

Mendelian randomization with pharmaceutically modifiable biomarkers

Society for Epidemiologic Research
June 15, 2022

Joy Shi, PhD

CAUSALab and Department of Epidemiology
Harvard T.H. Chan School of Public Health

Overview



Background:

Mendelian randomization (MR) studies of biomarkers

Objectives:



1. Describe the sources of bias that arise when using conventional methods to adjust for medication use.



2. Describe the causal estimands that can be targeted.



3. Demonstrate the use of g-methods to adjust for medication use.

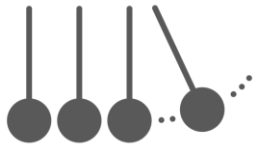
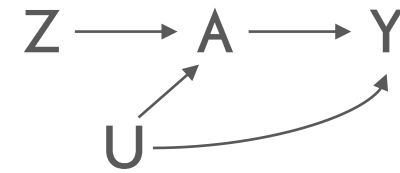


Mendelian randomization (MR) is an increasingly popular application of instrumental variable analysis



Genetic variants used as proposed instruments

- Three instrumental conditions



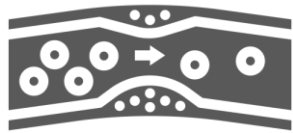
Estimate the effect of a non-genetic exposure on outcome

- Even with unmeasured confounding

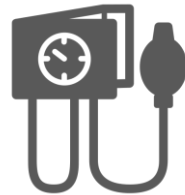


Exposures in MR studies

- ✖ Biomarkers that affect and are affected by medication use



LDL cholesterol
and statins



Blood pressure
and diuretics



Blood glucose
and metformin

- ✖ Adjusting for medication use or restricting to non-users introduces bias



Objective 1: Bias of conventional methods

Genetic variant (Z)

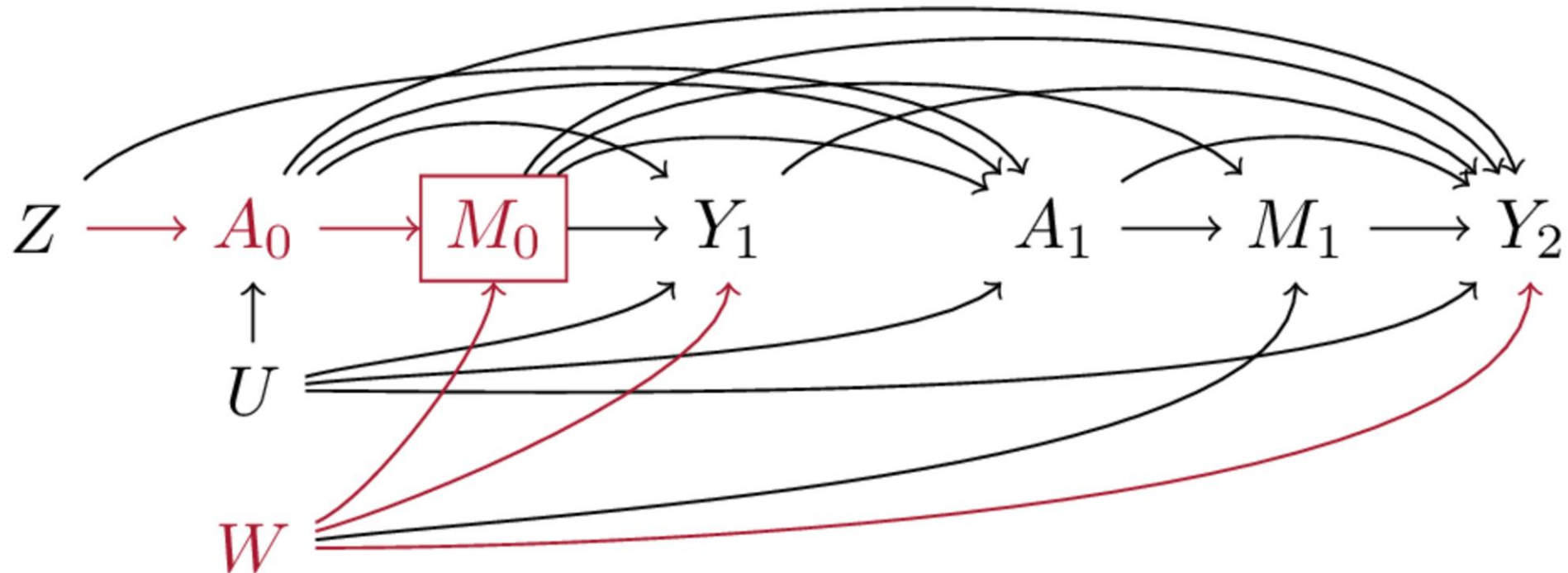
LDL cholesterol over time (A_k)

Statin use over time (M_k)

Coronary heart disease incidence (Y_{k+1})

LDL cholesterol-CHD confounders (U)

Statin-CHD confounders (W)



Objective 1: Bias of conventional methods

Genetic variant (Z)

