Multiple exposures

Arvid Sjölander

Department of Medical Epidemiology and Biostatistics
Karolinska Institutet

A short course on concepts and methods in Causal Inference

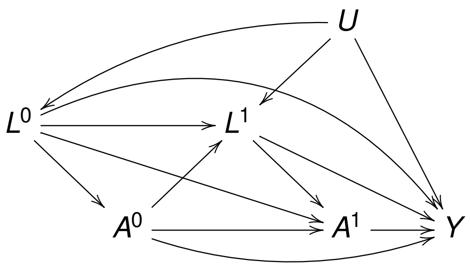
From single exposures to multiple exposures

- Until now, we have only considered single exposures
- Many epidemiological studies aim to examine multiple exposures
- Sometimes, we wish to estimate the causal effect of each exposure separately
- We may also wish to estimate
  - the joint effect of several exposures
  - the direct effect of one exposure, that is, the effect not mediated through another exposure
- The joint and direct effects are typically more challenging to estimate than effect of each exposure separately

Outline

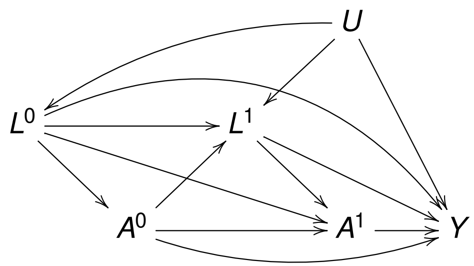
Outline

The DAG



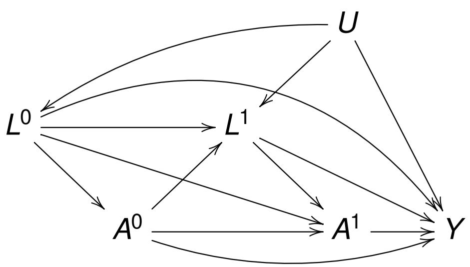
- For pedagogic reasons we will focus on two exposures
- We will restrict attention to the DAG above
- A^0 and A^1 are the two exposures of interest
  - possibly repeated measures of the same exposure
- L^0 and L^1 are two **measured** covariates
  - possibly repeated measures of the same covariate
- U are **unmeasured** covariates
- Y is the outcome

Important features



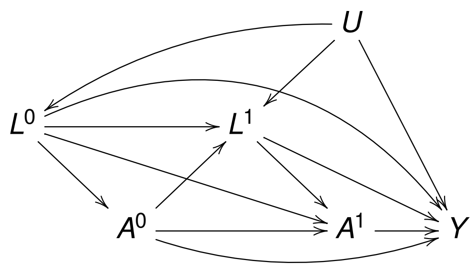
- The DAG has three important features
  - the ‘past’ affects the ‘present’
  - unmeasured common causes of (L^0, L^1) and Y
  - sequential exchangeability

The ‘past’ affects the ‘present’



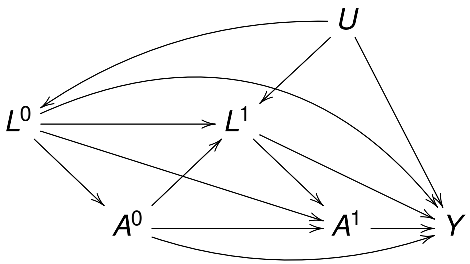
- The DAG allows for every variable to be affected by all preceding variables
  - A^0 is affected by L^0
  - L^1 is affected by L^0 and A^0
  - A^1 is affected by L^0, A^0 and L^1
  - Y is affected by L^0, A^0, L^1 and A^1
- This is a common feature in longitudinal studies with repeated measures

Unmeasured common causes of L and Y



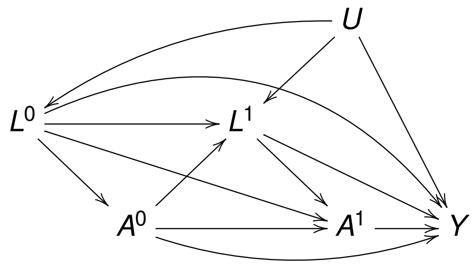
- The DAG allows for (L^0, L^1) and Y to have unmeasured common causes, U
- Such common causes are likely to be present in most observational studies

Sequential exchangeability



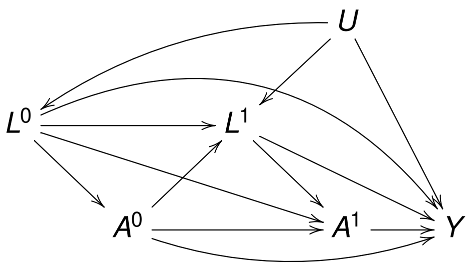
- The DAG assumes that adjustment for the 'observed past' blocks all non-causal paths between A^t and Y
  - observed past at $t = 0$: L^0
  - observed past at $t = 1$: L^0, A^0 and L^1
- When this holds, we say that we have **sequential exchangeability**

