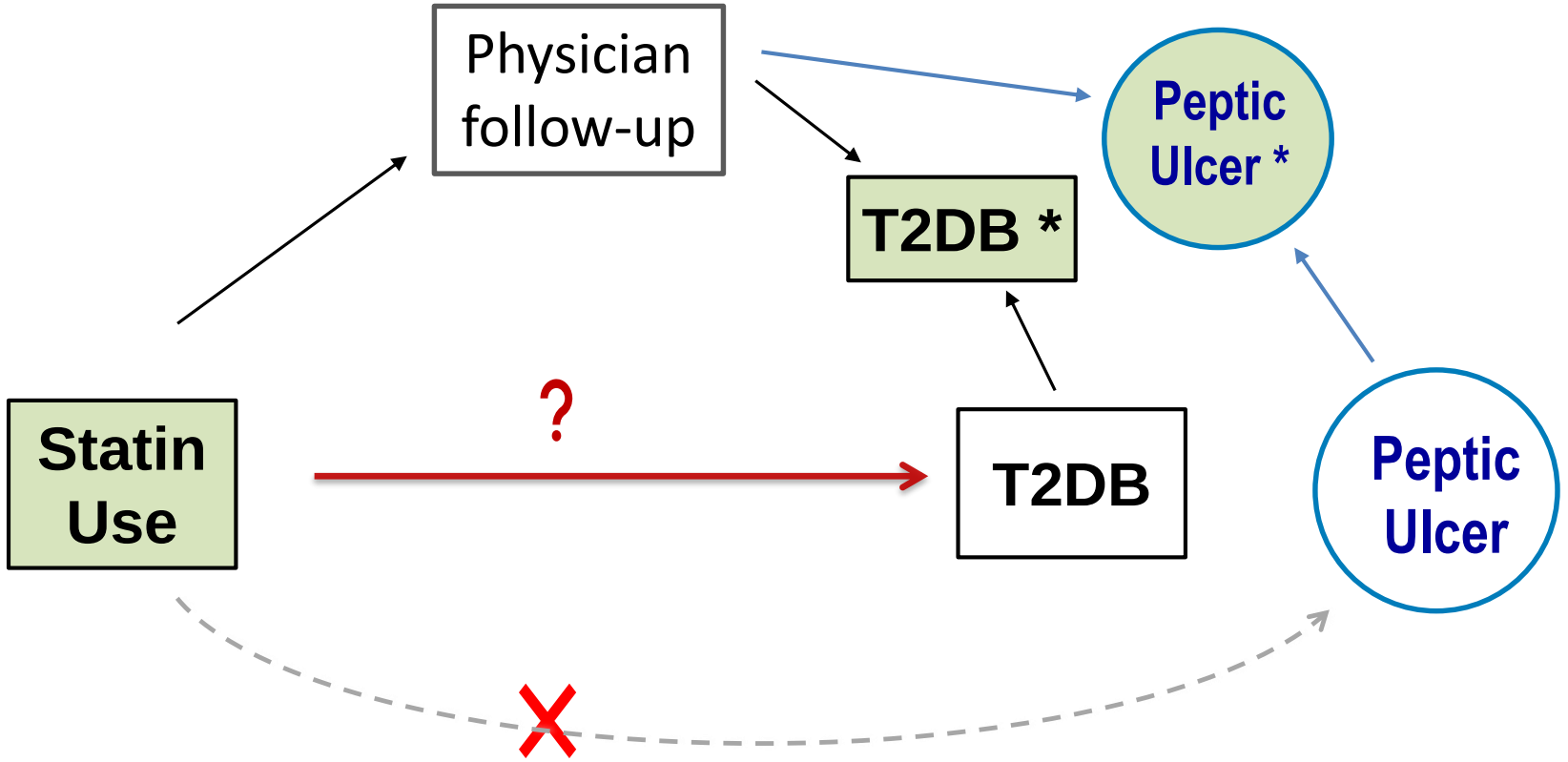
NCO: Negative control outcome

Key Assumptions:

- No causal association between NCE--O (or E--NCO).
- Ideally sharing the same causal structure as E-O association (U-comparable).
  - NCE should be an exposure in which the distribution of U in those exposed to NCE is comparable to the distribution in those exposed to E
  - NCO should be an outcome such that the set of common causes of E and O should be as identical as possible to the set of common causes of E and NCO

Adapted from: Lipsitch M, Tchetgen ET, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21(3):383.

Example of Other Utility of Negative Controls: Addressing Outcome Misclassification



Danaei G, García Rodríguez LA, Fernandez Cantero O, Hernán MA. Statins and risk of diabetes: an analysis of electronic medical records to evaluate possible bias due to differential survival. Diabetes care. 2013 May 1;36(5):1236-40.

Why Negative Controls?

- A method with broad utility
  - Aims to address unmeasured confounding, as well as information bias and selection bias
  - Can be used with various study designs: cohort study, case-control study, self-controlled case series study, etc.
  - Can be used to detect bias, reduce bias, correct bias, calibrate confidence interval or p-value
- Potential tool to routinely check for evidence of bias with adjustment to ensure high quality of RWE

Negative Controls: Challenges and Considerations

- Often requires subject matter knowledge
- Rely on strong assumptions
  - No causal association
  - Share common bias structure
- Need unique considerations in different study design scenarios
  - Type of study design (cohort study vs case-control study)
  - Selection of study population (e.g., healthy vs sick patients, Inpatient vs. outpatient setting)
  - Acute vs. chronic exposure (e.g., one-time vs long-term exposure)
  - Selection of comparator (e.g., active comparator vs non-use)
  - Outcomes of interest (acute vs long-term outcomes)
  - Objective ways of identifying negative controls for use in safety studies : automatic tool
- Interpret results with caution
  - What does a non-null association mean?

CDER Experience and Perspectives

- Regulatory submissions and internal literature review that involved negative controls
  - Mainly for detecting unmeasured confounding
  - Unverified assumptions:
    - ✓ Non-causality
    - ✓ Sharing common confounding structure
  - Often lack of sufficient interpretation and justification
- CDER scientific projects
  - Methods study to evaluate confounding control conducted in Sentinel
  - CERSI project in collaboration with University of Maryland

CDER Experience - Sentinel Study

American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2022.
This work is written by (a) US Government employee(s) and is in the public domain in the US.

Vol. 191, No. 5
https://doi.org/10.1093/aje/kwac020
Advance Access publication:
February 2, 2022

Practice of Epidemiology

Evaluating Confounding Control in Estimations of Influenza Antiviral Effectiveness in Electronic Health Plan Data

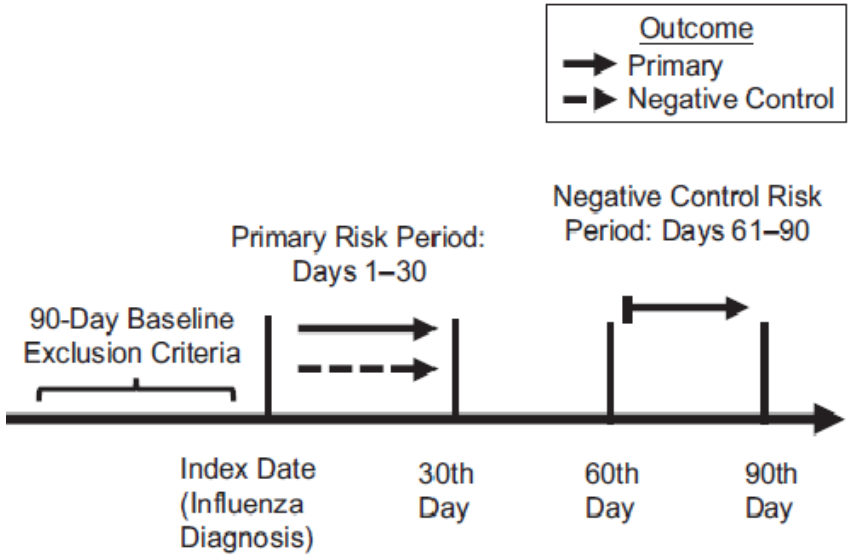
Phyo T. Htoo, Gregory Measer, Robert Orr, Justin Bohn, Alfred Sorbello, Henry Francis, Sarah K. Dutcher, Austin Cosgrove, Amanda Carruth, Sengwee Toh, and Noelle M. Cocoros*

*Correspondence to Dr. Noelle M. Cocoros, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 401 Park Drive, Suite 401E, Boston, MA 02215 (e-mail: noelle_cocoros@harvardpilgrim.org).

Initially submitted April 18, 2021; accepted for publication January 28, 2022.

Observational studies of oseltamivir use and influenza complications could suffer from residual confounding. Using negative control risk periods and a negative control outcome, we examined confounding control in a health-insurance-claims–based study of oseltamivir and influenza complications (pneumonia, all-cause hospitalization, and dispensing of an antibiotic). Within the Food and Drug Administration's Sentinel System, we identified individuals aged ≥18 years who initiated oseltamivir use on the influenza diagnosis date versus those who did not, during 3 influenza seasons (2014–2017). We evaluated primary outcomes within the following 1–30 days (the primary risk period) and 61–90 days (the negative control period) and nonvertebral fractures (the negative control outcome) within days 1–30. We estimated propensity-score–matched risk ratios (RRs) per season. During the 2014–2015 influenza season, oseltamivir use was associated with a reduction in the risk of pneumonia (RR = 0.72, 95% confidence interval (CI): 0.70, 0.75) and all-cause hospitalization (RR = 0.54, 95% CI: 0.53, 0.55) in days 1–30. During days 61–90, estimates were near-null for pneumonia (RR = 1.04, 95% CI: 0.95, 1.15) and hospitalization (RR = 0.94, 95% CI: 0.91, 0.98) but slightly increased for antibiotic dispensing (RR = 1.14, 95% CI: 1.08, 1.21). The RR for fractures was near-null (RR = 1.09, 95% CI: 0.99, 1.20). Estimates for the 2016–2017 influenza season were comparable, while the 2015–2016 season had conflicting results. Our study suggests minimal residual confounding for specific outcomes, but results differed by season.

antiviral agents; bias; confounding factors; epidemiologic methods; health-care administrative claims; human influenza; oseltamivir; pneumonia



- **Design:** New-user cohort study
- **Exposure:** Oseltamivir use
- **Outcome:** influenza complications (Pneumonia, all-cause hospitalization, antibiotic use)
- **Risk interval:** 1-30 days following oseltamivir initiation
- **NCO:** nonvertebral fractures
- **Negative control period:** 61-90 days following oseltamivir initiation

CDER Experience - CERSI Project

FDA Search Menu

IN THIS SECTION: Advancing Regulatory Science

← Advancing Regulatory Science

Evaluating the utility of negative controls in drug safety and effectiveness studies using real-world data

Share Tweet Linkedin Email Print

CERSI Collaborators: Zafar Zafari, M.Sc., PhD (Principal Investigator), Susan dosReis, PhD (co-Investigator), Chintal H. Shah, MS, Jeong-eun Park, MPH, Emily Gorman, MLIS, AHIP, University of Maryland Baltimore

FDA Collaborators: Fang Tian, PhD, MPH, MHS, Wei Hua, MD, PhD, Rita Ouellet-Hellstrom, PhD, MPH, Yong Ma, PhD, MS

Project Start Date: June 1, 2020

Regulatory Science Challenge

Randomized controlled trials (RCTs) are the gold standard for evaluating the efficacy and safety of a treatment. However, the high costs of conducting RCTs coupled with the limited applicability of the findings of RCTs to real-world settings have made the use of real-world data in long-term drug safety and effectiveness studies more popular. Nevertheless, a major concern in using real-world data is the presence of unaccounted biases (e.g., selection bias, information bias) and confounding (influences from multiple sources that cannot be separated) in the data, which may threaten the internal validity of the study results.

Project Description and Goals

Negative controls, defined as negative control exposures (or outcomes) that share the same potential source of bias with the primary exposure (or outcome) but are not causally related to the outcome (or exposure) of interest, are useful tools to address issues involving confounding and other unaccounted biases (e.g., selection bias, information bias), with potential for broad application.

In this project, researchers will:

1. Conduct a full literature review for the use of negative controls in epidemiological studies on drug safety and effectiveness, and
2. Evaluate assumptions behind the use of negative controls in these studies.

The main goal of this project is to build a methodological framework for evaluating the use of negative controls in drug safety and effectiveness studies using real-world data.

Content current as of:
06/28/2022

FDA Archive Visitor Information FOIA

PDUFA VII Commitment - (2023-2027)

2. Optimization of the Sentinel Initiative

ii. Use of Real-World Evidence – Negative Controls

FDA is building Sentinel/BEST methodology to improve understanding of robustness evaluations used to address the consistency of RWE with respect to study design, analysis, or variable measurement. FDA will develop new methods to support causal inference in Sentinel/BEST that could address product safety questions and advance our understanding of how RWE may be used for studying effectiveness.

(1) By September 30, 2023, FDA will hold a public workshop on use of negative controls for assessing the validity of non-interventional studies of treatment and the proposed Sentinel Initiative projects.

(2) FDA will initiate two methods development projects by September 30, 2024 to 1) develop an empirical method to automate the negative control identification process in Sentinel and integrate it into the Sentinel System tools; and 2) develop a method to use a double negative control adjustment to reduce unmeasured confounding in studying effectiveness of vaccines.

(3) By September 30, 2027, FDA will publish a report on the results of the development projects.

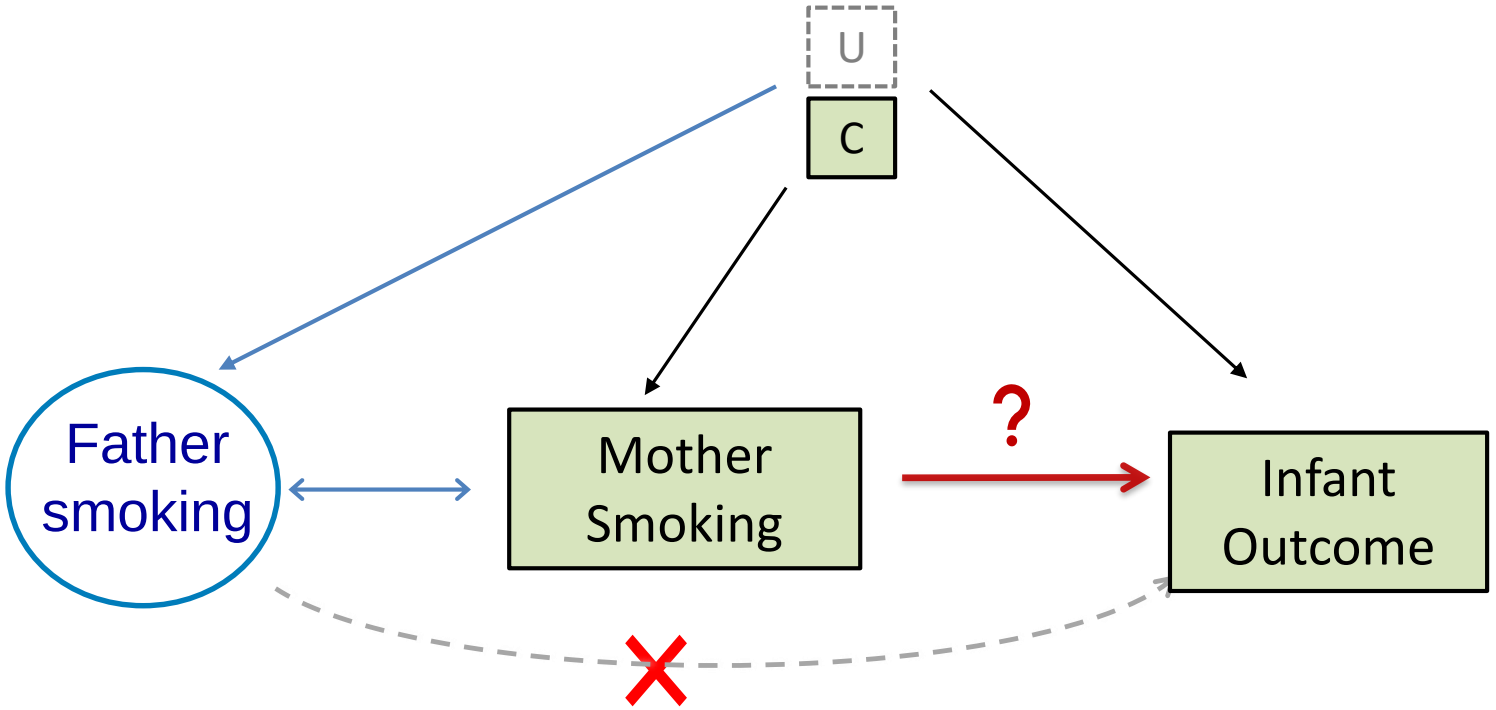
Acknowledgements

- Wei Hua, Division of Epidemiology I, OSE/CDER, FDA
- Yong Ma, Division of Biostatistics VII, OB/CDER, FDA
- Sarah Dutcher, Sentinel Core Team, OSE/CDER, FDA
- Katherine Scott, Sentinel Core Team, OSE/CDER, FDA
- Robert Ball, OSE/CDER, FDA


Thank
you


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Example: Violation of Non-causality Assumption

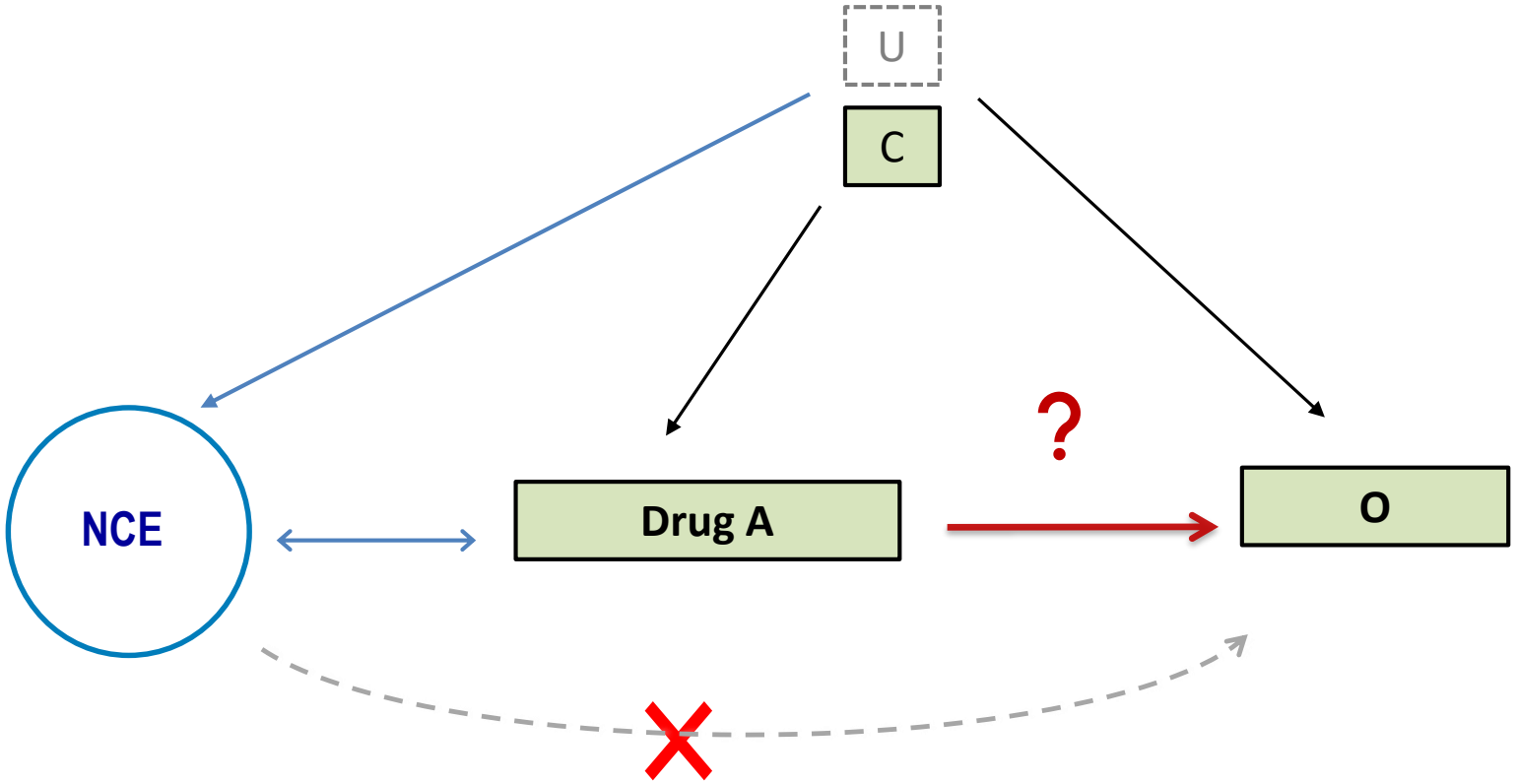


- **Design**: cohort study
- **Exposure**: Mother smoking during 1st trimester
- **Outcome**: Infant birth defect
- **Unmeasured confounders**: SES, environmental factors
- **NCE**: Father smoking during mother's 1st trimester

Negative Control Exposure to Address Unmeasured Confounding

Cohen JM, Wood ME, Hernández-Díaz S, Ystrom E, Nordeng H. Paternal antidepressant use as a negative control for maternal use: assessing familial confounding on gestational length and anxiety traits in offspring. Int J Epidemiol 2019;48:1665–72
Brew BK, Gong T. Modelling paternal exposure as a negative control. Int J Epidemiol 2020;49:1053–4.
Brew BK, Gong T, Williams DM, Larsson H, Almqvist C. Using fathers as a negative control exposure to test the Developmental Origins of Health and Disease Hypothesis: a case study on maternal distress and offspring asthma using Swedish register data. Scand J Public Health 2017;45:36–40.

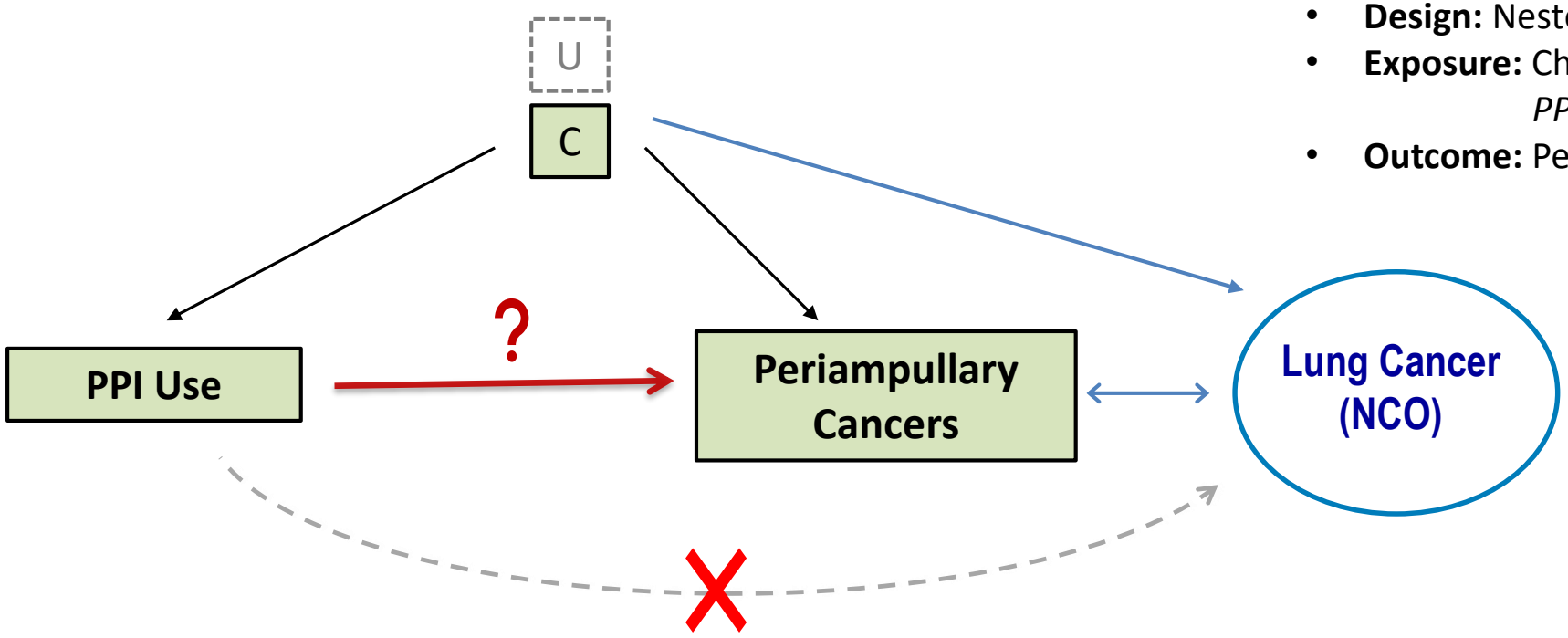
Example: Violation of U-comparable Assumption



- **Design:** cohort study
- **Exposure:** Drug A for indication X
- **Unmeasured confounder:** certain underlying condition that affects physician's prescription behavior
- **NCE:** Drug B for indication X

Negative Control Exposure to Address Unmeasured Confounding

Example: Selection of Negative Control Outcome in Consideration of Study Design Scenarios



- **Design:** Nested case-control study
- **Exposure:** Chronic PPI use (cumulative exposure)
  PPI: proton pump inhibitor
- **Outcome:** Periampulary Cancer (long latency)

Chien, Li-Nien, et al. "Proton pump inhibitors and risk of periampullary cancers- a nested case-control study." International journal of cancer 138.6 (2016): 1401-1409.

CBER experience using negative outcomes to account for the effect of unmeasured variables in vaccine effectiveness studies

Presented by:
Hector S. Izurieta, MD, MPH, PhD
Associate Director for Novel Clinical Investigations
Office of Vaccines Research and Review (OVRR), CBER, FDA

Study team leads:
CBER: Yun Lu, Hector S. Izurieta, Richard Forshee
Acumen LLC: Yoganand Chillarige, Michael Wernicke
CMS: Jeffrey Kelman

Virtual meeting: Exploring the Utility of Negative Controls for Causal Inference in the Sentinel Initiative
Robert J Margolis Center for Health Policy, Duke University, March 8, 2023


My comments are an informal communication and represent my own best judgement. They do not bind or obligate FDA.

Outline:

- Addressing unmeasured confounding: negative outcomes

- Studies in which we used negative outcomes:

  - Herpes zoster effectiveness
  - Influenza comparative effectiveness

- Summary
- Next steps

Addressing unmeasured confounding: negative outcomes

- There are multiple well-studied approaches to address measured confounders
- Assessing the effects of unmeasured confounders (not available in the database) is a challenge
- In these examples, we used multiple methods to account for unmeasured confounding associated with what we considered are important sets of potential confounding factors unmeasured in claims databases: physical fitness to care for self and other frailty measures, level of education, health seeking behavior
- Methods we used:
  - Restriction to populations similar with respect to health seeking (e.g. vaccinated comparator)
    - Linkages to Medicare Current Beneficiary Survey (MCBS)
    - Negative outcomes/falsification outcomes

Received: 8 February 2019 | Revised: 12 April 2019 | Accepted: 17 April 2019
DOI: 10.1002/pds.4801

ORIGINAL REPORT

WILEY

Zostavax vaccine effectiveness among US elderly using real-world evidence: Addressing unmeasured confounders by using multiple imputation after linking beneficiary surveys with Medicare claims

Hector S. Izurieta[1] | Xiyuan Wu[2] | Yun Lu[1] | Yoganand Chillarige[2] | Michael Wernecke[2] | Arnstein Lindaas[2] | Douglas Pratt[1] | Thomas E. MaCurdy[2,4] | Steve Chu[3] | Jeffrey Kelman[3] | Richard Forshee[1]

[1]Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland

[2]Acumen LLC, Burlingame, California

[3]Centers for Medicare and Medicaid Services, Washington, District of Columbia

[4]Department of Economics, Stanford University, Stanford, California

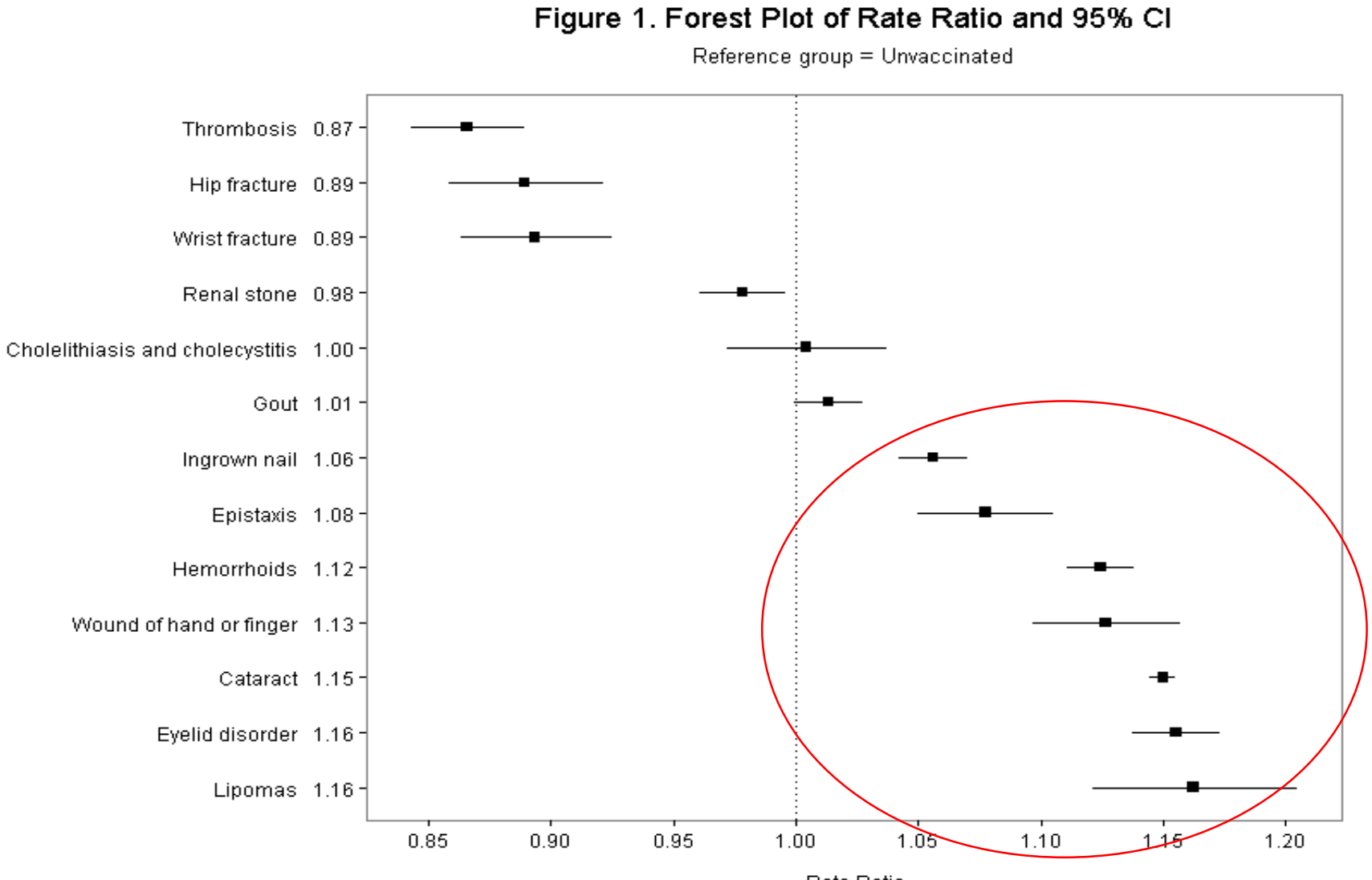
Correspondence
H. S. Izurieta, CBER/FDA, 10903 New Hampshire Avenue, Silver Spring, MD 20993.
Email: hector.izurieta@fda.hhs.gov

Abstract

Purpose: Medicare claims can provide real-world evidence (RWE) to support the Food and Drug Administration's ability to conduct postapproval studies to validate products' safety and effectiveness. However, Medicare claims do not contain comprehensive information on some important sources of bias. Thus, we piloted an approach using the Medicare Current Beneficiary Survey (MCBS), a nationally representative survey of the Medicare population, to (*a*) assess cohort balance with respect to unmeasured confounders in a herpes zoster vaccine (HZV) effectiveness claims-based study and (*b*) augment Medicare claims with MCBS data to include unmeasured covariates.