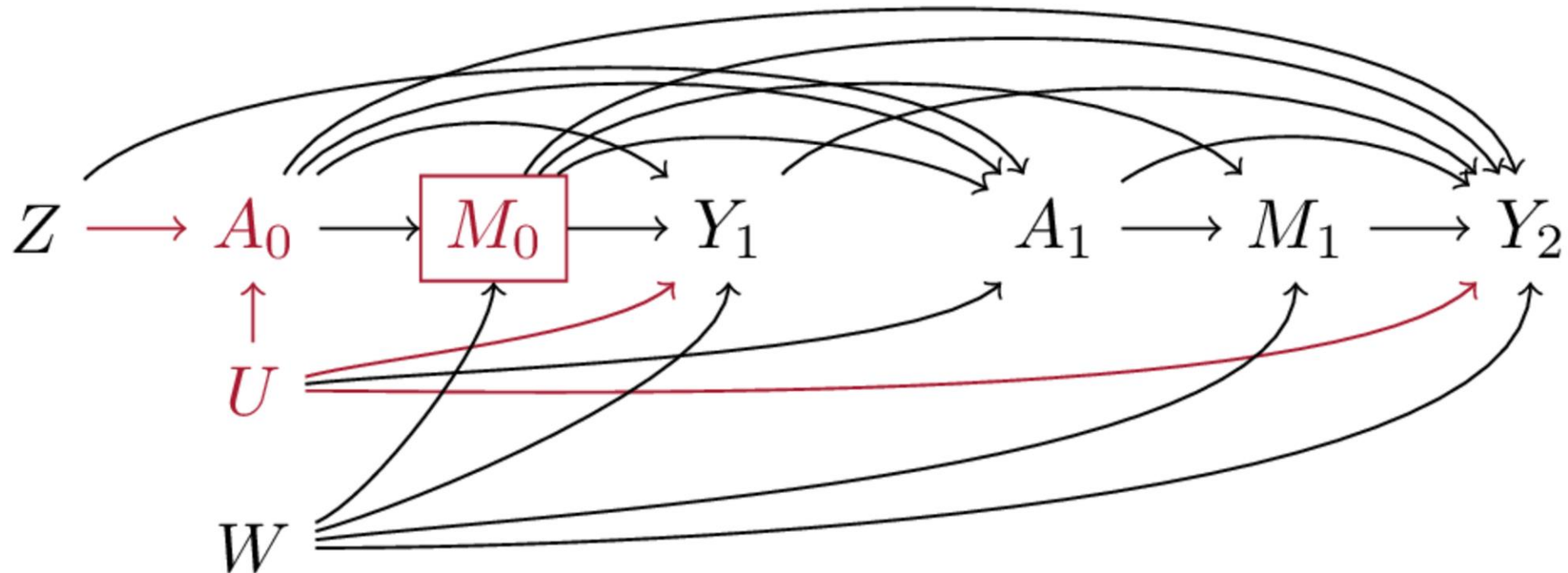
LDL cholesterol over time (A_k)

Statin use over time (M_k)

Coronary heart disease incidence (Y_{k+1})

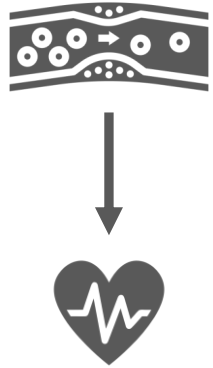
LDL cholesterol-CHD confounders (U)

Statin-CHD confounders (W)



Objective 2: Potential causal estimands of interest

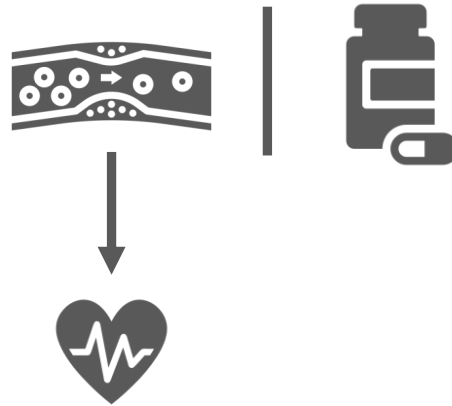
Total lifetime effect of LDL cholesterol



$$E[Y^{\bar{a}_{K+1}}] - E[Y^{\bar{a}_K}]$$

Unbiased

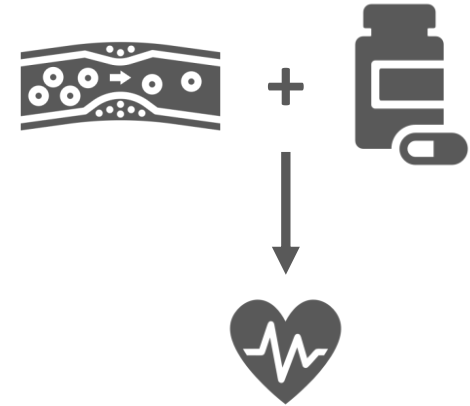
Lifetime effect of LDL cholesterol, **conditional on statin use**