Sequential exchangeability is not guaranteed



- The assumption of sequential exchangeability is not guaranteed to hold in observational studies
- To make the assumtion plausible we should try to measure all covariates that have a direct influence on the exposure
  - e.g. if A^t is a medical treatment, then we should try to measure all indications for this treatment

Sequential adjustment



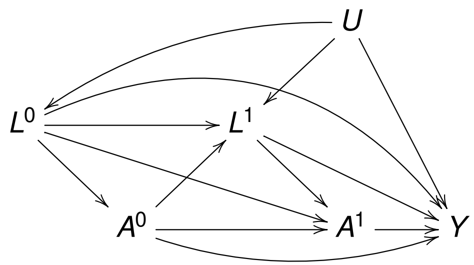
- Under sequential exchangeability we can estimate the causal effect of A^t by adjusting for the observed past:

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

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

- We say that we make **sequential adjustment**

Towards joint and direct effects



- Sequential adjustment gives the causal effect of A^0 and A^1 separately
- Often, we would like to estimate the joint effect of A^0 and A^1
- We could also be interested in the direct effect of A^0, not mediated through A^1

Outline

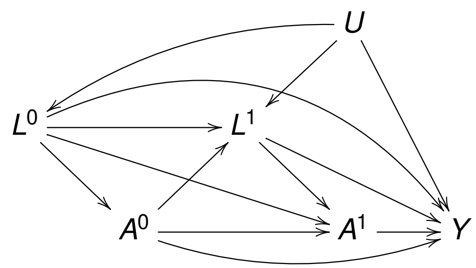
Potential outcomes under multiple exposures

- We define $Y_{a^0 a^1}$ as the potential outcome under joint exposures $A^0 = a^0$ and $A^1 = a^1$
  - e.g. Y_{11} is the potential outcome when exposed to $A^0 = 1$ and $A^1 = 1$
- $\Pr(Y_{a^0 a^1} = 1)$ is the proportion of subjects with $Y = 1$ if the whole population would receive $A^0 = a^0$ and $A^1 = a^1$

Joint effect

- The joint effect of A^0 and A^1 is a comparison between $\Pr(Y_{11} = 1)$ and $\Pr(Y_{00} = 1)$
  - e.g. the risk ratio $\Pr(Y_{11} = 1)/\Pr(Y_{00} = 1)$
- It measures the effect of simultaneously taking both exposures from 0 to 1

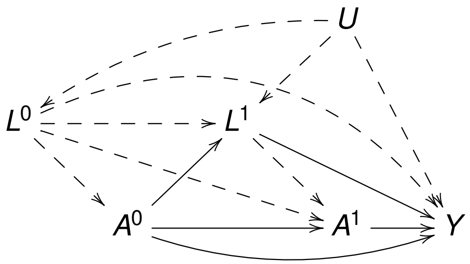
Joint effect in the DAG



- *What paths contribute to the joint effect of A^0 and A^1?*

Solution

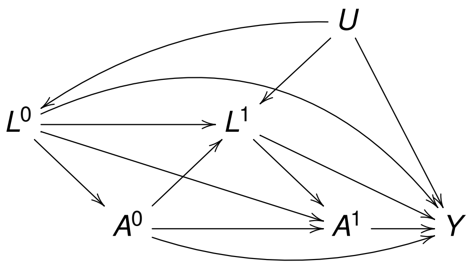
$$\Pr(Y_{11} = 1) \text{ vs } \Pr(Y_{00} = 1)$$



Direct effect

- The direct effect of A^0, at a fixed level $A^1 = a^1$, is a comparison between $\Pr(Y_{1a^1} = 1)$ and $\Pr(Y_{0a^1} = 1)$
  - e.g. the risk ratio $\Pr(Y_{1a^1} = 1)/\Pr(Y_{0a^1} = 1)$
- It measures the effect of taking A^0 from 0 to 1, while holding A^1 fixed at a^1
- By holding A^1 fixed, we ensure that the effect is not mediated through A^1

