Smarter studies
Global impact
Better health

Using causal graphs to understand estimands and estimation

Ian White <ian.white@ucl.ac.uk>

MRC Clinical Trials Unit at UCL

PSI one day event on Estimands and Causal Inference

Reading, 29th January 2019

Background

- Causal inference is a methodology for drawing causal conclusions (what happens if…?) from data
  - clarifies the assumptions needed
  - identifies when standard statistical methods fail
  - proposes new classes of statistical methods
- Directed acyclic graphs (the "causal graphs" of the title) are widely used in causal inference
- Causal inference is usually used in observational studies, but causal questions also arise in RCTs
  - especially around estimands: "hypothetical" and "principal stratum" strategies for defining estimands

Aim

- To discuss estimands and estimation
  - in RCTs with intercurrent events in the form of treatment changes
  - using directed acyclic graphs (causal graphs, DAGs)
- A new way of looking at existing methods
  - hopefully informative

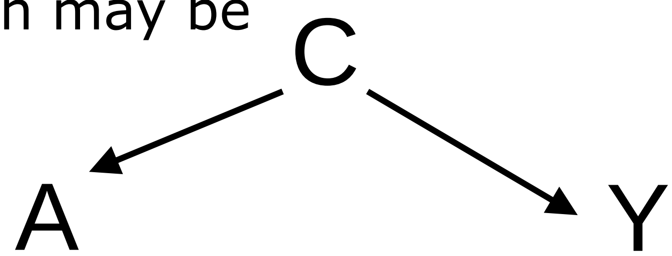
Plan

1. What is a DAG, and how do we use them?
2. DAG for a RCT with intercurrent events (treatment changes) just after baseline
3. Estimation methods for this setting
4. DAG for a RCT with intercurrent events (treatment changes) occurring over time
5. Estimation methods for this setting
6. Examples
7. Conclusions

Plan

1. **What is a DAG, and how do we use them?**
2. DAG for a RCT with intercurrent events (treatment changes) just after baseline
3. Estimation methods for this setting
4. DAG for a RCT with intercurrent events (treatment changes) occurring over time
5. Estimation methods for this setting
6. Examples
7. Conclusions

Idea of DAGs



- Graph: set of nodes (variables) which may be connected or unconnected
- Directed: connections are arrows
- Acyclic: no loops
- Absence of an arrow implies independence
  - here, A and Y are independent, conditional on C
- I'll assume a "causal DAG" in which arrows may be interpreted as causal effects
  - changing C affects A and Y
  - changing A doesn't affect Y
- Requires that "the common causes of any pair of variables in the graph are also in the graph"
  - Hernán MA, Robins JM (2019). Causal Inference. Chapman & Hall/CRC, forthcoming & on the web

Paths in DAGs imply potential associations

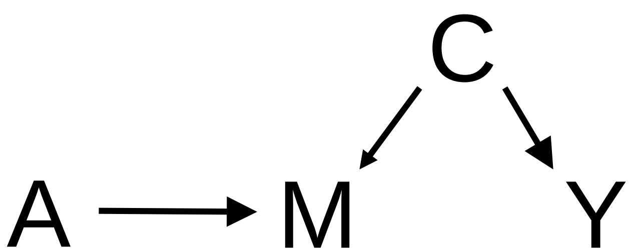
Front-door paths
from A to Y: start A→

Back-door path
from A to Y: starts A←

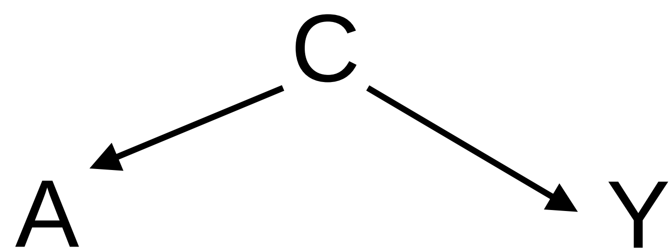
(a)

(b)

(c)



(d)



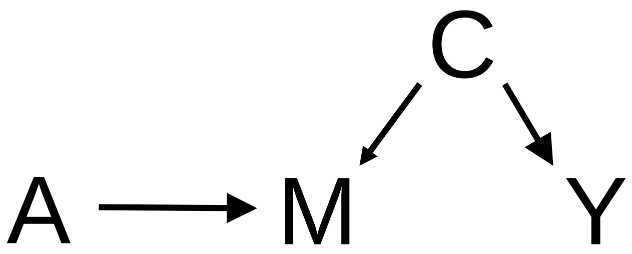
A front-door path is causal if all arrows point from A to Y (a,b but not c)

A back-door path is never causal

Colliders

- Here we have a path of form … → M ← …
- M is a collider on this path

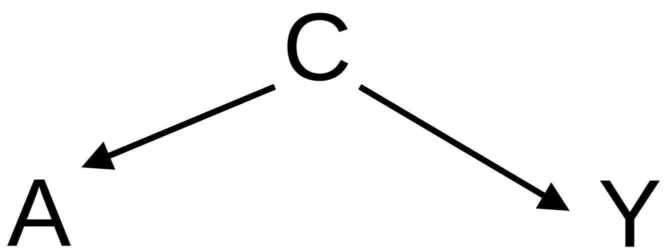
(c)



Open and blocked paths in DAGs

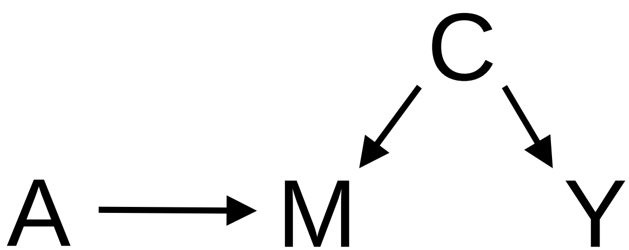
(a)

(d)



A path is open…

(c)

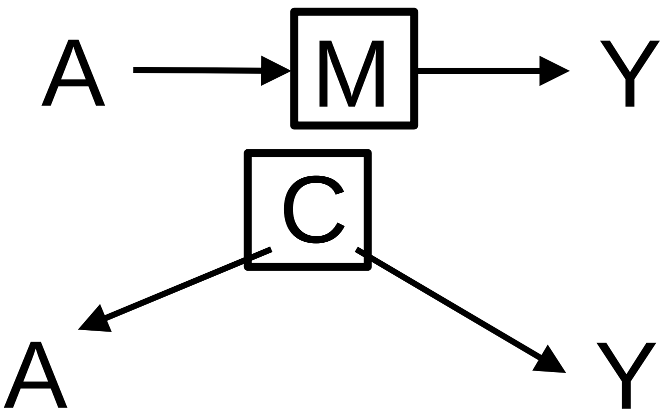


…unless it passes through a collider, when it is blocked.

Open paths contribute to associations.
Blocked paths don't.