$$E[Y^{\bar{a}_{K+1}} | M_k] - E[Y^{\bar{a}_K} | M_k]$$

Conditioning creates bias

Lifetime effect of a **joint intervention** on LDL cholesterol and statin use



$$E[Y^{\bar{a}_{K+1}, \bar{m}_K}] - E[Y^{\bar{a}_K, \bar{m}_K}]$$

Proposed approach



Proposed approach



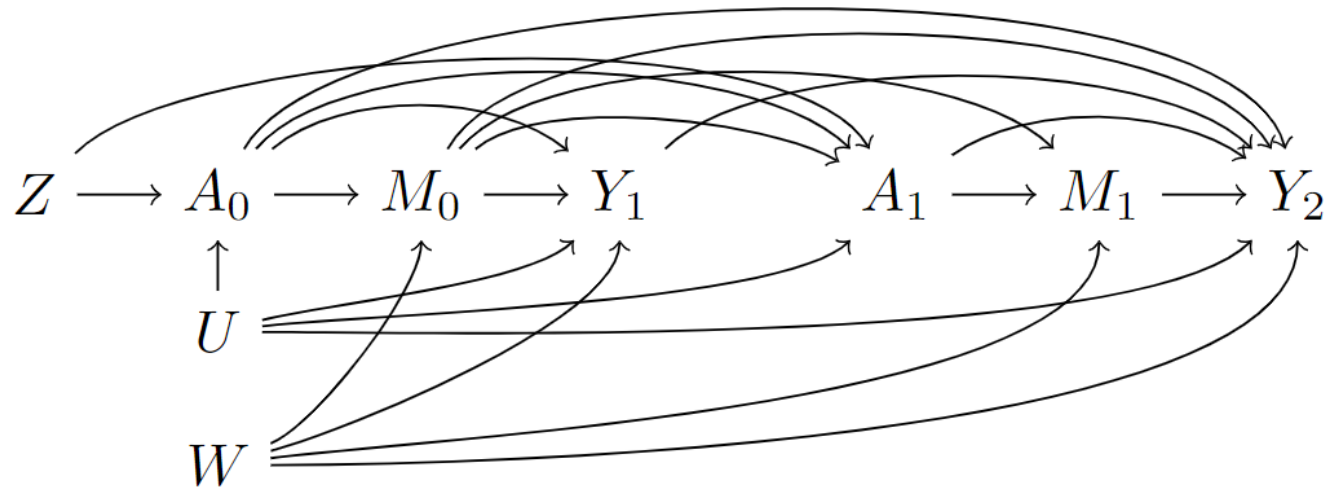
G-methods to model an intervention on statin use

- Without conditioning on statin use
- Generate data under a hypothetical statin intervention (e.g., never take statins)
- Conduct MR analysis (to assess the effect of an LDL cholesterol intervention) in the counterfactual data
- Estimating lifetime effect of a joint intervention



Objective 3: Data simulations

Data on LDL cholesterol (A_k), statin use (M_k) and CVD (Y_{k+1}) generated according to the DAG:



Three scenarios:

1. A_k and M_k under the null
2. A_k only under the null
3. Neither under the null



Analysis

Conventional MR design

- ✖ 2SLS (single measurement of the exposure)
- ✖ Varying age of participants at start of follow-up

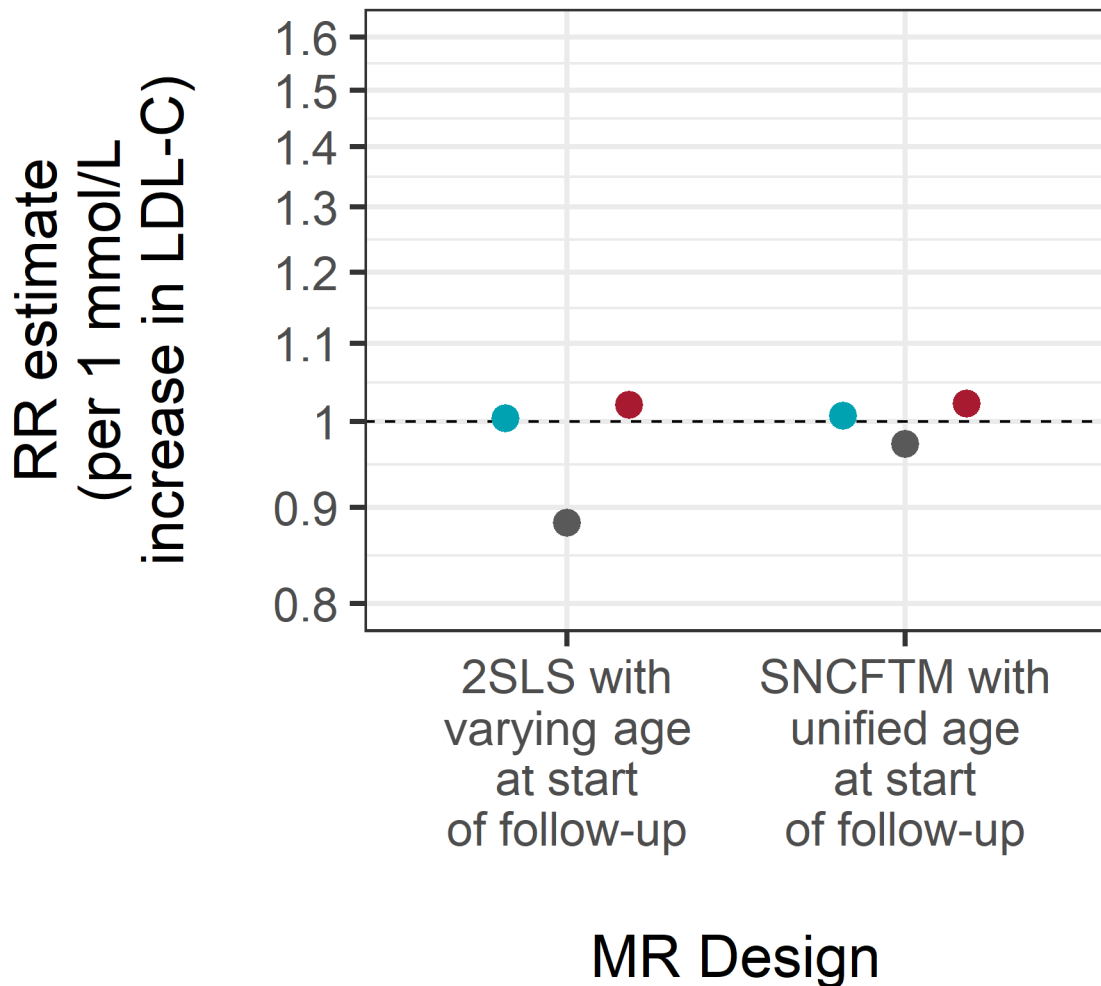
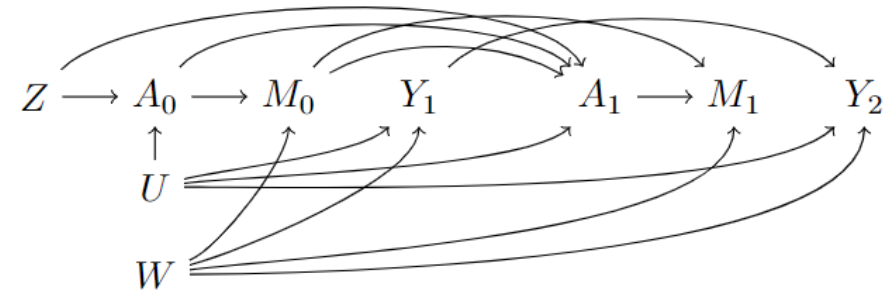
Ideal/proposed MR design