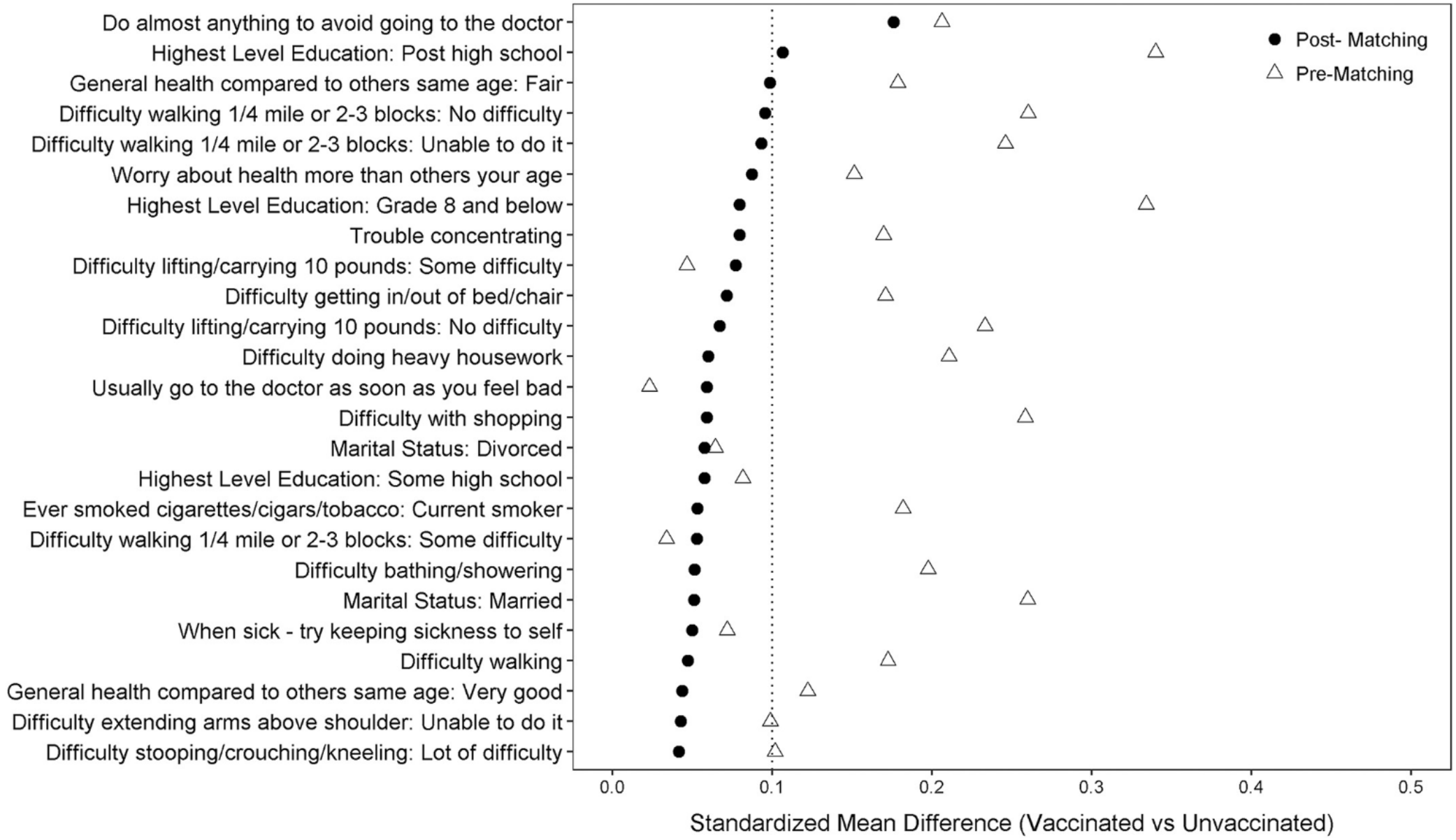
Izurieta, Wu, Lu et al, Pharmacoepidemiol Drug Safety 2019

Compared rates of negative endpoints between the vaccinated and unvaccinated cohort: potential health seeking behavior outcomes



Source: Izurieta, Wu, Lu et al, Pharmacoepidemiol Drug Safety 2019 (source for list of 13 endpoints: Tseng et al, JAMA 2011)

Used a linked outside source: Medicare Current Beneficiary Survey



Izurieta, Wu, Lu et al, Pharmacoepidemiol Drug Saf 2019

Clinical Infectious Diseases

MAJOR ARTICLE

Comparative Effectiveness of Influenza Vaccines Among US Medicare Beneficiaries Ages 65 Years and Older During the 2019–2020 Season

Hector S. Izurieta,[1] Michael Lu,[2] Jeffrey Kelman,[3] Yun Lu,[1] Arnstein Lindaas,[2] Julie Loc,[2] Douglas Pratt,[1] Yuqin Wei,[2] Yoganand Chillarige,[2] Michael Wernecke,[2] Thomas E. MaCurdy,[2,4] and Richard Forshee[1]

[1]Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland, USA, [2]Acumen LLC, Burlingame, California, USA, [3]Centers for Medicare and Medicaid Services, Washington, DC, USA, and [4]Stanford University Department of Economics, Stanford, California, USA

Source: Izurieta, Lu, Kelman et al, CID 2020

Compared effectiveness of all 5 influenza vaccines used in the elderly

Included data from >12 million influenza vaccinated Medicare beneficiaries such as the one shown here…

For cohort balance:
Inverse probability
of treatment weighting
(IPTW)

Used health seeking behavior negative endpoints to assess unmeasured confounding

Table 1. Distribution of Main Covariates Across Vaccine Cohorts for the 2019–20 Influenza Season After Implementing IPTW Weights: 5-Vaccine Comparison

	Cohort						
Covariate	HD-IIV3	aIIV3	IIV4	cIIV4	RIV4	Pre-IPTW Max SMD	Post-IPTW Max SMD
Base population, n	7 173 433	2 565 513	1 584 451	824 264	608 433		
Vaccinated at pharmacy							
Yes	45.6%	45.6%	46.2%	46.7%	42.8%	1.06	0.08
Age brackets							
65–74 years	50.1%	50.2%	50.4%	51.1%	50.5%	0.09	0.02
75–84 years	35.4%	35.2%	34.8%	34.5%	34.7%	0.04	0.02
≥85 years	14.6%	14.6%	14.8%	14.5%	14.7%	0.10	0.01
Gender							
Female	58.2%	58.3%	58.9%	58.9%	58.6%	0.03	0.01
Male	41.8%	41.7%	41.1%	41.1%	41.4%	0.03	0.01
Race							
White	88.6%	88.6%	89.2%	88.9%	88.8%	0.17	0.02
Black	4.3%	4.2%	4.0%	4.1%	4.4%	0.11	0.02
Asian	2.0%	1.9%	1.9%	1.9%	1.9%	0.14	0.01
Hispanic	0.8%	0.8%	0.8%	0.7%	0.8%	0.08	0.01
Other	4.3%	4.5%	4.1%	4.4%	4.2%	0.06	0.02
Dual eligible							
Yes	6.0%	5.8%	5.7%	5.6%	6.1%	0.24	0.02
All hospitalizations and observational stay							
At least 1	8.7%	8.6%	8.3%	8.7%	8.9%	0.12	0.02
Outpatient ER visits							
At least 1	13.8%	13.8%	13.6%	14.0%	14.2%	0.08	0.02

Used standardized mean differences (SMD) to confirm
that cohorts were well balanced for all measured covariates following IPTW

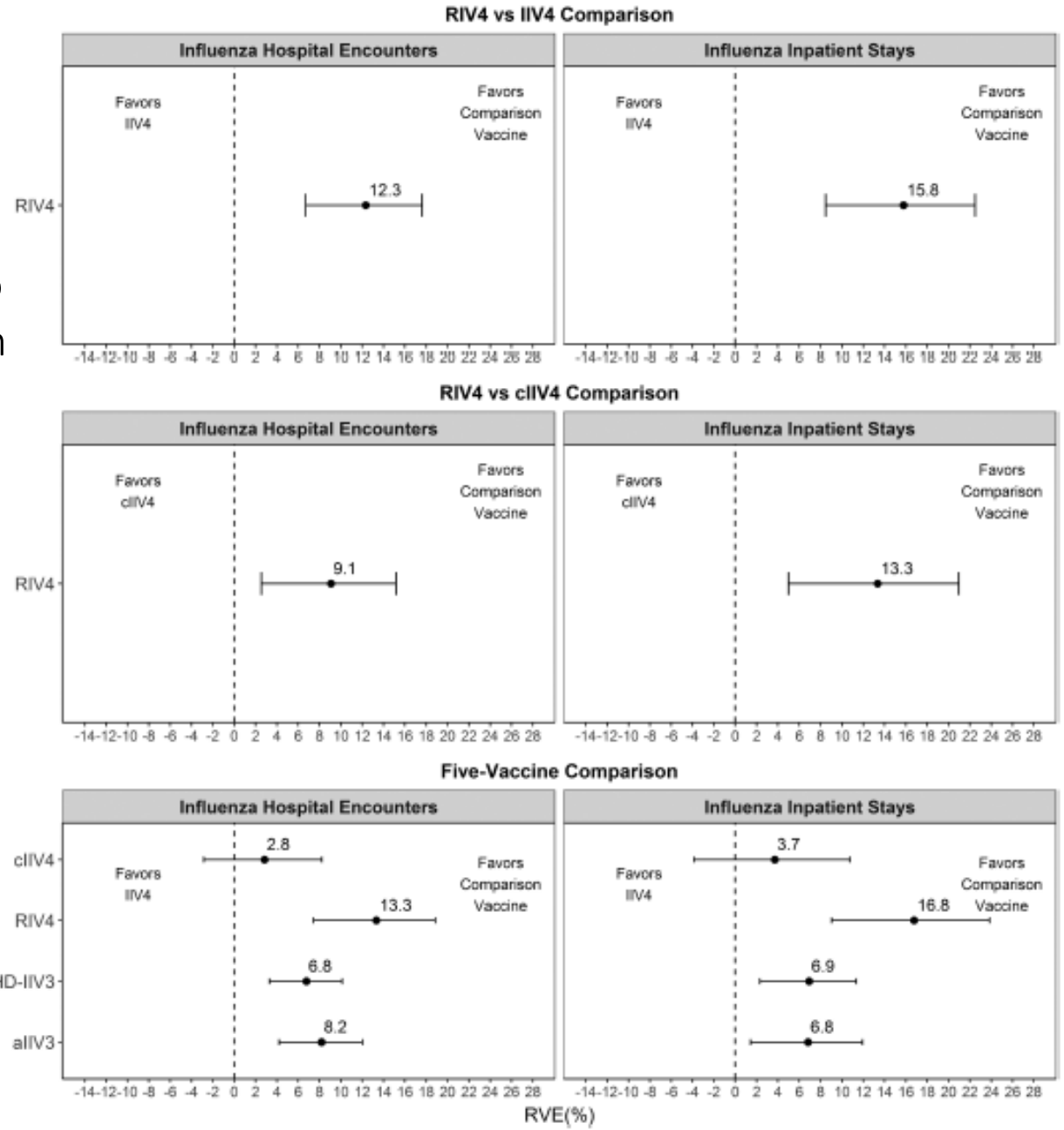
Table 4. Distribution of Health-Seeking Behavior Indicators across 5 Vaccine Cohorts for the 2019–20 Influenza Season After Implementing IPTW Weights

	Cohort						
Covariate	HD-IIV3	aIIV3	IIV4	cIIV4	RIV4	Pre-IPTW Max SMD	Post-IPTW Max SMD
Base population, n	7 173 433	2 565 513	1 584 451	824 264	608 433		
Health-seeking behavior indicators							
Cataracts	8.8%	8.8%	8.9%	8.7%	8.8%	0.04	0.01
Eyelid disorders	0.8%	0.8%	0.8%	0.8%	0.8%	0.02	0.01
Hemorrhoids	0.4%	0.4%	0.4%	0.4%	0.4%	0.01	0.00
Ingrown nail	1.1%	1.1%	1.1%	1.1%	1.0%	0.03	0.01
Lipomas	0.2%	0.2%	0.2%	0.2%	0.2%	0.01	0.00
UTI	4.4%	4.4%	4.4%	4.5%	4.5%	0.04	0.01
Wound of hand or finger	0.3%	0.2%	0.3%	0.3%	0.2%	0.00	0.00

"5-Vaccine" comparison refers to the comparison between the cIIV4, RIV4, HD-IIV3, aIIV3, and IIV4 cohorts, with IPTW weights truncated at 10. Medical office visit and OP non-ER claims with first-line diagnoses for the 7 health-seeking behavior indicators in the 183 days prior to vaccination date were included in the table. Abbreviations: aIIV3, egg-based adjuvanted trivalent; cIIV4, cell-cultured standard-dose quadrivalent; HD-IIV3, egg-based high-dose trivalent; IIV4, egg-based standard-dose quadrivalent; IPTW, inverse probability of treatment weighted; Max, maximum; RIV4, recombinant HA-only quadrivalent; SMD, standardized mean difference; UTI, urinary tract infection.

SMDs for all "health seeking behavior" outcomes (negative endpoints) we used to assess unmeasured confounding were, reassuringly, very low, all values were <0.10 even prior to IPTW

In June 23, 2022,* CDC's ACIP cited results from this (and another study by our team**) among data sources used to justify a preferential influenza vaccination recommendation for the elderly for the first time



*https://www.cdc.gov/flu/spotlights/2021-2022/specific-vaccines-seniors.htm

**Izurieta, Thadani, Kelman et al, Lancet Infect Dis 2015

Summary

- In observational studies, measured confounding can often be successfully addressed in the design (e.g. matching, Mahalanobis distance, IPTW, adjustments)
- Comparing the effectiveness of different vaccines with the same indication (provided there is equipoise) may help decrease the risk of measured and unmeasured confounding
- Identifying unmeasured confounding is challenging, negative outcomes are potentially useful (multiple methods to select negative outcomes are available)
- Linkages with surveys or other databases that provide data on unmeasured confounders should continue to be explored to assess multiple important confounders unmeasured in claims datasets, including but not limited to:
  - Physical fitness to care for self/other frailty measures
  - Level of education
  - Attitudes towards medical care/Health seeking behavior

Next steps: Double negative controls

- We plan to use double negative controls, including focusing on health seeking behavior bias, in claims studies of vaccines for influenza and other vaccine preventable diseases
  - Suggested negative health seeking behavior outcomes:
    - Medical encounters for 7 independent negative outcomes suggestive of health seeking behavior we consider as not associated with flu outcomes, selected from our prior studies:
      - Cataracts
      - Eyelid disorders
      - Hemorrhoids
      - Wound of hand or finger
      - Ingrown nail
      - Lipomas
      - Urinary track infection
  - Suggested negative health seeking behavior exposures :
    - Breast cancer screening (females)
    - Prostate cancer screening (males)

Acknowledgements

- Food and Drug Administration (FDA):

  Hector S. Izurieta, Yun Lu, Douglas Pratt, Richard Forshee

- ACUMEN LLC:

  Yoganand Chillarige, Michael Wernecke, Xiyuan Wu, Arnstein Lindaas, Steve Chu, Julie Loc, Yuqin Wei, Tom MaCurdy

- CMS:

  – Jeffrey Kelman

UNIVERSITY of MARYLAND
SCHOOL OF PHARMACY

The state of use and utility of negative controls in pharmacoepidemiologic studies

Zafar Zafari, MSc, PhD (zzafari@rx.umaryland.edu)
Jeong-eun Park, MPH (jpark1@umaryland.edu)
University of Maryland School of Pharmacy

March 8, 2023
Understanding the Use of Negative Controls to Assess the Validity of Non-Interventional Studies of Treatment Using Real-World Evidence – Public workshop

Acknowledgement

- **Funding**

  This work was supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award [Center of Excellence in Regulatory Science and Innovation grant to University of Maryland, U01FD005946] totaling $230, 801 with 100 percent funded by FDA/HHS. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the U.S. Government.

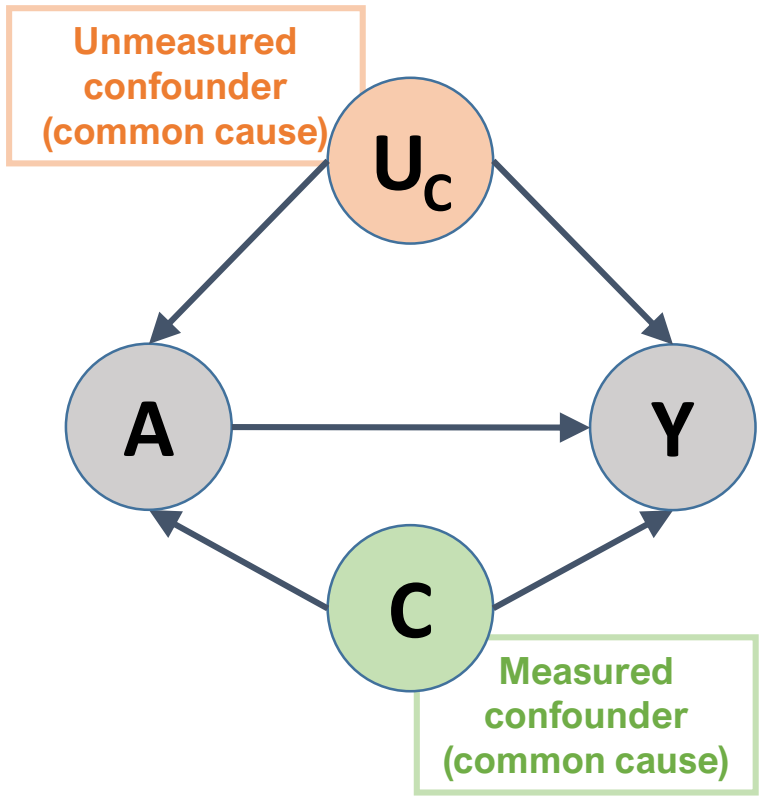
- **Collaborators**:
  - Susan dosReis, UMB
  - Chintal Shah, UMB
  - Emily Gorman, UMB
  - Wei Hua, FDA
  - Yong Ma, FDA
  - Fang Tian, FDA

Background

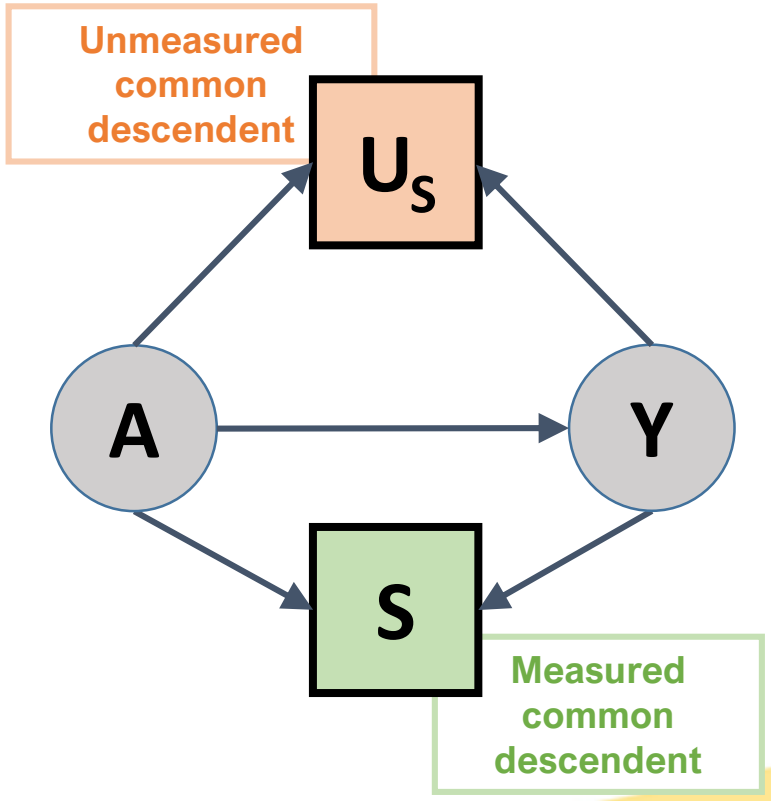
- Pharmacoepidemiologic studies often use real-world data (RWD), which has some advantage over randomized trials.
- RWD studies are challenged by various sources of bias, including confounding, selection, and information bias.
- Use of negative controls to address bias has recently received traction in pharmacoepidemiologic studies.
- Nonetheless, to what extent the use of negative controls methodologies have been applied in pharmacoepidemiologic studies is not well known.

Confounding bias, selection bias, and measurement error

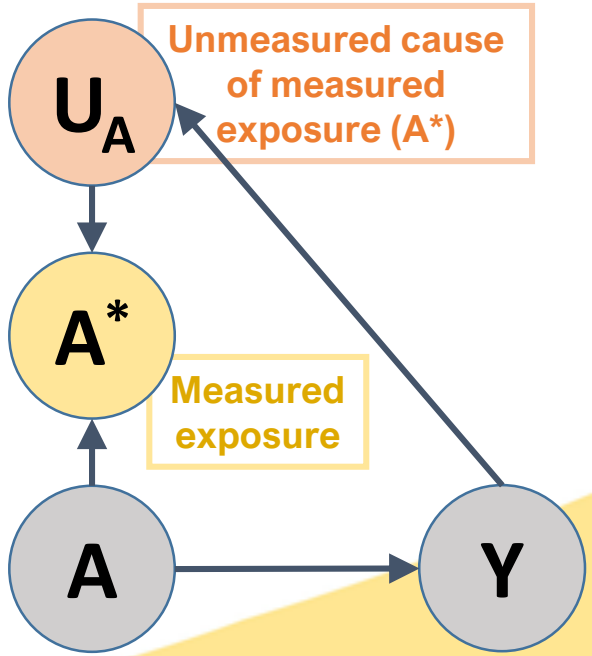
Confounding bias



Selection bias

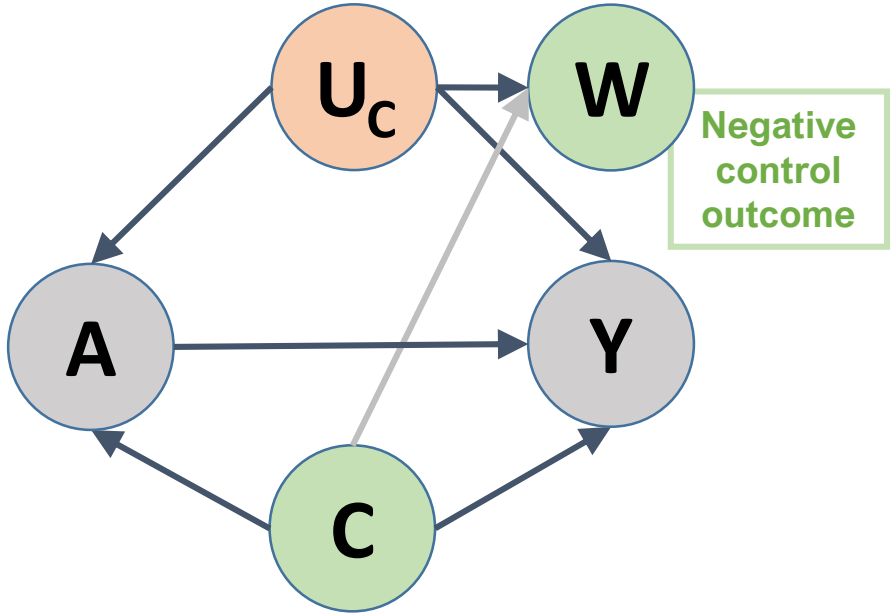


Differential exposure measurement error

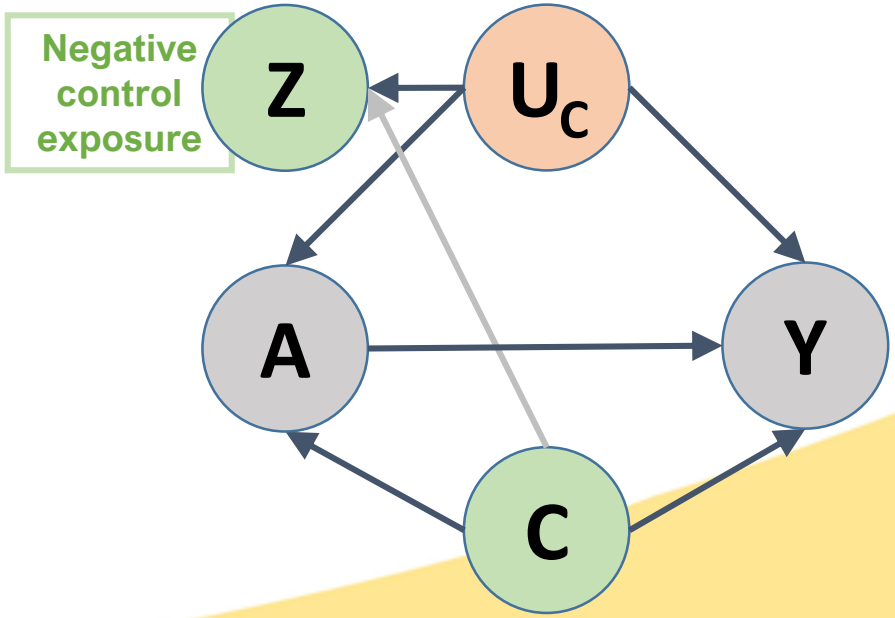


Negative control assumptions

- **Assumption 1**: The exposure does not cause the NCO / The NCE does not cause the primary outcome.
- **Assumption 2**: NC association shares a similar bias structure (U-comparability).



DAG for negative control outcome



DAG for negative control exposure

Objective

- We conducted a scoping review of the state of use, applications, and utility of negative controls in pharmacoepidemiologic studies.

Methods

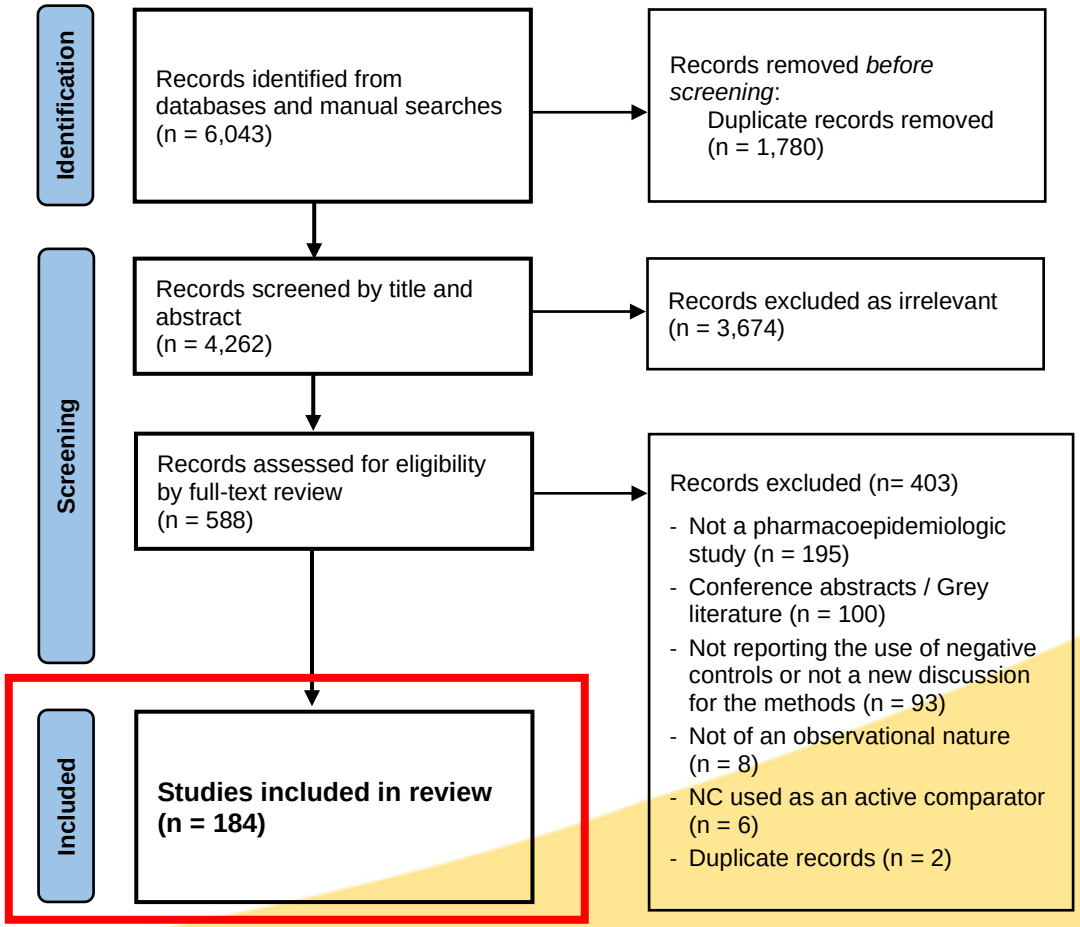
- **Data search**
  - Conducted systematic searches of published literature through Sep 2020
  - **Database searches** using search strategies constructed based on terms related to key concepts (i.e., negative control, bias, epidemiology) on PubMed, EMBASE, CINAHL, Cochrane Library, Scopus, Dissertations and Theses Global
  - **Manual searches** based on key articles identified at the planning phase of the research
- **Study selection**
  - **Selection criteria**

✓ Study **using negative control methods** to address residual bias
✓ Epidemiologic study with an **observational design**
✓ Study related to **utilization/safety/effectiveness of drugs**
✓ Not a conference abstract/grey literature

Results

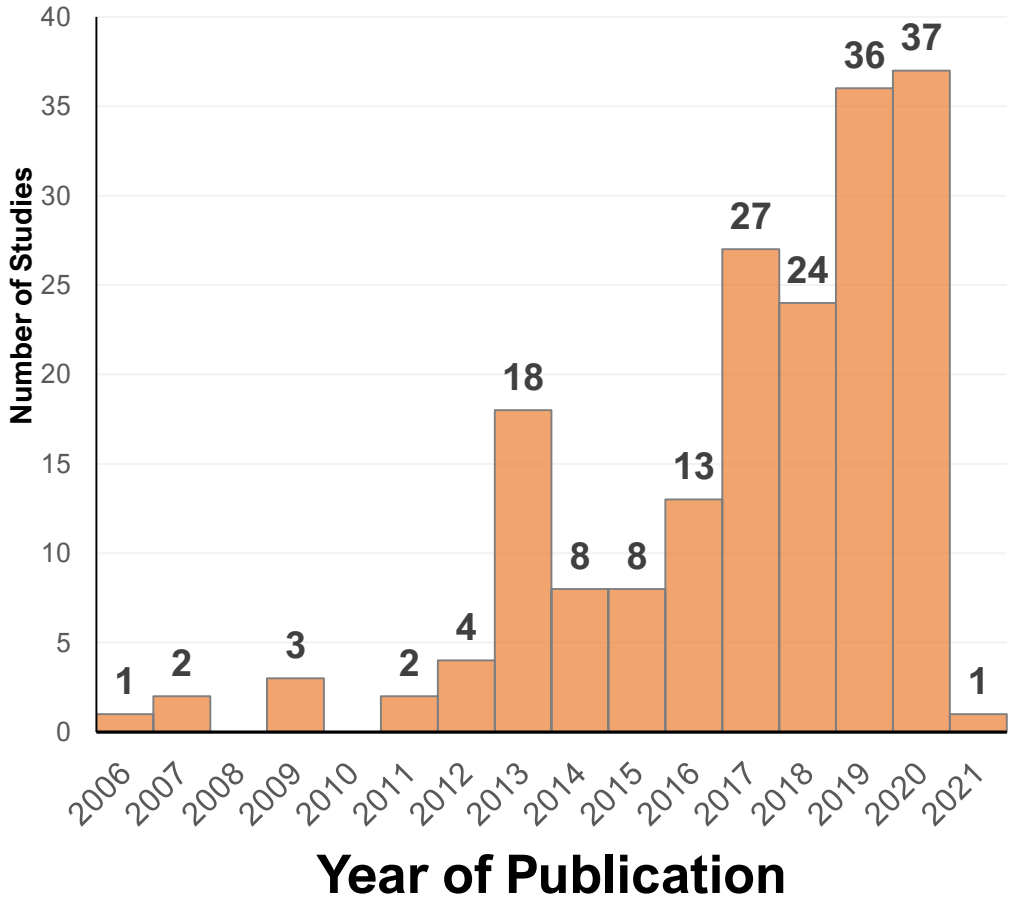
- **Search and selection process**
  - 6,043 records identified from database and manual searches
  - 4,262 records screened by titles and abstracts
  - 588 records assessed for eligibility by full-text review
  - **184 studies included in review**

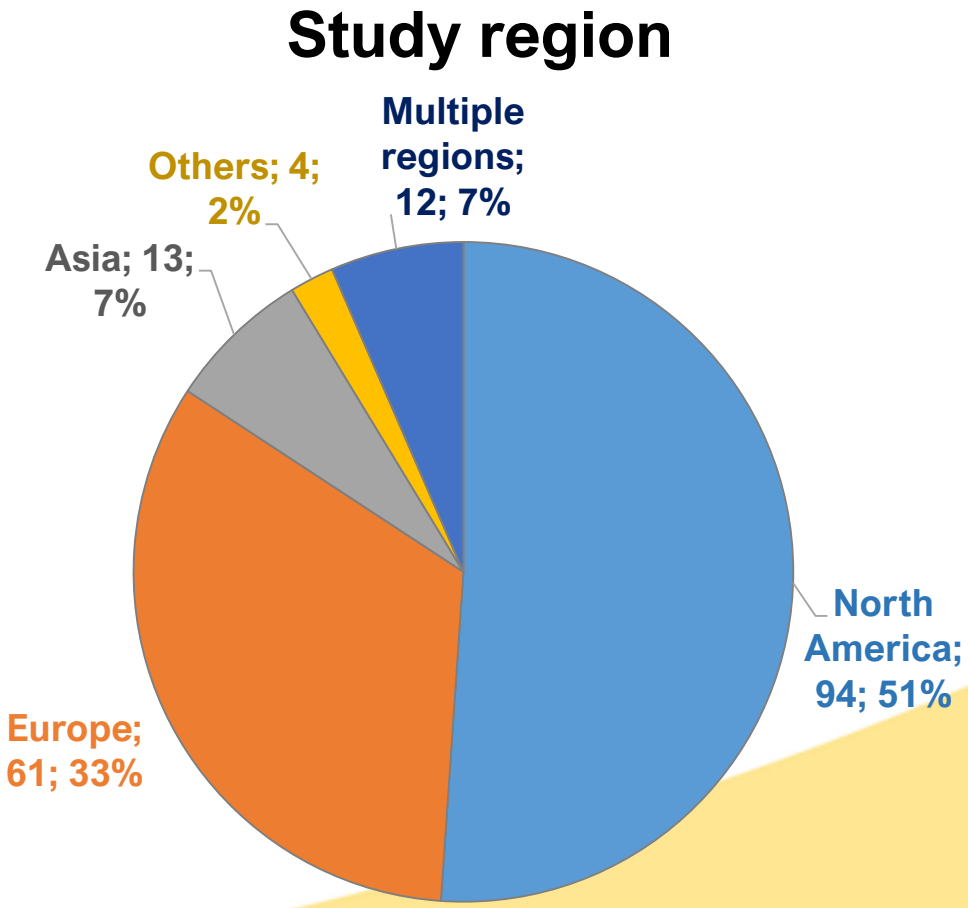
Figure 1. PRISMA flow diagram of literature search and selection process



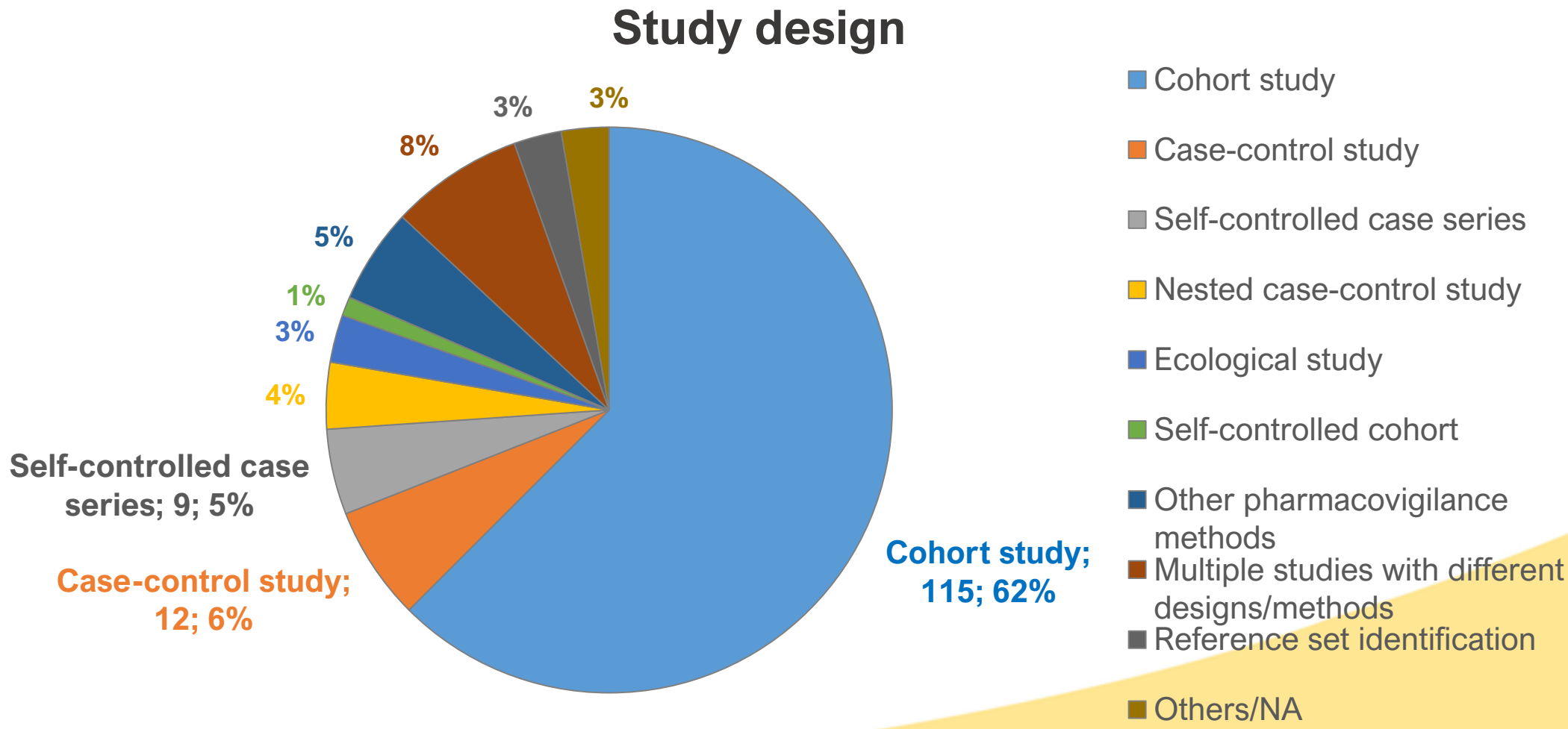
PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; NC, negative control.