Controlling for a variable

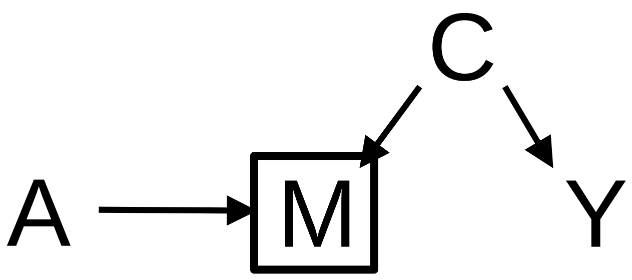
(a)
(d)



Box indicates controlling in analysis

Controlling for a non-collider blocks a path

(c)

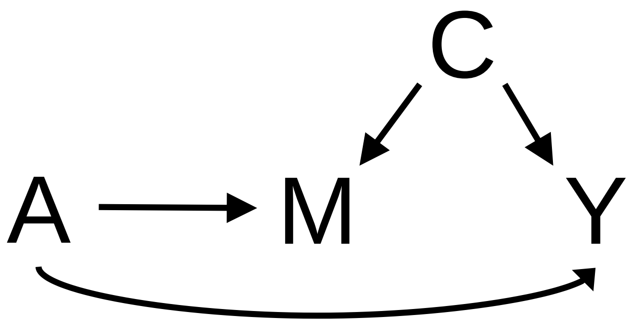


Controlling for a collider opens a path (if it's otherwise open)

Open paths contribute to associations.
Blocked paths don't.

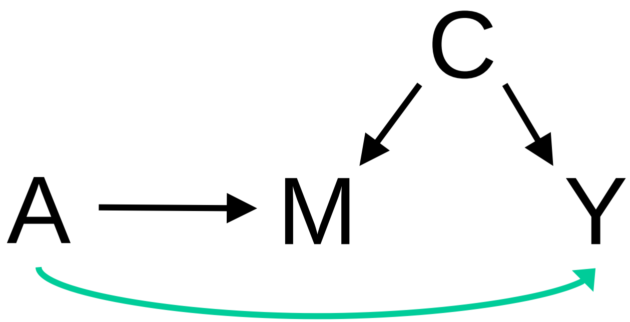
Key point

- An observed association has a causal interpretation if all open paths are causal
- Consider this DAG

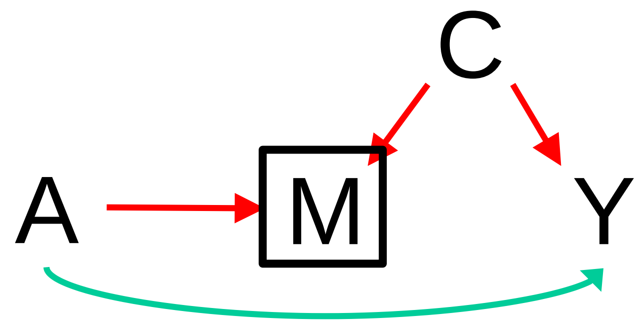


Key point

- An observed association has a causal interpretation if all open paths are causal
- Consider this DAG



One open path:
observed A-Y association
is causal



Two open paths:
observed A-Y association
is not causal

Plan

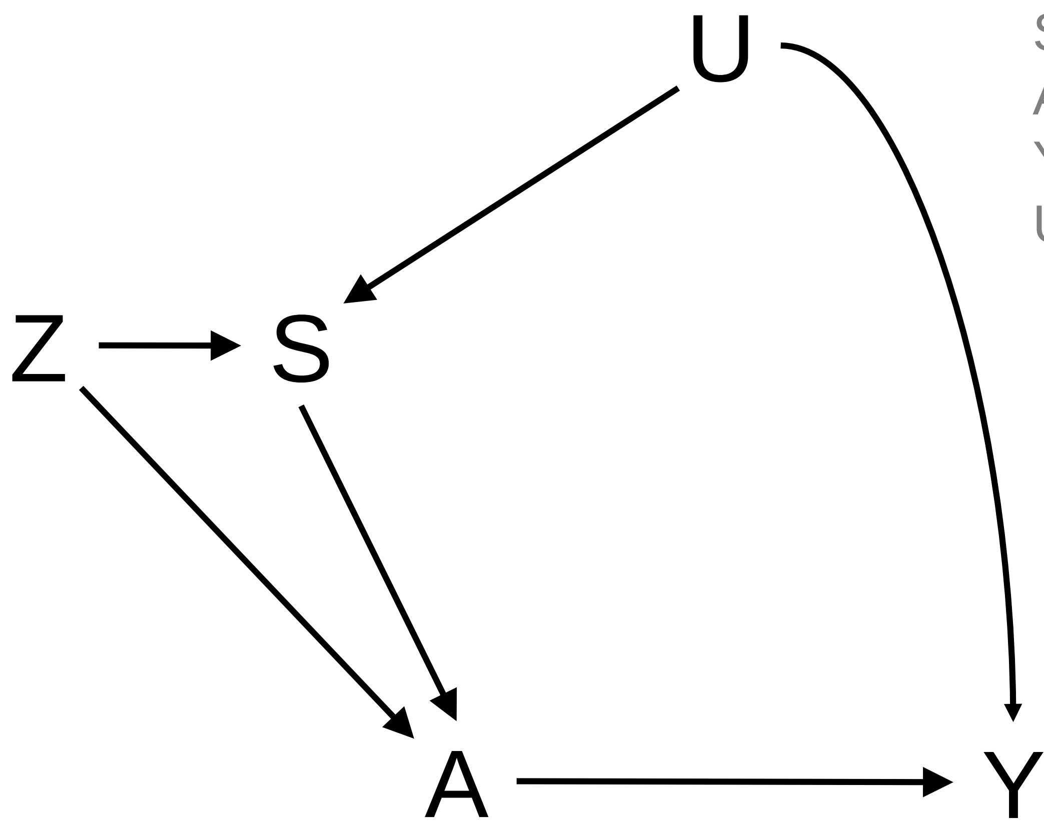
1. What is a DAG, and how do we use them?
2. **DAG for a RCT with intercurrent events (treatment changes) just after baseline**
3. Estimation methods for this setting
4. DAG for a RCT with intercurrent events (treatment changes) occurring over time
5. Estimation methods for this setting
6. Examples
7. Conclusions

Intercurrent events

Assume we have various types of intercurrent events (treatment changes) and we have decided how to handle them

- Events handled through a "treatment policy" strategy
  - we ignore these in DAG & analysis
- Events handled through a "composite strategy"
  - we include these in outcome Y
  - otherwise ignore them
- Events handled through a "hypothetical strategy"
  - we define S to indicate the occurrence of such events (S for switching, but could be any treatment change)
  - and A to be the actual treatment then followed
  - we include S and A in the DAG