- ✖ G-estimation of SNCFTM^a (longitudinal)
- ✖ Same age of participants at start of follow-up

^aShi et al. (BMC Medical Research Methodology 2021)



Simulation results for data generated under the null for A and M



Analytical approach

● Unconditional

● Conditional on M at start of follow-up

● Counterfactual data under no medication use

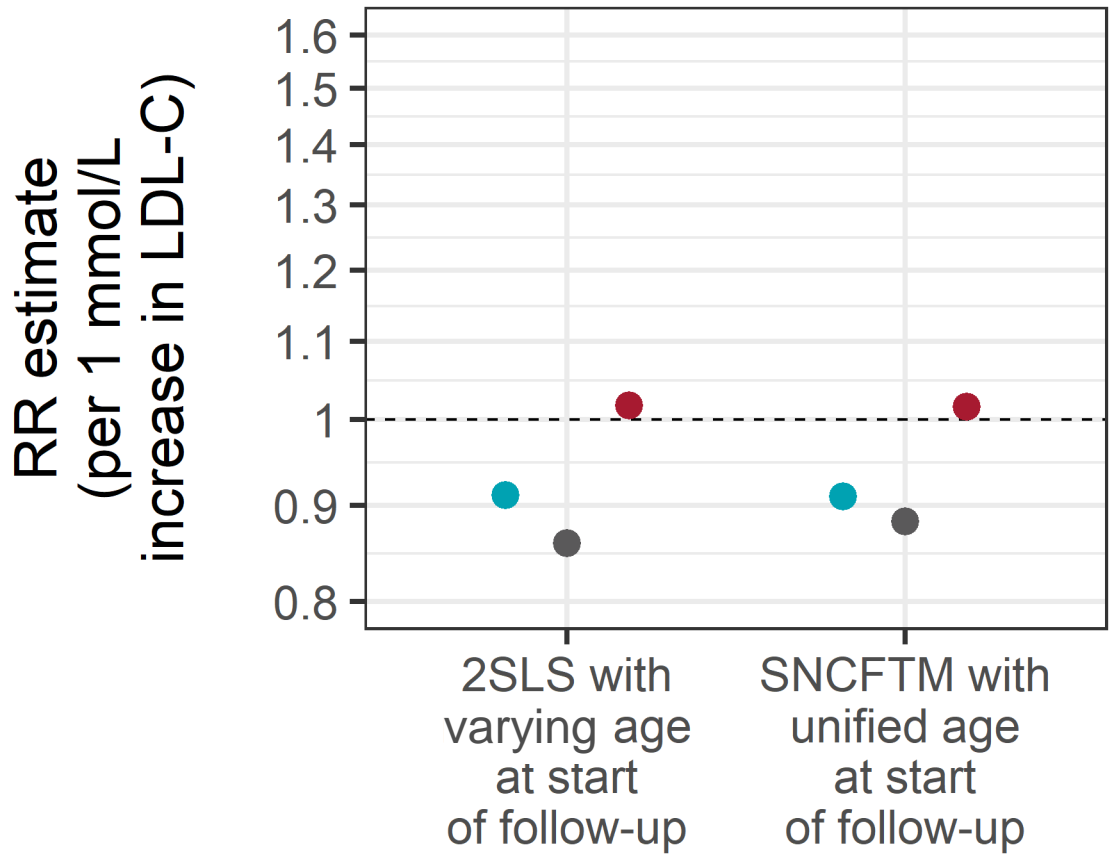
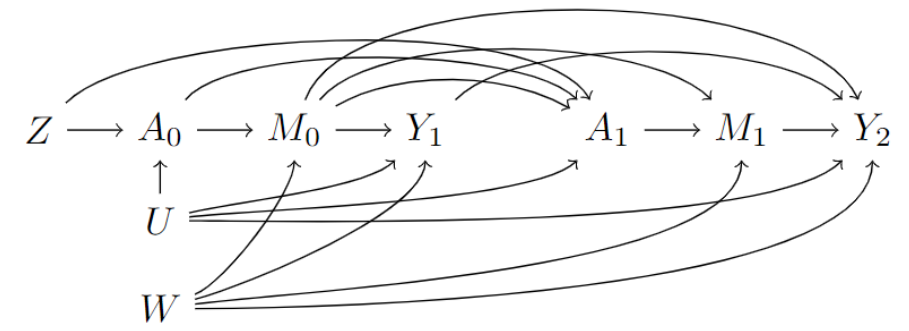
Unbiased