Characteristics of included studies
– study characteristics



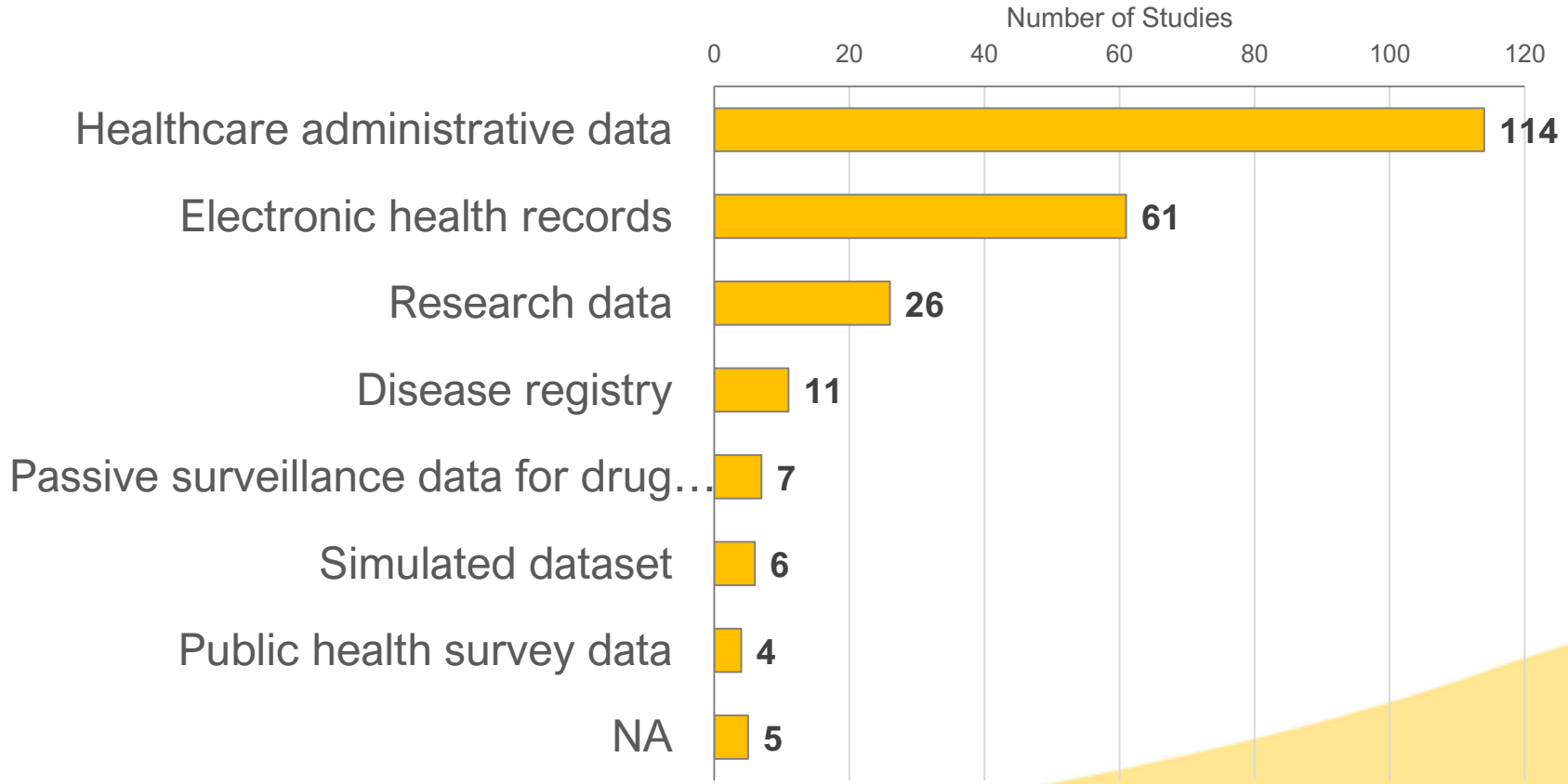


Characteristics of included studies
– study characteristics (cont.)



Characteristics of included studies
– study characteristics (cont.)

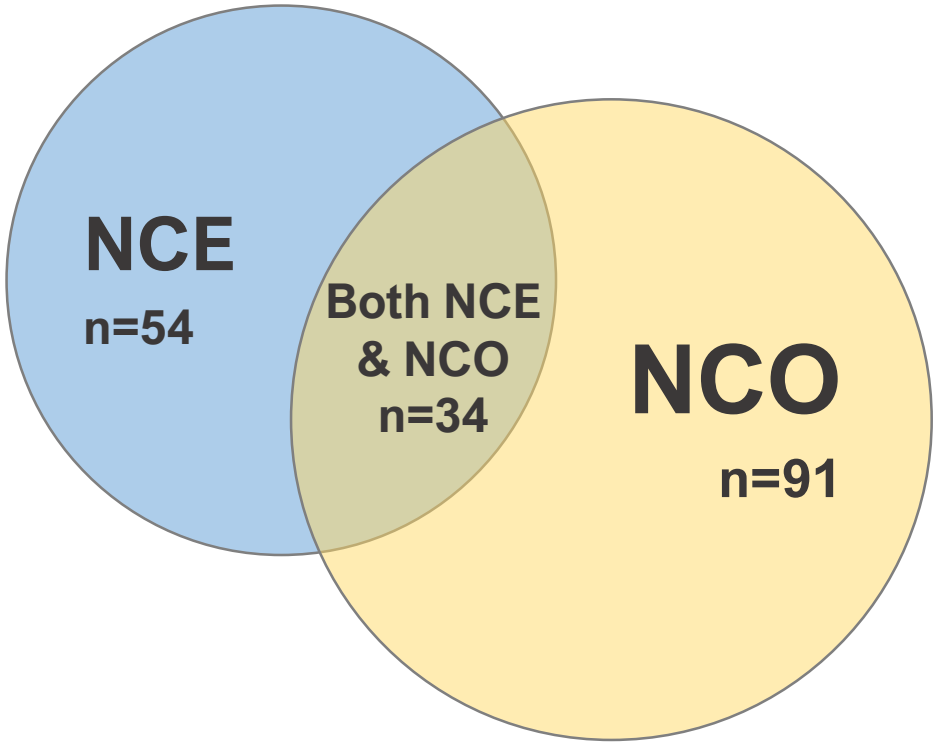
Data source*



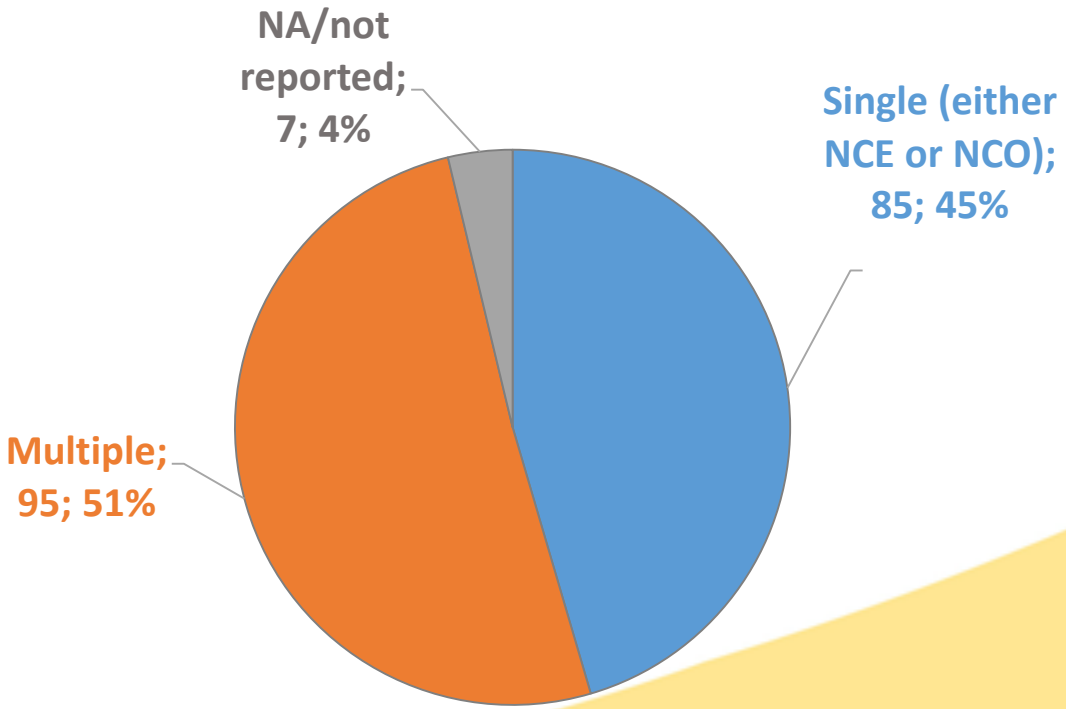
*Categories are not mutually exclusive.

Characteristics of included studies
– negative control characteristics

Type of NC*

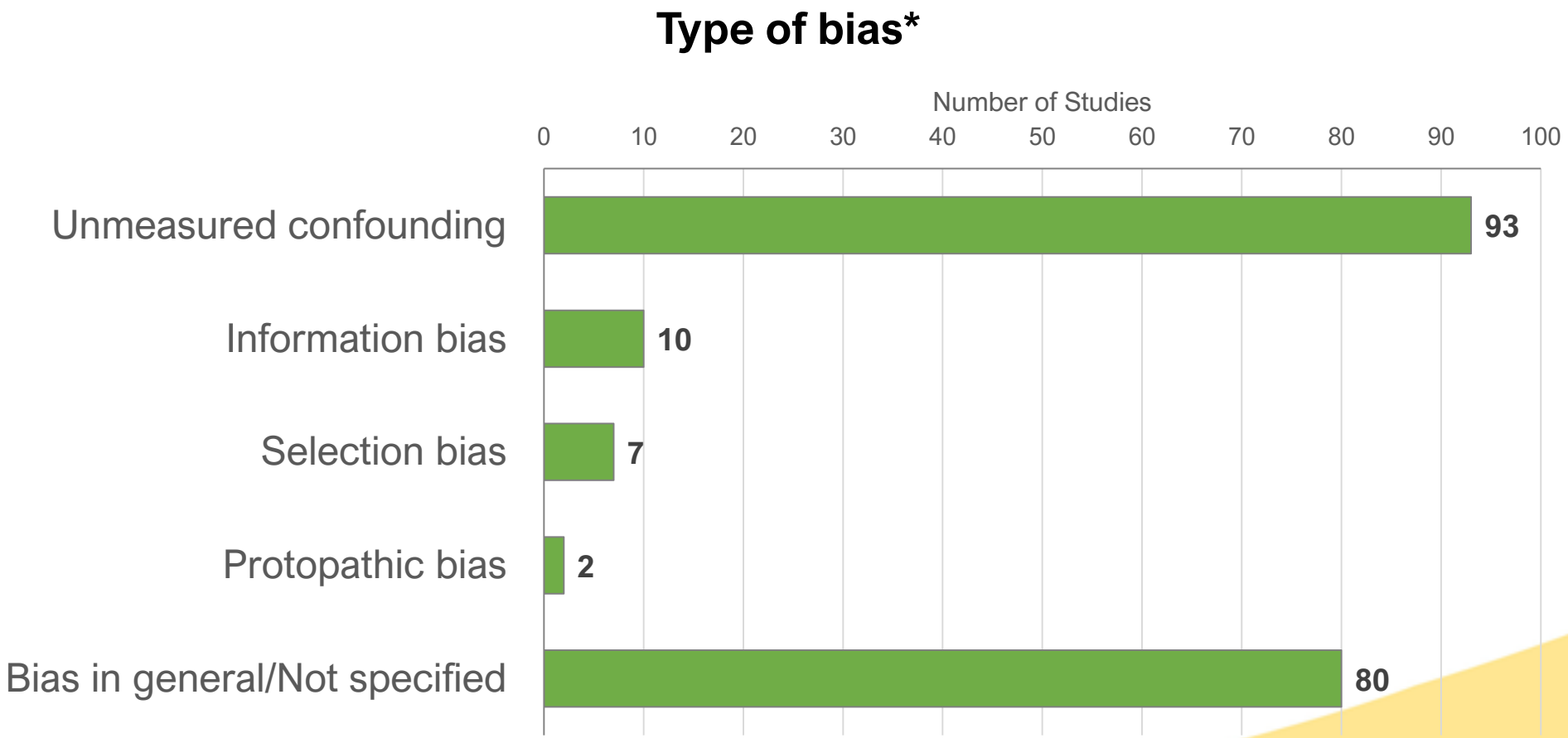


Number of NCs used



* 5 articles are not included (e.g., method papers for reference sets).

Characteristics of included studies
– negative control characteristics (cont.)



*Categories are not mutually exclusive.

Utility domains of negative control

- We classified the utility and application of negative controls into four domains:
  1) **Bias detection** – 149 studies
  2) **Bias correction in effect estimate** (e.g., correcting the point estimate, calibrating its confidence interval (CI)) – 16 studies
  3) **Calibration of *p*-value** – 8 studies
  4) **Performance assessment of different methods used in drug safety studies** – 26 studies

Utility Domain 1: Bias detection

- 149 (81%) out of 184 included studies used negative controls to detect potential bias in the causal association of investigation.
- Testing the negative control association against the null hypothesis.
  - Statistical significance (e.g., $p<0.05$) in negative control association was used to signal potential bias.
  - Different interpretation: A few studies concluded presence of bias based on the direction of the point estimate even when the negative control association was not significant.[1-3]
- Out of the 149 studies, 63 (42%) reported detecting a presence of bias.

1. Sinnott SJ, Smeeth L, Williamson E, et al. The comparative effectiveness of fourth-line drugs in resistant hypertension: An application in electronic health record data. *Pharmacoepidemiol Drug Saf*. 2019;28(9):1267-1277.
2. Brassard P, Wu JW, Ernst P, Dell'Aniello S, Smiechowski B, Suissa S. The effect of statins on influenza-like illness morbidity and mortality. Pharmacoepidemiol Drug Saf. 2017;26(1):63-70.
3. Lavikainen P, Helin-Salmivaara A, Eerola M, et al. Statin adherence and risk of acute cardiovascular events among women: A cohort study accounting for time-dependent confounding affected by previous adherence. BMJ Open. 2016;6(6).

Utility Domain 1: Bias detection

– a case of negative control exposure[1,2]

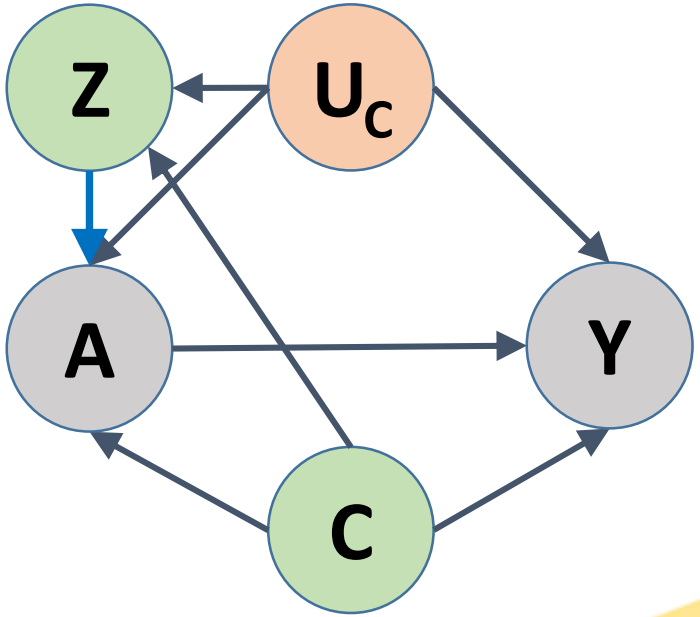
Primary analysis

$$Y = \beta_0 + \beta_1 A + \beta_2 C + \varepsilon$$

Negative control analysis

$$Y = \alpha_0 + \alpha_1 Z + \alpha_2 A + \alpha_3 C + \varepsilon$$

$H_0 : \alpha_1 = 0$
Conclude U exists *when* p<0.05



1. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative Controls: A tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21(3):383-388.
2. Flanders WD, Klein M, Darrow LA, et al. A method for detection of residual confounding in time-series and other observational studies. Epidemiology. 2011;22(1):59-67.

Utility Domain 1: Bias detection

– an example

- Yuan et al. 2020[1]
  - Retrospective cohort study using the Nurses' Health Study data on the effect of PPI use on incident rheumatoid arthritis
  - Used 3 NCOs (basal cell skin cancer, squamous cell skin cancer, and cervical cancer) to detect potential confounding bias by unknown factors
  - Used an NCE (H2 receptor blocker) to detect protopathic bias and confounding due to imbalance in the underlying diseases for acid suppressants use
  - No significant associations found in the negative control analyses

Received: 15 April 2020 | First decision: 20 April 2020 | Accepted: 13 May 2020
DOI: 10.1111/apt.15834

APT Alimentary Pharmacology & Therapeutics WILEY

Regular use of proton pump inhibitor and risk of rheumatoid arthritis in women: a prospective cohort study

Jinqiu Yuan[1] | Changhua Zhang[1] | Jeffrey A. Sparks[2] | Susan Malspeis[2] | Kelvin Kam-Fai Tsoi[3] | Jean H. Kim[3] | Benjamin A. Fisher[4] | Fang Gao[4] | Tim Sumerlin[3] | Yan Liu[1] | Yuxing Liu[1] | Yihang Pan[1] | Yulong He[1] | Joseph J.Y. Sung[3]

[1]Shenzhen, China
[2]Boston, MA, USA
[3]Shatin, Hong Kong
[4]Birmingham, UK

Correspondence
Yihang Pan, Precision Medicine Center, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, 518107, China.
Email: panyih@mail.sysu.edu.cn

Yulong He, Center for Digestive Disease, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, 518107, China.
Email: heyulong@mail.sysu.edu.cn

FUNDING
This work was supported by the Startup Fund for the 100 Top Talents Program, SYSU (392012), the National Institutes of Health grants (grant numbers: L30 AR066953, K23 AR069688, R01 AR049880, R03 AR075886, K24 AR052403, UM1 CA186107, R01 CA49449, UM1 CA176726, R01 CA67262), and the Fundamental Research Funds for the Central Universities (19ykzd03).

Summary
Background: Proton pump inhibitors (PPIs) have a significant impact on the gut microbiome, which in turn, might increase the risk of rheumatoid arthritis (RA).
Aim: To evaluate regular use of PPIs and risk of RA.
Methods: This is a prospective analysis of the US nurses who reported PPI use data, and were free of RA from the Nurses' Health Study (NHS 2002-2014) and NHS II (2003-2015). The exposure was regular use of PPI in the past 2 years, which was repeatedly evaluated in biennial surveys. RA was confirmed by the 1987 or 2010 American College of Rheumatology criteria. We estimated the hazard ratios (HRs) and confidence interval (CIs) with time-dependent Cox regression adjusting for potential confounders.
Results: We documented 421 cases of RA over 1 753 879 person-years of follow-up. Regular PPI users had a 44% higher risk of RA as compared with non-regular users (adjusted HR = 1.44; 95%CI, 1.10-1.89). The risk of RA increased with the total duration of PPI use (P-trend = 0.008). Compared with non-regular users, the adjusted HRs were 1.22 (95%CI, 0.93-1.62) for women with >0 to 4 years' use and 1.73 (95% CI, 1.14 to 2.61) for >4 years' use.
Conclusions: Regular use of PPI was associated with increased risk of RA in women, with a higher risk observed in individuals with a longer duration of PPI use. Due to the observational study design, large prospective trials are still required to confirm our finding.

The Handling Editor for this article was Professor Jonathan Rhodes, and it was accepted for publication after full peer-review.
The authors' complete affiliation are listed in Appendix 1.
Jinqiu Yuan and Changhua Zhang contributed equally to this study.

1. Yuan J, Zhang C, Sparks JA, et al. Regular use of proton pump inhibitor and risk of rheumatoid arthritis in women: a prospective cohort study. Aliment Pharmacol Ther. 2020;52(3):449-458.

Utility Domain 2: Bias correction

- 16 studies utilized negative controls to obtain bias-adjusted estimates (point estimates or CIs).
  - 14 out of 16 were published between 2017 and 2020.
- **Major approaches for correction**
  - Correction by subtracting the effect estimated from negative control analysis (6 studies)
  - Empirical calibration of CI (Schuemie et al. 2018; 6 studies)
  - Double negative control adjustment (Shi et al. 2020; 1 study)

Utility Domain 2: Bias correction

– the subtraction/division approaches

- Correction by subtracting the effect estimated from negative control analysis (division for ratio measures)
- Assumes the same bias (direction and magnitude) for the primary and the negative control analysis

Primary analysis

$$Y = \beta_0 + \boldsymbol{\beta_1}\mathrm{A} + \beta_2 C + \varepsilon$$

Negative control analysis

$$Y = \beta'_0 + \boldsymbol{\beta'_1}\mathrm{Z} + \beta'_2 C + \varepsilon$$

Measured (biased) effect: $\boldsymbol{\beta_1}$

Bias: $\boldsymbol{\beta'_1}$

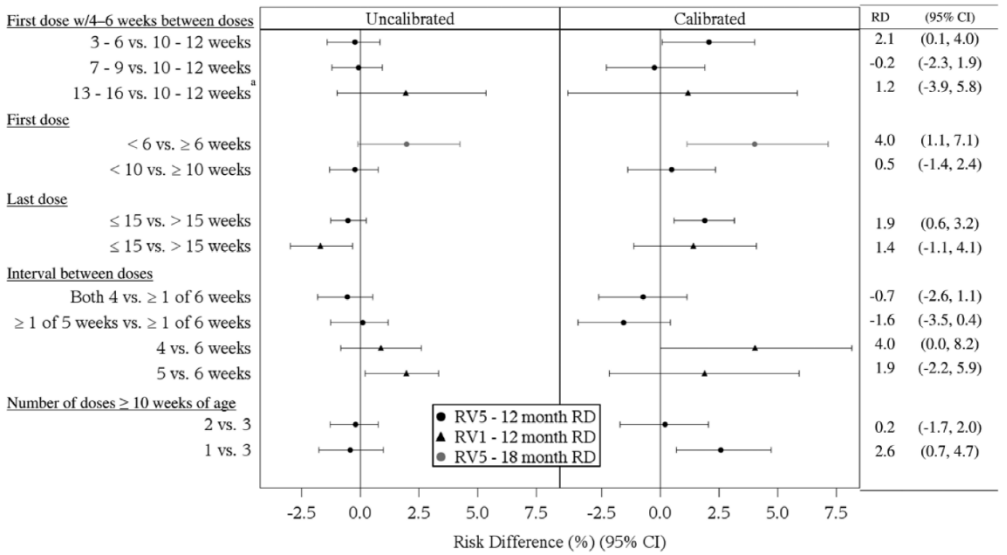
Corrected (unbiased) effect:

$\boldsymbol{\beta_1 - \beta'_1}$

Utility Domain 2: Bias correction

– examples for the subtraction/division approaches

- Gruber et al. 2018[1]
  - Cohort study using RCT data on the effectiveness of rotavirus vaccination on gastroenteritis incidence with different schedules
  - To calibrate the effect estimate (risk difference) for unmeasured confounding, the estimate for the placebo arm (NCE) was subtracted from the effect estimate for the vaccination arm.
  - A nonparametric bootstrap with replacement was used to obtain the point estimates and 95% CIs.
  - Significant results found for timing of the first dose (<6mo vs. >6mo. for RV5) after calibration.



[a] Due to differences in the covariate patterns between the placebo and vaccine arms, the calibrated estimate should be interpreted with caution.

FIGURE 2. Uncalibrated and calibrated RDs and 95% CIs.

1. Gruber JF, Becker-Dreps S, Hudgens MG, Alan Brookhart M, Thomas JC, Funk MJ. Timing of Rotavirus Vaccine Doses and Severe Rotavirus Gastroenteritis Among Vaccinated Infants in Low- and Middle-income Countries. Epidemiology. 2018;29(6):867-875.

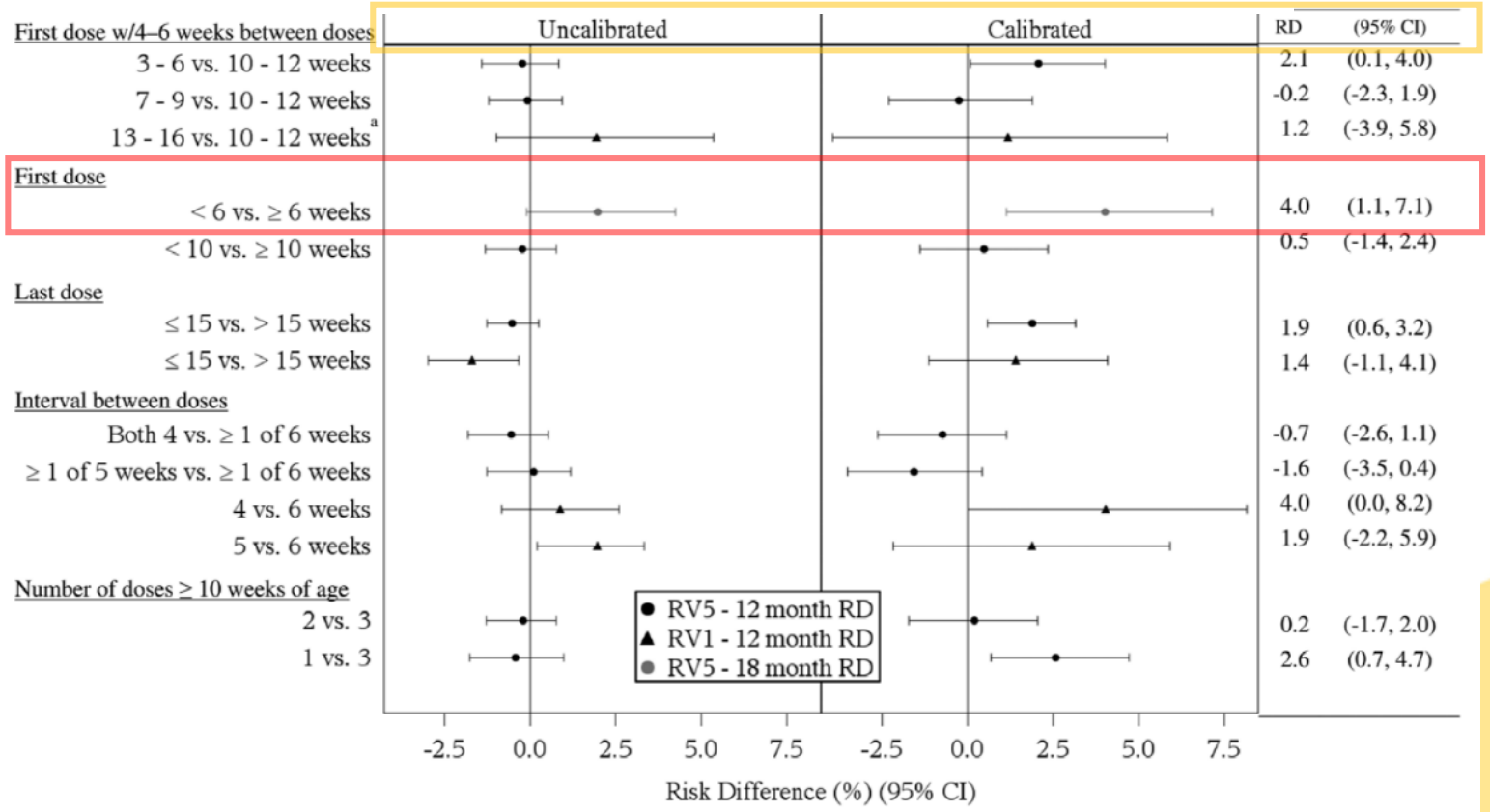
Utility Domain 2: Bias correction

– examples for the subtraction/division approaches

- Gruber et al. 2018[1]

Significant results found for timing of the first dose (<6mo vs. >6mo. for RV5) after calibration.

FIGURE 2. Uncalibrated and calibrated RDs and 95% CIs.



[a] Due to differences in the covariate patterns between the placebo and vaccine arms, the calibrated estimate should be interpreted with caution.

1. Gruber JF, Becker-Dreps S, Hudgens MG, Alan Brookhart M, Thomas JC, Funk MJ. Timing of Rotavirus Vaccine Doses and Severe Rotavirus Gastroenteritis Among Vaccinated Infants in Low- and Middle-income Countries. Epidemiology. 2018;29(6):867-875.

Utility Domain 2: Bias correction

– examples for the subtraction/division approaches

- Used in the drug-drug interaction (DDI) research (4 studies)
- Leonard et al. 2019[1]
  - Identified drugs (precipitant drugs) that increased serious bleeding risk when taken concomitantly with clopidogrel (an object drug)
  - Negative control drug (pravastatin) was used to account for inherent bleeding effect of precipitant drugs.
  - The ratio between the rate ratio (RR) from clopidogrel-precipitant association and the RR from pravastatin-precipitant association was estimated.

$$\boldsymbol{Ratio\ of\ RR} = \frac{\boldsymbol{RR_{clopidogrel-precipitant}}}{\boldsymbol{RR_{pravastatin\ (NCE)-precipitant}}}$$

1. Leonard CE, Zhou M, Brensinger CM, et al. Clopidogrel Drug Interactions and Serious Bleeding: Generating Real-World Evidence via Automated High-Throughput Pharmacoepidemiologic Screening. Clinical Pharmacology and Therapeutics. 2019;106(5):1067-1075.

Utility Domain 2: Bias correction

– examples for the empirical CI calibration

- Calibration of effect estimates and CIs using empirical null distributions constructed from multiple negative and synthetic positive controls[1]
- Lane et al. 2020[2]
  - Cohort study using healthcare administrative data and EHR data on the safety of hydroxychloroquine in patients with rheumatoid arthritis (effects on 16 severe adverse outcomes)
  - Used the empirical CI calibration to calibrate HR estimates for unmeasured confounding
  - 67 NCOs were used to construct the empirical null distribution and the synthetic positive controls.
  - Hydroxychloroquine use appears to be associated with excess cardiovascular mortality among patients with rheumatoid arthritis.

1. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. Proc Natl Acad Sci U S A. 2018;115(11):2571-2577.
2. Lane J, Weaver J, Kostka K, et al. Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study. Lancet Rheumatol. Published online 2020.

Utility Domain 2: Bias correction

– an example for the double negative control adjustment

- Estimating an unbiased average treatment effect (ATE) using a pair of an NCE and an NCO[1,2]
- Shi et al. 2020[2]
  - Expansion of the Control Outcome Calibration Approach (COCA)[3]
  - Demonstrated the performance of different estimators (including the multiply robust estimator) for ATE in conducting double negative control adjustments
  - Used an example cohort study on the risk of medically attended fever after receiving a DTaP-IPV-Hib vaccine in children (vs. other DTaP containing vaccines)
  - A pair of NCO (Injury or trauma) and NCE (Ringworm) was used to estimate ATE based on the double negative control adjustment method
  - Results indicated the slightly elevated risk of fever for DTaP-IPV-Hib vaccine relative to other DTap-containing comparator vaccines (the effect was not statistically significant).

1. Miao W, Geng Z, Tchetgen Tchetgen EJ. Identifying causal effects with proxy variables of an unmeasured confounder. Biometrika. 2018;105(4):987-993.
2. Shi X, Miao W, Nelson JC, Tchetgen Tchetgen EJ. Multiply robust causal inference with double-negative control adjustment for categorical unmeasured confounding. J R Stat Soc Ser B Stat Methodol. 2020;82(2):521-540.
3. Tchetgen Tchetgen E. The control outcome calibration approach for causal inference with unobserved confounding. Am J Epidemiol. 2014;179(5):633-640.

Utility Domain 3: Calibration of p-value

- Identified 8 studies which used the empirical p-value calibration method[1]

- Use multiple (n) NCs to estimate the empirical null distribution for the actual drug-outcome pair (y_{n+1})
- In the absence of residual bias, the null distribution for y_{n+1} should follow a normal distribution with mean of 0.

Bias

$$\theta_i \sim N(\mu, \sigma^2), \text{ and}$$

$$y_i \sim N(\theta_i, \tau_i^2)$$

Effect estimate for *i*th negative control drug-outcome pair

$$y_{n+1} \sim N(\hat{\mu}, \hat{\sigma}^2 + \tau_{n+1}^2)$$

Effect estimate for actual drug-outcome pair

1. Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in Medicine*. 2014;33(2):209-218.

Utility Domain 3: Calibration of p-value

– the empirical p-value calibration

- Estimation of the empirical null distribution
- Schuemie et al. 2014[1]
  - Demonstrated the empirical p-value calibration by replicating a cohort study on the effect of isoniazid on acute liver injury (Example A in the study).
  - Identified 37 negative control drugs for the outcome and used 34 drugs to estimate the empirical null distribution.
  - The estimated isoniazid effect was still significant after the p-value calibration.

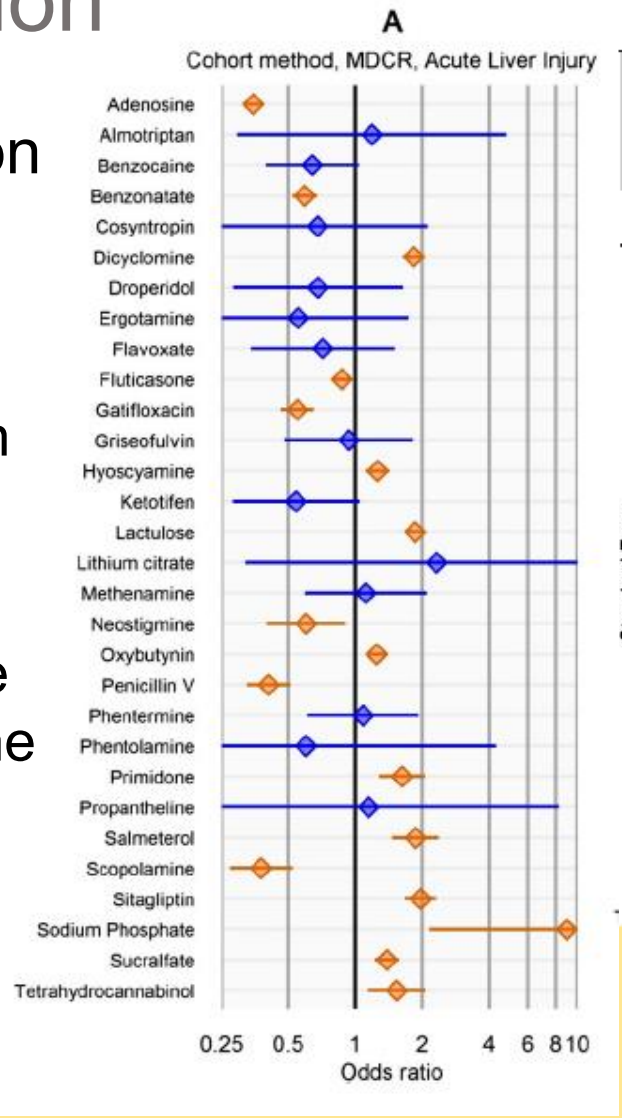
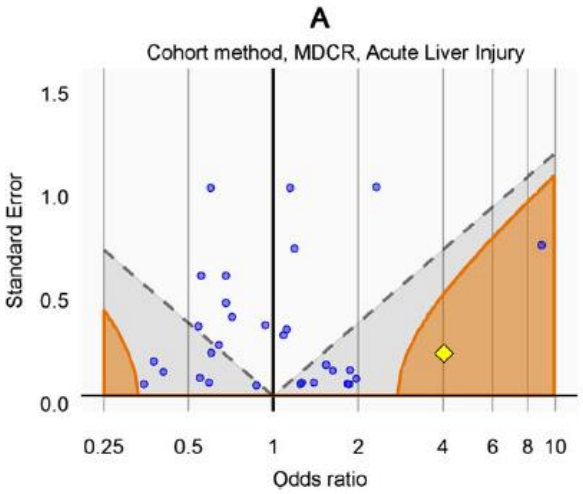


Table I. Estimated mean $\hat{\mu}$ and $\hat{\sigma}$ variance of the empirical null distribution for the three study designs.