DAG with treatment changes just after baseline

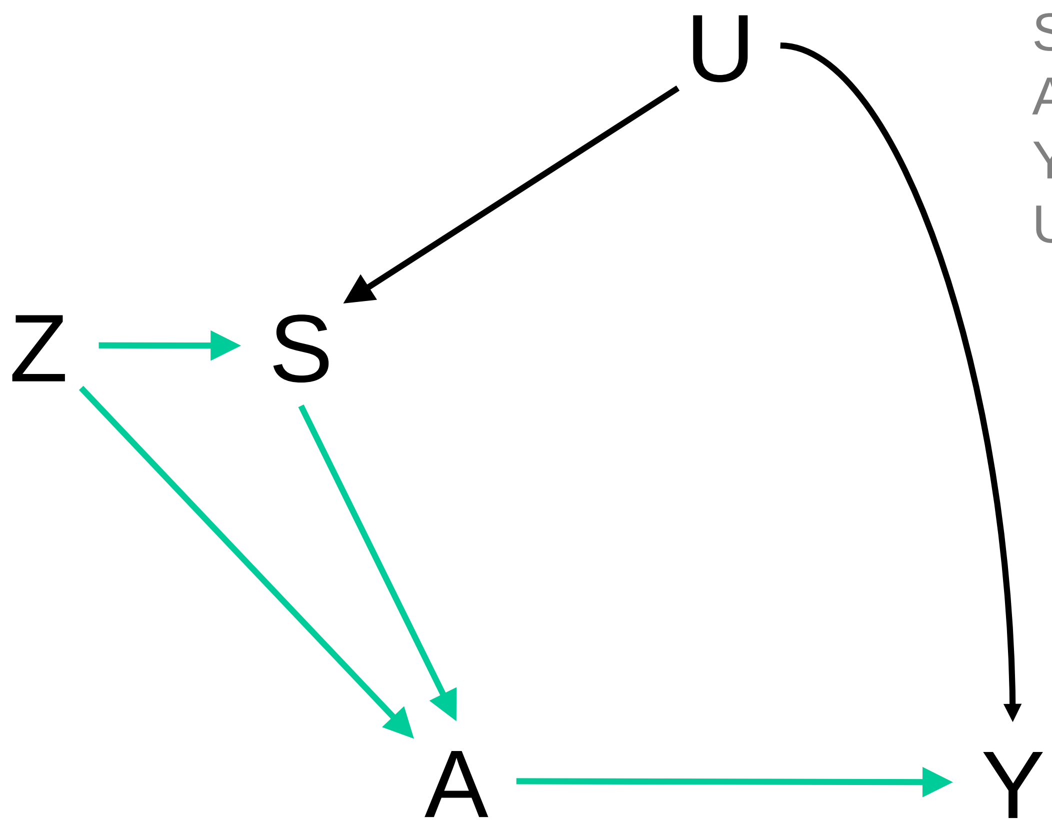


Z = Randomization
S = Switching treatment
A = Treatment
Y = Outcome
U = Unmeasured variable

We want causal effect of A on Y

But first we'll look at causal effect of Z on Y

Causal effect of Z on Y is easy



Z = Randomization
S = Switching treatment
A = Treatment
Y = Outcome
U = Unmeasured variable

ITT analysis: two open paths, both causal

But we want causal effect of A on Y

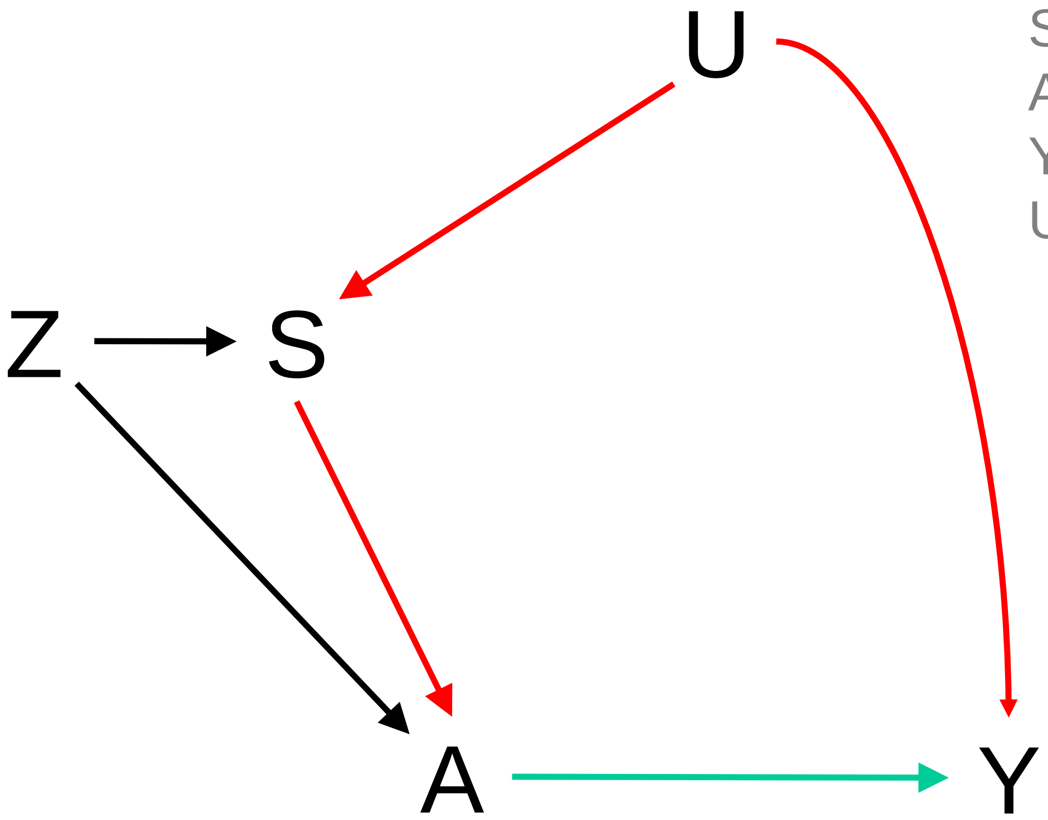
Plan

1. What is a DAG, and how do we use them?
2. DAG for a RCT with intercurrent events (treatment changes) just after baseline
3. **Estimation methods for this setting**
4. DAG for a RCT with intercurrent events (treatment changes) occurring over time
5. Estimation methods for this setting
6. Examples
7. Conclusions

Estimating causal effect of A on Y is harder

Regress Y on A?

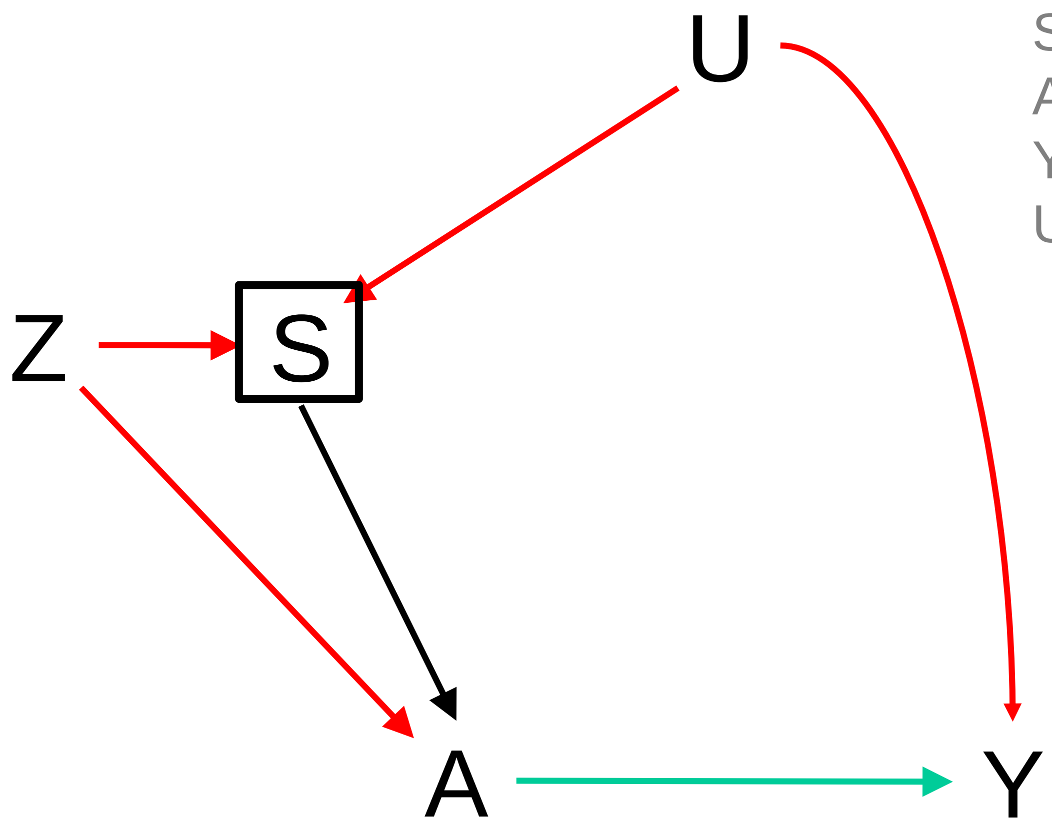
Z = Randomization
S = Switching treatment
A = Treatment
Y = Outcome
U = Unmeasured variable