Downward bias

Slightly upward biased



Simulation results for data generated under the null for A only



Analytical approach

- Unconditional
- Conditional on M at start of follow-up
- Counterfactual data under no medication use

Combined effect of A+M

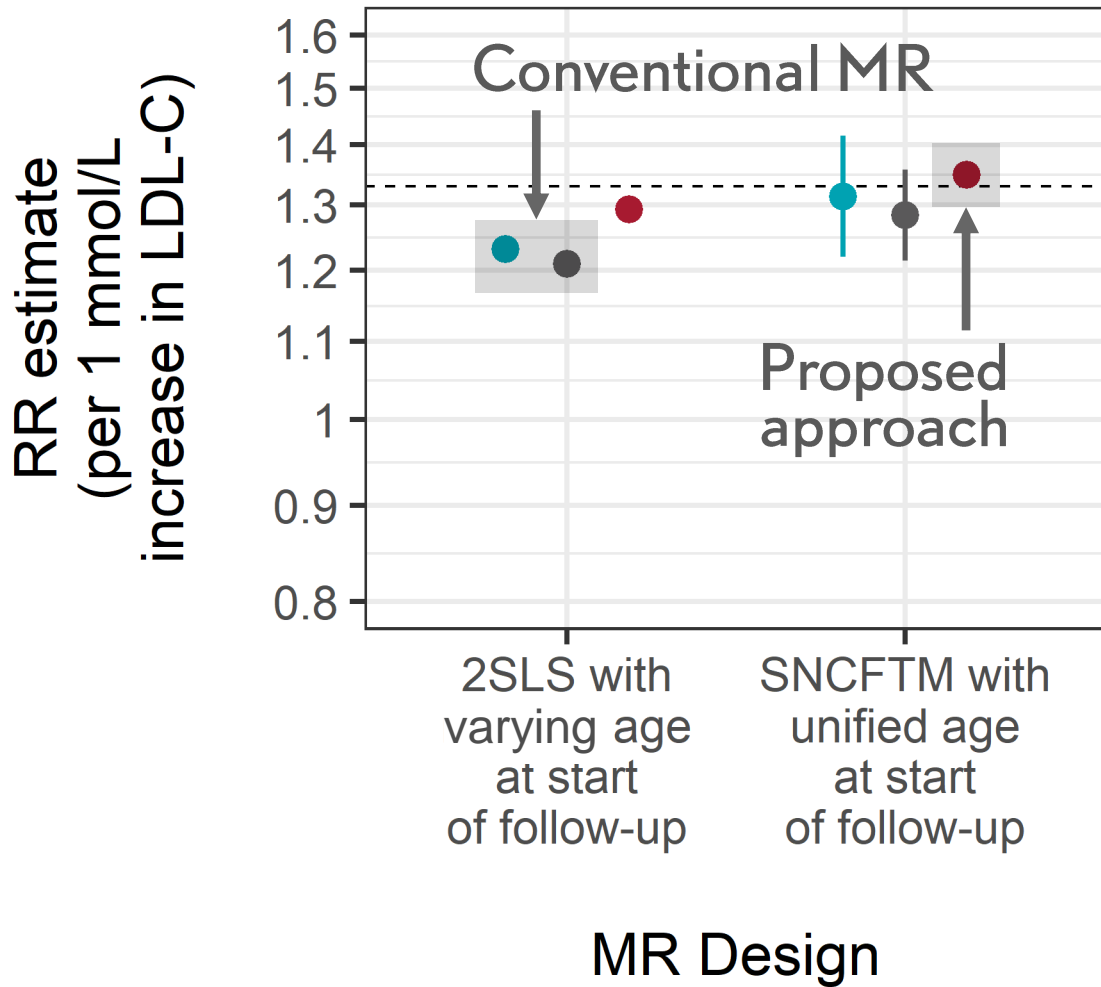
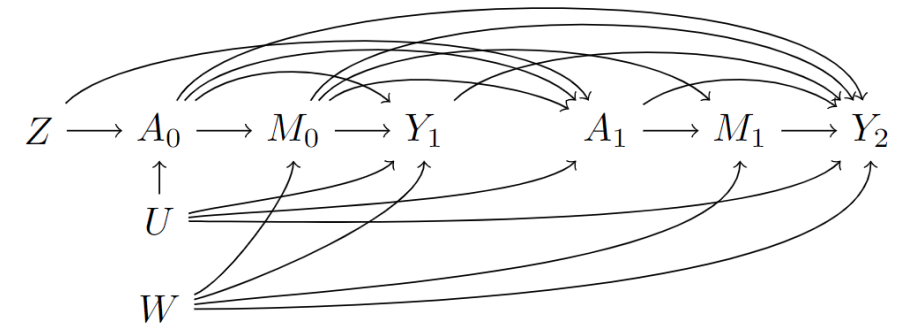
Downward bias

