Time varying exposures

Arvid Sjölander

Department of Medical Epidemiology and Biostatistics
Karolinska Institutet

A short course on concepts and methods in Causal Inference

From point exposures to time varying exposures

- Until now, we have only considered point exposures
  - i.e. exposures that occur at one single occasion
- Many epidemiological studies are longitudinal, and involve exposures that occur repeatedly
- Causal inference becomes more complex when the exposure varies with time

Motivating example (Robins, 1997)

- Aim: to estimate the effect of AZT on risk for HIV related infections for AIDS patients
- Design:
  - At $t = 0$: 32.000 subjects randomized to AZT ($A^0 = 1$) or placebo ($A^0 = 0$)
  - At $t = 1$:
    - CD4 count recorded; $L^1 = 0$ if 'low', $L^1 = 1$ if 'high'
    - all subjects with $L^1 = 0$ receive AZT ($A^1 = 1$)
    - subjects with $L^1 = 1$ are randomized to AZT ($A^1 = 1$) or placebo ($A^1 = 0$)
  - At $t = 2$: $Y = 1$ if no serious infection, $Y = 0$ else
  - No drop out or death during follow up

Data

A^0	L^1	A^1	$Y = 1$	$Y = 0$
0	0	0	0	0
0	0	1	10	6
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	4	4
1	1	0	1	3
1	1	1	3	1

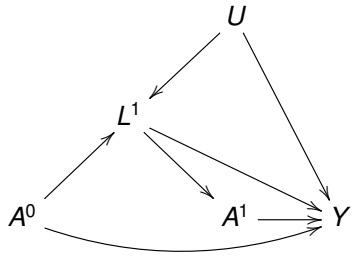
How do we analyze these data?

Outline

Outline

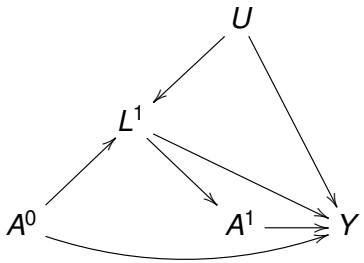
The DAG

- Before discussing analysis alternatives, we must agree on a plausible causal structure (i.e. a DAG)
  - We will assume the DAG



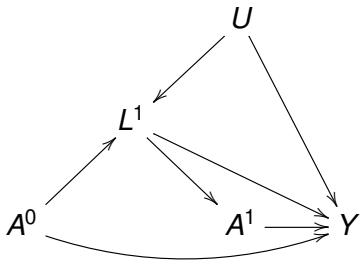
- *Which assumptions are encoded in this DAG? Are they reasonable?*

The causal effect of A^0 on Y



- Suppose we want to test whether there is a causal effect of A^0 on Y
- *What to adjust for?*

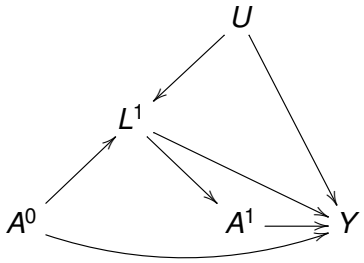
Solution



- Suppose we don't adjust for anything, and that we observe an association between A^0 and Y
- Three explanations
  - $A^0 \rightarrow Y$
  - $A^0 \rightarrow L^1 \rightarrow Y$
  - $A^0 \rightarrow L^1 \rightarrow A^1 \rightarrow Y$

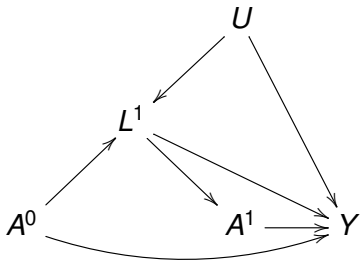
  all causal

Solution, cont'd



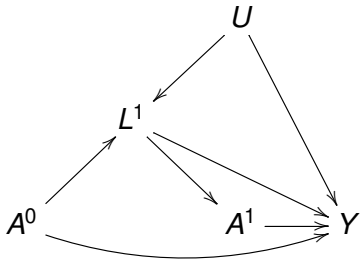
- If the aim is to test for a causal effect of A^0 on Y, then we should not adjust for anything
- In terms of potential outcomes
  - exposed ($A^0 = 1$) and unexposed ($A^0 = 0$) are unconditionally exchangeable