No: there is a non-causal back-door path

Per-protocol (PP) analysis doesn't gives effect of A on Y

Regress Y on A, conditioning on S=0 (PP analysis)?

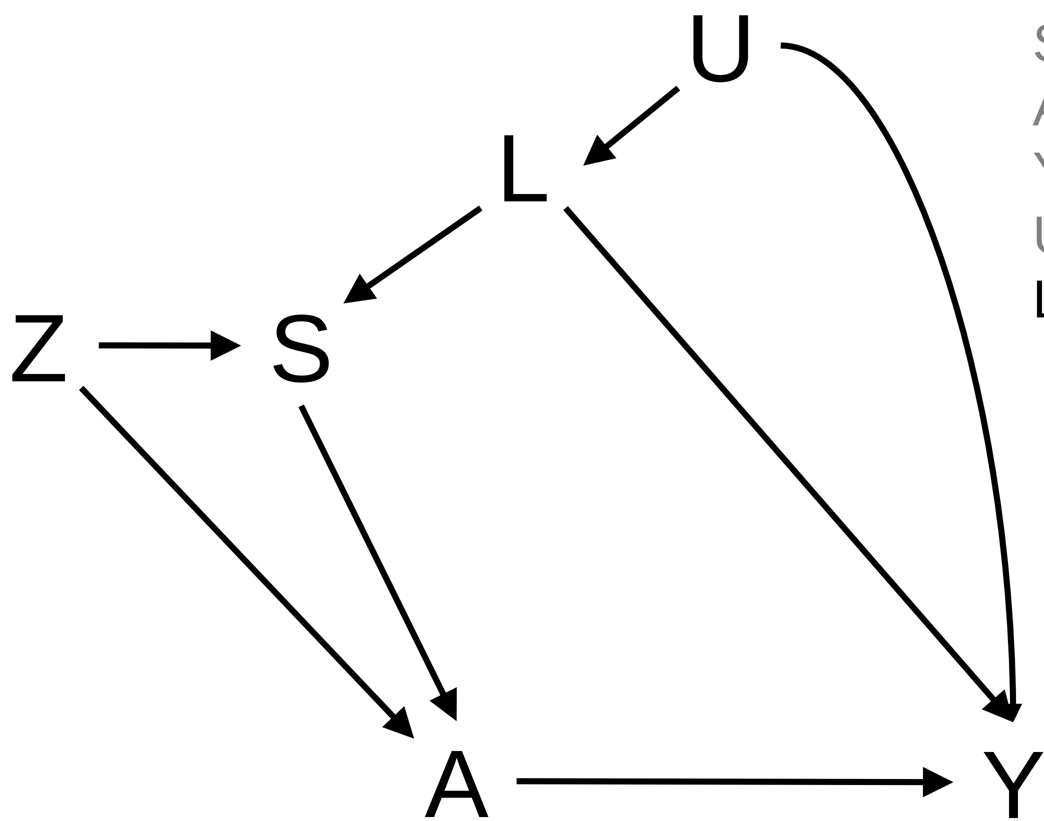


Z = Randomization
S = Switching treatment
A = Treatment
Y = Outcome
U = Unmeasured variable

No: opens a non-causal back-door path

But what if we measure all confounders L?

New DAG using L

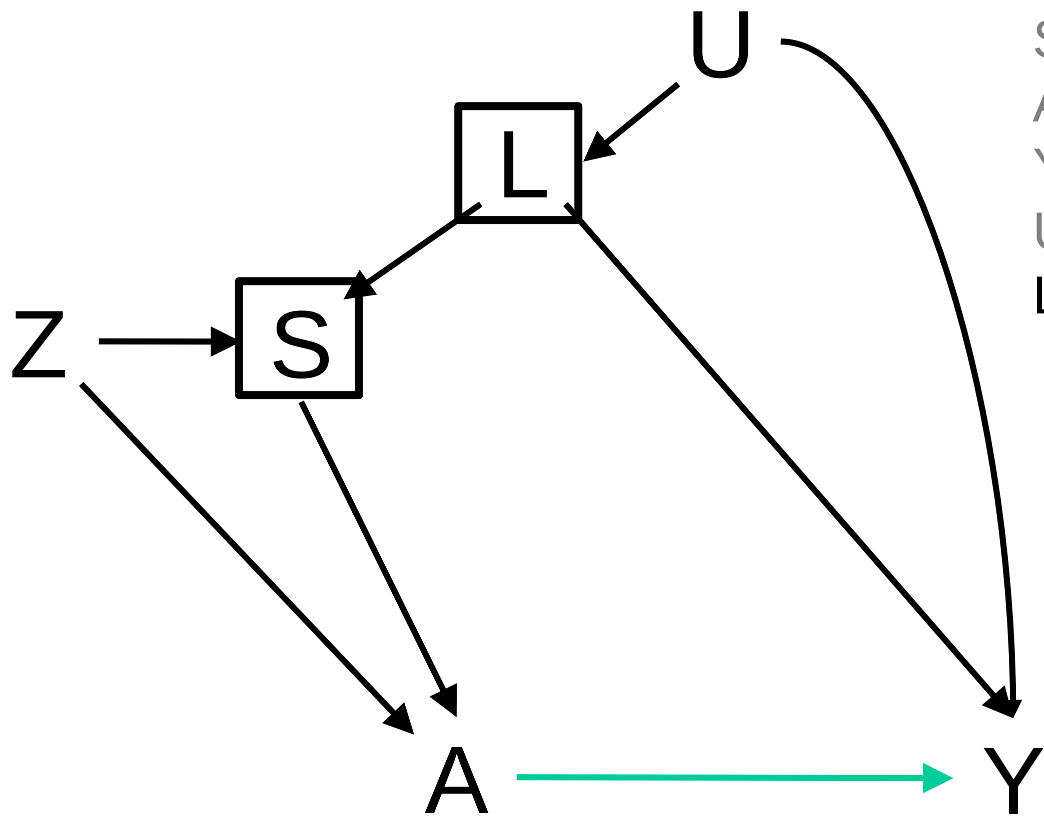


Z = Randomization
S = Switching treatment
A = Treatment
Y = Outcome
U = Unmeasured variable
L = Confounder

NB no direct path U→S

PP is OK if we measure all confounders

Regress Y on A, conditioning on S=0 and L (adjusted PP analysis)?