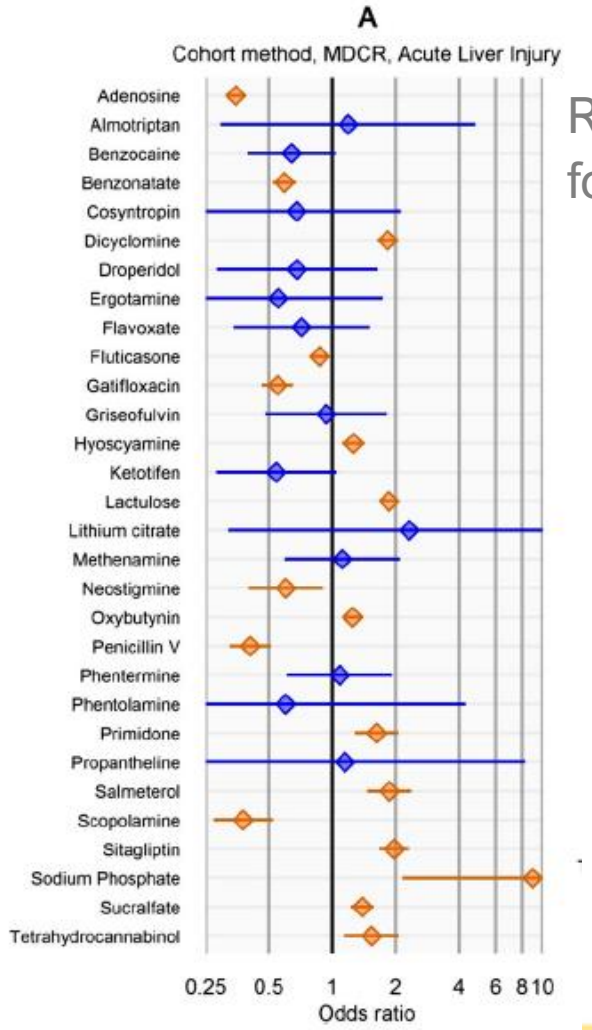
Design	$\hat{\mu}$	$\hat{\sigma}$
Cohort	−0.05	0.54



1. Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in Medicine*. 2014;33(2):209-218.

Utility Domain 3: Calibration of p-value

– the empirical p-value calibration



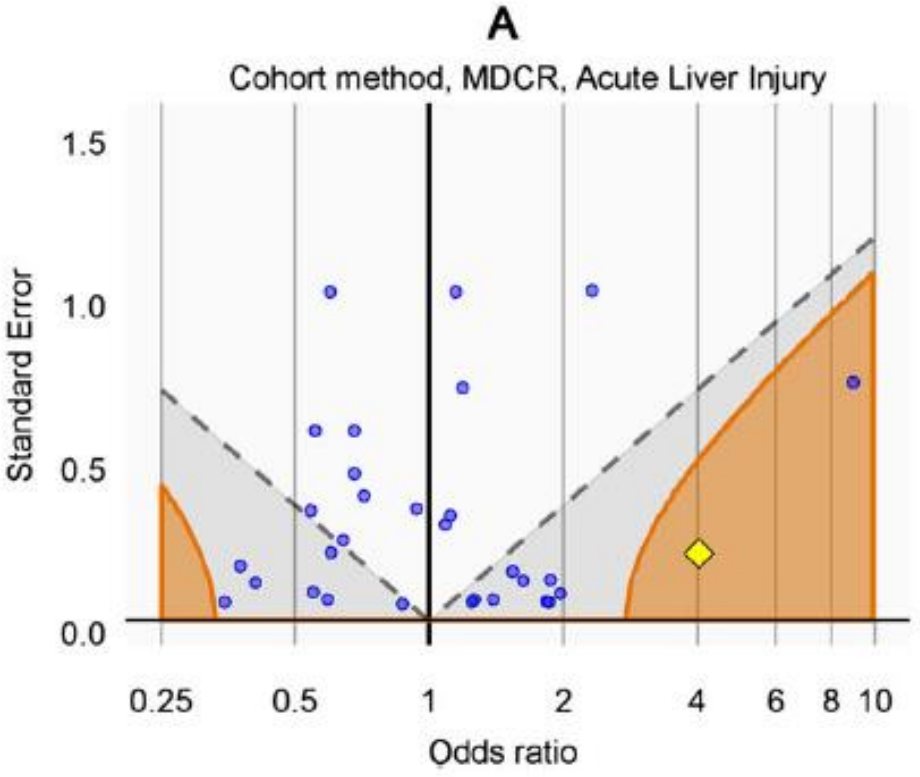
Results of negative control analyses for 34 drugs

Estimated parameters for the empirical null distribution

Table I. Estimated mean $\hat{\mu}$ and $\hat{\sigma}$ variance of the empirical null distribution for the three study designs.

Design	$\hat{\mu}$	$\hat{\sigma}$
Cohort	−0.05	0.54

Significant results under the empirical null distribution (orange area)

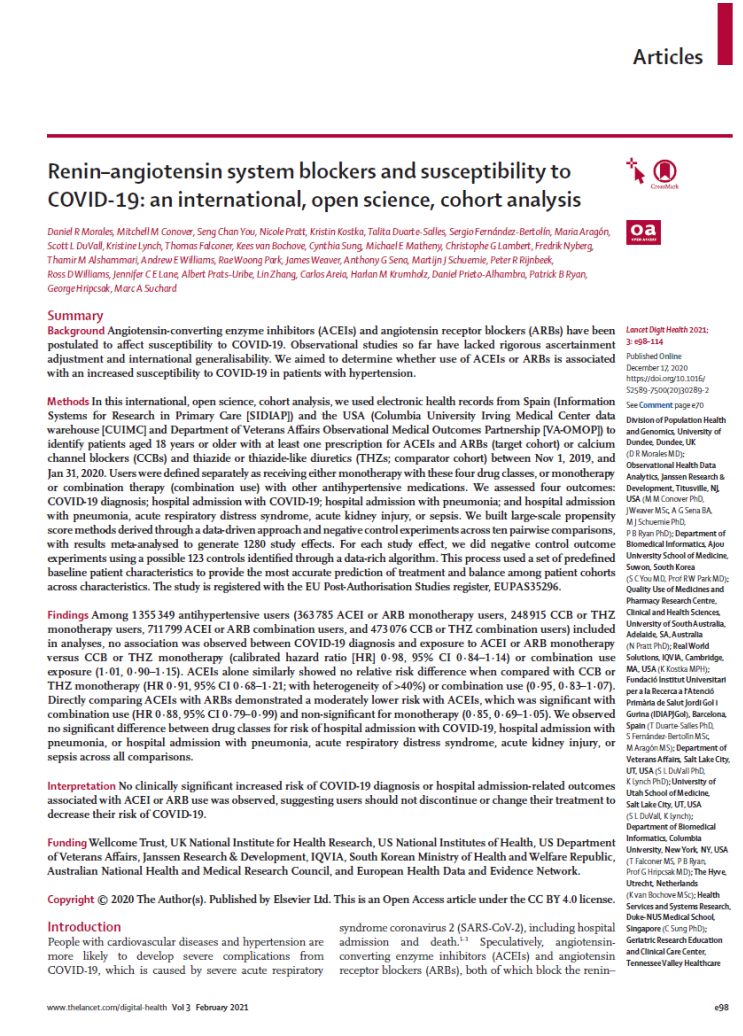


1. Schuemie MJ, Ryan PB, Dumouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in Medicine*. 2014;33(2):209-218.

Utility Domain 3: Calibration of p-value

– an example for the empirical p-value calibration

- Morales et al. 2021[1]
  - Investigated effects of use ACEIs and ARBs on COVID-19 diagnosis and COVID-related outcomes (vs. other classes of antihypertensive drugs)
  - Cohort study using multiple healthcare utilization databases (EHR, administrative databases)
  - Used the empirical calibration method to calibrate p-values and CIs (reported calibrated HRs)
    - 123 NCOs were used for the calibration.
  - Found no significant effects of ACEIs or ARBs on COVID-related outcomes.



Articles

Renin–angiotensin system blockers and susceptibility to COVID-19: an international, open science, cohort analysis

Daniel R Morales, Mitchell M Conover, Seng Chan You, Nicole Pratt, Kristin Kostka, Talita Duarte-Salles, Sergio Fernández-Bertolín, María Aragón, Scott L DuVall, Kristine Lynch, Thomas Falconer, Kees van Bochove, Cynthia Sung, Michael E Matheny, Christophe G Lambert, Fredrik Nyberg, Thamir M Alshammari, Andrew E Williams, Rae Woong Park, James Weaver, Anthony G Sena, Martijn J Schuemie, Peter R Rijnbeek, Ross D Williams, Jennifer C E Lane, Albert Prats-Uribe, Lin Zhang, Carlos Areia, Harlan M Krumholz, Daniel Prieto-Alhambra, Patrick B Ryan, George Hripcsak, Marc A Suchard

Summary

Background Angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin receptor blockers (ARBs) have been postulated to affect susceptibility to COVID-19. Observational studies so far have lacked rigorous ascertainment adjustment and international generalisability. We aimed to determine whether use of ACEIs or ARBs is associated with an increased susceptibility to COVID-19 in patients with hypertension.

Methods In this international, open science, cohort analysis, we used electronic health records from Spain (Information Systems for Research in Primary Care [SIDIAP]) and the USA (Columbia University Irving Medical Center data warehouse [CUIMC] and Department of Veterans Affairs Observational Medical Outcomes Partnership [VA-OMOP]) to identify patients aged 18 years or older with at least one prescription for ACEIs and ARBs (target cohort) or calcium channel blockers (CCBs) and thiazide or thiazide-like diuretics (THZs; comparator cohort) between Nov 1, 2019, and Jan 31, 2020. Users were defined separately as receiving either monotherapy with these four drug classes, or monotherapy or combination therapy (combination use) with other antihypertensive medications. We assessed four outcomes: COVID-19 diagnosis; hospital admission with COVID-19; hospital admission with pneumonia; and hospital admission with pneumonia, acute respiratory distress syndrome, acute kidney injury, or sepsis. We built large-scale propensity score methods derived through a data-driven approach and negative control experiments across ten pairwise comparisons, with results meta-analysed to generate 1280 study effects. For each study effect, we did negative control outcome experiments using a possible 123 controls identified through a data-rich algorithm. This process used a set of predefined baseline patient characteristics to provide the most accurate prediction of treatment and balance among patient cohorts across characteristics. The study is registered with the EU Post-Authorisation Studies register, EUPAS35296.

Findings Among 1 355 349 antihypertensive users (363 785 ACEI or ARB monotherapy users, 248 915 CCB or THZ monotherapy users, 711 799 ACEI or ARB combination users, and 473 076 CCB or THZ combination users) included in analyses, no association was observed between COVID-19 diagnosis and exposure to ACEI or ARB monotherapy versus CCB or THZ monotherapy (calibrated hazard ratio [HR] 0·98, 95% CI 0·84–1·14) or combination use exposure (1·01, 0·90–1·15). ACEIs alone similarly showed no relative risk difference when compared with CCB or THZ monotherapy (HR 0·91, 95% CI 0·68–1·21; with heterogeneity of >40%) or combination use (0·95, 0·83–1·07). Directly comparing ACEIs with ARBs demonstrated a moderately lower risk with ACEIs, which was significant with combination use (HR 0·88, 95% CI 0·79–0·99) and non-significant for monotherapy (0·85, 0·69–1·05). We observed no significant difference between drug classes for risk of hospital admission with COVID-19, hospital admission with pneumonia, or hospital admission with pneumonia, acute respiratory distress syndrome, acute kidney injury, or sepsis across all comparisons.

Interpretation No clinically significant increased risk of COVID-19 diagnosis or hospital admission-related outcomes associated with ACEI or ARB use was observed, suggesting users should not discontinue or change their treatment to decrease their risk of COVID-19.

Funding Wellcome Trust, UK National Institute for Health Research, US National Institutes of Health, US Department of Veterans Affairs, Janssen Research & Development, IQVIA, South Korean Ministry of Health and Welfare Republic, Australian National Health and Medical Research Council, and European Health Data and Evidence Network.

Introduction

People with cardiovascular diseases and hypertension are more likely to develop severe complications from COVID-19, which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), including hospital admission and death.[1–3] Speculatively, angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin receptor blockers (ARBs), both of which block the renin–

Lancet Digit Health 2021; 3: e98–114

Published Online December 17, 2020 https://doi.org/10.1016/S2589-7500(20)30289-2

See Comment page e70

Division of Population Health and Genomics, University of Dundee, Dundee, UK (D R Morales MD); Observational Health Data Analytics, Janssen Research & Development, Titusville, NJ, USA (M M Conover PhD, J Weaver MSc, A G Sena BA, M J Schuemie PhD, P B Ryan PhD); Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea (S C You MD, Prof R W Park MD); Quality Use of Medicines and Pharmacy Research Centre, Clinical and Health Sciences, University of South Australia, Adelaide, SA, Australia (N Pratt PhD); Real World Solutions, IQVIA, Cambridge, MA, USA (K Kostka MPH); Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain (T Duarte-Salles PhD, S Fernández-Bertolín MSc, M Aragón MSc); Department of Veterans Affairs, Salt Lake City, UT, USA (S L DuVall PhD, K Lynch PhD); University of Utah School of Medicine, Salt Lake City, UT, USA (S L DuVall, K Lynch); Department of Biomedical Informatics, Columbia University, New York, NY, USA (T Falconer MS, P B Ryan, Prof G Hripcsak MD); The Hyve, Utrecht, Netherlands (K van Bochove MSc); Health Services and Systems Research, Duke-NUS Medical School, Singapore (C Sung PhD); Geriatric Research Education and Clinical Care Center, Tennessee Valley Healthcare

1. Morales DR, Conover MM, You SC, et al. Renin–angiotensin system blockers and susceptibility to COVID-19: an international, open science, cohort analysis. The Lancet Digital Health. 2021;3(2):e98-e114.

Utility Domain 4: Performance assessment of different methods used in drug safety studies

- Identified 26 studies
  A. Studies used negative controls to evaluate and validate methods used to identify potential ADRs for pharmacovigilance purposes (24 studies)
  B. Studies used negative controls to compare the performance of different analytic methods to mitigate bias in pharmacoepidemiologic studies (2 studies)

Utility Domain 4: Performance assessment of different methods used in drug safety studies

A. Studies used negative controls to <u>evaluate and validate methods used to identify potential ADRs</u>

- Various analytical methods are routinely used to identify risk of adverse drug reaction in large observational databases.
- Validity of these methods can be tested against a set of known associations (known negative and positive control drug-outcome pairs) as reference.
- Common accuracy measures include area under the receiver operating curve (AUC), sensitivity, specificity, etc.
- E.g., Studies by the Observational Medical Outcomes Partnership (OMOP)

- **Case-control (CC)** compares the rate of exposure prior to outcomes with the rate of exposure in patients without outcomes [7].
- **Cohort method (CM)** is a new-user cohort design. New users of the target drug are identified using a predefined minimum period of non-use, and are compared to new users of a comparator drug or group of drugs Relative risk can be adjusted for baseline covariates through various strategies, including propensity score matching [8].
- **Disproportionality methods (DP)** are a suite of methods borrowed from data-mining in spontaneous reports, including proportional reporting ratios (PRR), reporting odds ratios (ROR), BCPNN (Bayesian Confidence Propagation Neural Networks) and MGPS (Multi-item Gamma Poisson Shrinker) [9].
- **Information Component Temporal Pattern Discovery (ICTPD)** compares the disproportionality of events during a post-exposure period with the disproportionality of events during one or more pre-exposure periods to produce a self-controlled-adjusted measure of temporal association [10].
- **Longitudinal Gamma Poisson Shrinker (LGPS)** compares the incidence rate during exposure to the drug of interest to the background incidence rate, optionally applying Bayesian shrinkage. This method is often combined with **Longitudinal Evaluation of Observational Profiles of Adverse events Related to Drugs (LEOPARD)**, a technique for detecting and discarding spurious signal caused by protopathic bias [11].
- **Self-controlled cohort design (SCC)** estimates the strength of association by comparing the post-exposure incidence rate with the pre-exposure incidence rate among the patients exposed to the target drug of interest [12].
- **Self-controlled case series (SCCS)** focuses on time exposed/unexposed to target drug and occurrences of target condition. It is basically a Poisson regression conditioned on the person [13].

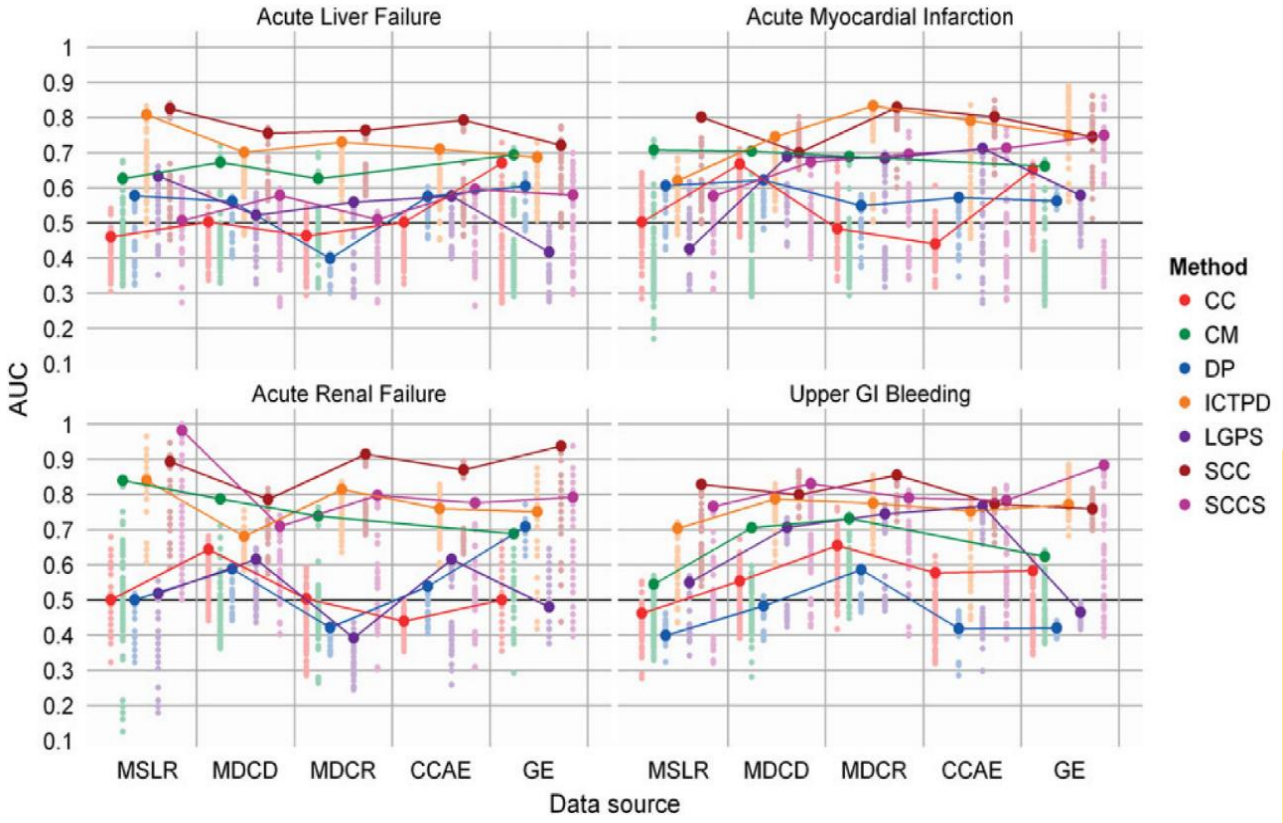
Schuemie et al. 2013[1]

1. Schuemie MJ, Gini R, Coloma PM, et al. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. Drug Safety. 2013;36:159-169.

Utility Domain 4: Performance assessment of different methods used in drug safety studies

- Ryan et al. 2013[1]
  - Compared the performance of 7 risk identification methods in 5 large observational databases
  - Used 399 drug-outcome pairs as reference standards for the evaluation
    - 165 positive and 234 negative control pairs across 4 outcomes
  - Self-controlled methods (SCC, SCCS, ICTPD) had higher predictive accuracy.

Fig. 1 AUC values for all analytic methods by database and method



1. Ryan PB, Stang PE, Overhage JM, et al. A comparison of the empirical performance of methods for a risk identification system. Drug Safety. 2013;36:143-158.

Utility Domain 4: Performance assessment of different methods used in drug safety studies

B. Studies used negative controls to <u>compare the performance of different analytic methods to mitigate bias</u>

- Davies et al. 2017[1]
  - Demonstrated the utility of negative control analysis and bias component plots using an example study for the comparative safety of smoking cessation therapies
    - Effects of varenicline vs. nicotine replacement products on suicide and self-harm – potential risk of healthy user bias (confounding by indication)
  - Used both the instrumental variable (IV) approach and the conventional regression method and compared the relative bias of the approaches
  - Used NCO (urinary tract infection) and NCE (patients prescribed drugs unrelated to smoking from the same GP) to measure the relative bias in the effect estimates
  - No bias was detected for the IV approach while conventional regression approach showed signs of bias.

1. Davies NM, Thomas KH, Taylor AE, et al. How to compare instrumental variable and conventional regression analyses using negative controls and bias plots. International Journal of Epidemiology. 2017;46(6):2067-2077.

Strategies to validate negative controls

1) Assumption of lack of causality

- Studies identified negative control associations that are:
  1. Logically impossible (e.g., causally implausible timing of the event, spatial constraints, etc.)
  2. Clinically implausible
  3. No previous evidence of potential causal association.
- Use of reference sets to validate negative controls
  - Reference sets of negative control associations based on the current best evidence by combining information from multiple sources (drug labels, spontaneous adverse drug event reporting systems, literature review by Tisdale et al., etc.)
  - e.g., EU-ADR reference set[1], OMOP reference set[2], Algorithm to identify a reference set proposed by OHDSI[3]

1. Coloma PM, Avillach P, Salvo F, et al. A Reference Standard for Evaluation of Methods for Drug Safety Signal Detection Using Electronic Healthcare Record Databases. *Drug Saf*. 2013;36(1):13-23.
2. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. Drug Safety. 2013;36:33-47.
3. Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. Accuracy of an automated knowledge base for identifying drug adverse reactions. Journal of Biomedical Informatics. 2017;66:72-81.

Strategies to validate negative controls

2) Assumption of shared bias structure

- Comparison of distributions of major measured confounders between the primary exposure/outcome and the negative control exposure/outcome (8 studies)
  - In practice, the ideal negative controls should share a similar set of confounders including measured confounders.[1]
- An example: Re-weighting of the NCE sample based on the covariate distribution in the sample for the primary analysis [2]

1. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative Controls: A tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21(3):383-388. doi:10.1097/EDE.0b013e3181d61eeb
2. Dillon P, Smith SM, Gallagher PJ, Cousins G. Association between gaps in antihypertensive medication adherence and injurious falls in older community-dwelling adults: A prospective cohort study. BMJ Open. 2019;9(3).

Discussion

- Primary use of negative controls has been to detect bias in the literature (149 studies, 81%).
- Other utility domains of negative control applications, such as bias correction, p-value calibration, and performance evaluation of epidemiological methods were identified.
- Application of several key methodological developments for bias correction using negative controls has been growing in recent years.
  - Empirical calibration methods
  - Double negative control adjustment for vaccine effectiveness studies

Q & A

Thank you

- **Questions?**

- Correspondence:
  - zzafari@rx.umaryland.edu
  - jpark1@umaryland.edu

Additional examples

1) Bias detection

- Yates et al. 2017[1]
  - Retrospective cohort study on the protective effects of lansoprazole on risk of tuberculosis (compared to other PPIs) using UK CPRD data
  - Used myocardial infarction as an NCO to detect confounding due to overall health status
  - Used herpes zoster as a second NCO to detect confounding due to impaired immunity
  - No protective effect of lansoprazole found on both NCOs

PLOS | MEDICINE

RESEARCH ARTICLE

Lansoprazole use and tuberculosis incidence in the United Kingdom Clinical Practice Research Datalink: A population based cohort

Tom A. Yates[1], Laurie A. Tomlinson[2], Krishnan Bhaskaran[2], Sinead Langan[2], Sara Thomas[2], Liam Smeeth[2], Ian J. Douglas[2]*

1 Institute for Global Health, University College London, Institute of Child Health, London, United Kingdom, 2 London School of Hygiene & Tropical Medicine, London, United Kingdom

* ian.douglas@lshtm.ac.uk

Check for updates

OPEN ACCESS

Citation: Yates TA, Tomlinson LA, Bhaskaran K, Langan S, Thomas S, Smeeth L, et al. (2017) Lansoprazole use and tuberculosis incidence in the United Kingdom Clinical Practice Research Datalink: A population based cohort. PLoS Med 14 (11): e1002457. https://doi.org/10.1371/journal.pmed.1002457

Academic Editor: John Z. Metcalfe, University of California San Francisco, UNITED STATES

Received: February 22, 2017

Accepted: October 23, 2017

Published: November 21, 2017

Data Availability Statement: The data were obtained from the Clinical Practice Research Datalink (CPRD). CPRD data governance does not allow us to distribute patient data to other parties. Researchers may apply for data access at http://www.CPRD.com/. The codes used to produce the data for this study are provided in the Supporting Information.

Funding: TAY was funded via an MRC PhD studentship while undertaking this work (MR/

Abstract

Background

Recent in vitro and animal studies have found the proton pump inhibitor (PPI) lansoprazole to be highly active against *Mycobacterium tuberculosis*. Omeprazole and pantoprazole have no activity. There is no evidence that, in clinical practice, lansoprazole can treat or prevent incident tuberculosis (TB) disease.

Methods and findings

We studied a cohort of new users of lansoprazole, omeprazole, or pantoprazole from the United Kingdom Clinical Practice Research Datalink to determine whether lansoprazole users have a lower incidence of TB disease than omeprazole or pantoprazole users. Negative control outcomes of myocardial infarction (MI) and herpes zoster were also studied. Multivariable Cox proportional hazards regression was used to adjust for potential confounding by a wide range of factors. We identified 527,364 lansoprazole initiators and 923,500 omeprazole or pantoprazole initiators. Lansoprazole users had a lower rate of TB disease ($n = 86$; 10.0 cases per 100,000 person years; 95% confidence interval 8.1–12.4) than omeprazole or pantoprazole users ($n = 193$; 15.3 cases per 100,000 person years; 95% confidence interval 13.3–17.7), with an adjusted hazard ratio (HR) of 0.68 (0.52–0.89). No association was found with MI (adjusted HR 1.04; 95% confidence interval 1.00–1.08) or herpes zoster (adjusted HR 1.03; 95% confidence interval 1.00–1.06). Limitations of this study are that we could not determine whether TB disease was due to reactivation of latent infection or a result of recent transmission, nor could we determine whether lansoprazole would have a beneficial effect if given to people presenting with TB disease.

Conclusions

In this study, use of the commonly prescribed and cheaply available PPI lansoprazole was associated with reduced incidence of TB disease. Given the serious problem of drug

1. Yates TA, Tomlinson LA, Bhaskaran K, et al. Lansoprazole use and tuberculosis incidence in the United Kingdom Clinical Practice Research Datalink: A population-based cohort. PLoS Medicine. 2017;14(11):1-15.

Panel Discussion

Moderator: **Rachele Hendricks-Sturrup,** Duke-Margolis Center for Health Policy

Panelists:

- **Fang Tian**, U.S. Food and Drug Administration
- **Hector Izurieta**, U.S. Food and Drug Administration
- **Zafar Zafari and Jeong-eun Park**, University of Maryland School of Pharmacy

Overview of Analytic Techniques: A review of Negative Controls

Statistical Validation for Disconnected Negative Control Models

Erich Kummerfeld

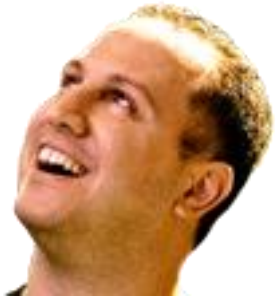
Institute for Health Informatics
University of Minnesota

March 8, 2023

UNIVERSITY OF MINNESOTA
Driven to Discover®

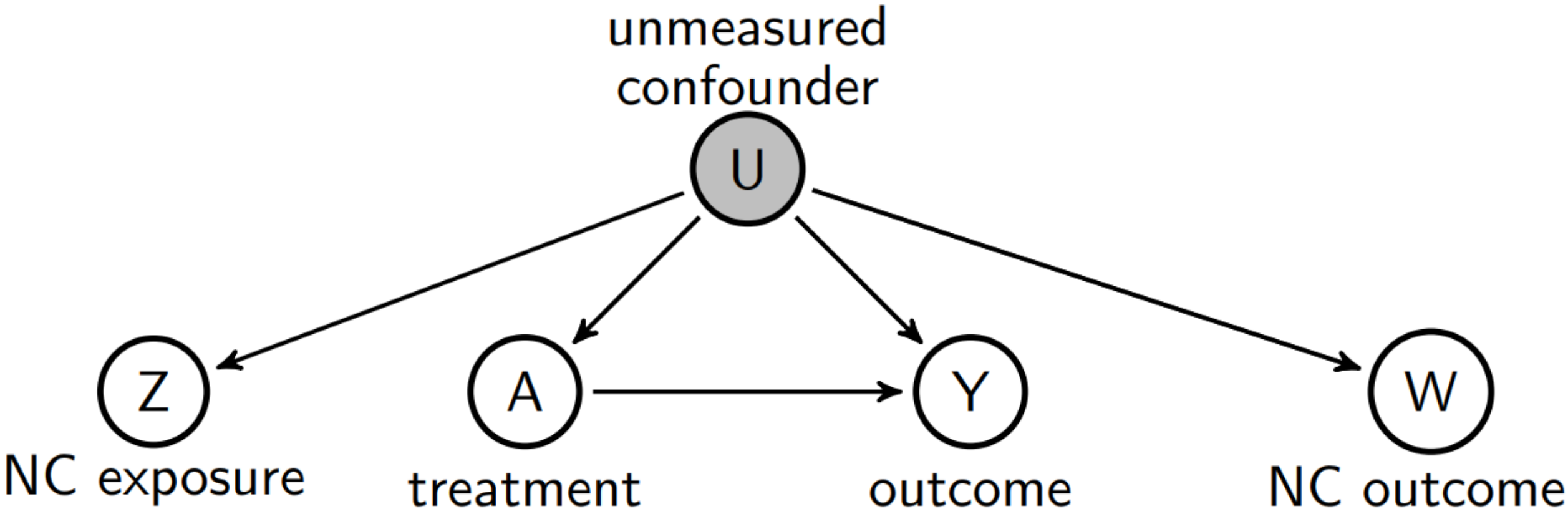
Challenge

+ Negative controls enable **powerful** causal inference methods that overcome unmeasured confounding.



- Estimation can be **biased** if the proposed negative controls do not satisfy certain criteria
- These criteria include causal structure requirements with the **unmeasured** confounder

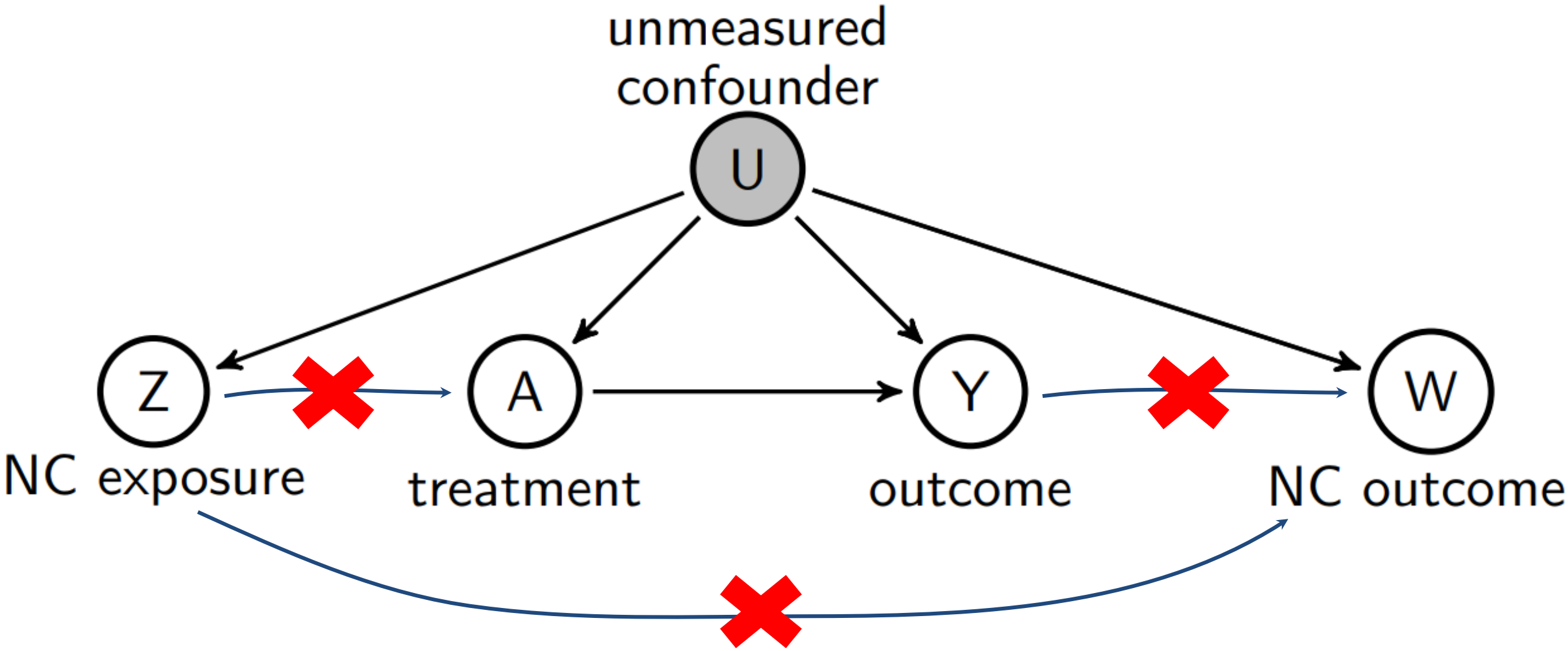
Disconnected negative controls



Disconnected negative controls

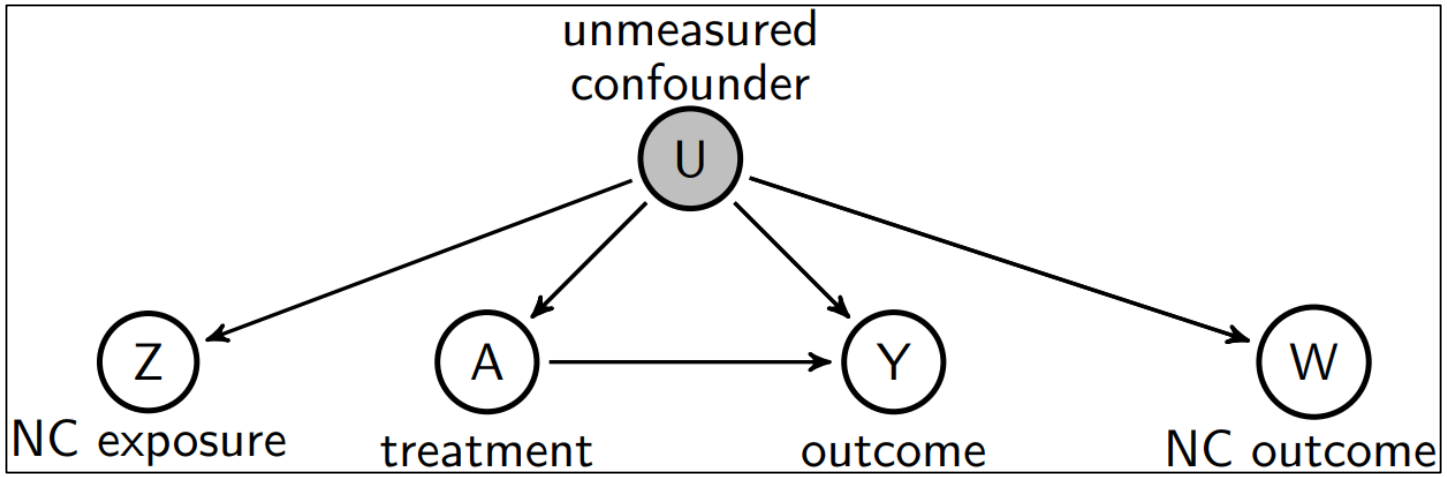
- Disconnected negative controls (DNCs) require: **all treks** connecting a DNC to…
  - …other DNCs…
  - …the exposure…
  - …the outcome…
- …go **through the unmeasured confounder**.

