Emulating target trials for nutritional studies:

Instrumental variable methods

Society for Epidemiologic Research
June 15, 2022

Joy Shi, PhD
CAUSALab and Department of Epidemiology
Harvard T.H. Chan School of Public Health

CAUSALab
A Center to Learn What Works

HARVARD
T.H. CHAN
SCHOOL OF PUBLIC HEALTH

Overview

Instrumental variable (IV) methods

Mendelian randomization (MR) as an application of IV

Alcohol intake and CVD
  Simple analogy between MR and randomized controlled trials
  Target trial emulation using MR
  Threats to emulation success

Instrumental variable conditions

Need a variable (i.e., an instrument, Z) that meets three conditions:

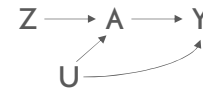1. **Relevance**
   Associated with exposure

2. **Exclusion restriction**
   Does not affect outcome except through potential effect on exposure

3. **Independence**
   No common causes (or other sources of lack of exchangeability) with outcome
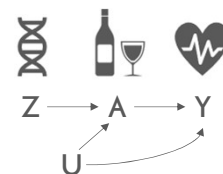
$Z \longrightarrow A \longrightarrow Y$
U

Canonical IV DAG

To estimate a point effect, need **homogeneity** or **monotonicity**

Mendelian randomization (MR) studies of nutritional factors as an application of IV

- Increasingly popular application of IV
- Genetic variants used as proposed instruments
- Leverage genetic data to estimate the effect of a non-genetic exposure on an outcome

$Z \longrightarrow A \longrightarrow Y$
U

Z: genetic variant(s) related to alcohol intake

A: alcohol intake

Y: coronary heart disease

The appeal of MR studies

Identify causal effects even with unmeasured confounding
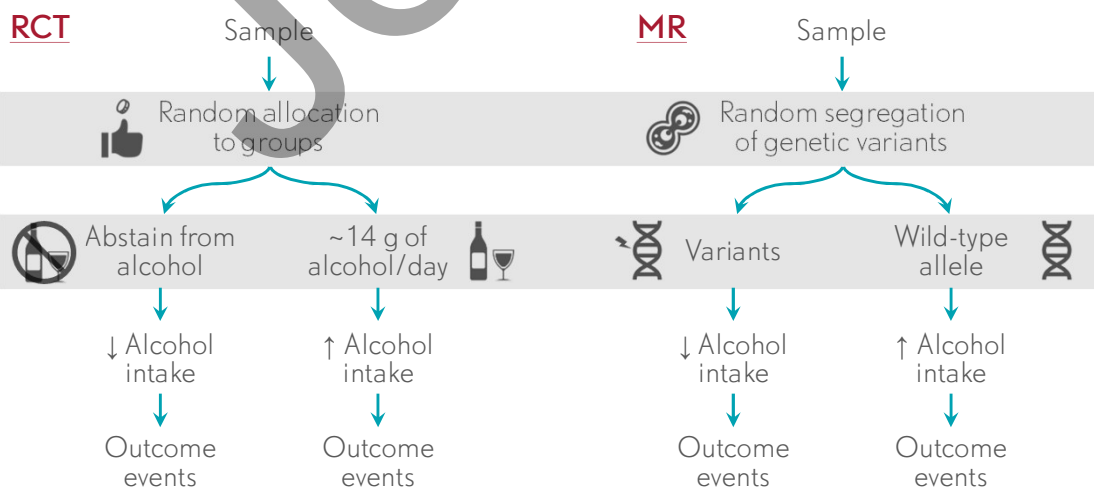
"Natural analogue" of a randomized controlled trial

Estimate the long-term ("lifetime") effects of exposures

A simplistic analogy between MR and RCTs

RCT	Sample	MR	Sample

Random allocation to groups   Random segregation of genetic variants

Abstain from alcohol	~14 g of alcohol/day	Variants	Wild-type allele
↓ Alcohol intake	↑ Alcohol intake	↓ Alcohol intake	↑ Alcohol intake
Outcome events	Outcome events	Outcome events	Outcome events

See Sanderson et al. (Nat Rev Methods Primers 2022), Larsson (Curr Opin Lipidol 2021), Smith & Ebrahim (BMJ 2005)

The appeal of MR studies

The New York Times

Does Moderate Drinking Protect Your Heart? A Genetic Study Offers a New Answer.