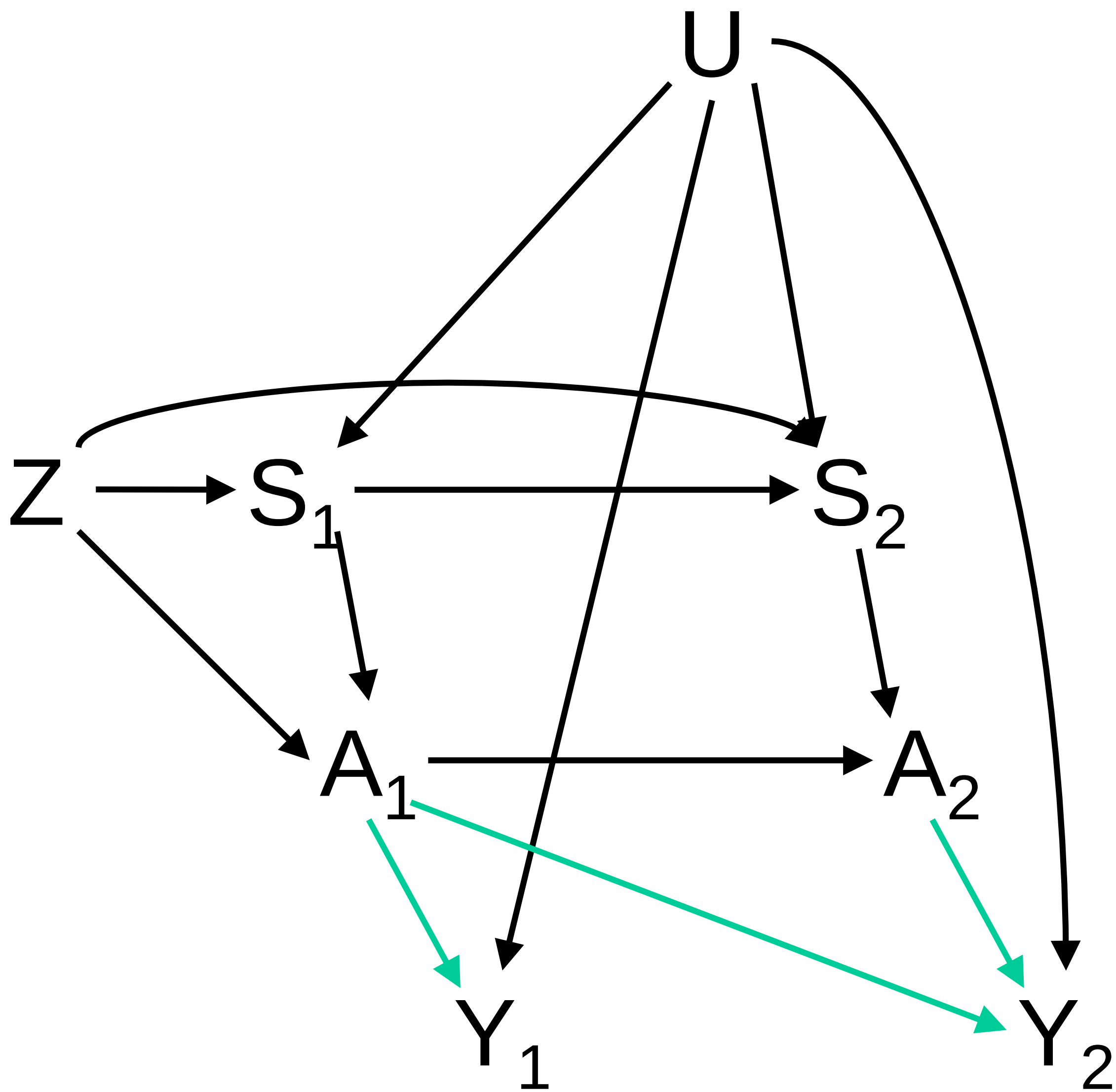
Z = Randomization
S = Switching treatment
A = Treatment
Y = Outcome
U = Unmeasured variable
L = Confounder

OK!
Back-door path is blocked by conditioning on L as well as S=0

Plan

1. What is a DAG, and how do we use them?
2. DAG for a RCT with intercurrent events (treatment changes) just after baseline
3. Estimation methods for this setting
4. **DAG for a RCT with intercurrent events (treatment changes) occurring over time**
5. Estimation methods for this setting
6. Examples
7. Conclusions

DAG with treatment changes over time



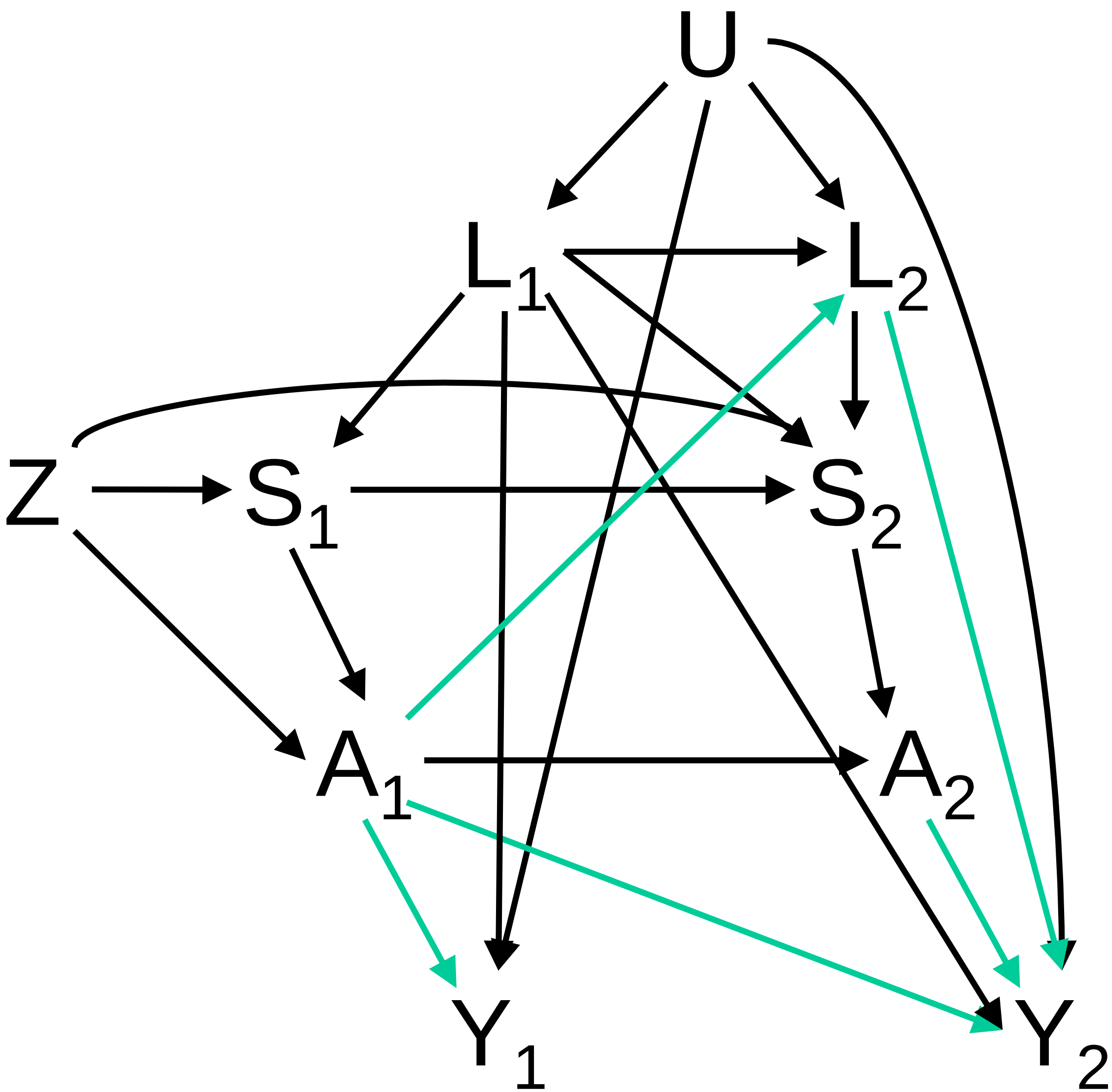
Here illustrated with only 2 intervals, for clarity