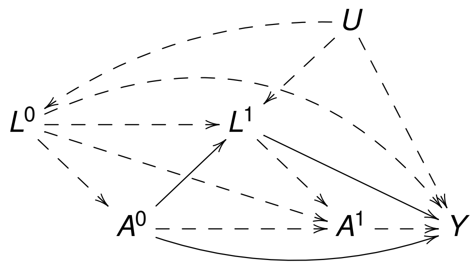
Direct effect in the DAG



- *What paths contribute to the direct effect of A^0, at a fixed level of A^1?*

Solution

$$\Pr(Y_{1a^1} = 1) \text{ vs } \Pr(Y_{0a^1} = 1)$$



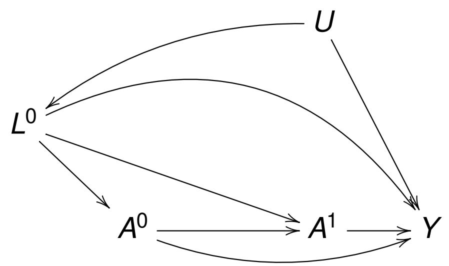
- The term ‘direct’ is relative to A^1, the direct effect of A^0 could still be mediated through L^1

Technical note

- In the literature, there are several definitions of direct effects
  - the controlled direct effect
  - the natural/pure direct effect
  - the principal stratum direct effect
- Our definition gives the controlled direct effect; the other definitions are beyond the scope of this course

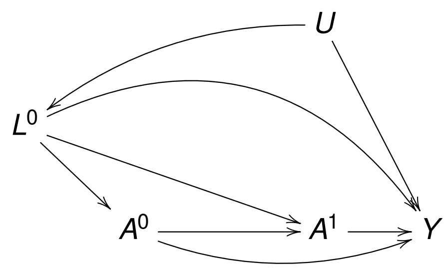
Outline

The DAG



- In the general case, joint and direct effects are quite complicated to estimate
- We will first consider the simple scenario where L^1 is absent, as in the DAG above
  - in terms of a longitudinal study; only baseline covariates

Standard adjustment



- In this simple scenario, standard adjustment for L^0 gives conditional joint and direct effects, given L^0:

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

- Heuristically, because L^0 blocks all non-causal paths between the 'joint' exposure (A^0, A^1) and Y

Example

L^0	A^0	A^1	$Y=1$	$Y=0$
0	0	0	20	20
0	0	1	60	20
0	1	0	10	10
0	1	1	30	10
1	0	0	30	5
1	0	1	4	4
1	1	0	5	3
1	1	1	8	1

- *Compute the conditional joint effect of A^0 and A^1, given $L^0 = 0$, as a risk ratio*
- *Compute the conditional direct effect of A^0, at $A^1 = 1$, given $L^0 = 0$, as a risk ratio*

Solution

L^0	A^0	A^1	$Y=1$	$Y=0$
0	0	0	20	20
0	0	1	60	20
0	1	0	10	10
0	1	1	30	10

$$\begin{aligned}\Pr(Y_{11}=1|L^0=0) &= \Pr(Y=1|L^0=0,A^0=1,A^1=1)\\ &= 30/(30+10)=3/4\end{aligned}$$

$$\begin{aligned}\Pr(Y_{00}=1|L^0=0) &= \Pr(Y=1|L^0=0,A^0=0,A^1=0)\\ &= 20/(20+20)=1/2\end{aligned}$$

- Conditional joint effect of A^0 and A^1, given $L^0 = 0$, as a risk ratio:

$$\Pr(Y_{11}=1|L^0=0)/\Pr(Y_{00}=1|L^0=0)=1.5$$

Solution, cont'd

L^0	A^0	A^1	$Y=1$	$Y=0$
0	0	0	20	20
0	0	1	60	20
0	1	0	10	10
0	1	1	30	10

$$\begin{aligned}\Pr(Y_{11}=1|L^0=0) &= \Pr(Y=1|L^0=0,A^0=1,A^1=1)\\ &= 30/(30+10)=3/4\end{aligned}$$

$$\begin{aligned}\Pr(Y_{01}=1|L^0=0) &= \Pr(Y=1|L^0=0,A^0=0,A^1=1)\\ &= 60/(60+20)=3/4\end{aligned}$$

- Conditional direct effect of A^0, at $A^1 = 1$, given $L^0 = 0$, as a risk ratio:

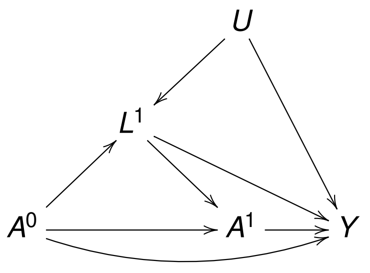
$$\Pr(Y_{11}=1|L^0=0)/\Pr(Y_{01}=1|L^0=0)=1$$