The causal effect of A^1 on Y



- Suppose we want to test whether there is a causal effect of A^1 on Y
- *What to adjust for?*

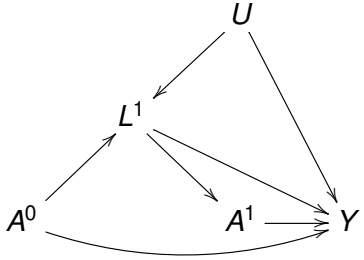
Solution



- Suppose we don't adjust for anything, and that we observe an association between A^1 and Y
- Four explanations:
  - $A^1 \rightarrow Y$
  - $A^1 \leftarrow L^1 \rightarrow Y$
  - $A^1 \leftarrow L^1 \leftarrow A^0 \rightarrow Y$
  - $A^1 \leftarrow L^1 \leftarrow U \rightarrow Y$

  only the first is causal

Solution, cont'd



- The three non-causal paths

$$A^1 \leftarrow L^1 \rightarrow Y,$$

$$A^1 \leftarrow L^1 \leftarrow A^0 \rightarrow Y,$$

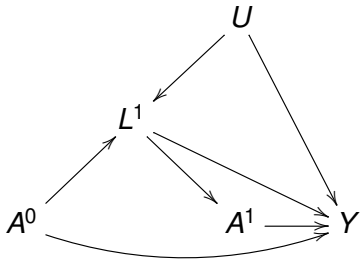
and

$$A^1 \leftarrow L^1 \leftarrow U \rightarrow Y$$

can be blocked by adjusting for L^1 and A^0

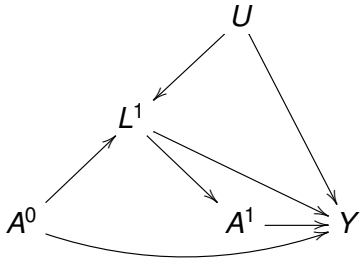
- In fact, adjusting for L^1 is enough

Solution, cont'd



- If the aim is to test for a causal effect of A^1 on Y, then we should adjust for A^0 and L^1 (or just L^1)
- In terms of potential outcomes
  - exposed ($A^1 = 1$) and unexposed ($A^1 = 0$) are conditionally exchangeable, given A^0 and L^1

Sequential adjustment and sequential exchangeability



- To summarize, adjusting for the 'observed past' at each time point produces exchangeability
  - The observed past at $t = 0$: nothing
  - The observed past at $t = 1$: A^0 and L^1
- We say that **sequential adjustment** produces **sequential exchangeability**

Data analysis

A^0	L^1	A^1	$Y=1$	$Y=0$
0	0	0	0	0
0	0	1	10	6
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	4	4
1	1	0	1	3
1	1	1	3	1

$$\begin{aligned}\Pr(Y=1|A^0=1) &= 8/16 \\ &= 0.5\end{aligned}$$

$$\begin{aligned}\Pr(Y=1|A^0=0) &= 10/16 \\ &= 0.63\end{aligned}$$

- A^0 has a negative causal effect on Y
  - Rather counterintuitive, explanation later

Data analysis, cont'd

A^0	L^1	A^1	$Y=1$	$Y=0$
0	0	0	0	0
0	0	1	10	6
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	4	4
1	1	0	1	3
1	1	1	3	1

$$\begin{aligned}\Pr(Y=1|A^0=1,L^1=1,A^1=1) \\ = 3/4 = 0.75\end{aligned}$$

$$\begin{aligned}\Pr(Y=1|A^0=1,L^1=1,A^1=0) \\ = 1/4 = 0.25\end{aligned}$$

- A^1 has a positive conditional causal effect on Y, given $(A^0=1,L^1=1)$
  - The only stratum for which the effect can be calculated