Disconnected negative controls



Why define such a restrictive NC?

- DNCs have **simple** and **testable** implications!
- DNC model:

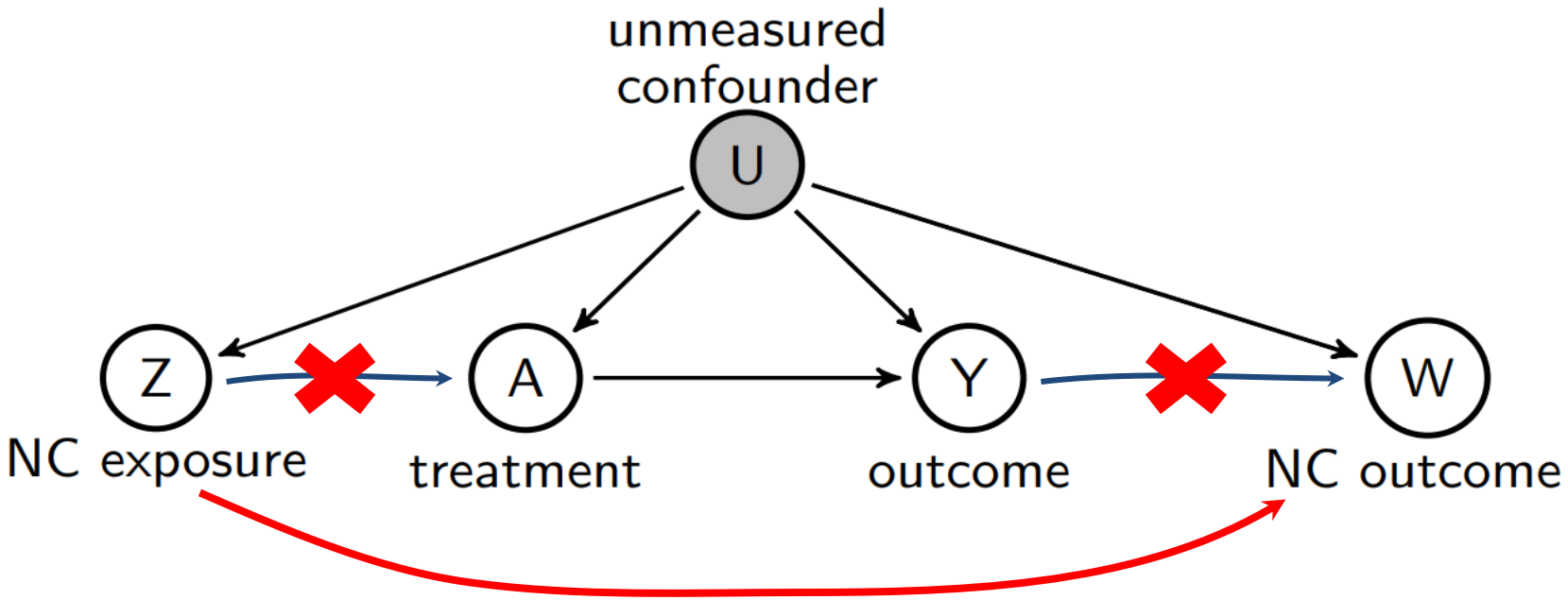


- For example, this
- Is rank deficient!

$$\Sigma_{\{W,Z\},\{Y,A\}} = \begin{pmatrix} cov(W,Y) & cov(W,A) \\ cov(Z,Y) & cov(Z,A) \end{pmatrix}$$

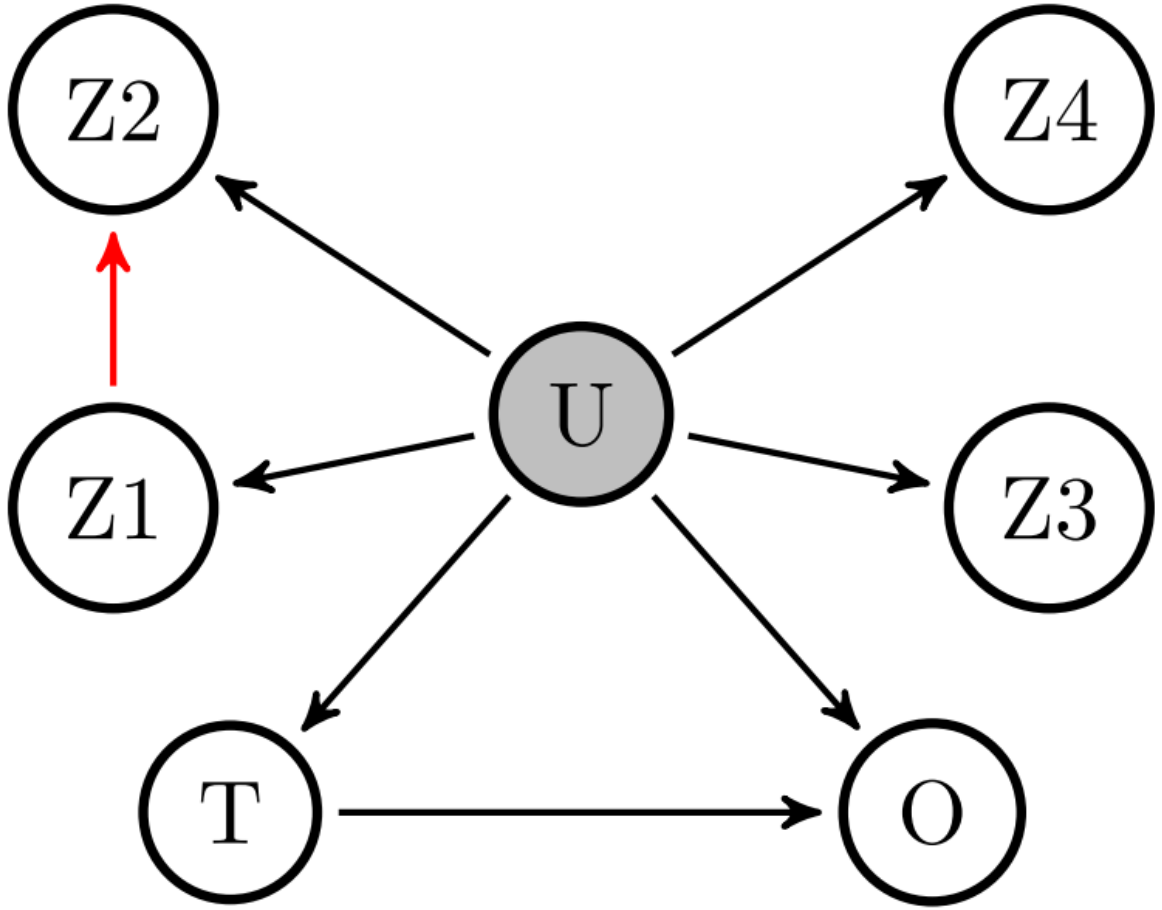
Two DNCs is not enough for testing

- That one rank deficiency does not rule out the Z-W edge



But we CAN validate DNC **triplets**

- **V**alid DNC triplets:
  {Z1, Z4, Z3}
  {Z2, Z4, Z3}
- **E**ach valid triplet implies six testable rank deficiencies
- **5** distinct DNC pairs



Three candidate DNCs, Six tests

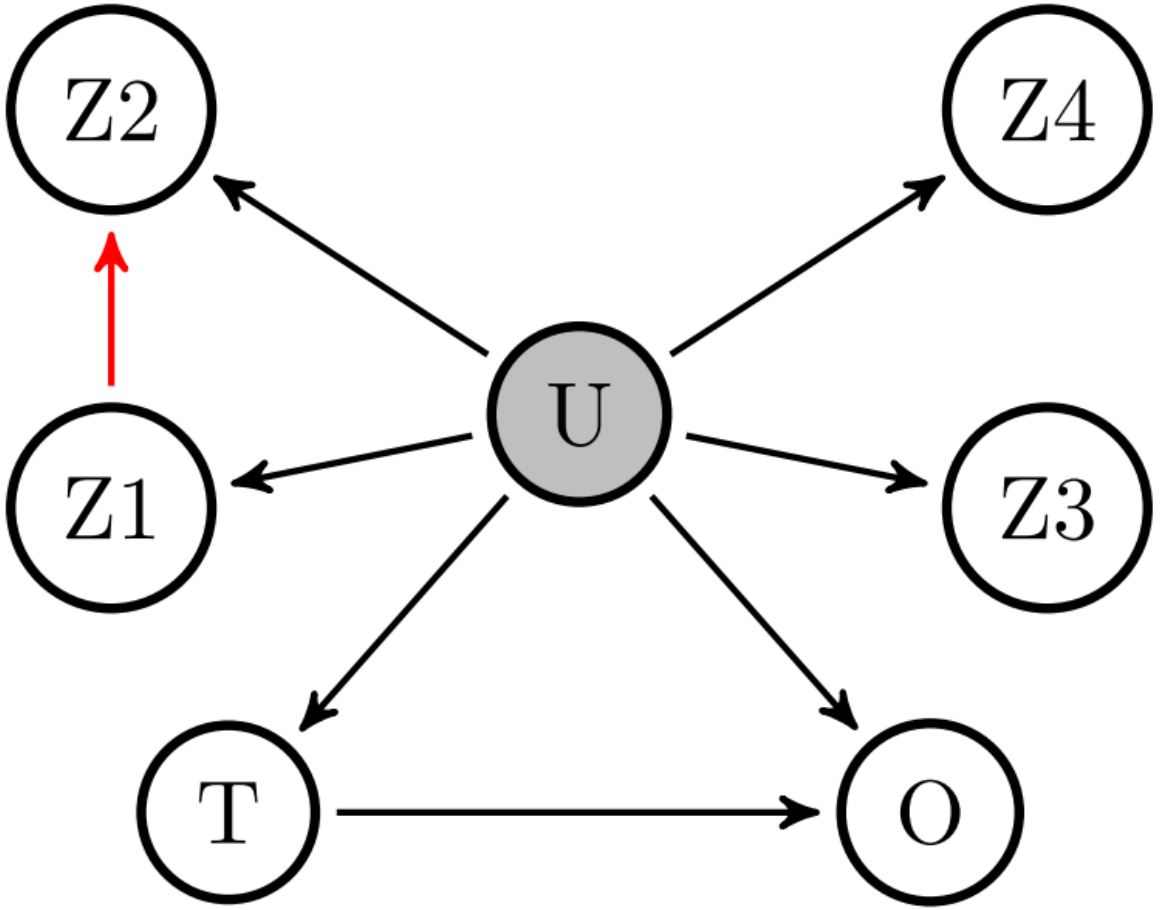
Let A, B, and C be any three candidate DNCs. Test if these sub-covariance matrices have determinant 0.

1. {A,B} x {C,T}
2. {A,C} x {B,T}
3. {C,B} x {A,T}
4. {A,B} x {C,O}
5. {A,C} x {B,O}
6. {C,B} x {A,O}

These are also called tetrads. If the determinant is 0, it is a “vanishing tetrad”. A statistical test can be used to determine if an empirical determinant is significantly different from 0, e.g. the Wishart test.

Example application of the test

- {Z2,Z4} x {Z1,T} is not rank deficient, so {Z1,Z2,Z4} is not a valid DNC triplet



Correctness proved from 3 assumptions

Assumption 1 *The data is generated by a SNCM with* $|\boldsymbol{M}| \geq 5$. *This implies that there are at least three candidate negative controls.*

Assumption 2 Tetrad Faithfulness. *In the data distribution implied by the SNCM, tetrads vanish only if they are implied to vanish by the structure of the SNCM. In other words, tetrads do not vanish as an "accident" of the model's particular coefficients.*

Assumption 3 *The data is generated by a SNCM which is linear and acyclic among its measured variables.*

Theorem 5 *Let the data be generated by a SNCM,* $\mathcal{G}$, *and assume Assumptions 1, 2, and 3. The DNCT validation test will return TRUE for any DNCT in* $\mathcal{G}$, *and FALSE otherwise.*

Brute force search DNC classifier

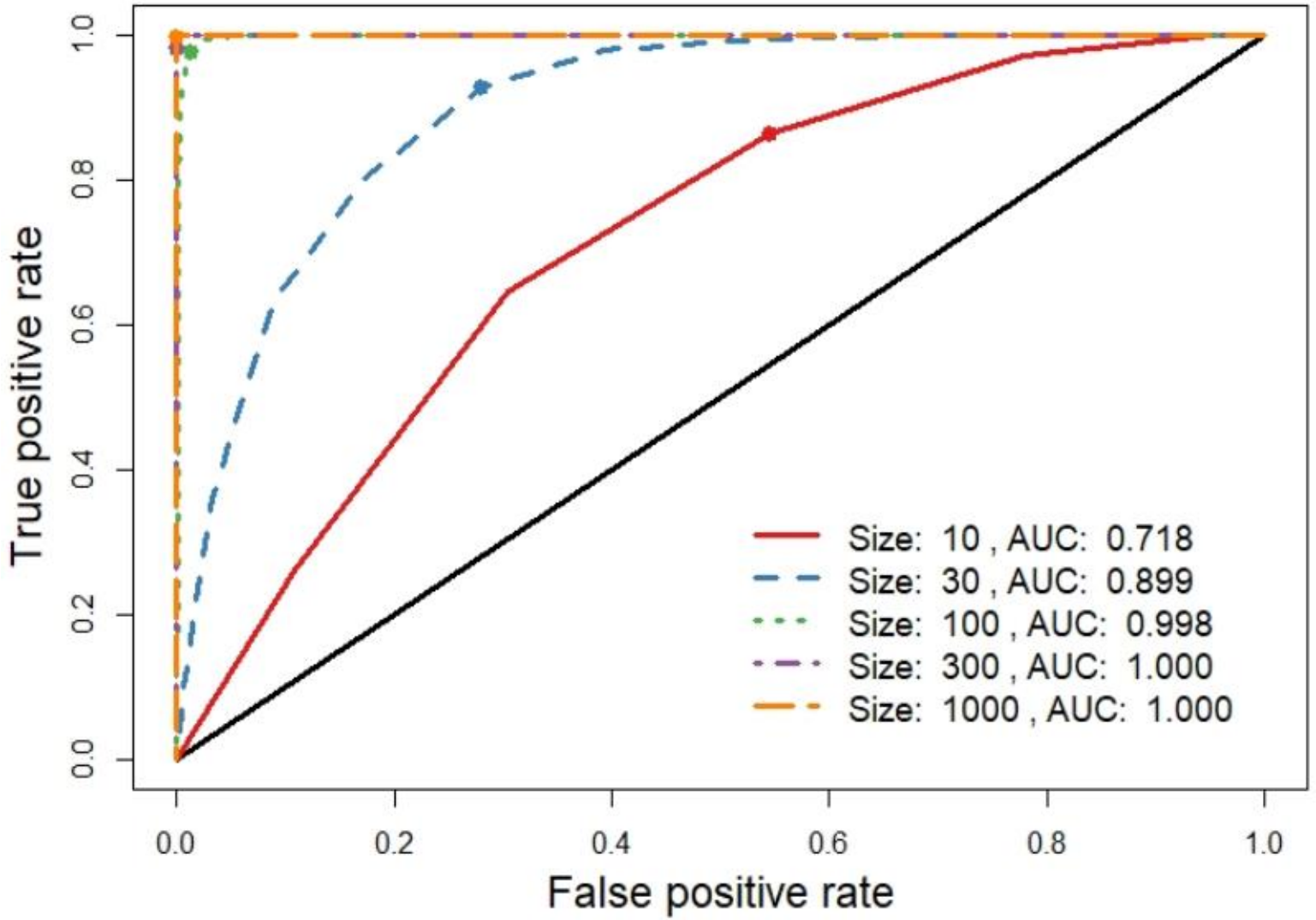
Algorithm 1: Find Negative Controls (FNC)

Data: $Data, CandidateNegativeControls, T, O, vt, alpha$

Result: A collection of validated negative control triplets (DNCTs)

```
1 DNCTs ← ∅
2 for X, Y, Z ∈ CandidateNegativeControls, X ≠ Y ≠ Z do
3     if DNCT(Data, X, Y, Z, T, O, alpha, vt) then
4         Output ← Output ∪ {X, Y, Z}
5 return(DNCTs)
```

This works well for classifying DNCs



Implemented and tested with Wishart test

Find, Estimate, Aggregate Estimates

Algorithm 3: Data-driven Automated Negative Control Estimation (DANCE)

Data: $Data, CandidateNegativeControls, T, O, vt, alpha$

Result: A point estimate and confidence interval for the effect of T on O

$Estimates \leftarrow \varnothing$

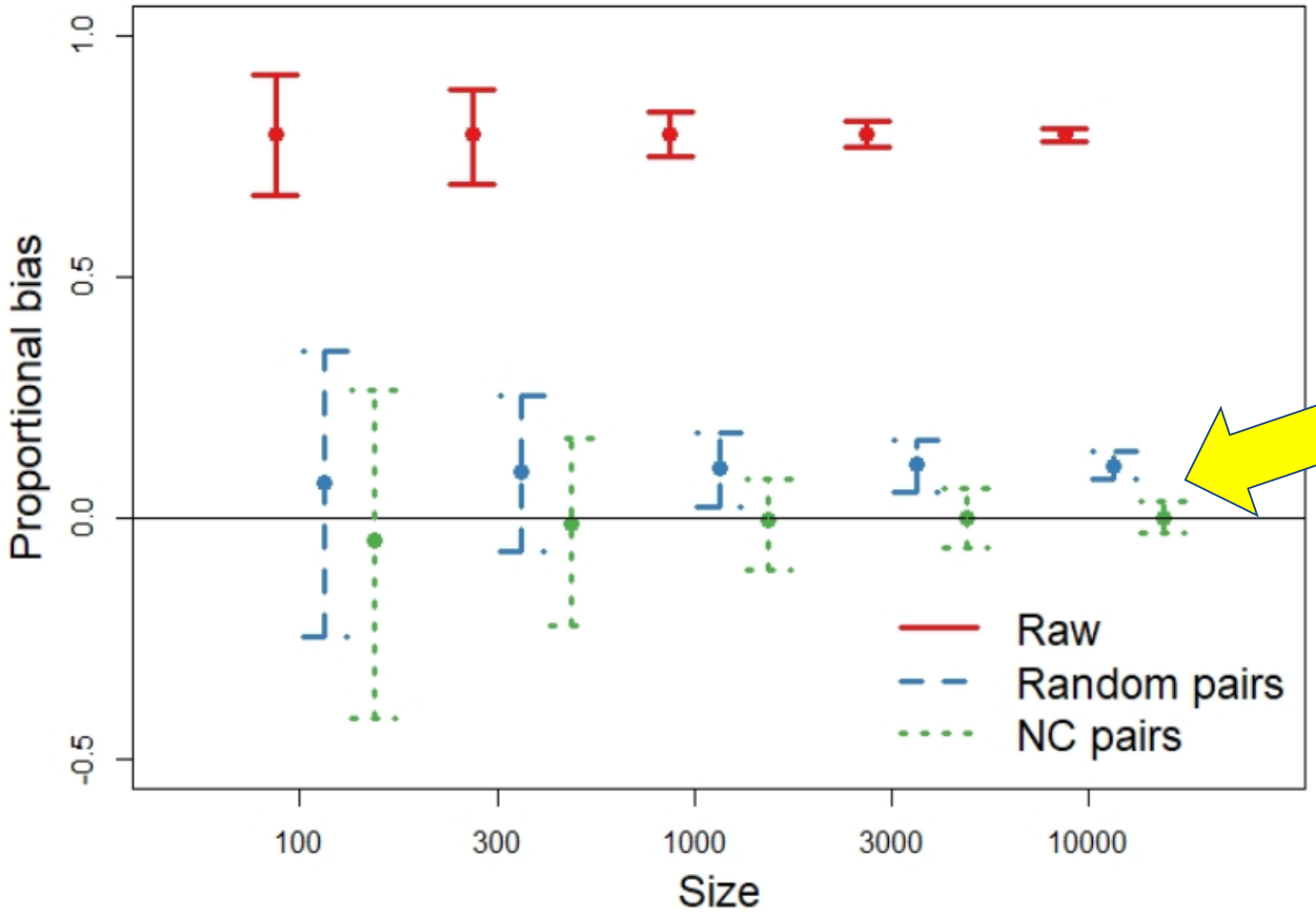
$DNCTs \leftarrow FNC(Data, CandidateNegativeControls, T, O, vt, alpha)$

$\langle pointestimate, confidenceinterval \rangle \leftarrow ANCE(Data, DNCTs, T, O, \theta)$

$return(\langle pointestimate, confidenceinterval \rangle)$

DANCE works very well

Test all candidate DNC triplets, only use validated pairs



Random pairs can be <u>biased</u>. **V**alidated pairs are <u>not</u> biased.

Pros and Cons

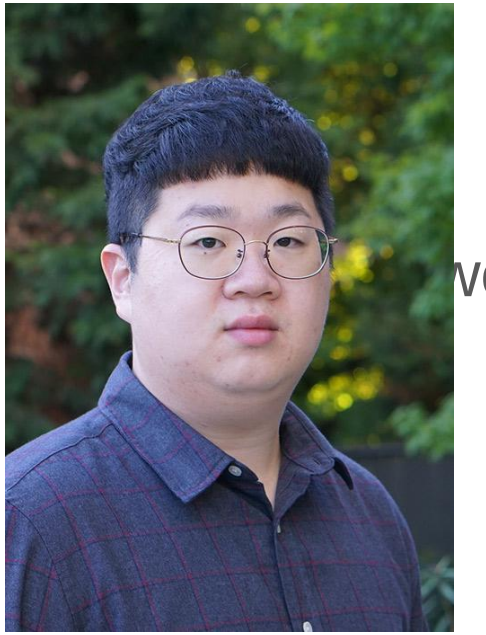
- Cons:
  - Need DNC triplets for validation, when only pairs required for inference
  - DNCs are more restrictive than other NC types
- Pros:
  - Statistical validation of the NCs
  - Can be automated on large set of candidate NCs
  - Potentially identifies many valid NCs

Other uses for set of valid DNC triplets?

- DANCE algorithm computes aggregation of estimates from validated DNC triplets
- Could a large set of validated DNC triplets be used in other ways….??

Thanks to my collaborators!

Xu Shi



won Lim

UNIVERSITY OF MINNESOTA

Driven to Discover®

Crookston Duluth Morris Rochester Twin Cities

An Introduction to Negative Controls

Eric Tchetgen Tchetgen

University of Pennsylvania

March 8, 2023

Acknowledgment

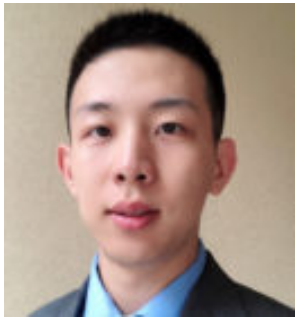
Yifan Cui
National U of Singapore

Oliver Dukes
U of Pennsylvania

Kendrick Li
U of Michigan

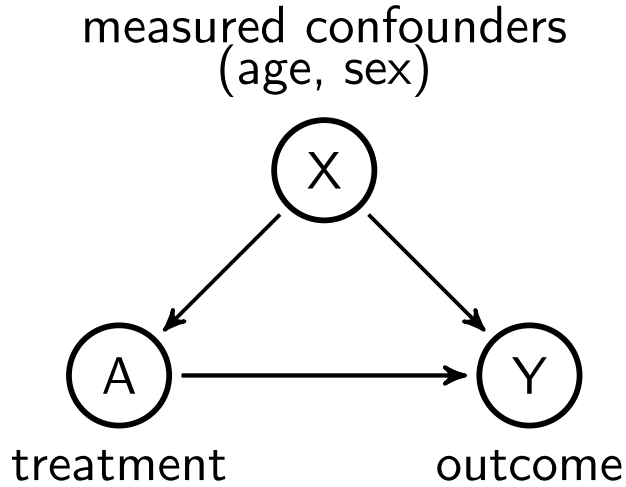
Wang Miao
Peking U

Xu Shi
U of Michigan



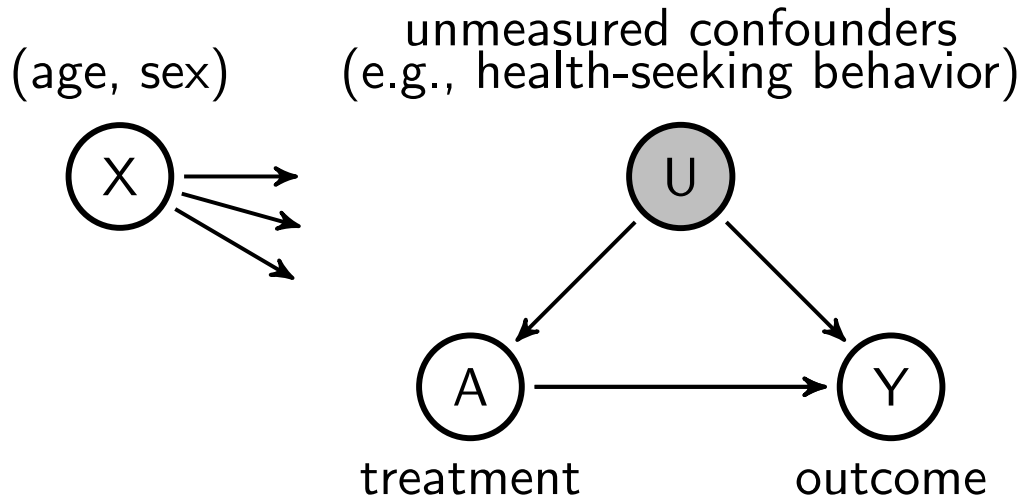
Andrew Ying
U of Pennsylvania

The "randomized" scenario in causal inference



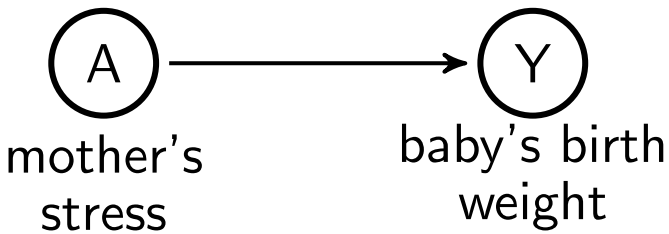
- Estimand: the average treatment effect ATE $= E[Y(1)] - E[Y(0)]$
- Key identification assumption: no unmeasured confounding
  - "Randomized" within each stratum of X
  - Not empirically verifiable

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

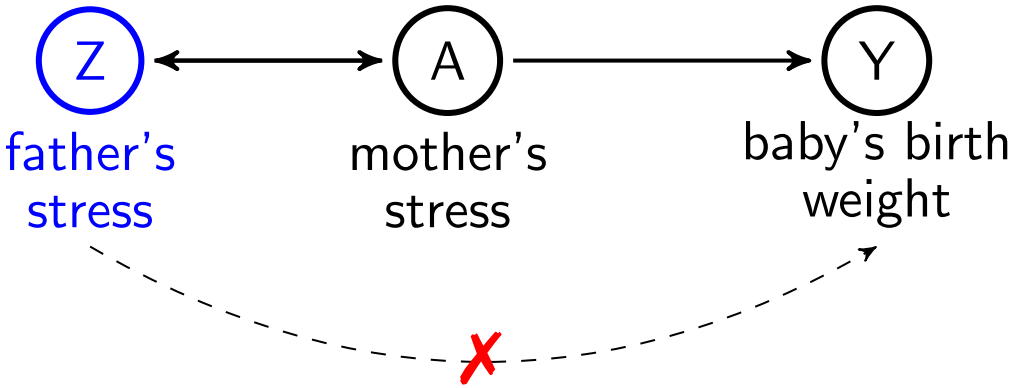
Does stress during pregnancy affect birth weight?



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

Davey Smith 2008, 2012

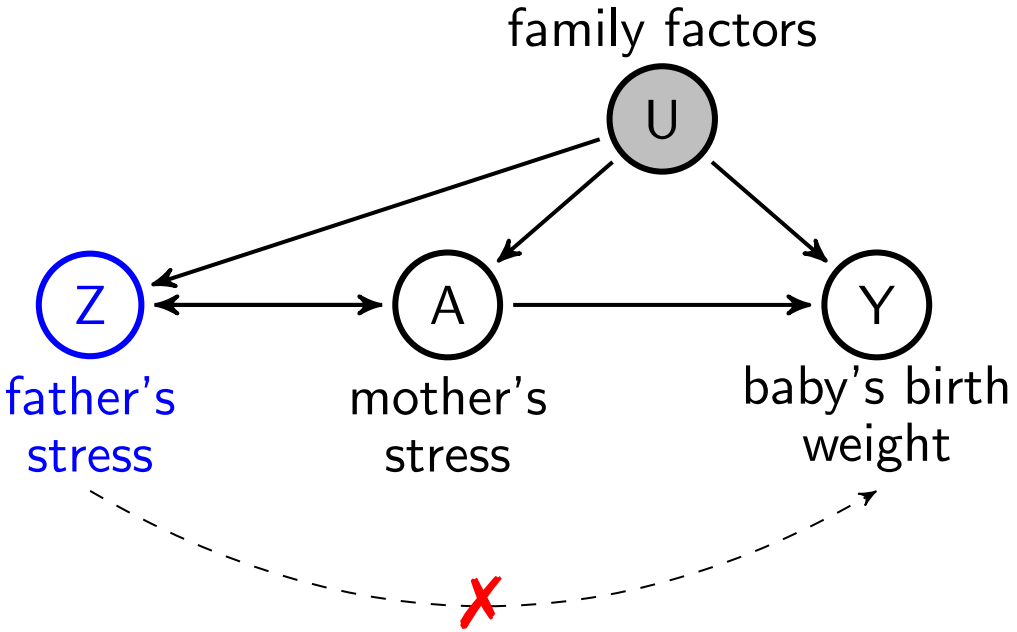
Does stress during pregnancy affect birth weight?



- 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

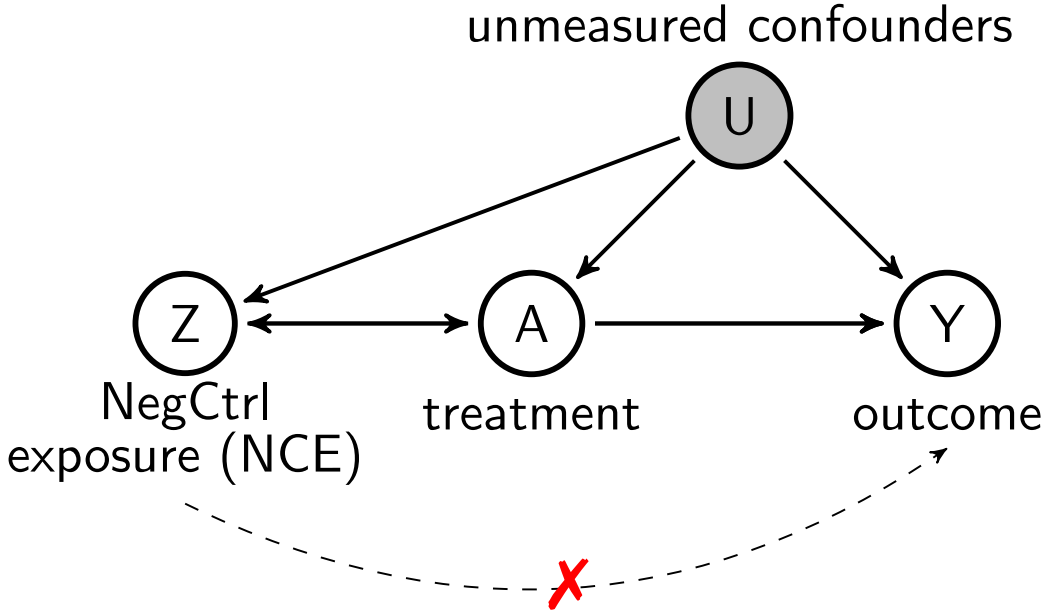
Does stress during pregnancy affect birth weight?



- 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

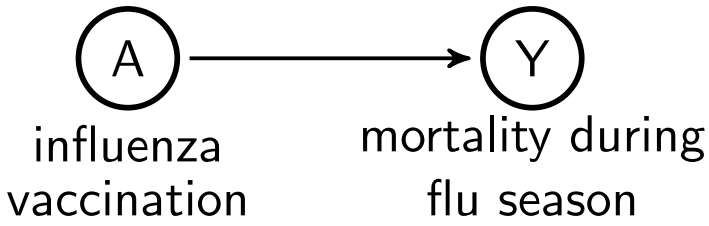
Davey Smith 2008, 2012

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

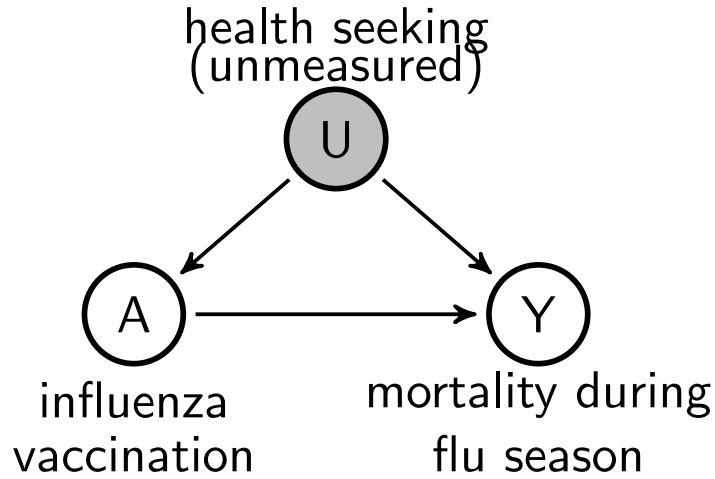
Does flu shot prevent 50% death in the elderly?



- 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?