"Because the variants are distributed randomly in a population, they can serve in a study as the equivalent of randomly assigning people to abstain or to drink at varying levels."

Why explicitly emulate a target trial for MR studies?

Ideal   Reality

Distinguish between the target trials:
- Ideally like to emulate
- Can realistically emulate using MR

Avoid methodological pitfalls

Inform the plans to use or avoid a MR approach

4

Consider a target trial for alcohol intake and CVD

Protocol component	Target trial specification	Target trial emulation
Eligibility criteria	• ≥50 years old • No history of CVD, alcohol/substance abuse • No intolerance or allergy to alcohol • No chronic kidney or liver disease	Same, plus • Must have available genetic data
Treatment strategies	1. Never drink 2. One standard serving (~14 g) daily of a preferred alcohol beverage	Same
Assignment procedures	Randomly assigned at baseline; unblinded	Randomly assigned at conception, assuming IV conditions hold

Consider a target trial for alcohol intake and CVD

Protocol component	Target trial specification	Target trial emulation
Follow-up period	Start: at randomization End: development of the outcome, loss to follow-up, or administrative end of follow-up	Start: time of eligibility criteria/outcome recording End: same
Outcome	Time to first occurrence of CVD event	Same
Causal contrasts of interest	Per-protocol effect	Same
Analysis plan	Per-protocol analysis accounting for time-varying adherence	Same

The follow-up period in the target trial vs. MR

👍 Target trial

Unified "time zero"

(1) Randomization /
treatment assigned
(2) Eligibility criteria applied
(3) Outcome recording begins

*Time since
start of
follow-up*

0

*End of
follow-up*
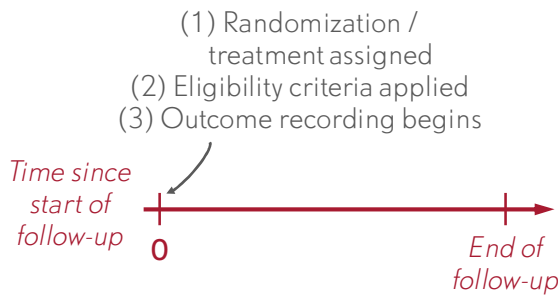
🧬 Mendelian randomization

Misaligned "time zero"
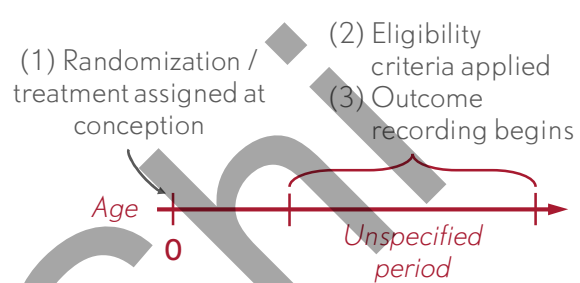
(1) Randomization /
treatment assigned at
conception

(2) Eligibility
criteria applied
(3) Outcome
recording begins

Age

0

*Unspecified
period*

Swanson et al. (Epidemiology 2017)

The follow-up period in the target trial vs. MR

Emulate RCTs with a lag between time of randomization and time of eligibility/outcome recording

❯ Lag time can be years or decades

❯ Average age of 50+ years for participants in MR studies of nutritional factors

Post-randomization selection may introduce bias

❯ Genetic factor (instrument) related to survival, eligibility, etc.