Limitations of sequential adjustment

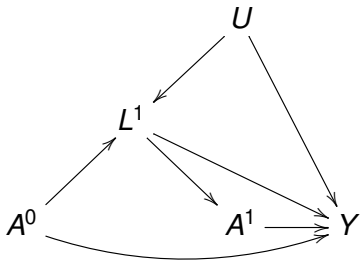
- Sequential adjustment gives
  - the marginal (population) causal effect of A^0
  - the conditional causal effect of A^1, given (A^0, L^1)
- Suppose we want to know which combination of (A^0, A^1) is most favorable; $(0,0)$, $(0,1)$, $(1,0)$, or $(1,1)$
- This cannot be answered by sequential adjustment
  - Sequential adjustment gives the effect of A^0 and A^1 separately, for different subpopulations
  - The question concerns the joint effect of (A^0, A^1), for the same population

Outline

An incorrect argument

- We have observed a negative effect of A^0 and a positive effect of A^1, so the best combination is $(A^0 = 0, A^1 = 1)$
- *Why is this argument not valid?*

Solution



- The argument is invalid because the effect of A^0 is partly mediated through A^1
  - AZT at $t = 0 \Rightarrow$ increased CD4 count at $t = 1$
  - $\Rightarrow$ decreased chance of getting AZT at $t = 1$
  - $\Rightarrow$ increased risk for infection at $t = 2$
- But A^0 could have a positive 'direct' effect on Y
  - i.e. not mediated through A^1
  - then the best combination could be $(A^0 = 1, A^1 = 1)$

Potential outcomes under more than one exposure

- Before estimating joint causal effects, we need to define the concept
  - Towards this end we need to consider potential outcomes under more than one exposure
- $Y_{a^0 a^1}$ = the potential outcome under joint exposures $A^0 = a^0$ and $A^1 = a^1$
  - E.g. Y_{10} = the potential outcome when receiving AZT at $t = 0$, but not at $t = 1$

Potential outcomes under more than one exposure, cont'd

- $\Pr(Y_{a^0 a^1} = 1)$ = the proportion of subjects without infection if the whole population receives $A^0 = a^0$ and $A^1 = a^1$
  - e.g. $\Pr(Y_{10} = 1)$ = the proportion of subjects without infection if the whole population cohort receives AZT at $t = 0$, but not at $t = 1$

Joint causal effects

- The four counterfactual probabilitites

$$\Pr(Y_{11} = 1), \quad \Pr(Y_{10} = 1), \quad \Pr(Y_{01} = 1), \quad \Pr(Y_{00} = 1)$$

  together describe how the population reacts to all four combinations of A^0 and A^1
- Together, they quantify the joint causal effect of (A_0, A_1), for the whole population
- How do we estimate these probabilities?

Unadjusted analysis

A^0	L^1	A^1	$Y = 1$	$Y = 0$
0	0	0	0	0
0	0	1	10	6
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	4	4
1	1	0	1	3
1	1	1	3	1

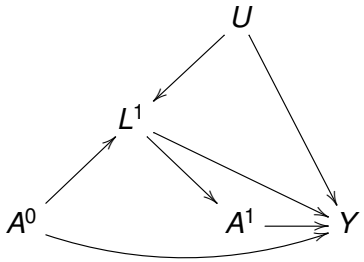
- Lets compare those who factually received $(1, 0)$ and $(1, 1)$ without adjusting

$$\Pr(Y = 1|A^0 = 1, A^1 = 0) = 1/4 = 0.25$$

$$\Pr(Y = 1|A^0 = 1, A^1 = 1) = 7/12 = 0.58$$

- Can we conclude that (1,1) beats (1,0)?

Problem with the unadjusted analysis



- By not adjusting for L^1, we leave the non-causal path $A^1 \leftarrow L^1 \rightarrow Y$ open
  - An unadjusted association between A^1 and Y cannot be given a causal interpretation

Adjusted analysis

A^0	L^1	A^1	$Y = 1$	$Y = 0$
0	0	0	0	0
0	0	1	10	6
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	4	4
1	1	0	1	3
1	1	1	3	1