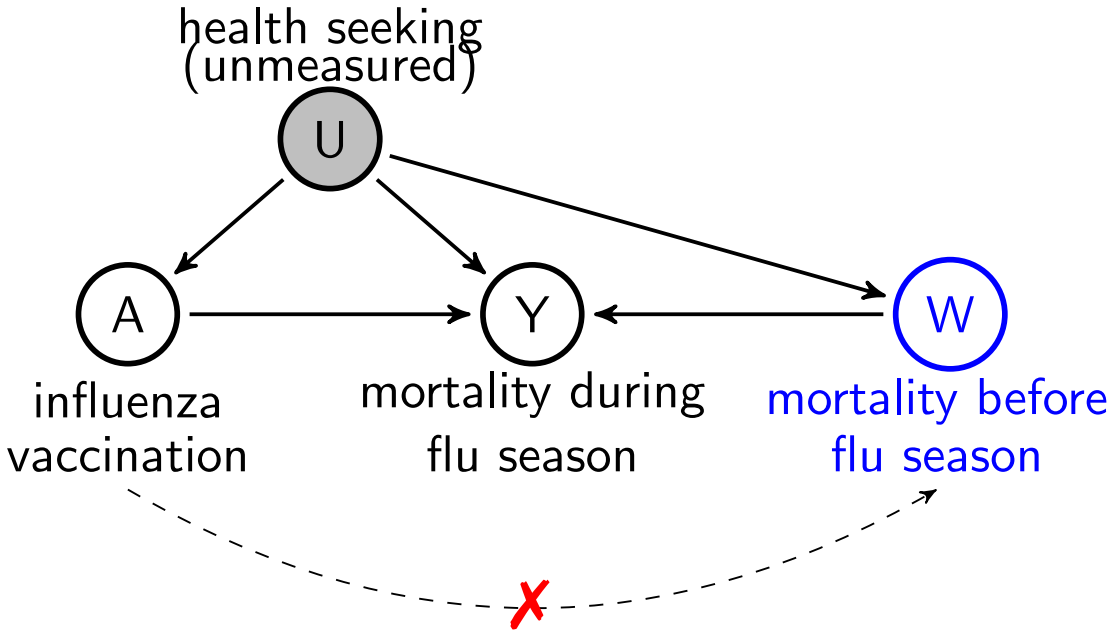


- Observational study on flu vaccine effectiveness
  - found 50% reduction in risk of all cause mortality during winter
- Potential unmeasured confounding by health seeking behavior

---

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

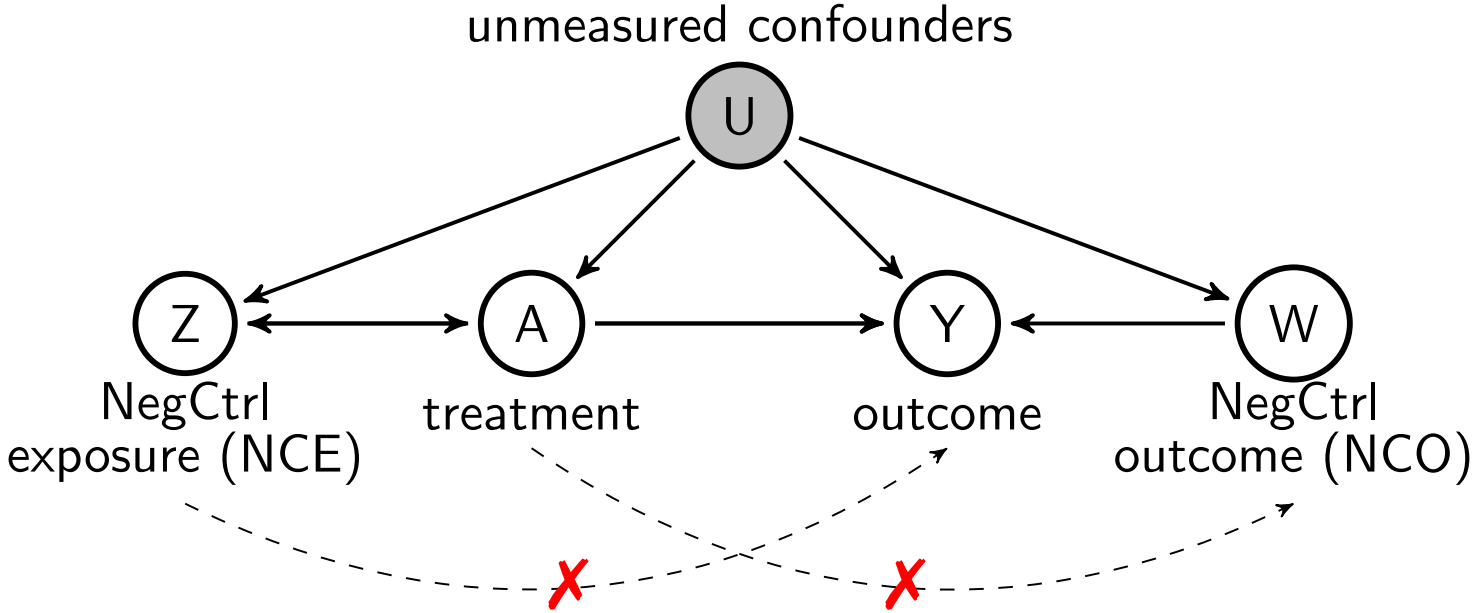
Does flu shot prevent 50% death in the elderly?



- 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

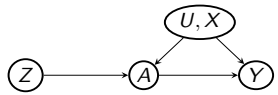
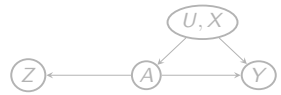
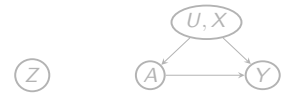
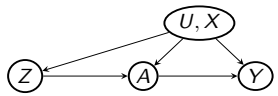
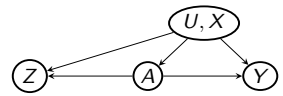
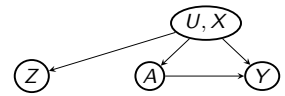
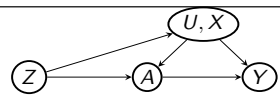
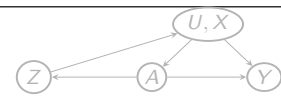
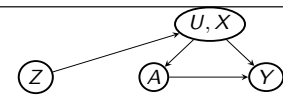
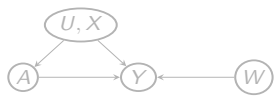
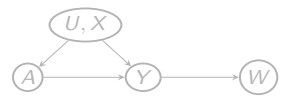
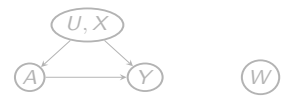
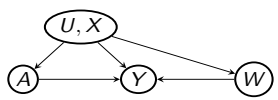
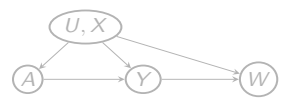
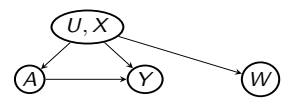
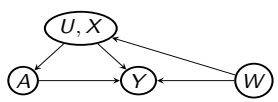
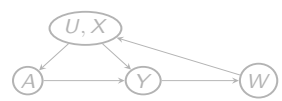
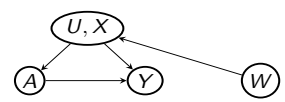
Jackson et al. 2006; also considered using injury/trauma hospitalization to detect unmeasured confounding bias

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

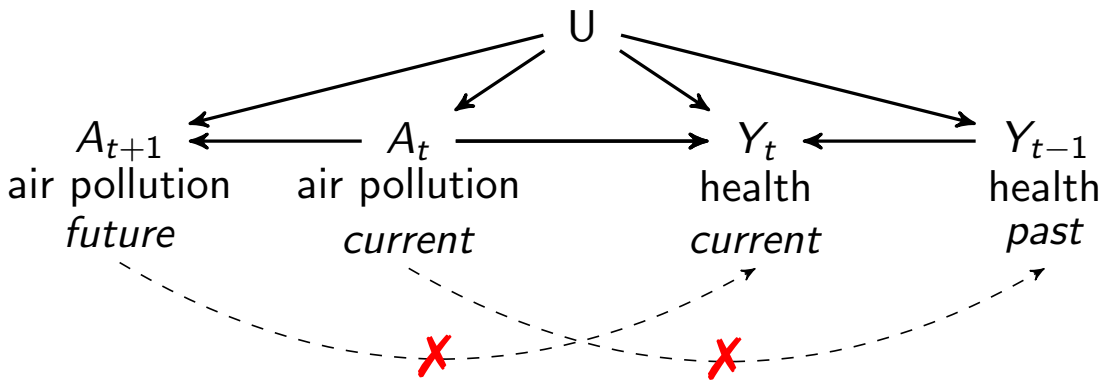
More examples that encode the NC assumptions

Examples of NCE			
	$Z \to A$ (pre-treatment)	$A \to Z$ (post-treatment)	$Z \perp\!\!\!\perp A$
No arrow between U and Z (may violate U-comparability)	Instrumental variable (IV) 		
$U \to Z$	Invalid IV 	Post-treatment proxy of U 	Proxy of U 
$Z \to U$	May violate Assumptions if there is $W \to U$ 		
Examples of NCO			
	$W \to Y(a)$	$Y(a) \to W$	$Y(a) \perp\!\!\!\perp W \mid (U, X)$
No arrow between U and W (violate U-comparability)		Violate NCO definition 	
$U \to W$		Violate NCO definition 	
$W \to U$	May violate Assumption if there is $Z \to U$ 	Violate NCO definition 	

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

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; Tchetgen Tchetgen, Park, and Richardson 2023.

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

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

Control Outcome Calibration Approach

- We presume that one has measured a valid so-called negative control outcome, or more broadly a proxy W, which is known a priori to satisfy two conditions.
  - First W must be associated with the mechanism leading to residual confounding in the sense that it stands for a proxy for the latter;
  - Secondly, W cannot be causally impacted by the treatment of interest.

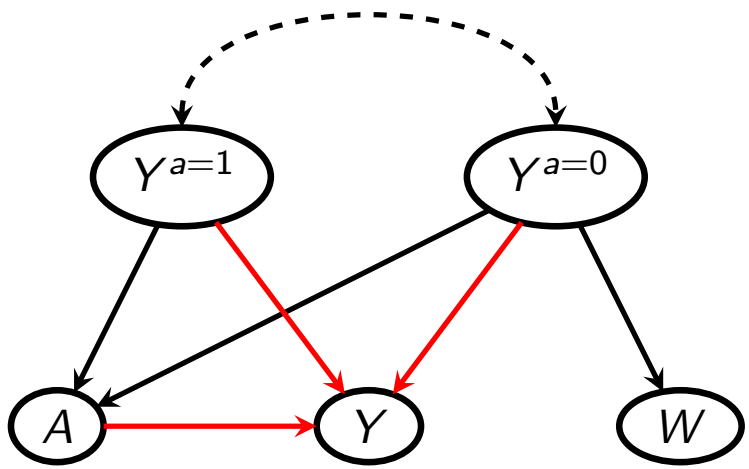


Figure: Red arrows depict the deterministic relationship between Y and $(Y^{a=1}, Y^{a=0}, A)$, as established by the consistency assumption $Y = AY^{a=1} + (1 - A)Y^{a=0}$.

Control Outcome Calibration Approach

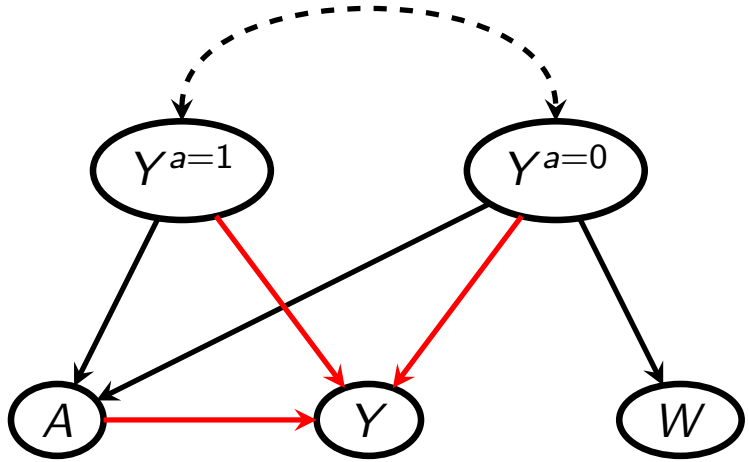


Figure: Red arrows depict the deterministic relationship between Y and $(Y^{a=1}, Y^{a=0}, A)$, as established by the consistency assumption $Y = AY^{a=1} + (1 - A)Y^{a=0}$.

- Inspired by Tchetgen Tchetgen (2014), we formalize these conditions as followed:

$$W^a = W; \tag{1}$$

where W^a is the potential negative control outcome under an external intervention that sets $A = a$; and

$$W \amalg Y^{a=0}; \tag{2}$$

$$W \amalg A | Y^{a=0}. \tag{3}$$

Control Outcome Calibration Approach

- For identification and estimation in the case of a continuous exposure, Tchetgen Tchetgen (2014) further assumed the rank preserving structural model:

$$Y = Y^{a=0} + \psi_0 A, \tag{4}$$

which by consistency, implies constant individual level causal effect $\psi_0 = Y^{a=1} - Y^{a=0}$.

- Under this model, he noted that upon defining $Y(\psi) = Y - \psi A$, then one can deduce from conditions (1)-(3) that

$$W \amalg A | Y(\psi)$$

if and only if $\psi = \psi_0$.

- Then one can identify ψ_0 by fitting say a linear regression of W on $(Y(\psi), A)$ for different values of ψ and finding the value that makes the coefficient for A null. In Tchetgen Tchetgen, Park, and Richardson 2023 we establish nonparametric identification using COCA without assumption of constant treatment effect.

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
- Select work by the proximal causal inference group

Bibliography I

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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/.

Bibliography II

[11] 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.

[12] 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.

[13] 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.

[14] 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.**

[15] 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.

[16] 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.**

[17] 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.

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

[19] 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.

[20] 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`.

Bibliography III

[21] Xu Shi, Wang Miao, and Eric J Tchetgen Tchetgen. "A selective review of negative control methods in epidemiology". In: Current Epidemiology Reports (2020), pp. 1–13.

[22] Eric Tchetgen Tchetgen, Chan Park, and David Richardson. "Single Proxy Control". In: arXiv:2302.06054 (2023).

Questions?

Double Negative Control Methods and Proximal Causal Inference

Xu Shi

Department of Biostatistics
University of Michigan

Background

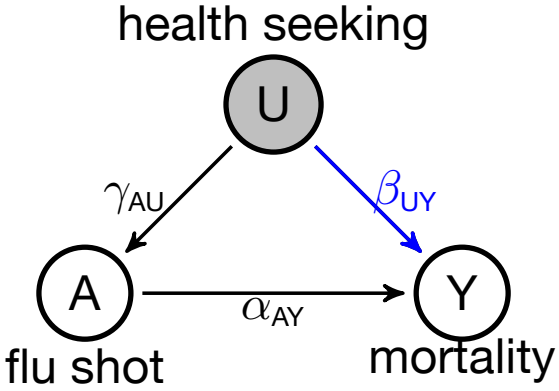
- Gaps in literature
  - Mainly focus on bias detection and bias attenuation[1]
  - Estimation relies on relatively strong assumptions[2]
  - Use NCE alone or NCO alone
- The double negative control method
  - Use one NCE and one NCO to completely remove bias
  - Nonparametric identification without model restrictions

[1] Rosenbaum 1989; Lipsitch, Tchetgen Tchetgen, and Cohen 2010; Flanders et al. 2011; Lipsitch, Tchetgen Tchetgen, and Cohen 2012; Flanders, Strickland, and Klein 2017

[2] Gagnon-Bartsch and Speed 2012; Tchetgen Tchetgen 2014; Sofer et al. 2016

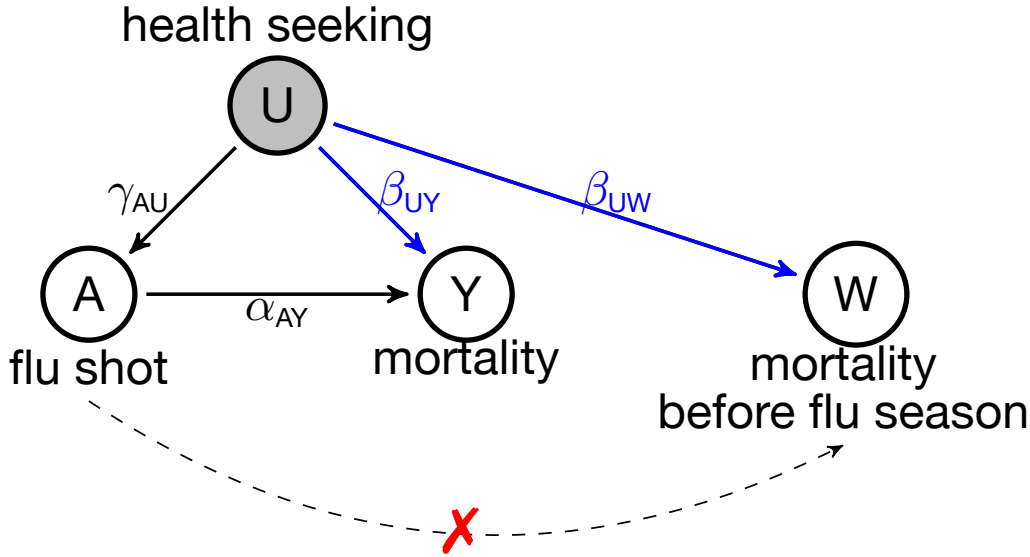
Double Negative Control Method

Intuition behind identification under linear additive models



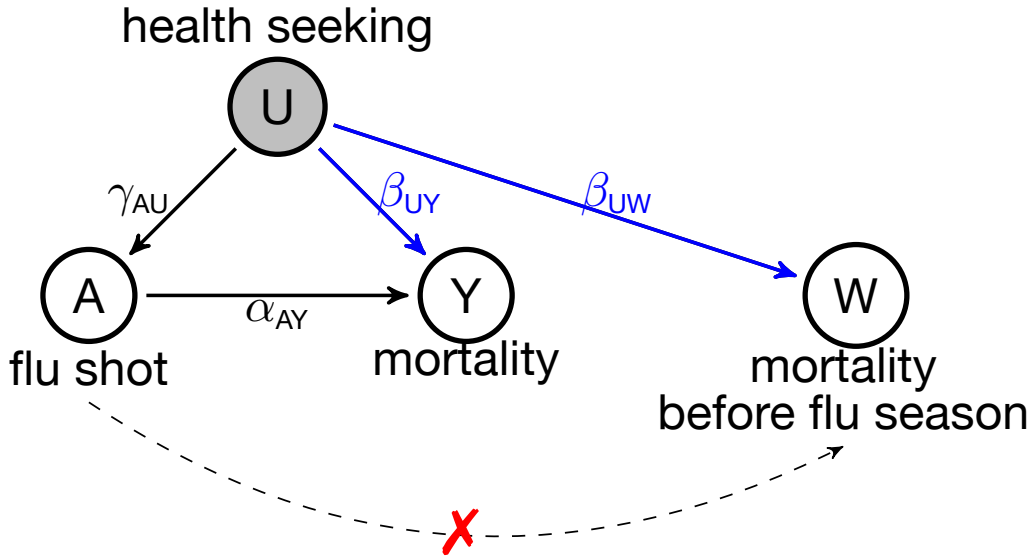
- Confounding bias is a product of *U*-*A* and *U*-*Y* association ($\gamma_{AU}\beta_{UY}$)

Intuition behind identification under linear additive models



- 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

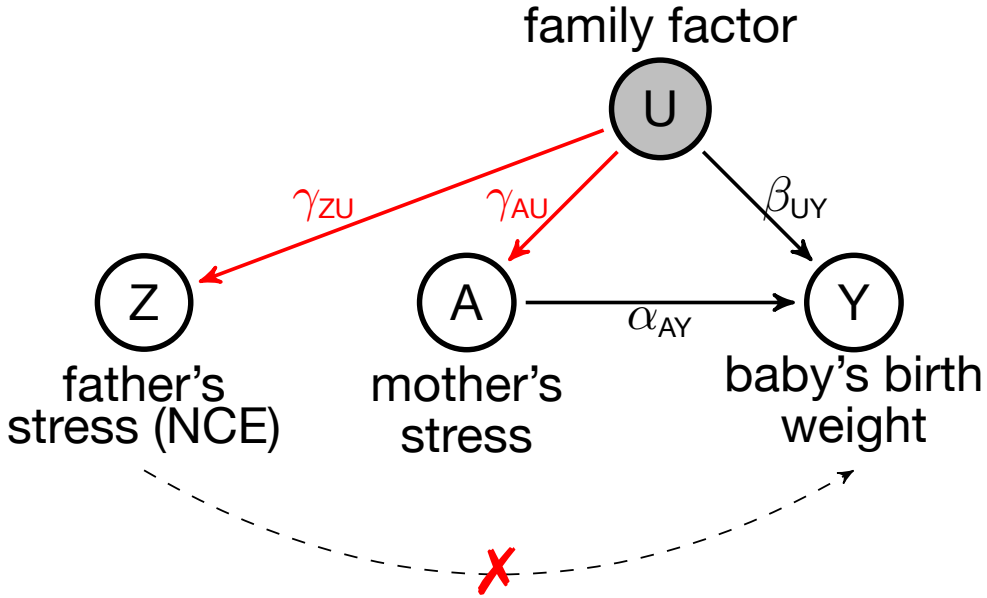
Intuition behind identification under linear additive models



- 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

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

Intuition behind identification under linear additive models

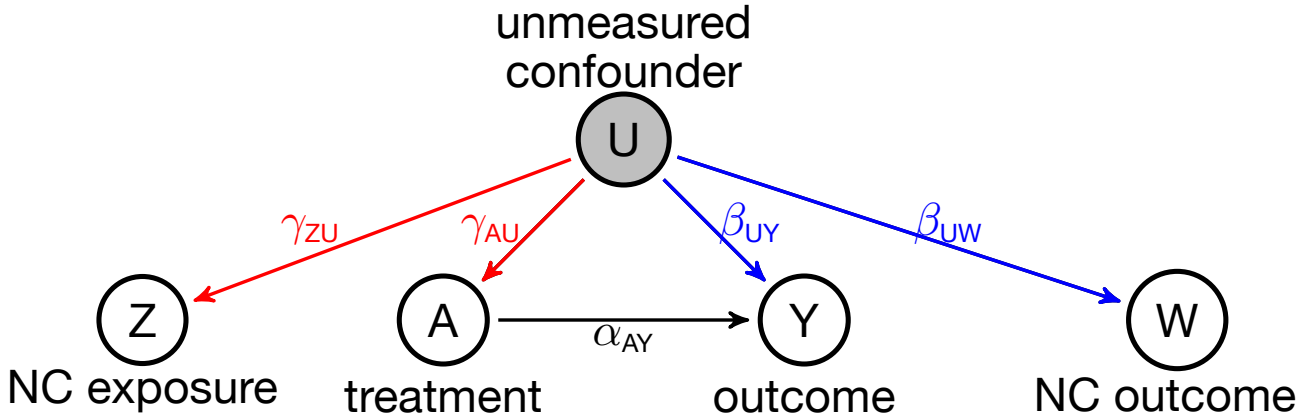


- Confounding bias is a product of *U-A* and *U-Y* association ($\gamma_{AU}\beta_{UY}$)
  - *Z-Y* association is a product of *U-Z* and *U-Y* association ($\gamma_{ZU}\beta_{UY}$)
  - Problem solved if *U-A* association = *U-Z* association
- Regress $\underbrace{Y \text{ on } A}_{\alpha_{AY} + \gamma_{AU}\beta_{UY}}$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

γ_{ZU} is the coef of Z in $E[U \mid Z]$

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

Intuition behind identification under linear additive models

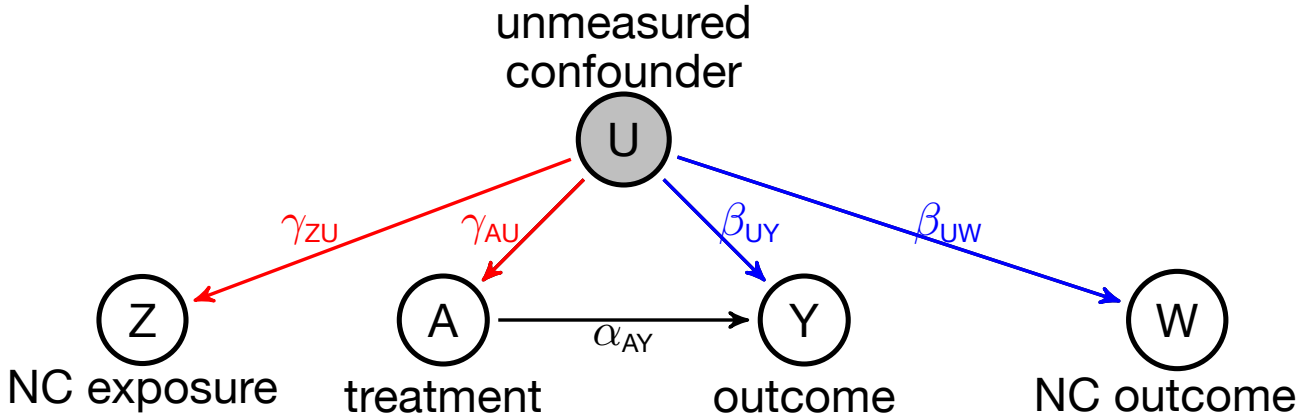


- What if $\beta_{UY} \neq \beta_{UW}$ and $\gamma_{AU} \neq \gamma_{ZU}$?
  - *A*-*W* association ($\gamma_{AU}\beta_{UW}$) recovers the confounding bias ($\gamma_{AU}\beta_{UY}$) up to a scale $\frac{\beta_{UY}}{\beta_{UW}}$
  - We cannot identify either β_{UY} or β_{UW}, but can we identify the ratio?

γ_{AU} and γ_{ZU} are the coefs of A and Z in $E[U \mid A, Z]$, respectively

Miao, Geng, and Tchetgen Tchetgen 2018; Miao, Shi, and Tchetgen Tchetgen 2019; Shi, Miao, and Tchetgen Tchetgen 2020a,b; Tchetgen et al. 2020

Intuition behind identification under linear additive models

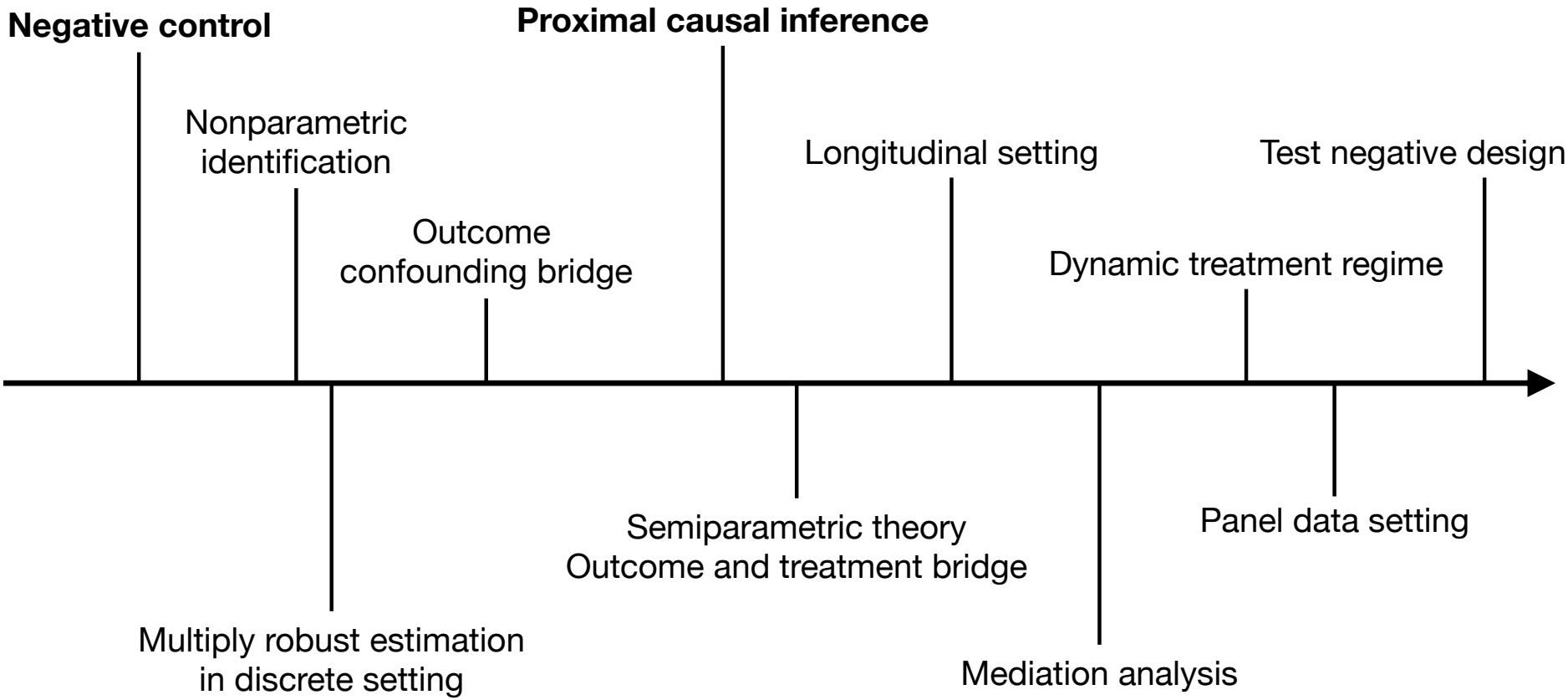


- What if $\beta_{UY} \neq \beta_{UW}$ and $\gamma_{AU} \neq \gamma_{ZU}$?
  - *A*-*W* association ($\gamma_{AU}\beta_{UW}$) recovers the confounding bias ($\gamma_{AU}\beta_{UY}$) up to a scale $\frac{\beta_{UY}}{\beta_{UW}}$
  - We cannot identify either β_{UY} or β_{UW}, but can we identify the ratio?
- Double negative control: use both an NCE and an NCO
  - 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
  - $\gamma_{ZU}\beta_{UW} \neq 0$ requires *Z* and *W* to be associated with *U*

γ_{AU} and γ_{ZU} are the coefs of *A* and *Z* in $E[U \mid A, Z]$, respectively

Miao, Geng, and Tchetgen Tchetgen 2018; Miao, Shi, and Tchetgen Tchetgen 2019; Shi, Miao, and Tchetgen Tchetgen 2020a,b; Tchetgen et al. 2020

The double negative control literature



Lipsitch, Tchetgen Tchetgen, and Cohen 2010; Miao, Geng, and Tchetgen Tchetgen 2018; Miao, Shi, and Tchetgen Tchetgen 2019; Cui et al. 2020; Shi et al. 2020; Shi, Miao, and Tchetgen Tchetgen 2020a; Tchetgen et al. 2020; Dukes, Shpitser, and Tchetgen 2021; Shi et al. 2021; Ying et al. 2021

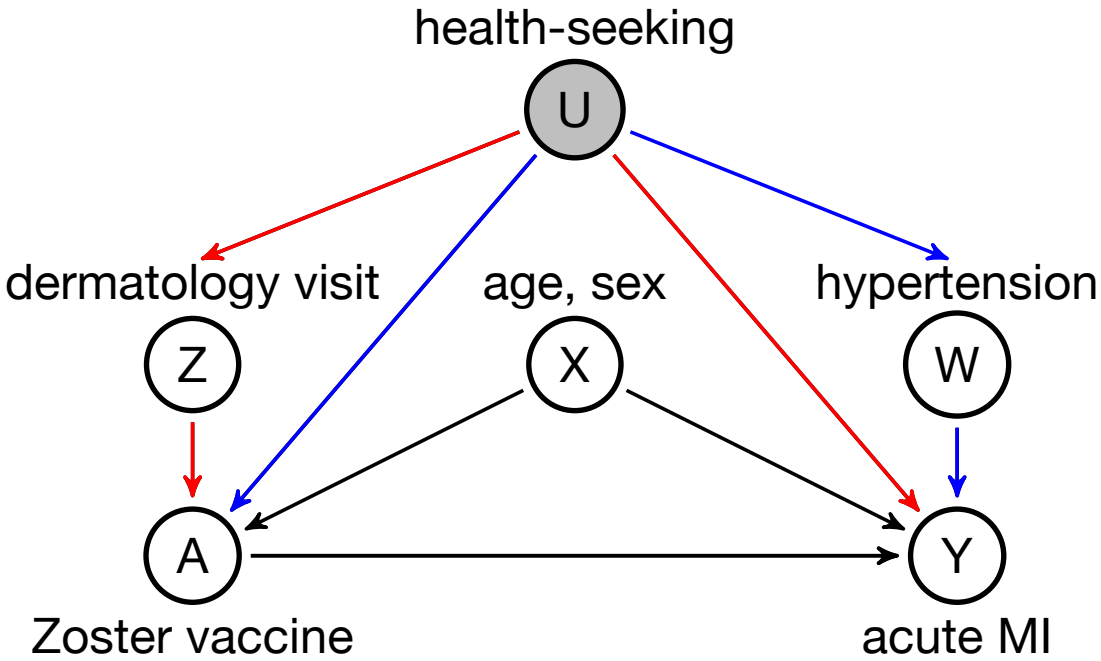
Proximal Causal Learning

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

- The “no unmeasured confounding (NUC)” assumption depends on investigator’s ability to accurately measure covariates
- Hard to eliminate measurement error
  - Covariates are at best proxies of true confounders
- Easier to get to the right kind of measurement error
  - Leverage different types of proxies without assuming NUC