❯ Violation of instrumental conditions
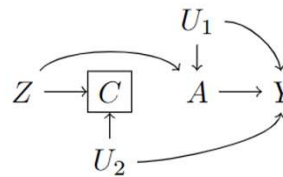
$$Z \rightarrow \boxed{C} \rightarrow A \rightarrow Y$$

with U_1 and U_2

Swanson et al. (Epidemiology 2017)

Analysis plan:
(Adherence to a given) diet changes over time…

👍 <u>Target trial</u>

Expect per-protocol analysis to incorporate repeated measures of adherence

Estimate long-term effects of a given dietary intervention

🧬 <u>Conventional Mendelian randomization</u>

Generally conducted on a single measurement of "adherence"/exposure

MR estimate often interpreted as a "lifetime effect"

Requires strong, unverifiable assumptions about genetic effects over time[†]
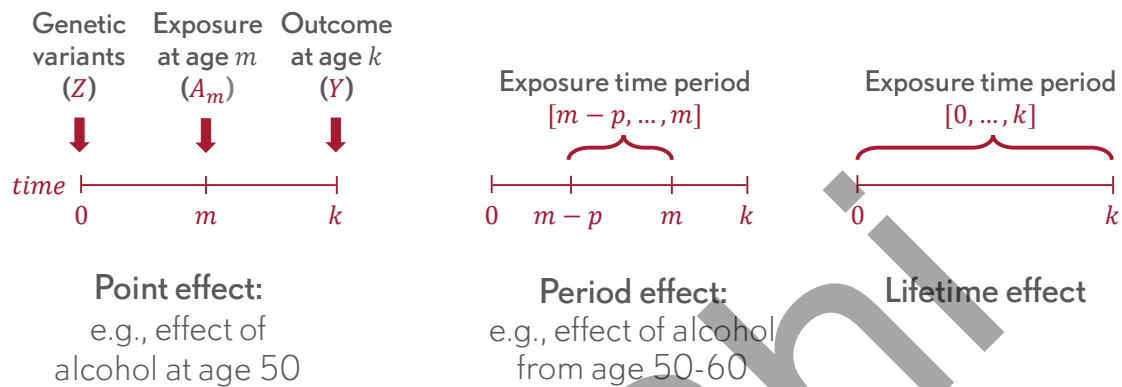
[†]Labrecque and Swanson (AJE 2019)

Analysis plan:
(Adherence to a given) diet changes over time…

❯ Methods which can accommodate repeated measures of a time-varying exposure, e.g.

   o G-estimation of structural nested models[†]

❯ Be explicit about the causal estimand of interest

   o Multiple causal estimands can be targeted

   o Different assumptions needed

$$Z \rightarrow A_0 \rightarrow A_1 \cdots A_k \rightarrow Y$$
$$U$$
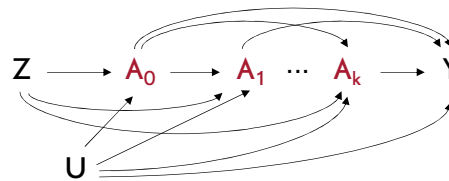
[†]Shi et al. (Epidemiology 2022), Shi et al. (BMC Med Res Methodol 2022)

What types of causal estimands can we target?

Genetic variants (Z) — Exposure at age m (A_m) — Outcome at age k (Y)

$time$ | 0 — m — k

Point effect:
e.g., effect of alcohol at age 50

Exposure time period $[m - p, ..., m]$

0 — $m - p$ — m — k

Period effect:
e.g., effect of alcohol from age 50-60

Exposure time period $[0, ..., k]$

0 — k

Lifetime effect

15

Takeaways

Mendelian randomization (and IV methods) are an appealing alternative to confounding-adjustment methods

Given the complexities of studying nutritional factors, more explicit target trial emulation in MR studies can clarify:

↘ Sources of potential biases (e.g., misalignments in time zero)

↘ An appropriate analytical strategy that accommodates the time-varying nature of diet (e.g., g-estimation of SMMs)

↘ The causal estimand of interest

16

Acknowledgements

Sonja A. Swanson

Miguel A. Hernán

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

CAUSALab
A Center to Learn What Works

HARVARD T.H. CHAN
SCHOOL OF PUBLIC HEALTH

17

Questions

Email: joyshi@hsph.harvard.edu

@joy_shi1

CAUSALab
A Center to Learn What Works

HARVARD
T.H. CHAN
SCHOOL OF PUBLIC HEALTH