Y_1, Y_2,… could be quantitative or binary (for time-to-event outcome)

We want
- effect of A_1 on Y_1
- effects of A_1 and A_2 on Y_2

3 causal paths

DAG with treatment changes over time + no unmeasured confounders



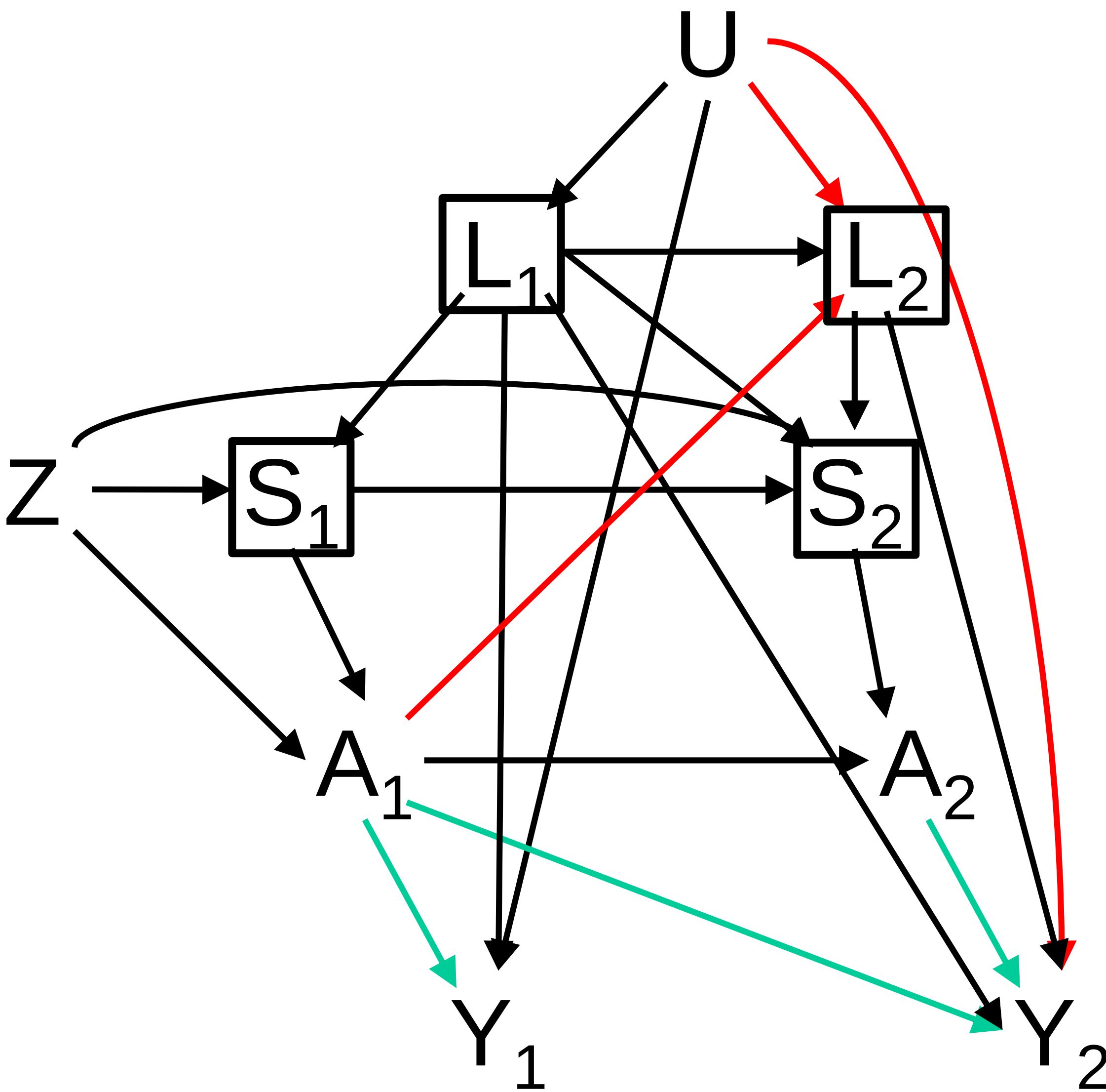
Z = Randomization
S = Switching treatment
A = Treatment
Y = Outcome
U = Unmeasured variable
L = Confounder