Browning and Crossley 2009

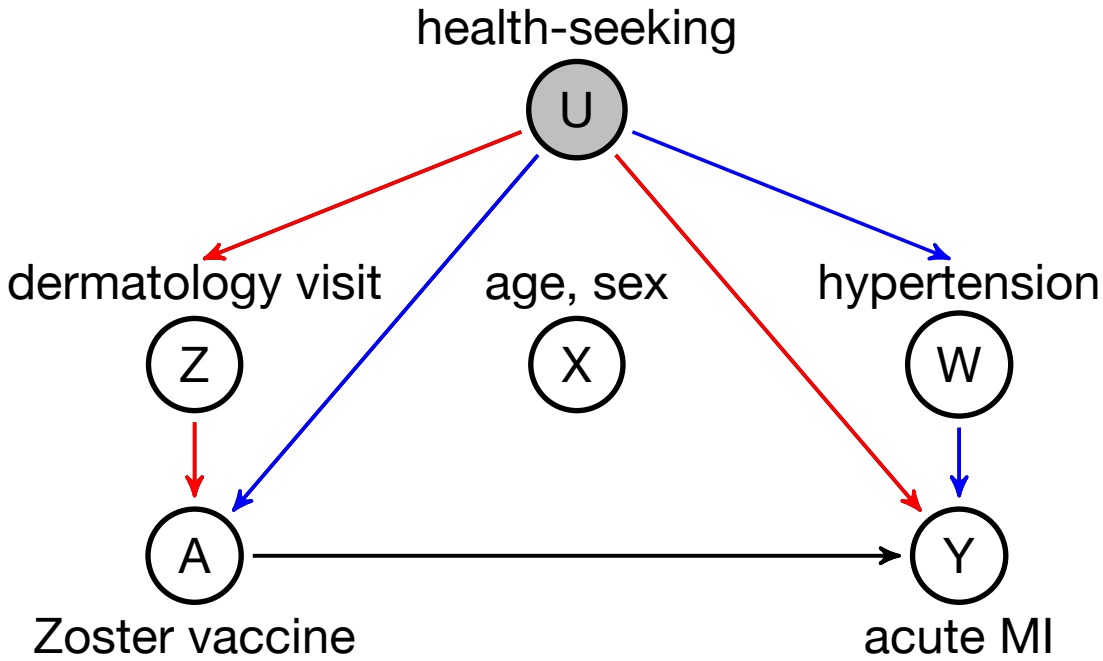
An example in vaccine safety study



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

Tchetgen et al. 2020

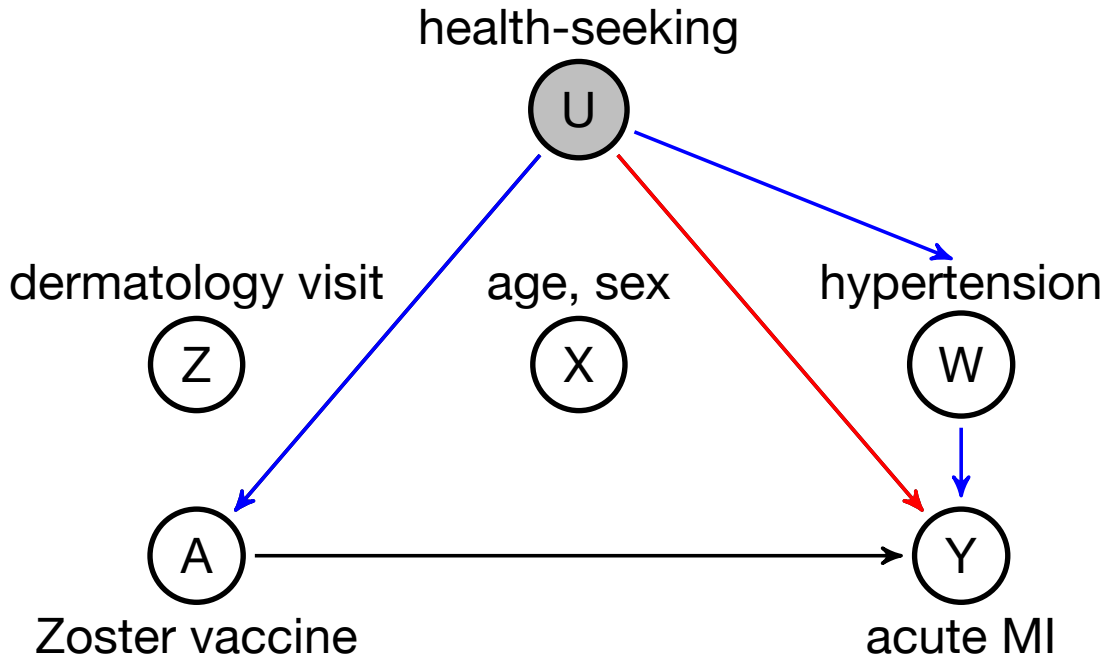
An example in vaccine safety study



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

Tchetgen et al. 2020

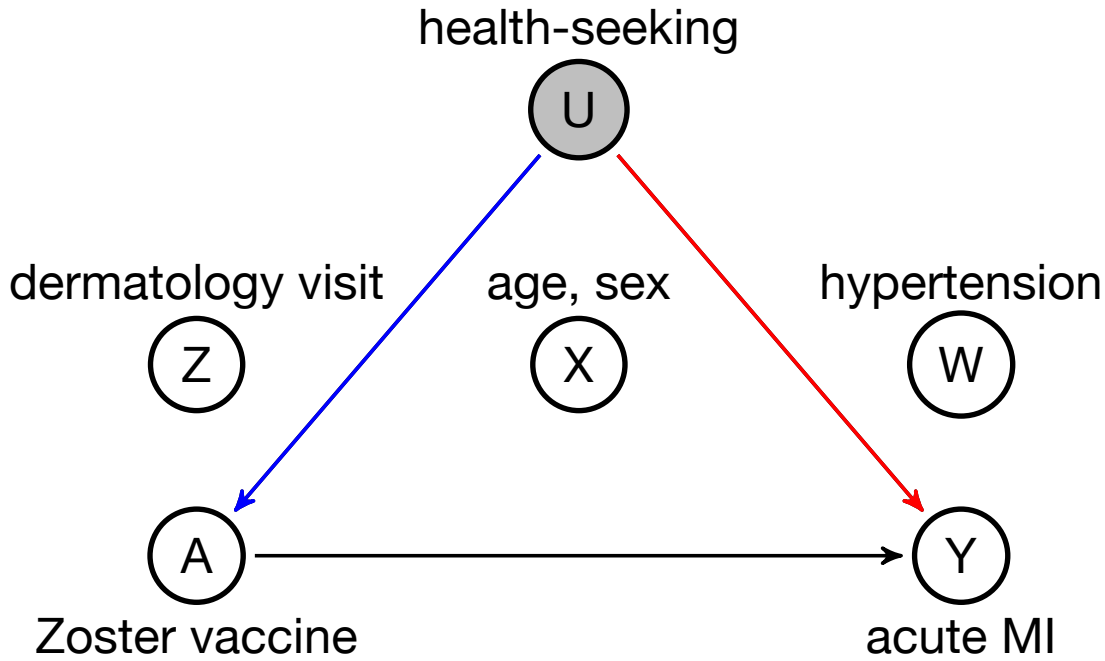
An example in vaccine safety study



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

Tchetgen et al. 2020

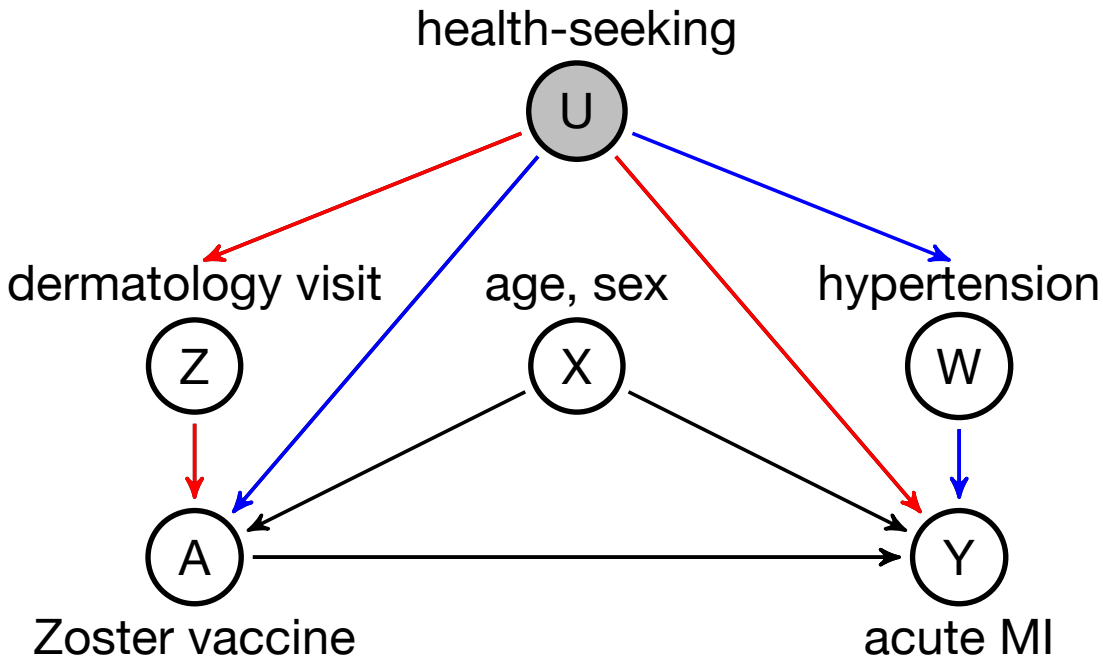
An example in vaccine safety study



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

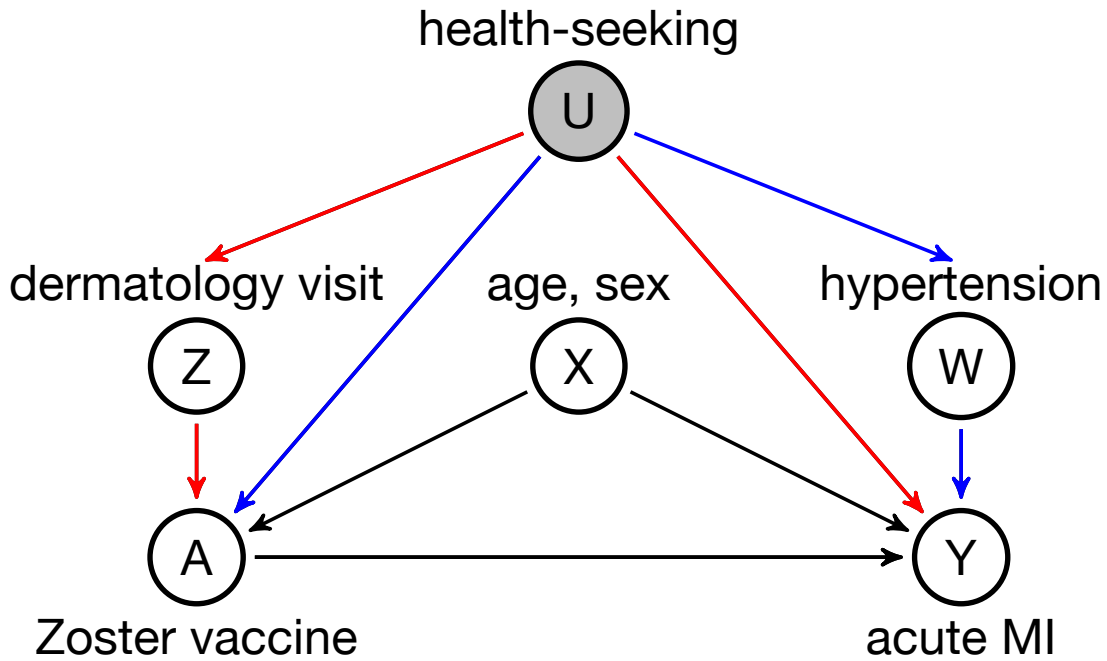
Tchetgen et al. 2020

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

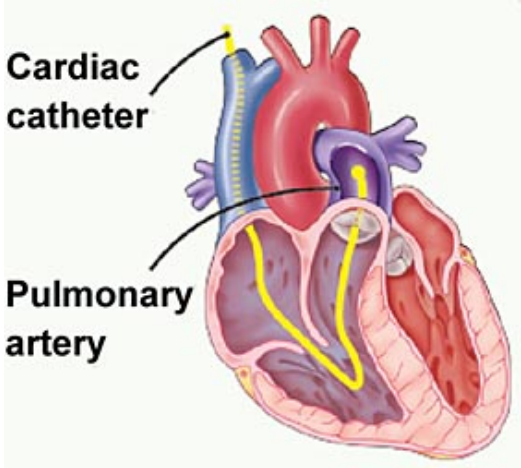
Covariates	**Classical**	Treat all equally as confounders {X,W,Z}: age, sex, hypertension, dermatology visit
	Proximal	Divide into three buckets X=age, sex; W=hypertension, Z=dermatology visit
g-formula	**Classical**	$E[Y(a)] = E[m(a, W, X)]$ $m()$ is the outcome model
	Proximal	$E[Y(a)] = E[h(a, W, X)]$ $h()$ is the outcome confounding bridge function

Discussion

- We can go beyond bias detection and adjust for confounding bias leveraging the double negative control
  - NCO recovers bias up to a scale; NCE recovers that scale
  - Requires NCs to be sufficiently informative about confounding
- Proximal causal inference is a generalization of classical methods
  - Acknowledge that covariates are imperfect proxies of confounders
  - Leverage proxies to adjust for suspected unmeasured confounding

Questions?

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

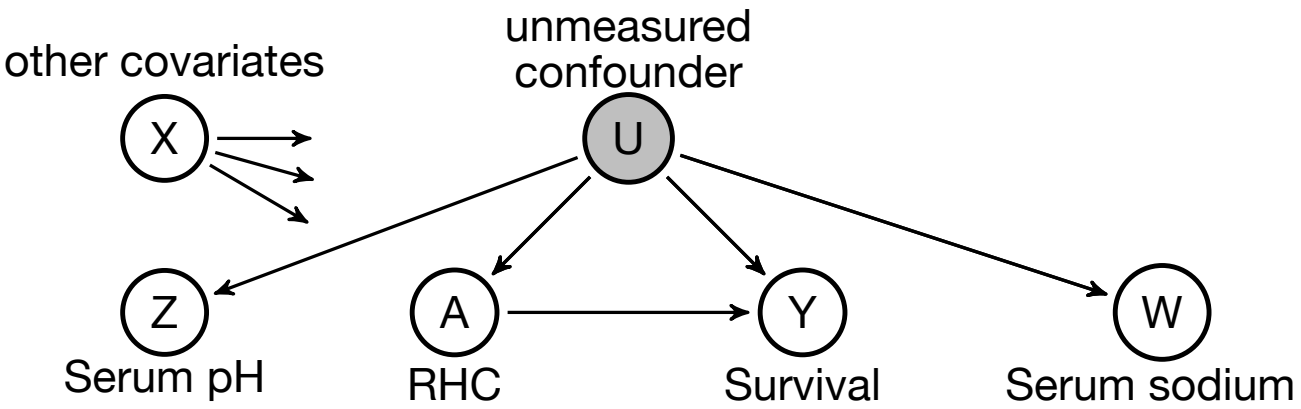
A controversial result

- 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

---

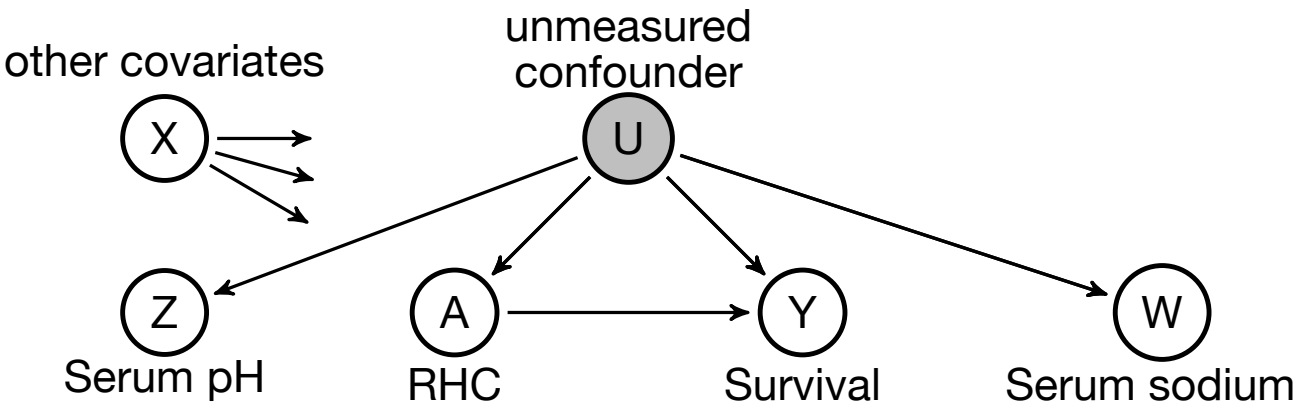
Lin, Psaty, and Kronmal 1998; Tan 2006; Li, Morgan, and Zaslavsky 2018; Mao and Li 2020

Candidate proxies in the SUPPORT study



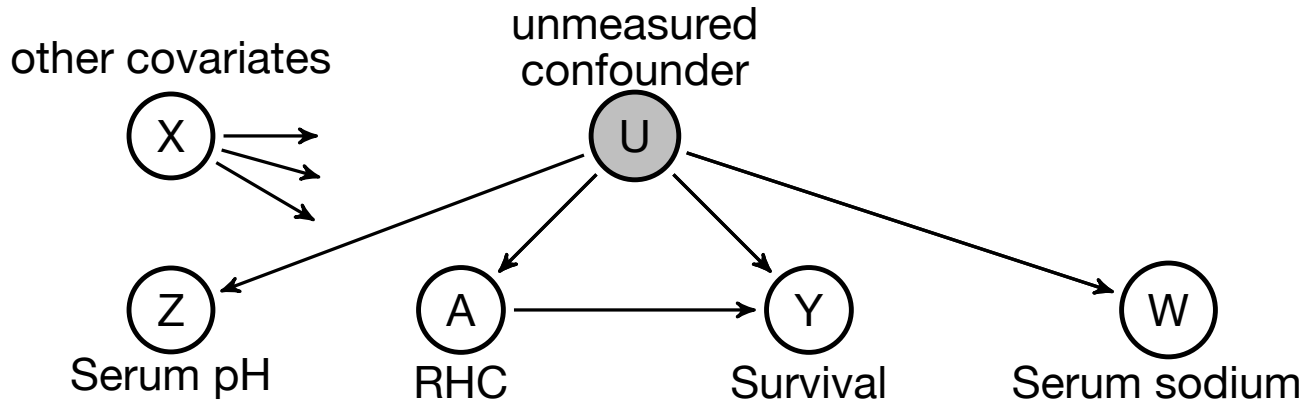
- The SUPPORT study collected 72 covariates including
  - demographics, comorbidity, vital signs, functional status
  - physiological status measured from a blood test during the initial 24 hours in the ICU $\Rightarrow$ 10 candidate proxies
- We applied our DANCE algorithm to find valid proxies
  -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

Bibliography I

[1] Paul R Rosenbaum. “The role of known effects in observational studies”. In: Biometrics 45.2 (1989), pp. 557–569.

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

[5] 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.

[6] Marc Lipsitch, Eric J Tchetgen Tchetgen, and Ted Cohen. “Negative controls: a tool for detecting confounding and bias in observational studies”. In: Epidemiology 21.3 (2010), pp. 383–388.

[7] 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.

[8] 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.

[9] Marc Lipsitch, Eric J Tchetgen Tchetgen, and Ted Cohen. “Negative control exposures in epidemiologic studies (author reply)”. In: Epidemiology 23.2 (2012), pp. 351–352.

[10] 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.

[11] 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.

Bibliography II

[12] 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.

[13] 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/.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] Wang Miao, Xu Shi, and Eric J Tchetgen Tchetgen. “A Confounding Bridge Approach for Double Negative Control Inference on Causal Effects”. In: (2019). In progress, a prior version can be found at https://arxiv.org/abs/1808.04945.

[20] Yifan Cui, Hongming Pu, Xu Shi, Wang Miao, and Eric J Tchetgen Tchetgen. “Semiparametric proximal causal inference”. In: arXiv preprint arXiv:2011.08411 (2020).

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

Bibliography III

[22] Xu Shi, Wang Miao, Jennifer C Nelson, 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.

[23] Xu Shi, Wang Miao, and Eric J Tchetgen Tchetgen. "A selective review of negative control methods in epidemiology". In: Current Epidemiology Reports (2020), pp. 1–13.

[24] Xu Shi, Wang Miao, and Eric J Tchetgen Tchetgen. "Multiply robust causal inference with double negative control adjustment for unmeasured confounding". In: Journal of the Royal Statistical Society: Series B (Statistical Methodology) (2020). In press, online version available at `https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/rssb.12361`.

[25] 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.**

[26] Oliver Dukes, Ilya Shpitser, and Eric J Tchetgen Tchetgen. "Proximal mediation analysis". In: arXiv preprint arXiv:2109.11904 (2021).

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

[28] Andrew Ying, Wang Miao, Xu Shi, and Eric J Tchetgen Tchetgen. "Proximal causal inference for complex longitudinal studies". In: arXiv preprint arXiv:2109.07030 (2021).

Identifying negative control sets to estimate systematic error for study diagnostics and empirical calibration

Patrick Ryan, PhD

Vice President, Observational Health Data Analytics, Johnson & Johnson

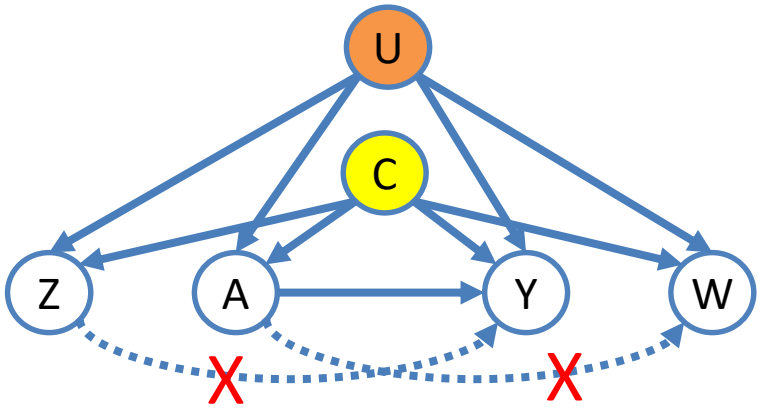
Assistant Professor, Adjunct, Department of Biomedical Informatics, Columbia University Irving Medical Center

Martijn Schuemie, PhD

Senior Director, Observational Health Data Analytics, Johnson & Johnson

Research Fellow, Biostatistics, UCLA

Conceptual model for negative controls



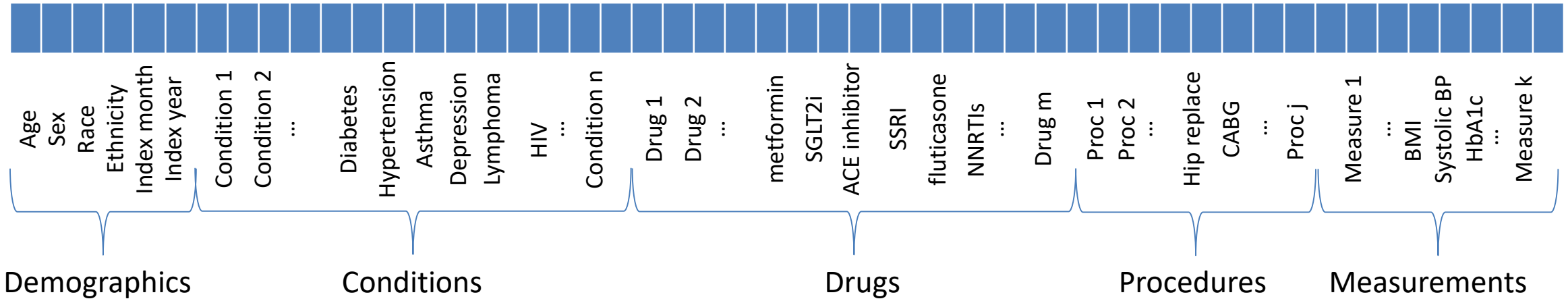
Primary criteria: no causal effect

- Negative control exposure: Z does not cause Y
- Negative control outcome: A does not cause W

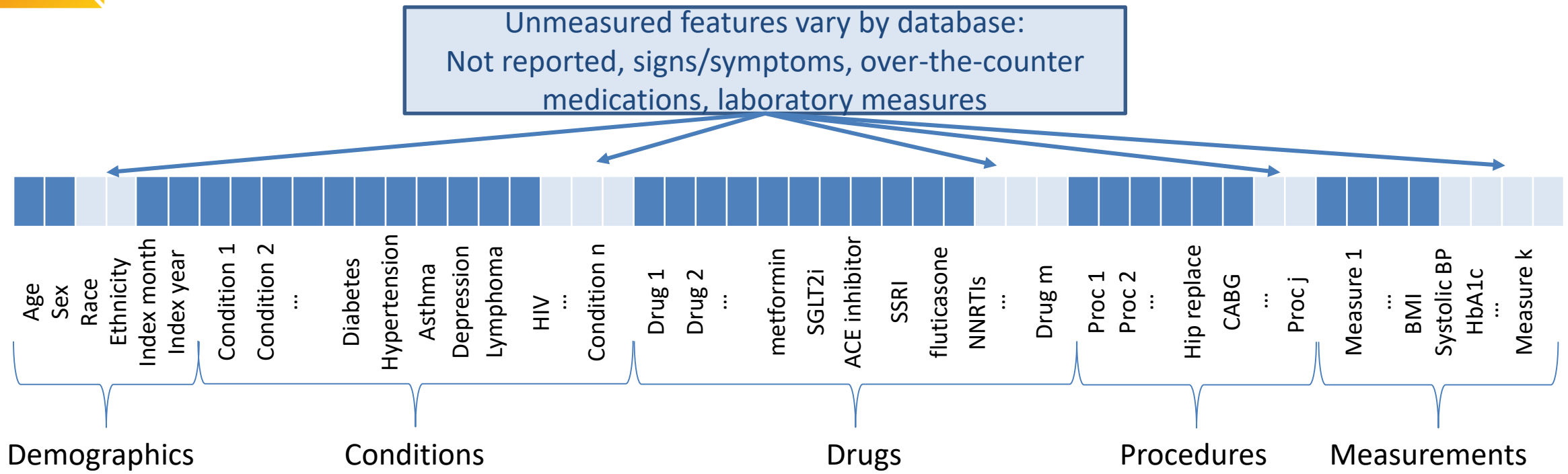
Methods are aiming to find error due to unmeasured confounding U, under the assumption that all measured confounding C has been addressed.

We're going to discuss an approach to find error due to confounding, whether it was explicitly adjusted for or not

Baseline characteristics (candidate confounders) in healthcare data

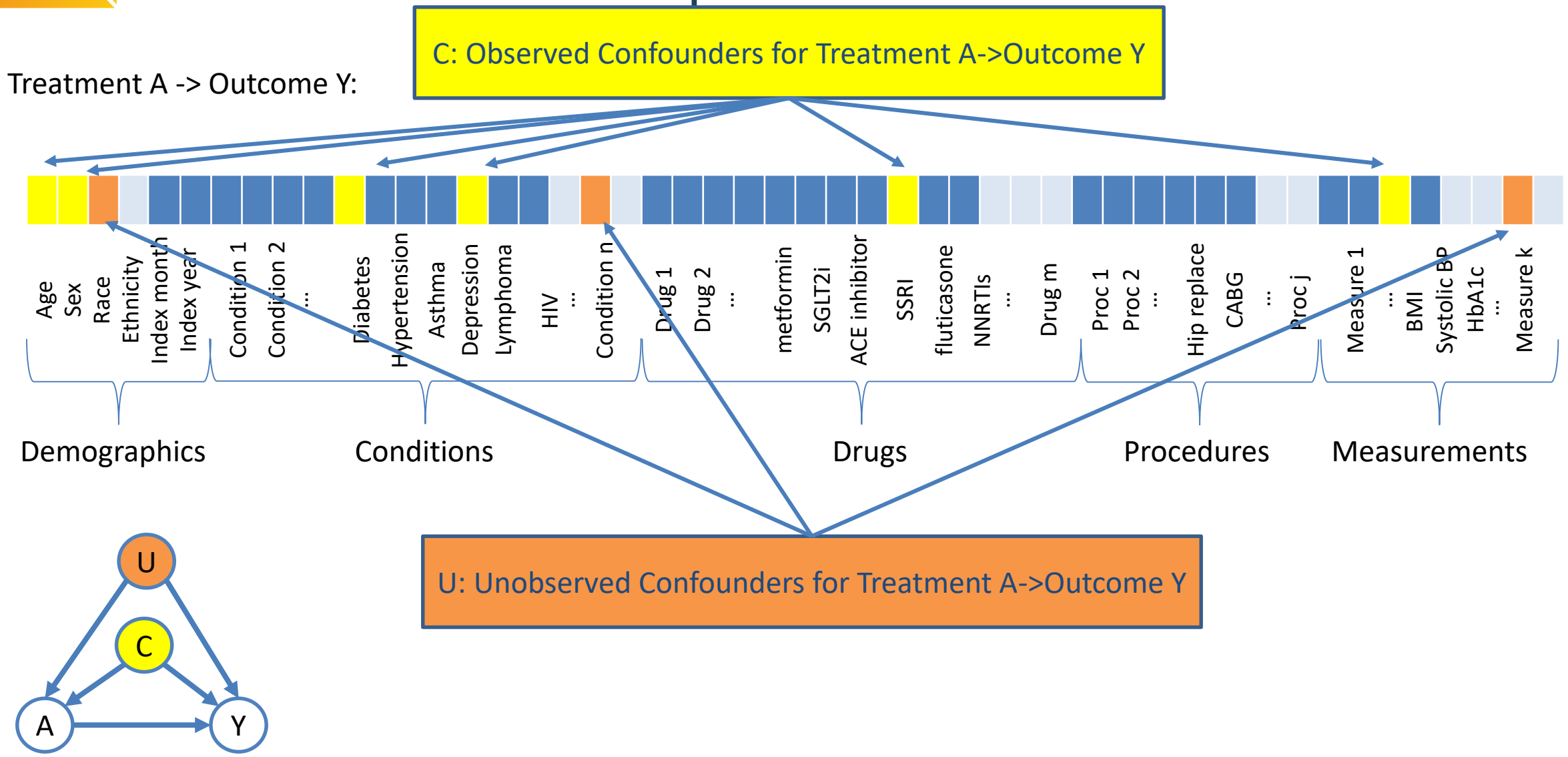


Not all baseline characteristics are measured in all databases



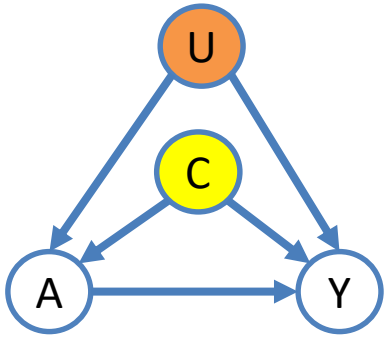
- Most databases have some unmeasured features that are desired for any given analysis
- Measured features often total >10,000s in claims or EHR datasets
- Some unmeasured features may be 'indirectly measured' if correlated with other measured features (Zhang JBI 2022)

Confounders are baseline characteristics associated with exposure and outcome

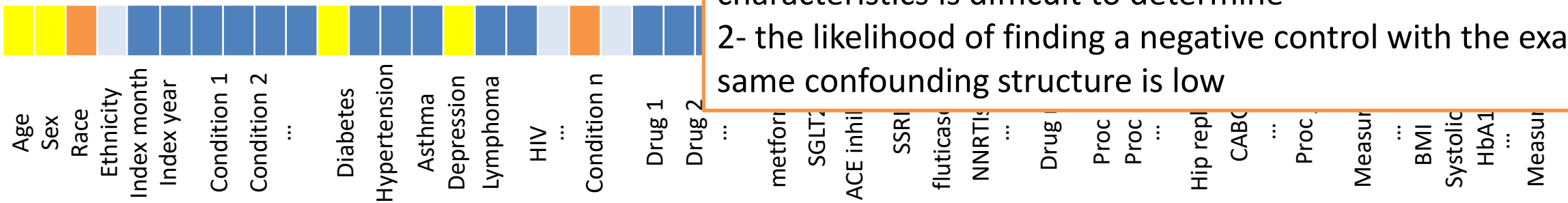


Desired exchangeability of confounding with negative controls

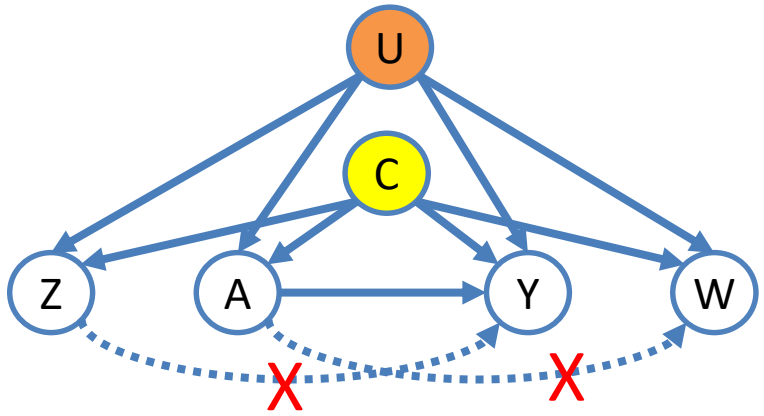
Treatment A -> Outcome Y:



It'd be great to find negative controls that follow this DAG, but:
1- the full confounding structure across all baseline characteristics is difficult to determine
2- the likelihood of finding a negative control with the exact same confounding structure is low

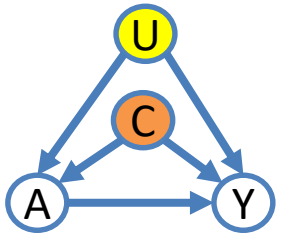


Negative control exposure Z → Outcome Y
OR
Treatment A → Negative control outcome W
OR
Negative control exposure Z → Negative control outcome W

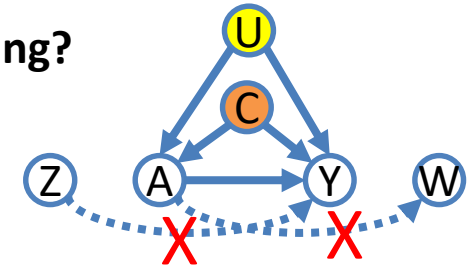


Relaxing assumption about confounding structure, we can still learn about the reliability of our method

Treatment A -> Outcome Y:

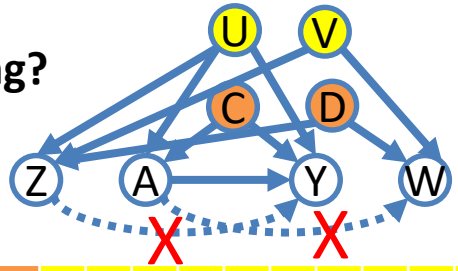


What if negative control has no confounding?



If your method produces a biased estimate in this case, wouldn't you be concerned about your target estimate?

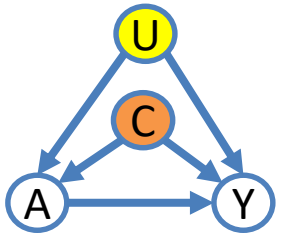
What if negative control has more confounding?



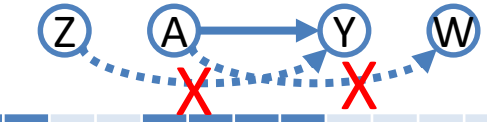
If your method produces an unbiased estimate in this case, wouldn't you be reassured about your target estimate?

Value of using a large set of negative controls: confounding structure is unknown and can vary

Treatment A -> Outcome Y:



Negative control exposure Z → Outcome Y OR Treatment A → Negative control outcome W:



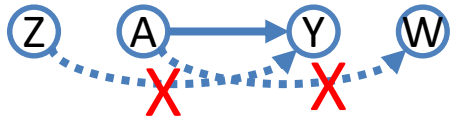
NC1
NC2
NC3
NC4
NC5
NC6
…
NC*n*

…

Estimating error across a range of confounding structures can provide general understanding of the residual error that could be expected for method

Identifying sets of negative controls is easy to do when you don't require identical confounding

- Primary criteria: no casual effect
  - Negative control exposure: Z does not cause Y
  - Negative control outcome: A does not cause W



- Candidate negative controls: exposure → outcome relationships which have:
  - No evidence in published literature (co-occurrence in PubMed abstracts/MeSH indexing)
  - No evidence in FDA structured product labeling (drug is neither indicated for nor has warning or adverse reaction listed for outcome on any label on DailyMed)
  - No disproportionality analysis alert (ex: PRR>2) from FAERS spontaneous adverse event reporting

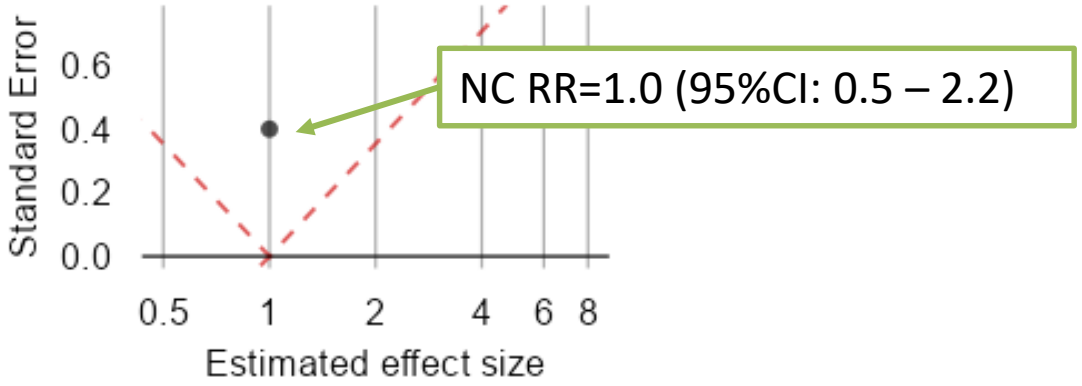
  Absence of evidence of effect doesn't mean evidence of absence, but it can serve as a useful proxy
- Candidate list can be generated automatically using OHDSI's Common Evidence Model via ATLAS, but then undergo manual review to remove questionable candidates (Voss JBI 2017)
- Assumption is that all negative control exposure→outcome pairs have true RR=1, but estimating error distribution is robust even if there is misclassification among some controls within the set (Schuemie ICPE 2014)

How to use negative controls once you've identified them

- **Diagnostics:** determine if the method has residual error
  - Not observing error could be reassuring
  - Observing large bias should make you question your design, potentially stop analyses and not unblind results

- **Calibration:** correct the method's estimates for the residual error
  - Provide an estimate with uncertainty reflecting both random and systematic error

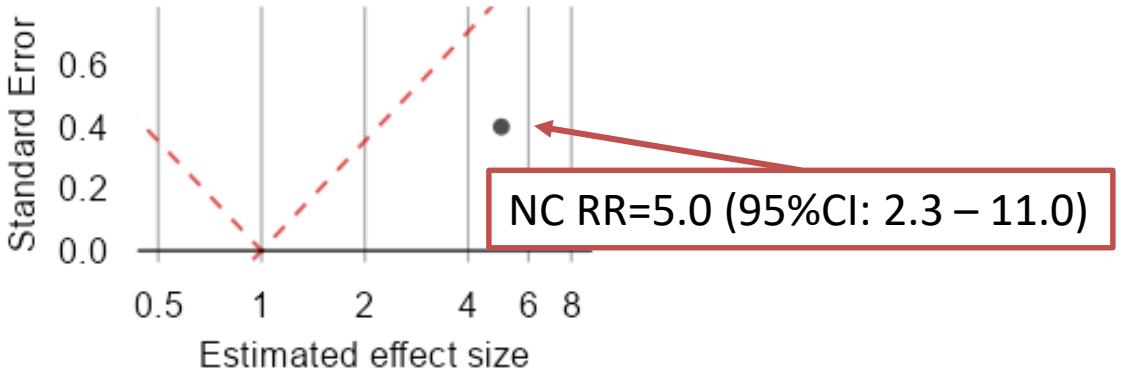
How can one negative control help as a diagnostic?