Outline

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

The DAG



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

Sequential adjustment: the causal effect of A^0

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

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

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

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

Sequential adjustment: the causal effect of A^1

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

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

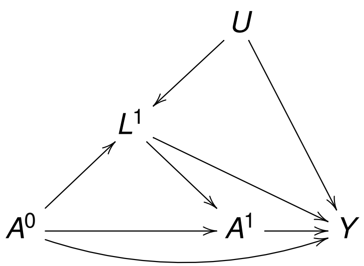
$$\begin{aligned}&\Pr(Y_{a^1=0}=1|A^0=1,L^1=1)\\&=\Pr(Y=1|A^0=1,L^1=1,A^1=0)\\&=1/4=0.25\end{aligned}$$

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

Conclusion

- We have observed a negative effect of A^0 and a positive effect of A^1
- Now, suppose we want to figure out the optimal combination of (A^0, A^1)
- *Can we not just conclude that* $(A^0=0, A^1=1)$ *is optimal?*

Solution

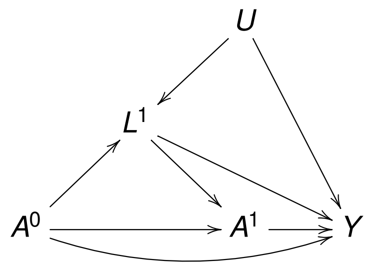


- No, since the effect of A^0 is partly mediated through A^1
  - AZT at $t=0 \Rightarrow$ increased CD4 count at $t=1$
  - $\Rightarrow$ decreased chance of getting AZT at $t=1$
  - $\Rightarrow$ increased risk for infection at $t=2$
- But A^0 could have a positive direct effect on Y, not mediated through A^1
  - then the optimal combination could well be $(A^0=1, A^1=1)$

Separate vs joint effects

- Sequential adjustment gives the effect of A^0 and A^1 separately:
  - the marginal causal effect of A^0
  - the conditional causal effect of A^1, given (A^0, L^1)
- But to figure out the optimal combination of (A^0, A^1) we really need to consider A^0 and A^1 jointly

Adjusting or not adjusting



- To estimate the effect of A^1 we need to adjust for L^1
  - if we don't, then the non-causal paths $A^1 \leftarrow L^1 \rightarrow Y$ and $A^1 \leftarrow L^1 \leftarrow U \rightarrow Y$ are open
- To estimate the effect of A^0 we **should not** adjust for L^1
  - if we do, then the non-causal path $A^0 \rightarrow L^1 \leftarrow U \rightarrow Y$ is open
- **Standard adjustment for L^1 does not give joint and direct effects**

Sequential standardization

- We can use **sequential standardization** to compute $\Pr(Y_{a^0 a^1} = 1)$ as

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

- Also known as the **G-formula**
- Sequential standardization gives joint and direct effects, marginal over L^1

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

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

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

Solution

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

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

Solution

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

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

Solution

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

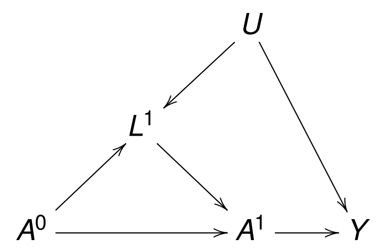
$$\begin{aligned}\Pr(Y_{00}=1) &= \overbrace{?}^{\Pr(Y=1|A^0=0,L^1=0,A^1=0)} \times \overbrace{1}^{\Pr(L^1=0|A^0=0)} \\ &+ \underbrace{?}_{\Pr(Y=1|A^0=0,L^1=1,A^1=0)} \times \underbrace{0}_{\Pr(L^1=1|A^0=0)} \quad =?\end{aligned}$$

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
  - we cannot estimate the joint effect of (A^0, A^1)

Direct effects

- A comparison between $\Pr(Y_{11}=1)$ and $\Pr(Y_{01}=1)$ is the direct effect of A^0, at $A^1=1$
- $\Pr(Y_{11}=1)=\Pr(Y_{01}=1)=0.625$; no direct effect of A^0, at $A^1=1$:



- 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 $t=2$

Sequential standardization for arbitrary many time points

- Arbitrary T:

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

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

- $T = 0$ (just one time point):

$$\Pr(Y_a = 1) = \sum_L \Pr(Y = 1 | A = a, L) \Pr(L)$$

More complex outcomes

- 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
- Sequential standardization can be used for repeated outcomes and survival outcomes as well
  - analysis and interpretation get more complex
  - beyond the scope of this course

Summary

- **Sequential exchangeability** means that the observed past blocks all non-causal paths between A^t and Y
- **Sequential adjustment** gives the conditional causal effect of A^t, given the observed past
- If there are only 'baseline covariates', L^0, then standard adjustment for L^0 gives conditional joint and direct effects, given L^0
- If there are also covariates L^1 that are affected by previous exposures, and affect later exposures, then standard adjustment does not give joint and direct effects
- **Sequential standardization** gives marginal joint and direct effects