Direct effect of A only

MR Design



Simulation results for data generated not under the null for A or M



Analytical approach

- Unconditional
- Conditional on M at start of follow-up
- Counterfactual data under no medication use

Combined effect of A+M

Downward bias

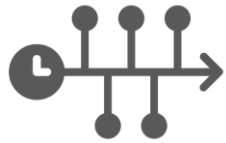
Direct effect of A only



Conclusions



Conditioning on variables downstream of the exposure can introduce **bias**



Need to consider **time-varying** nature of the exposure in MR (and other IV) studies



Combining g-methods and IV can mitigate bias of conventional approaches



Future steps: real data analysis in the Million Veterans Program



Acknowledgements



Miguel A. Hernán



Sonja A. Swanson

- Elizabeth W. Diemer
- JP Casas
- CAUSALab
- VA-CAUSAL Methods Core (CSP #2032)



A Center to
Learn What Works



HARVARD
T.H. CHAN

SCHOOL OF PUBLIC HEALTH



Questions

Email: joyshi@hsph.harvard.edu

 @joy_shi1

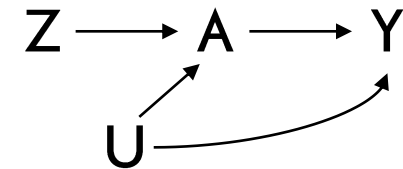


Supplementary Slides



Instrumental conditions

1. **Relevance:** the instrument (genetic variant) is associated with the exposure
2. **Exclusion restriction:** the instrument (genetic variant) does not affect the outcome except through its potential effect on the exposure
3. **No confounding for Z :** The instrument (genetic variant) and the outcome do not share common causes



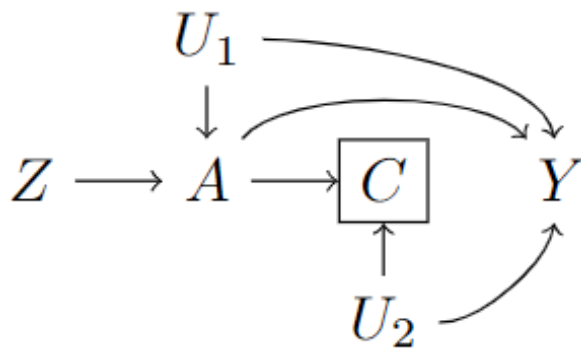
To estimate a point effect, need a fourth assumption of homogeneity or monotonicity.



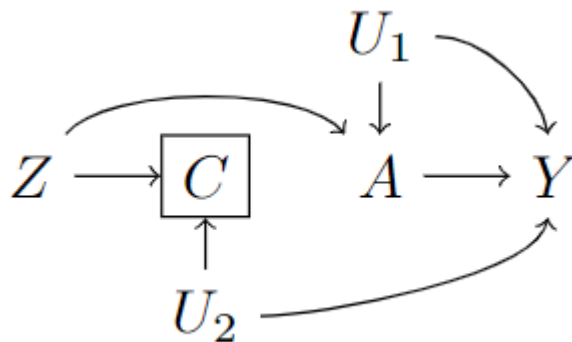
Selection bias in instrumental variable analyses

The instrumental conditions are violated in the presence of selection bias. For example, for a time-fixed exposure:

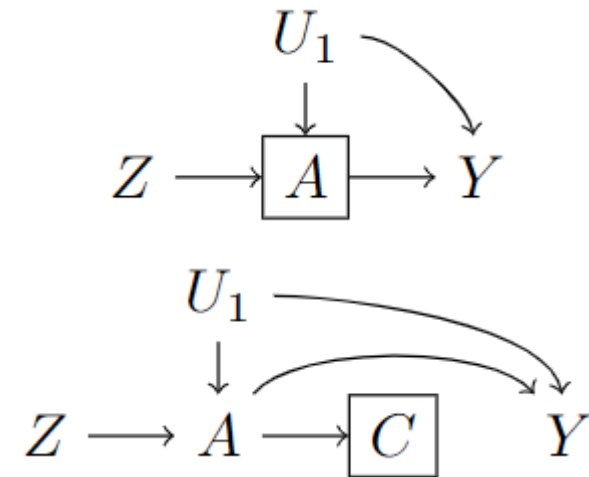
Loss to follow-up:



Misalignment of t_0 :



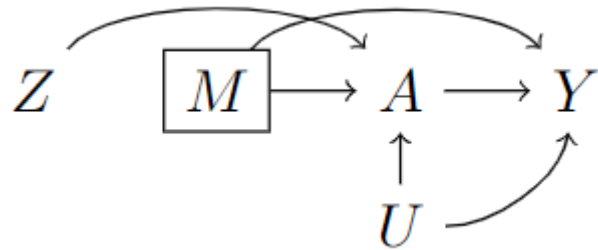
Conditioning on the exposure:



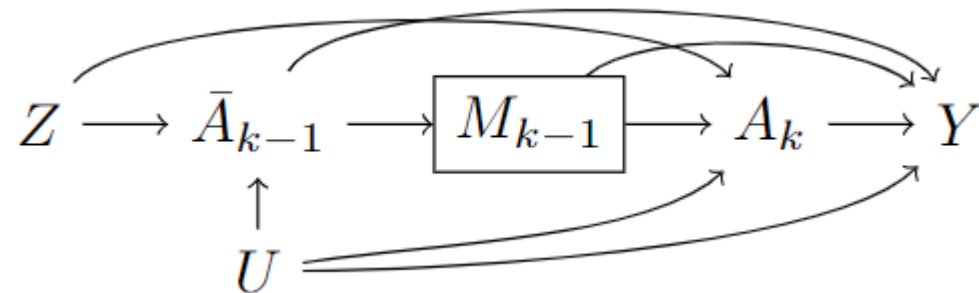
Mendelian randomization studies are often conditioning on variables downstream of the exposure without realizing it

- Most MR studies are interested in the effects of a time-varying exposure, but
 - Consider only a single measure of the exposure in the analysis
 - Conceptualize the exposure as time-fixed
- Conditioning on a “pre-baseline” variable could introduce selection bias
 - If this variable is affected by prior exposure

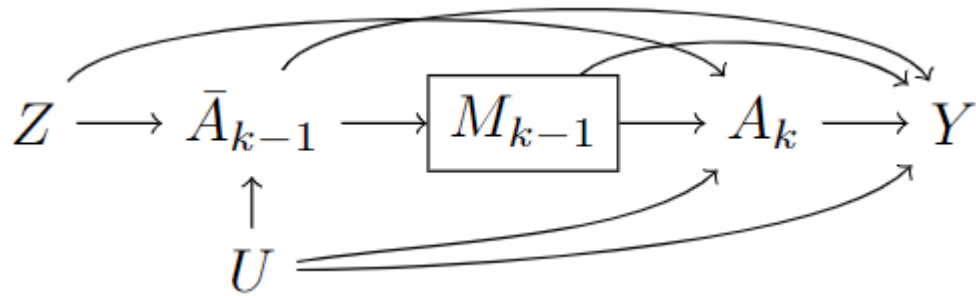
i.e., the DAG is not:



but rather more like:



Consider a MR study of LDL cholesterol and CVD



Z : LDL cholesterol-related genetic variant

\bar{A}_{k-1} : history of LDL cholesterol

M_{k-1} : statin use at time $k - 1$

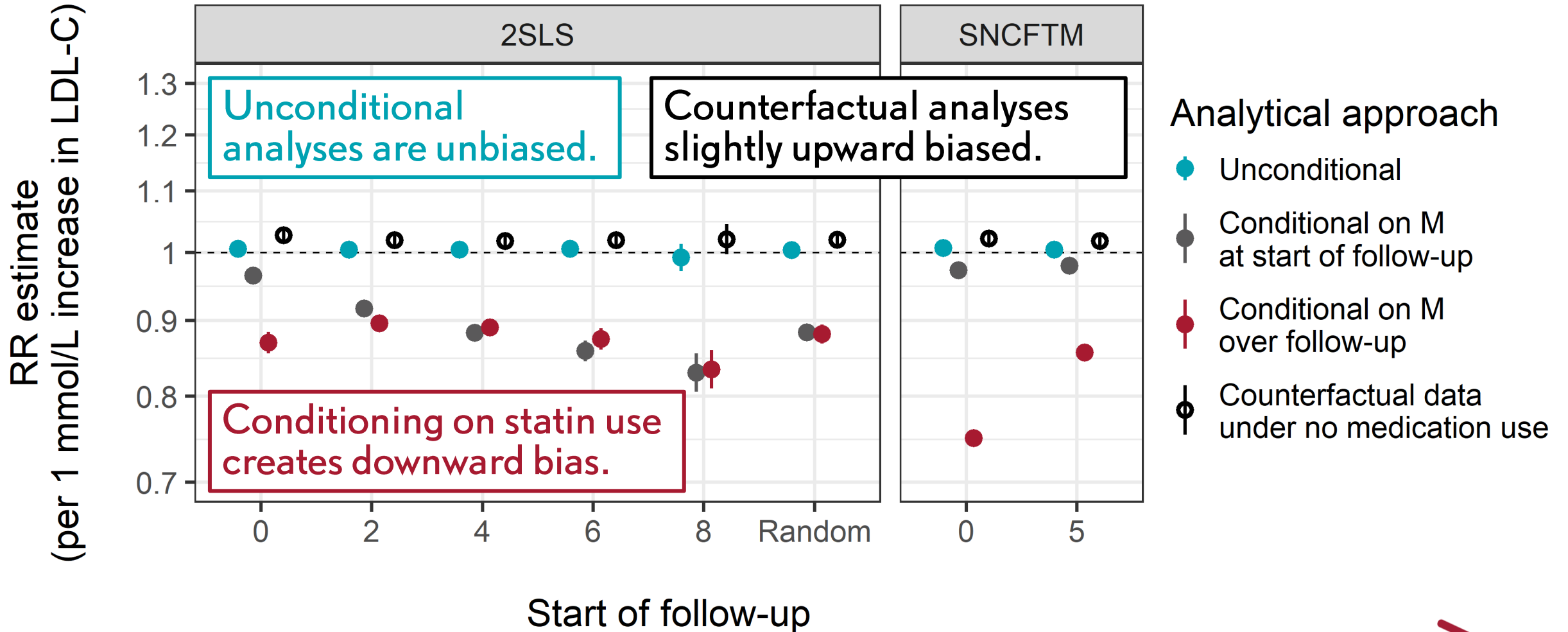
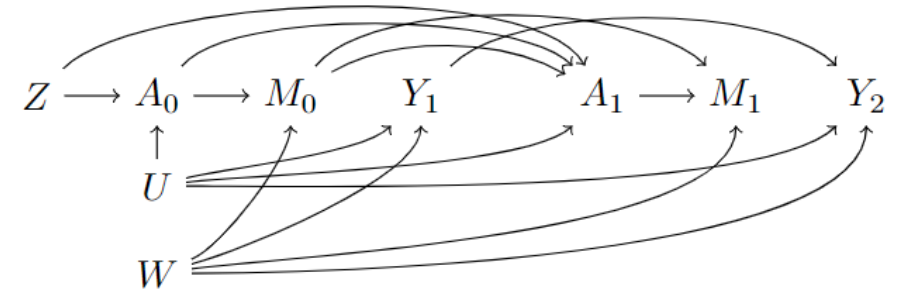
A_k : LDL cholesterol at time k