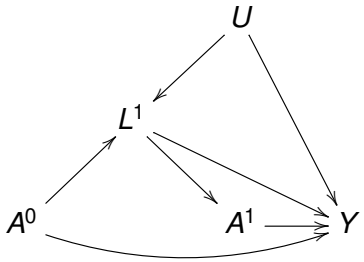
- Lets compare those who factually received $(0, 1)$ and $(1, 1)$, stratifying on $L^1 = 0$

$$\Pr(Y = 1|A^0 = 0, L^1 = 0, A^1 = 1) = 10/16 = 0.625$$

$$\Pr(Y = 1|A^0 = 1, L^1 = 0, A^1 = 1) = 4/8 = 0.5$$

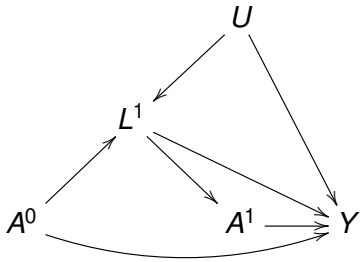
- Can we conclude that (0,1) beats (1,1), for subjects with $L^1 = 0$?

Problem with the adjusted analysis



- By adjusting for L^1, we open the non-causal path $A^0 \rightarrow L^1 \leftarrow U \rightarrow Y$ open
  - An association between A^0 and Y, adjusted for L^1, cannot be given a causal interpretation

Conclusion



- Damned if you do (adjust), damned if you don't!
- Fortunately there is a solution to this problem

Sequential standardization

- Under sequential exchangeability, we can use **sequential standardization** to calculate $\Pr(Y_{a^0a^1} = 1)$

$$\Pr(Y_{a^0a^1} = 1) = \sum_{L^1} \Pr(Y = 1 | A^0 = a^0, L^1, A^1 = a^1) \Pr(L^1 | A^0 = a^0)$$

  - It will be clear later why it is called 'sequential standardization'
  - Also known as the **G-formula**

Analysis

A^0	L^1	A^1	$Y = 1$	$Y = 0$
0	0	0	0	0
0	0	1	10	6
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	4	4
1	1	0	1	3
1	1	1	3	1

$$\Pr(Y_{01} = 1) = \overbrace{10/16}^{\Pr(Y=1|A^0=0,L^1=0,A^1=1)} \times \overbrace{1}^{\Pr(L^1=0|A^0=0)} + \underbrace{?}_{\Pr(Y=1|A^0=0,L^1=1,A^1=1)} \times \underbrace{0}_{\Pr(L^1=1|A^0=0)} = 0.625$$

- *Compute* $\Pr(Y_{11} = 1)$, $\Pr(Y_{10} = 1)$, and $\Pr(Y_{00} = 1)$

Solution

$$\Pr(Y_{11}=1) = \overbrace{4/8}^{\Pr(Y=1|A^0=1,L^1=0,A^1=1)} \times \overbrace{8/16}^{\Pr(L^1=0|A^0=1)} + \underbrace{3/4}_{\Pr(Y=1|A^0=1,L^1=1,A^1=1)} \times \underbrace{8/16}_{\Pr(L^1=1|A^0=1)} = 0.625$$

$$\Pr(Y_{10}=1) = \overbrace{?}^{\Pr(Y=1|A^0=1,L^1=0,A^1=0)} \times \overbrace{8/16}^{\Pr(L^1=0|A^0=1)} + \underbrace{1/4}_{\Pr(Y=1|A^0=1,L^1=1,A^1=0)} \times \underbrace{8/16}_{\Pr(L^1=1|A^0=1)} =?$$

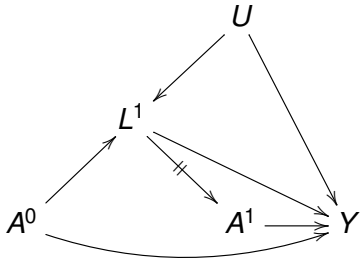
$$\Pr(Y_{00}=1) = \overbrace{?}^{\Pr(Y=1|A^0=0,L^1=0,A^1=0)} \times \overbrace{1}^{\Pr(L^1=0|A^0=0)} + \underbrace{?}_{\Pr(Y=1|A^0=0,L^1=1,A^1=0)} \times \underbrace{0}_{\Pr(L^1=1|A^0=0)} =?$$

Conclusion

- $\Pr(Y_{11}=1) = 0.625$
- $\Pr(Y_{01}=1) = 0.625$
- $\Pr(Y_{10}=1) =?$
- $\Pr(Y_{00}=1) =?$
- The combinations (0,1) and (1,1) perform equally well
- The combinations (1,0) and (0,0) cannot be evaluated given the observed data