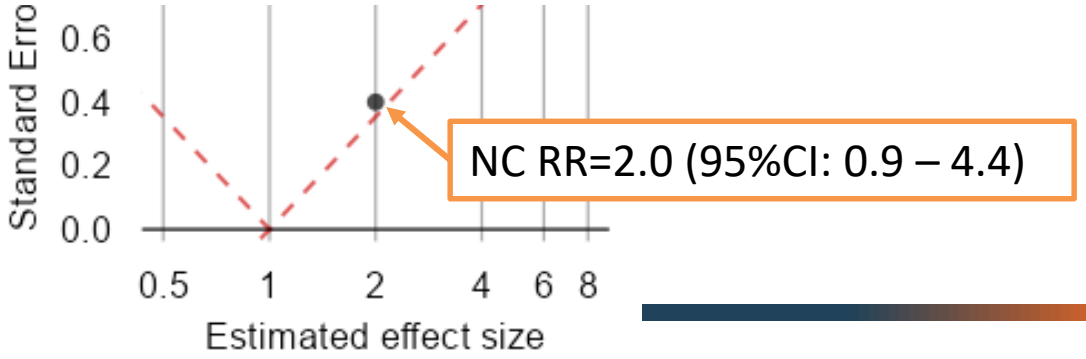
No evidence of error observed:
Is this sufficient evidence to trust your estimate of interest?

Probably, but:

1) Uncertainty in the NC estimate
2) Uncertainty in the exchangeability of NC to the target question

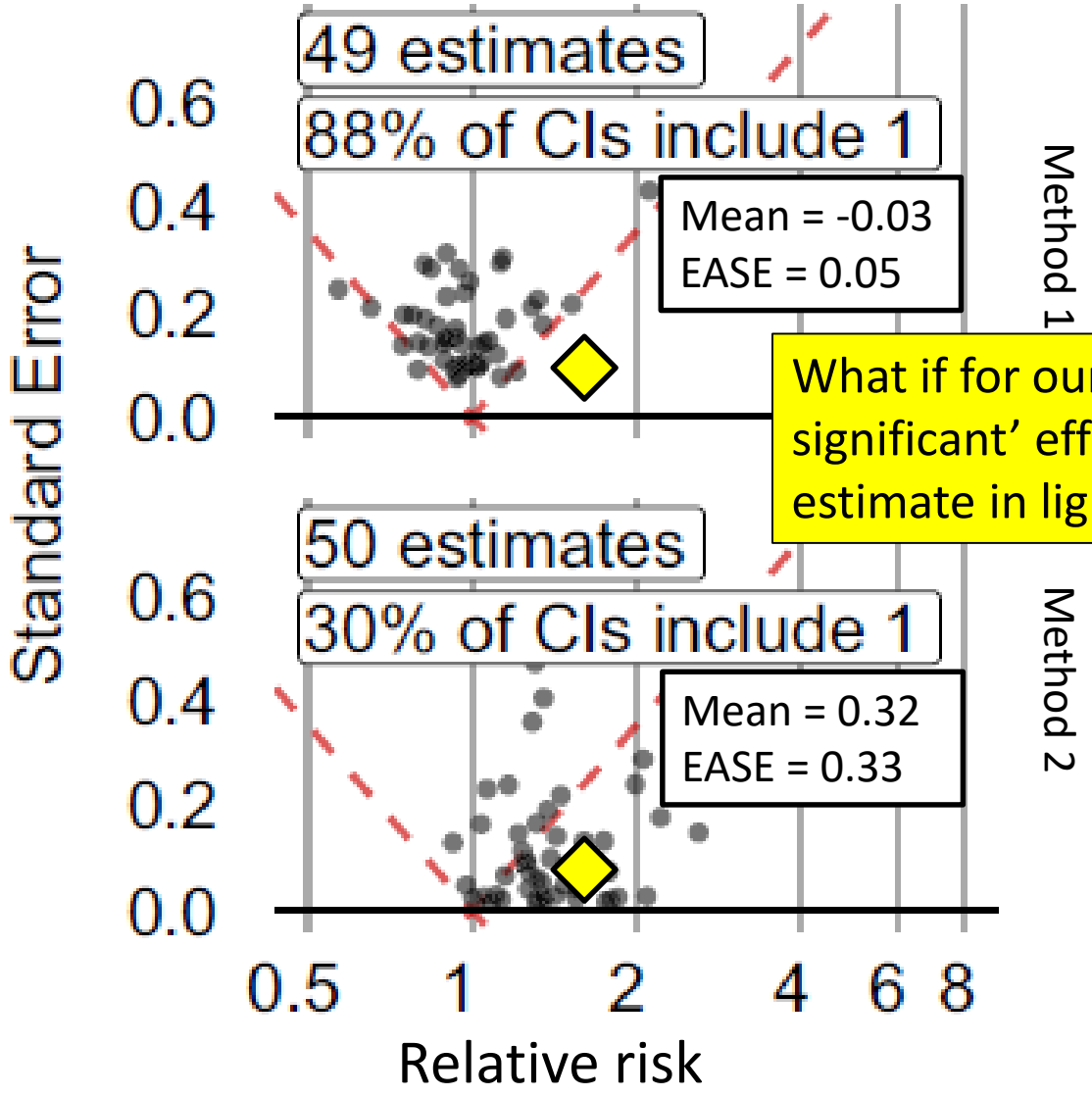


Large statistically significant error observed:
Is this sufficient evidence to disregard your estimate of interest?



Positive non-significant error observed:
What do we do now?

How can a large set of negative controls help as a diagnostic?



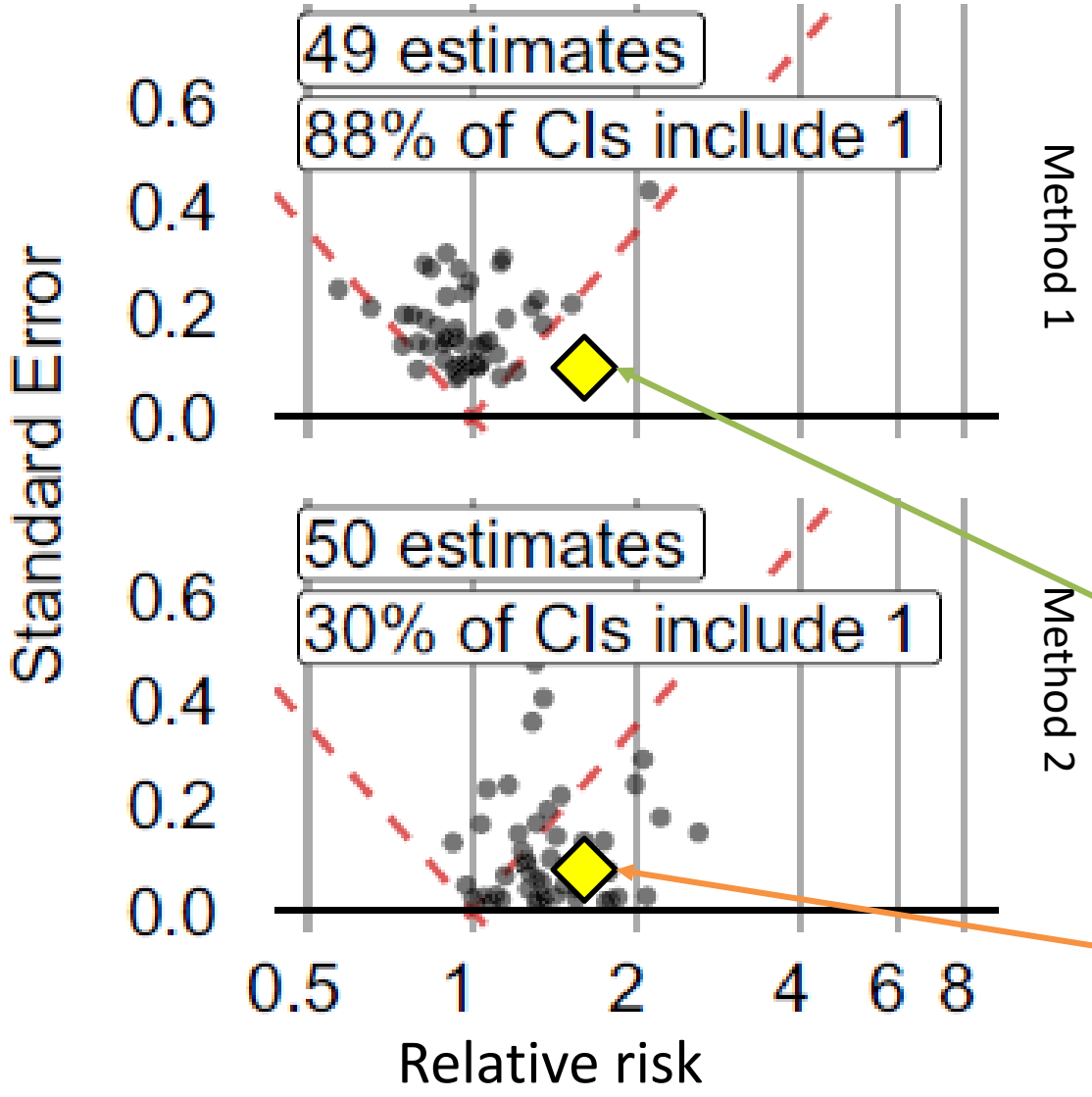
- A set of negative controls provides a distribution of errors that can be generated by the method
  - More negative controls reduce both types of uncertainties: random error, exchangeability
  - Here, each method has >=1 negative control

What if for our question of interest, we observe a nominally 'statistically significant' effect of RR=1.50 (1.22-1.83) (SE=0.1), do we still trust that estimate in light of these distributions of negative controls?

consistently robust to systematic error
  - Method 1 shows little systematic error across all negative controls
  - Method 2 shows consistent positive bias
- Objective diagnostic with pre-defined thresholds can be applied to this distributions
  - Metrics: mean of error distribution, false positive rate, Expected Absolute Systematic Error (EASE) >= 0.25

Schuemie et al PNAS 2017

How can a large set of negative controls help for calibration?



- Our current methods typically produce estimation statistics that only reflect random error
- A large set of negative controls provide a distribution that can help us understand the magnitude and uncertainty of systematic error
- Calibration is integrating the systematic error into your estimation statistics (RR, 95%CI, p-value)

Method 1:
Uncalibrated RR=1.5 (1.2-1.8)
Calibrated RR = 1.6 (1.2-2.0)
Calibration only slightly moves original estimates when little systematic error observed

Method 2:
Uncalibrated RR=1.5 (1.2-1.8)
Calibrated RR = 1.1 (0.7-1.7)
Calibration shifts point estimate and increases uncertainty due to observed systematic error

p-value calibration: Schuemie et al Stat Med 2016 ; CI Calibration: Schuemie et al PNAS 2017

Concluding thoughts

- Modeling confounding from the complexity of health care data is difficult, so finding a negative control that satisfies strict modeling assumptions with certainty can be challenging and may still be insufficient
- Identifying negative control exposure→outcome pairs that can be assumed to have no causal relationship is feasible, and can help to estimate systematic error from observed and unobserved confounding
- A set of negative controls can estimate an error distribution for a given method applied to a given database, which can be used for both objective diagnostics and empirical calibration

Panel Discussion

Moderator: **Rachele Hendricks-Sturrup,** Duke-Margolis Center for Health Policy

Panelists:

- **Eric Tchetgen Tchetgen,** University of Pennsylvania
- **Xu Shi,** University of Michigan
- **Erich Kummerfeld,** University of Minnesota
- **Patrick Ryan,** OHDSI
- **Martijn Schuemie,** OHDSI

Break

We will reconvene at 12:45 p.m. ET for the next panel discussion.

Utilizing Negative Controls in Safety and Effectiveness: Methods development and key considerations

Sentinel

Automated Negative Control Identification in the Sentinel Setting

Xu Shi, PhD
Richard Wyss, PhD
Shirley Wang, PhD
Rishi Desai, PhD

Background

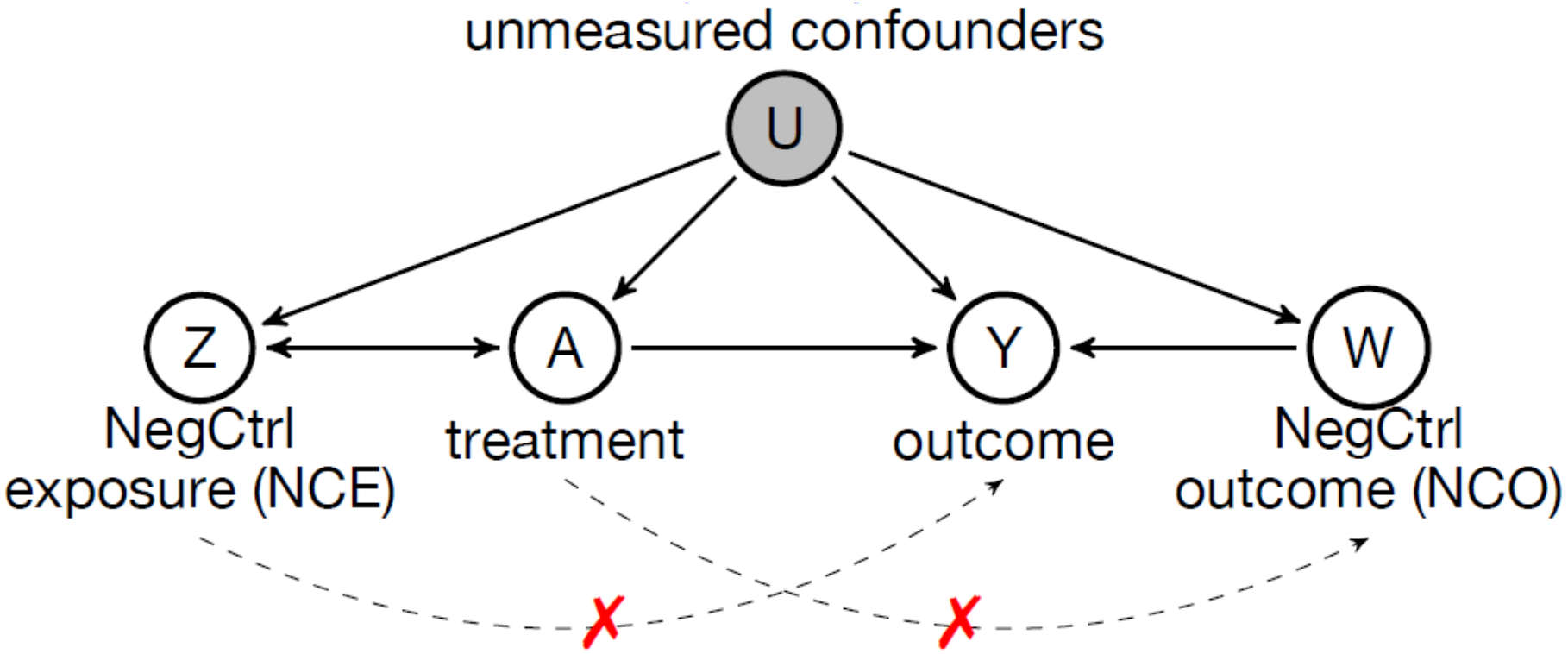
Background: Challenges for Confounding Control in RWE Studies

- Confounding arising from non-randomized treatment choices remains a fundamental challenge for extracting valid evidence to help guide treatment and regulatory decisions.

- Failure to reproduce research findings has become more common, largely due to the inherent risk of hidden bias from intractable confounding.

Background: Negative Controls

- A powerful tool increasingly recognized to mitigate unmeasured confounding is the negative control (Lipsitch 2010).
- Negative controls are variables associated with the unmeasured confounders but not causally related to either the treatment or outcome variables of interest.



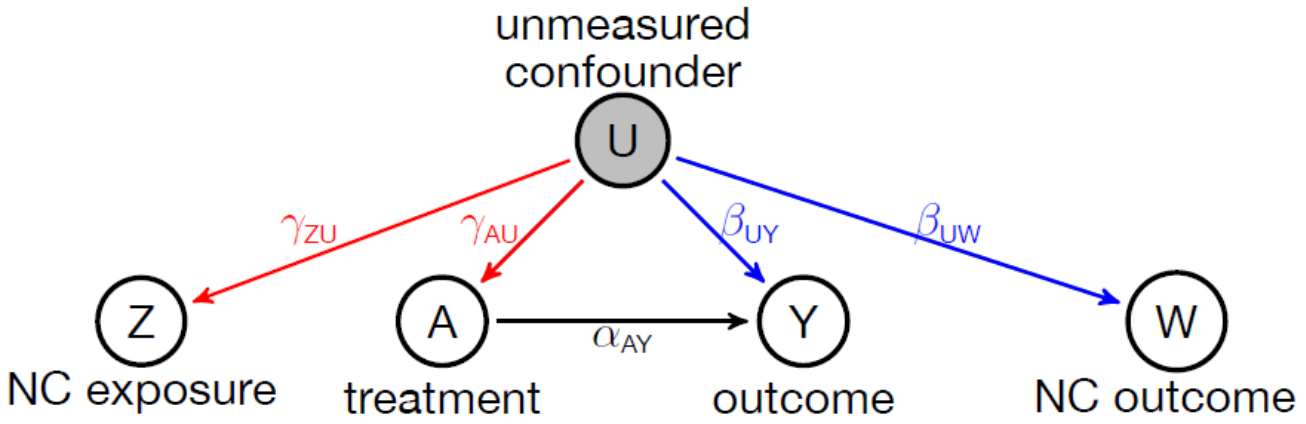
Background: Identifying Negative Controls

- Typically, negative control variables must be identified laboriously from background knowledge, and it also has to be assumed that the identified variables were genuine negative controls as no validation test existed.

- Recently, Kummerfeld, Lim, and Shi (2022) developed a validation test for discovering negative controls of a special type -- disconnected negative controls -- that can serve as surrogates of the unmeasured confounder
  - A disconnected negative control is a negative control that is causally related to neither the treatment nor the outcome (while negative controls can be causally related to one of the treatment and outcome).

- An automated approach to find disconnected negative controls: Data-driven Automated Negative Control Estimation (**DANCE**) algorithm.

How to Find a Candidate Negative Control Variable?



- Data-driven Automated Negative Control Estimation (DANCE)
  - Identifies triplets of negative control variables
  - Aggregates average treatment effects obtained from all pairs of negative controls
  - Limitation: can only detect a special type of negative control
- Rationale: all paths from {W, Z} and {Y, A} are rank deficient
  - 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

Objectives

Study Objectives

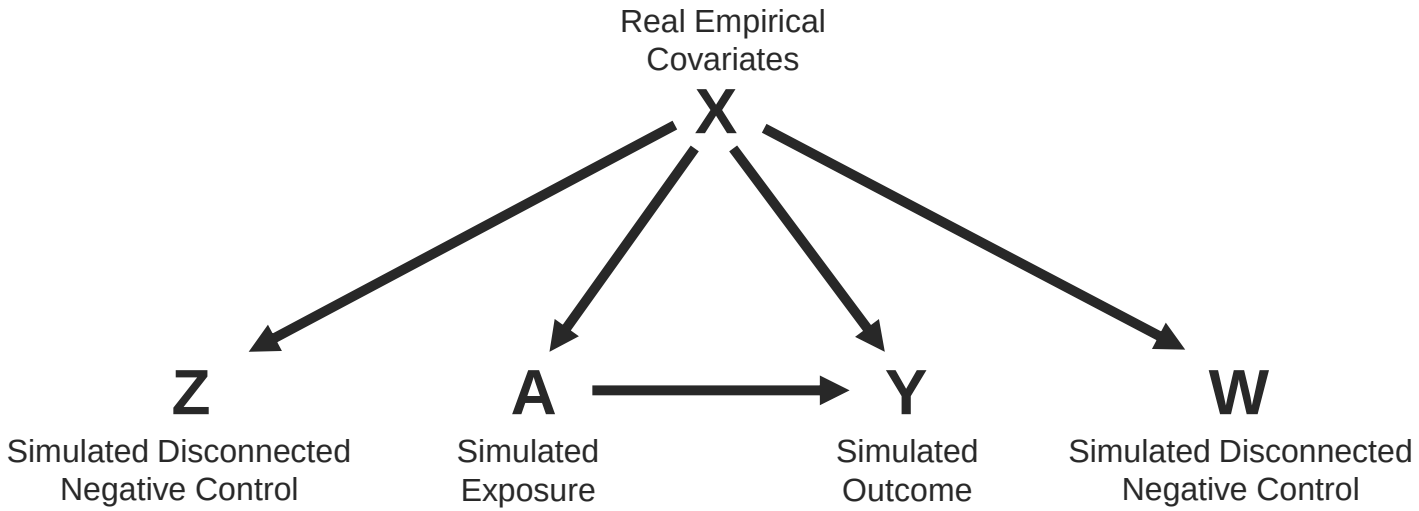
- To extend, test, and adapt the DANCE algorithm to large-scale healthcare data reflective of Sentinel data environments.
  - Little is known about how DANCE performs in the setting of large healthcare data sources where thousands of variables could be considered as candidate negative controls.
  - The Sentinel Innovation Center data assets, including linked claims with EHR data structures, provide an ideal environment to investigate and adapt the DANCE algorithm for practical use.

- Develop software packages within the R and SAS programming environments to make computer modules freely available and easily accessible to applied researchers.

Approach

Approach

- We will design plasmode simulation studies constructed based on empirical cohorts reflective of the complexities observed in real world data and a known truth to evaluate DANCE in a controlled setting.



Approach

- We will design a sequence of plasmode simulation studies based on the empirical example with a range of tightly controlled parameters and known truth
  - We will simulate true disconnected negative controls to evaluate DANCE's ability to identify them, and we will control the true value of the causal effect to evaluate DANCE's ability to estimate it.
  - We will look into the impact of high-dimensional covariates on the computational efficiency of DANCE.
  - The simulation studies will allow us to adapt and tailor DANCE to the unique challenges when dealing with complex and large health databases.

Methods: Data Source

- Mass General Brigham (MGB) Research Patient Data Registry (RPDR)
  - The electronic health records (EHR) of all the patients aged 65 and above identified in the Mass General Brigham (MGB) Research Patient Data Registry (RPDR) linked to Medicare claims data

- We will construct Plasmode simulations using 2 cohort studies generated from linked RPDR-Medicare claims (details on later slide).

Methods: Example Study Cohorts for Plasmode Simulation

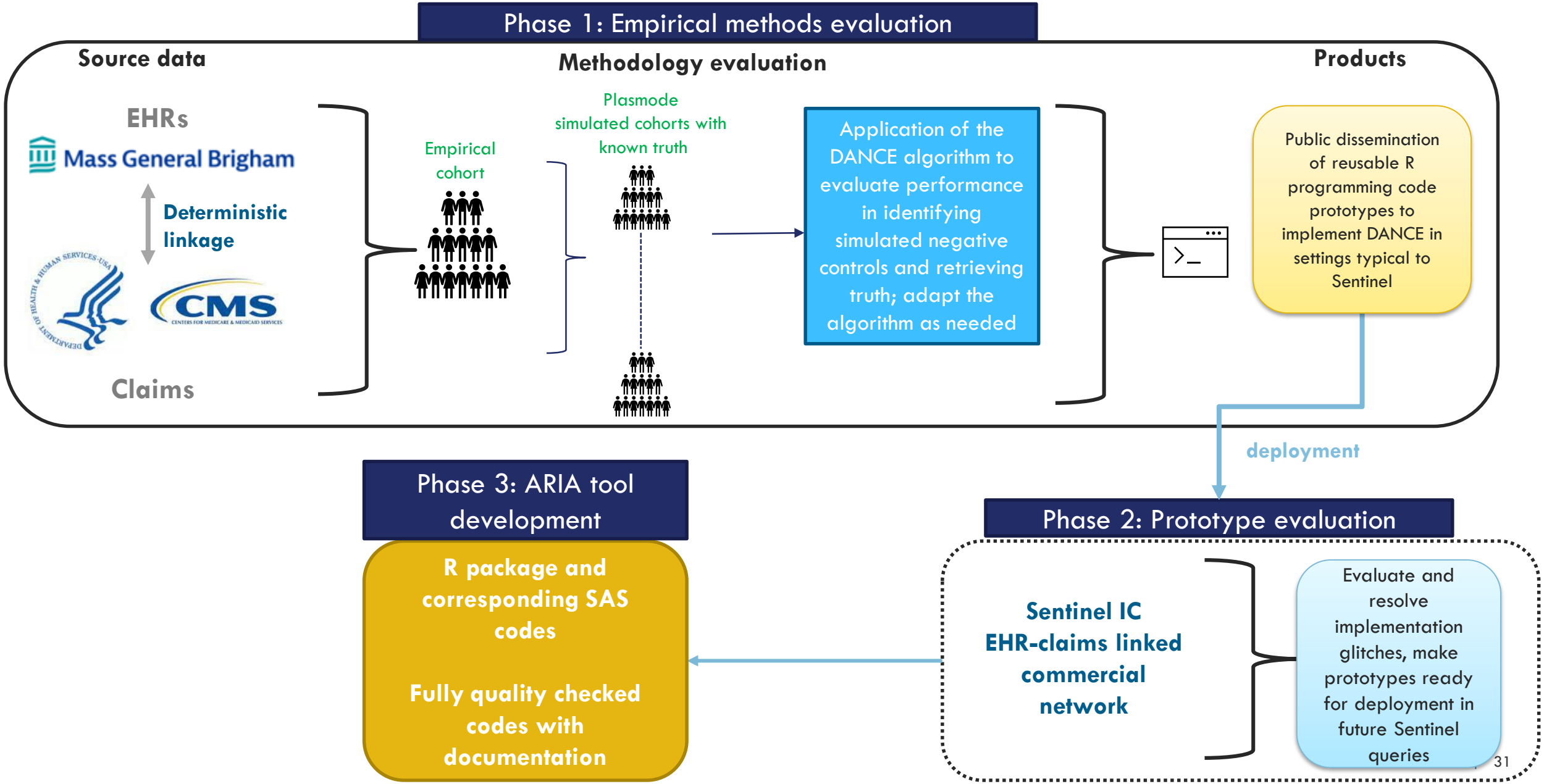
Table 1. Characteristics for Studies 1 and 2 that will be used to generate 'Plasmode Simulations'							
	Sample Size			Outcome	Baseline Covariates*		
Cohort	N_{Total}	$N_{Treated}$	$N_{Comparator}$	N_{Total}	N_{Total}	$N_{Predefined}$	N_{codes}
Study 1:[A]	21,343	13,576 (63.6%)	7,767 (36.4%)	899 (4.2%)	14,937	91	14,846
Study 2:[B]	35,031	12,872 (36.7%)	22,159 (63.3%)	251 (0.7%)	12,464	91	12,373

[A] Study 1: effect of NSAIDs versus opioids on acute kidney injury

[B] Study 2: effect of high vs low-dose proton pump inhibitors (PPIs) on gastrointestinal bleeding

***Number of claims and EHR features after screening those with prevalence <0.005**

Summary


Phase 1: Empirical methods evaluation
Source data
Methodology evaluation
Products
EHRs
Mass General Brigham
Deterministic linkage
CMS
Claims
Empirical cohort
Plasmode simulated cohorts with known truth
Application of the DANCE algorithm to evaluate performance in identifying simulated negative controls and retrieving truth; adapt the algorithm as needed
Public dissemination of reusable R programming code prototypes to implement DANCE in settings typical to Sentinel
deployment
Phase 2: Prototype evaluation
Sentinel IC EHR-claims linked commercial network
Evaluate and resolve implementation glitches, make prototypes ready for deployment in future Sentinel queries
Phase 3: ARIA tool development
R package and corresponding SAS codes
Fully quality checked codes with documentation

Thank you

Double Negative Control Adjustment for Real-World Vaccine Effectiveness Studies

Yun Lu, Ph.D.

Office of Biostatistics and Pharmacovigilance (OBPV)

FDA/Center for Biologics Evaluation and Research (CBER)

PDUFA VII Negative Controls Public Workshop, March 8, 2023

Disclaimer

- This presentation reflects the views of the author and should not be construed to represent FDA's views or policies.

Background

Real-world evidence (RWE) and real-world data (RWD) have been increasingly used to answer scientific and regulatory questions.

It is important to enhance the capabilities to use RWE and RWD for regulatory decision-making.

Real-World Vaccine Safety and Effectiveness Studies

- Data reflect the exposure and outcome experiences during routine clinical practice
- Subjects have a wider range of health conditions than in randomized controlled trials
- Large dataset provides power to detect small but clinically relevant differences and analyze rare serious outcomes
- Bias will not be improved with increasing sample size

FDA Draft 'RWE' Guidance – Sep-Dec 2021

Guidance for Industry

DRAFT GUIDANCE

Sep 2021

Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

Oct 2021

Data Standards for Drug and Biological Product Submissions Containing Real-World Data

Nov 2021

Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products
Guidance for Industry

Dec 2021

Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products

https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

‘Electronic Health Record (EHR)/claims data’ draft guidance: Sep 2021

- “In general, EHR and medical claims data **do not systematically capture** the use of **nonprescription drugs** or drugs that are **not reimbursed under health plans**, or **immunizations offered in the workplace**. ”

https://www.fda.gov/media/152503/download

COVID-19 Mass Vaccination Site

Picture from NY Times

Potential Sources of Bias in Real-World Vaccine Safety and Effectiveness Studies

- Variable misclassification Example: Underreporting of vaccination status
  - COVID-19 vaccinations received at mass vaccination site
  - Influenza vaccinations received at workplace
- Confounding Example: Health seeking behavior
  - Vaccinated individuals may tend to seek more health care than unvaccinated individuals
  - COVID-19 pandemic impacted health seeking behavior over time (time-varying confounding)

PDUFA VII – Use of RWE - Negative Controls

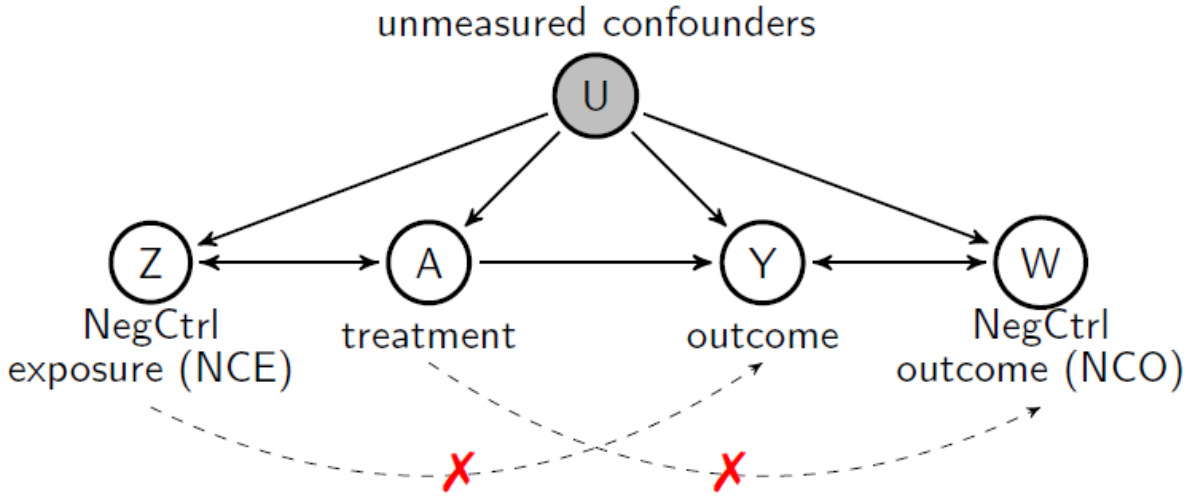
- **PDUFA VII commitment letter**

"FDA will initiate two methods development projects by September 30, 2024 to

...

2) develop a method to use a double negative control adjustment to reduce unmeasured confounding in studying effectiveness of vaccines."

Negative Control Pairs



- Z is an NCE if
  - (1) It does not directly affect Y
  - (2) It is associated with U

Z is associated with Y only through A and U

- W is an NCO if
  - (1) It is not directly affected by the treatment
  - (2) It is associated with U

W is associated with A or Z only through U

Slides from Xu Shi's September 2018 presentation

Double Negative Control Adjustment

- Suppose we correctly fit the following models
  - $E[Y|A,Z] = \alpha_0 + \delta_A^Y A + \delta_Z^Y Z$
  - $E[W|A,Z] = \beta_0 + \delta_A^W A + \delta_Z^W Z$
- Then the bias-adjusted ATE is simply $\boxed{\delta_A^Y - \frac{\delta_Z^Y}{\delta_Z^W} \cdot \delta_A^W}$

Double negative controls approach uses a pair of negative control exposure (NCE) Z and negative control outcome (NCO) W as a proxy for unmeasured confounder U, and provides bias-adjustment.

Formula from Xu Shi's September 2018 presentation

Proposed Vaccine Effectiveness Study

- Data source: medical claims data
- Exposure: Herpes Zoster Vaccine (exposures reliably captured in claims data)
- Outcome: Herpes zoster, hospitalized herpes zoster, post-herpetic neuralgia (outcomes reliably captured in claims data)
- Two previously published studies to measure vaccine effectiveness and duration of effectiveness for Zostavax and Shingrix vaccines

Confounders

- Measured covariates: Individual-level characteristics including demographics, socioeconomic status, health status, etc.
- Unmeasured confounder: Health seeking behavior
  - Vaccinated individuals may tend to seek more health care than unvaccinated individuals
  - Health seeking behavior tend to impact mild and moderate health outcomes more

Control for Measured Covariates

- Standardized mean differences (SMDs) will be used to determine cohort balance for covariates
- Inverse probability of treatment weighting (IPTW) will be used to address imbalance in all measured covariates
- Doubly robust approach will be applied to include all covariates in both the weighting and outcome model.

Adjust for Unmeasured Confounding using Double Negative Controls

- Negative Control Exposure: does not directly affect zoster outcome, but associated with health seeking behavior
- Negative Control Outcome: is not directly affected by herpes zoster vaccines, but associated with health seeking behavior
- Need to identify appropriate negative control exposure and negative control outcome
- Need to consider negative control outcomes with different disease severity

U.S. FOOD & DRUG ADMINISTRATION

Any suggestions or questions?


FDA
U.S. FOOD & DRUG
ADMINISTRATION

Panel Discussion

Moderator: **Mark McClellan,** Duke-Margolis Center for Health Policy

Panelists:

- **Rohini Hernandez**, Amgen
- **Josh Gagne**, Johnson & Johnson
- **Susan Gruber**, Putnam Data Sciences
- **Daniel Morales**, European Medicines Agency

Key Stakeholder Perspectives

Key Stakeholder Perspectives

Moderator: **Mark McClellan,** Duke-Margolis Center for Health Policy

Panelists:

- **Rima Izem,** Novartis
- **Leah McGrath,** Pfizer
- **Alan Brookhart,** Duke University
- **George Hripcsak,** Columbia University

Closing Remarks

Mark McClellan, MD, PhD

Director, Duke-Margolis Center for Health Policy

Thank You!

Contact Us

healthpolicy.duke.edu

Subscribe to our monthly newsletter at
dukemargolis@duke.edu

1201 Pennsylvania Avenue, NW, Suite 500
Washington, DC 20004

DC office: 202-621-2800
Durham office: 919-419-2504

Follow Us

DukeMargolis

@DukeMargolis

@DukeMargolis

Duke Margolis