Y : CVD

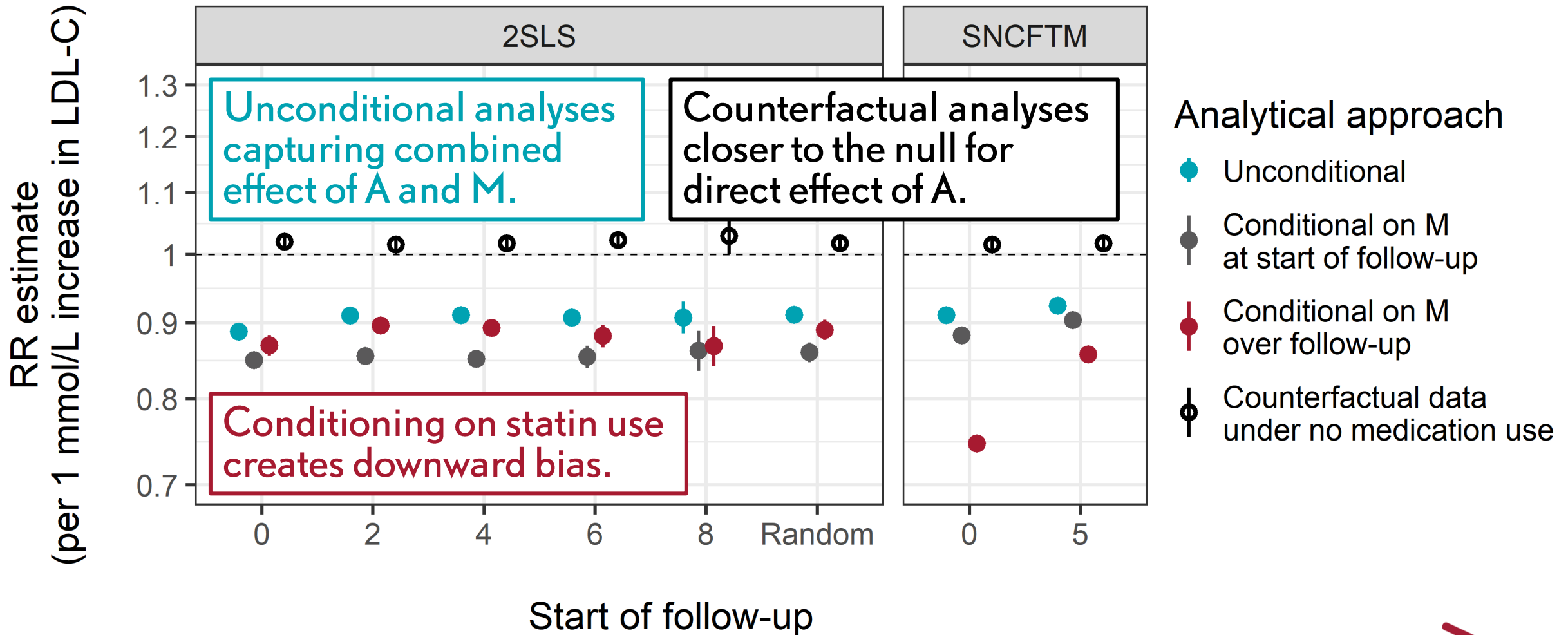
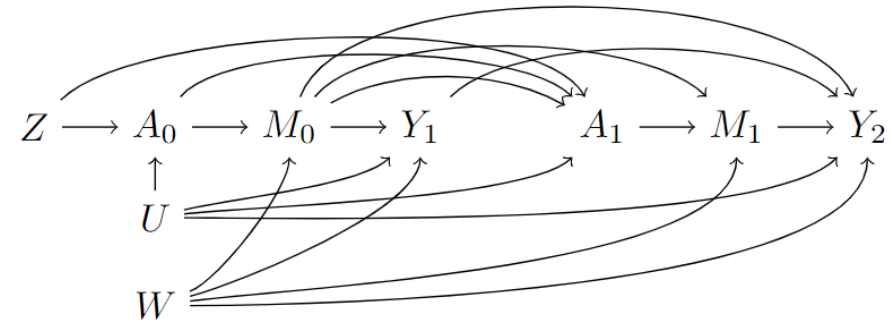
A MR analysis of LDL cholesterol which conditions on statin use will introduce selection bias.



Simulation results for data generated under the null for A and M



Simulation results for data generated under the null for A only



Simulation results for data generated not under the null for A or M