We want

- effect of A_1 on Y_1
- effects of A_1 and A_2 on Y_2

4 causal paths

DAG with treatment changes over time + no unmeasured confounders



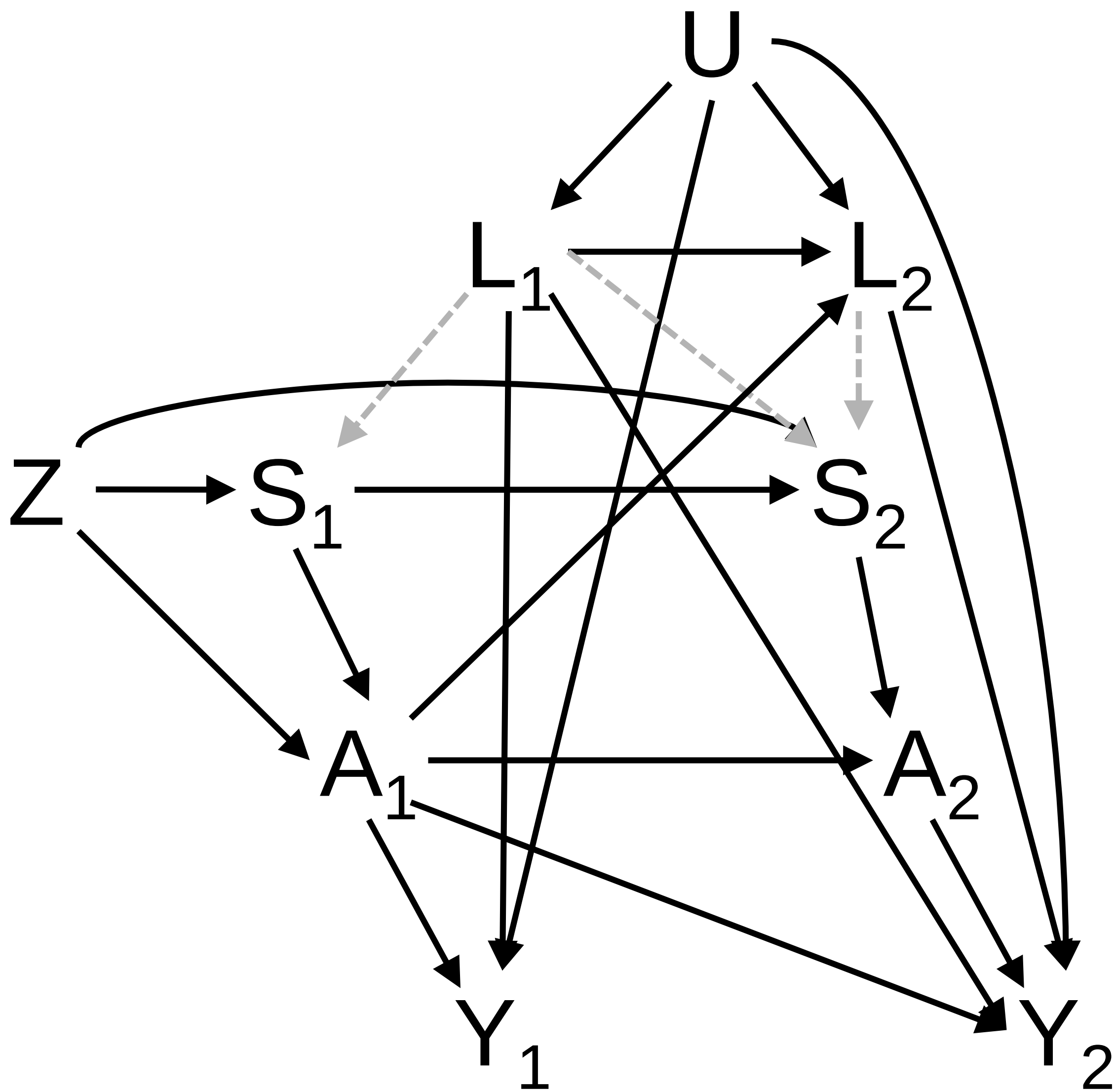
PP analysis controlling for L_1 and L_2?

- No!
- Opens non-causal path via U
- Blocks causal path via L_2
- Wrong estimation of the effects of A_1 and A_2 on Y_2
- "Time-varying confounding"

Plan

1. What is a DAG, and how do we use them?
2. DAG for a RCT with intercurrent events (treatment changes) just after baseline
3. Estimation methods for this setting
4. DAG for a RCT with intercurrent events (treatment changes) occurring over time
5. **Estimation methods for this setting**
6. Examples
7. Conclusions