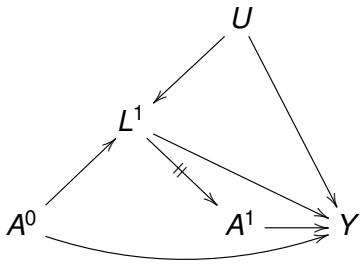
Outline

Direct causal effects



- A comparison between $\Pr(Y_{11}=1)$ and $\Pr(Y_{01}=1)$ is a 'direct' causal effect of A^0, not mediated through A^1
  - By holding A^1 fixed at 1, we eliminate the path $L^1 \to A^1$
  - Any remaining effect goes through $A^0 \to Y$ or $A^0 \to L^1 \to Y$

Analysis



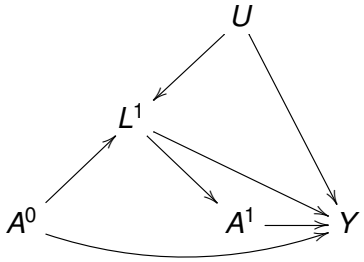
- We observed no direct effect of A^0, at $A^1 = 1$
  - $\Pr(Y_{11} = 1) = \Pr(Y_{01} = 1) = 0.625$
- This may explain the negative total effect of A^0
  - All that happens if AZT is received at $t = 0$ is that the chances of getting AZT at $t = 1$ decreases, which in turn increases the infection risk at $t = 2$

Outline

Generality of the scenario: data and aim

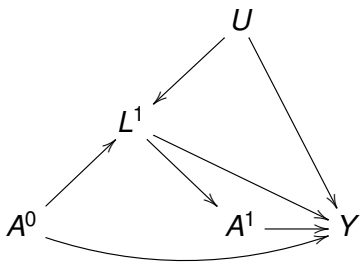
- Longitudinal studies typically generate repeated measures of
  - a time varying exposure: A^0, A^1, ..., A^T
  - a set of time varying covariates L^0, L^1, ..., L^T
  - an outcome Y, possibly time varying
- The aim is typically to estimate the causal effect of the exposure on the outcome
- Same data, same aim

Generality of the scenario: the past affects the present



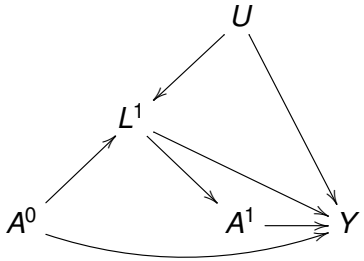
- We have allowed for each variable to be directly affected by each preceding variable
  - Likely in most longitudinal studies
  - In our specific example, A^1 was not directly affected by A^0

Generality of the scenario: U



- We have allowed for unmeasured common causes of L^1 and Y
  - Likely in most longitudinal studies

Generality of the scenario: sequential exchangeability



- In our example, sequential exchangeability followed by design (randomization)
- In well designed observational studies, this assumption may hold to a reasonable approximation
- It should be viewed as the 'gold standard' of a longitudinal observational study

The general G-formula (sequential standardization)

$$\Pr(Y_{a^0 \ldots a^T}) = \Big\{ \sum_{L^0 \ldots L^T} \Pr(Y = 1 | L^0, A^0 = a^0, \ldots, L^T, A^T = a^T) \prod_{t=0}^{T} \Pr(L^t | L^0, A^0 = a^0, \ldots, L^{t-1}, A^{t-1} = a^{t-1}) \Big\}$$

More complex situations

- In our example, we assumed
  - infections were only measured once, at the end of follow up
  - all subjects survived and no subject dropped out during follow up
- In real studies
  - outcomes are often measured repeatedly
  - the survival time (often censored) is often the main target of analysis
- The concepts and methods in this lecture apply to repeated outcomes and survival outcomes as well
  - However, analysis and interpretation get more complex
  - Beyond the scope of this course

Summary

- Causal inference in longitudinal studies is non-trivial
- Sequential adjustment gives separate effects of each exposure, for different subpopulations
- Sequential standardization gives joint (and direct) effects, for the whole population