DAG with treatment changes over time + no unmeasured confounders: IPCW

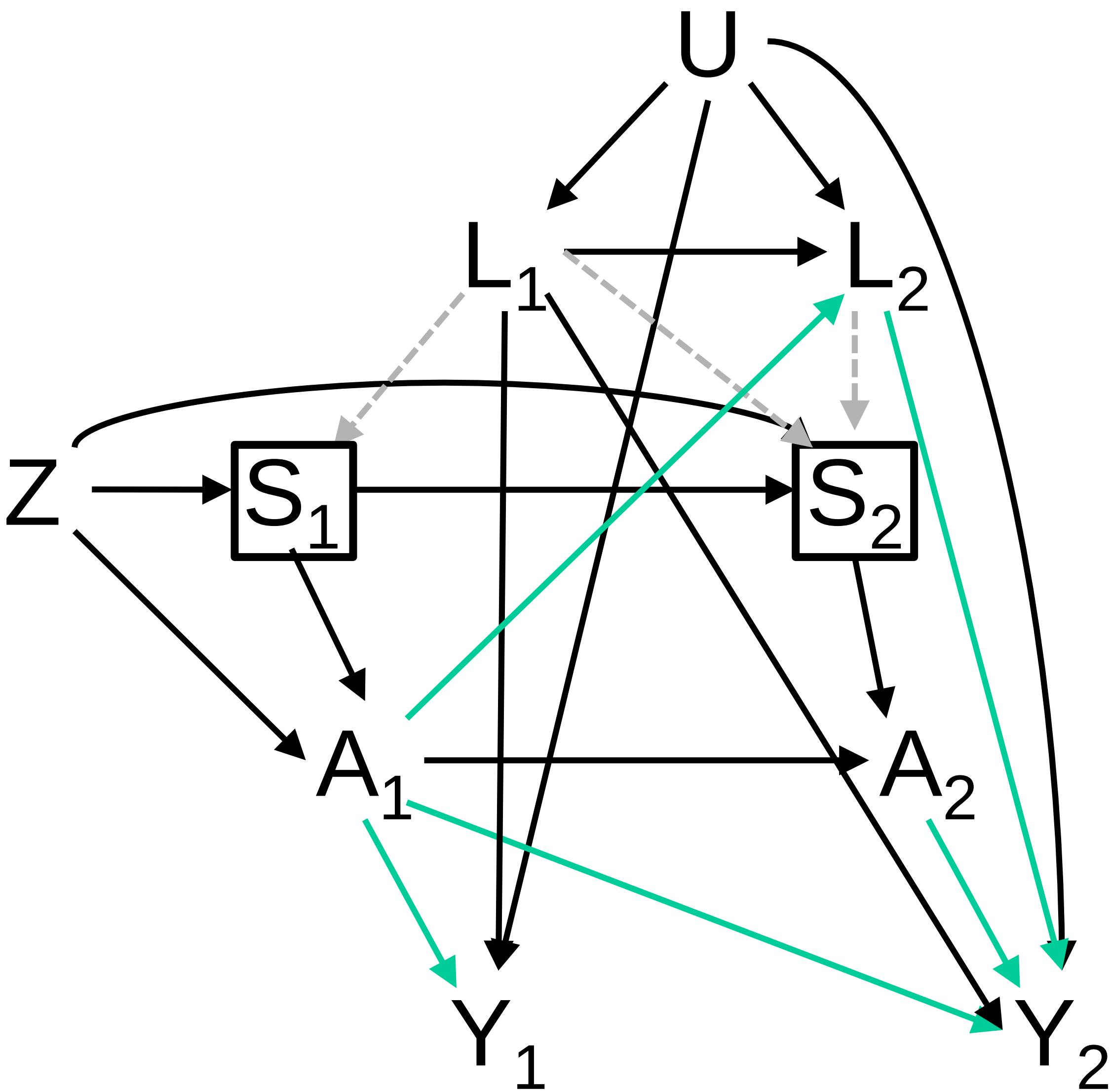


IPCW = inverse-probability-of-censoring weighting

Weight by $1/p(S=0|\text{past } L\text{'s})$

- removes dashed arrows L→S

DAG with treatment changes over time + no unmeasured confounders: IPCW



IPCW = inverse-probability-of-censoring weighting

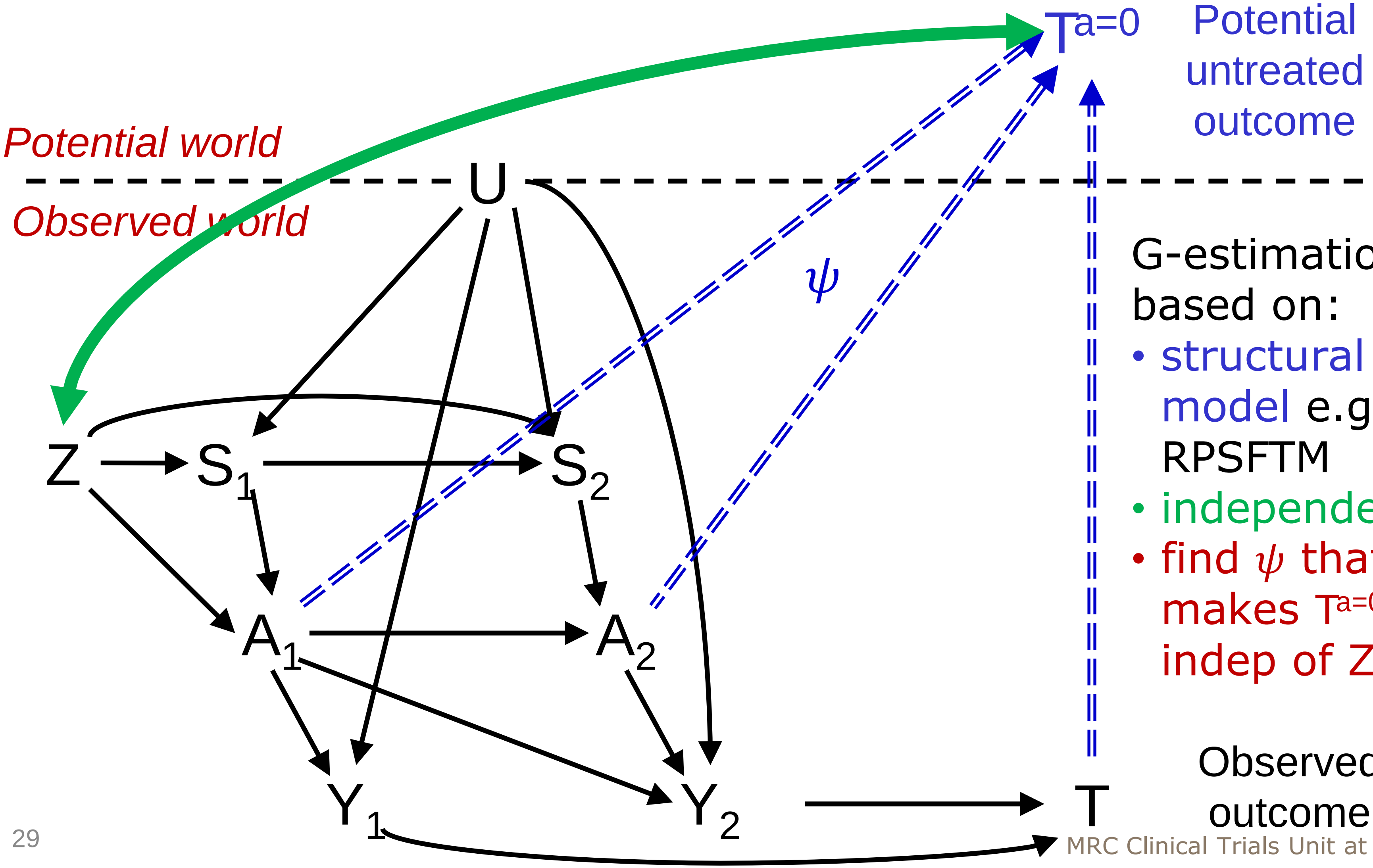
Weight by 1/p(S=0|past L's)

- removes dashed arrows L→S

and use PP (condition on S=0 i.e. censor at S)

- No open back-door paths now

DAG with treatment changes over time and unmeasured confounders: instrumental variable (IV) / G-estimation method



G-estimation is based on:

- structural model e.g. RPSFTM
- independence
- find ψ that makes $T^{a=0}$ indep of Z

Modelling challenges

- PP and IPCW condition on S=0
  - so A=Z in the data used
  - so we only have to model the effect of randomised treatment
- Structural model / RPSFTM approach doesn't condition on S=0
  - have to model A → outcome
  - gets harder as A gets more complex

Two-stage method

- Another method that works well is the two-stage method (Latimer et al 2014)
  - uses observational analysis within arms to estimate effect of treatment in S=1
- Like IPCW, based on no unmeasured confounders
- Like IV method, requires modelling A → outcome
- DAGs omitted here

Plan

1. What is a DAG, and how do we use them?
2. DAG for a RCT with intercurrent events (treatment changes) just after baseline
3. Estimation methods for this setting
4. DAG for a RCT with intercurrent events (treatment changes) occurring over time
5. Estimation methods for this setting
6. **Examples**
7. Conclusions

A late-stage cancer trial

- Consider a trial of SoC+new drug vs. SoC+placebo
- Outcomes: PFS (prog-free survival), OS (overall surv)
- Intercurrent events for OS analysis:
  1. Patients may stop new drug
     - use treatment-policy estimand
  2. Placebo patients often start new drug after progression
     - use hypothetical estimand
- We may be confident in defining a causal model that applies in both arms
  - use RPSFTM?
- We may be confident that we've measured all time-dependent confounders
  - use IPCW / two-stage?

An epilepsy trial

- Consider a trial of a new anti-epilepsy medicine
- Outcome: time to seizure-free 12 months
- Intercurrent events:
  - patients with inadequate control may switch to any of 4 other drugs
    - use hypothetical estimand
- We're unlikely to be confident about a causal model for all drugs (& doses)
- Important to measure all time-dependent confounders and use IPCW / two-stage

Plan

1. What is a DAG, and how do we use them?
2. DAG for a RCT with intercurrent events (treatment changes) just after baseline
3. Estimation methods for this setting
4. DAG for a RCT with intercurrent events (treatment changes) occurring over time
5. Estimation methods for this setting
6. Examples
7. **Conclusions**

Acknowledgements

- Uwe Siebert & Feli Kühne (UMIT, Austria)
- Nick Latimer (Sheffield)
- We give a course, "Causal Inference in Observational Studies and Clinical Trials Affected by Treatment Switching: A Practical Hands-on Workshop"
  - next: 18-21 March 2019, Hall-in-Tirol, Austria
  - search for "UMIT Causal"

Conclusions

- DAGs may be a useful way to understand why analysis methods do / don't work
- Including intercurrent events in the DAG means adopting a hypothetical estimand for them