

L10: Simple time-varying exposures

Jess Edwards,

Epidemiology, UNC Chapel Hill

EPID 722

Spring 2021

Where are we?

Reading: Lu H, Cole SR, Westreich D, Hudgens MG, Adimora AA, Althoff KN, Silverberg MJ, Buchacz K, Li J, Edwards JK, Rebeiro PF. Clinical Effectiveness of Integrase Strand Transfer Inhibitor–Based Antiretroviral Regimens Among Adults With Human Immunodeficiency Virus: A Collaboration of Cohort Studies in the United States and Canada. *Clinical Infectious Diseases*. 2020 Aug 11.

Today: Introduction to (simple) time varying exposures

Thursday: Exercise and discussion

Why do exposures vary over time?

Interventions are often static:

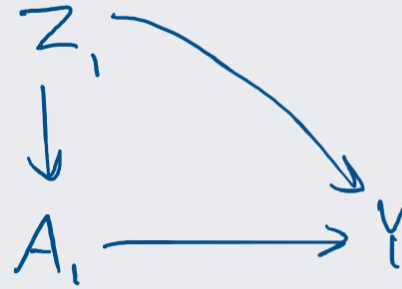
- Ban chemical X
- Treat condition Y using drug Z for the duration of illness

But when estimating the effects of these (often hypothetical) interventions using observational data, we often see that exposure varies over time.

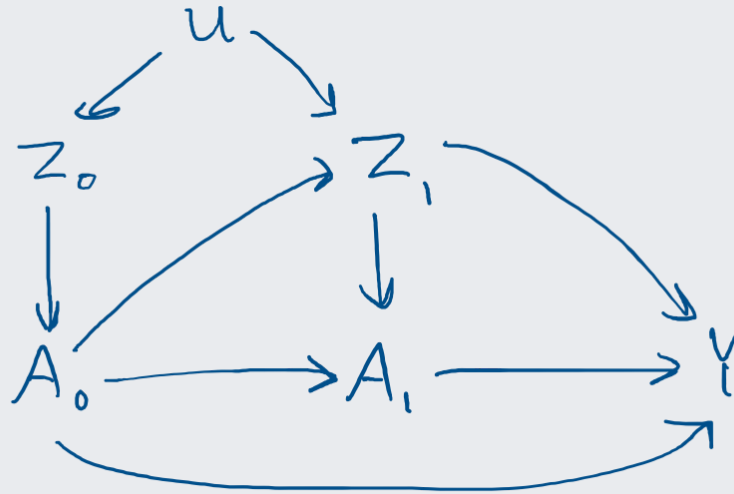
For example:

- Exposure to chemical X changes when people move or get a new job
- People discontinue drug Z due to side effects

DAGs for time-varying exposures



DAGs for time-varying exposures



A simple scenario

This week, we will begin to think about time varying exposures in a simple setting:

You are studying a possible treatment for disease X, which currently has no available treatments.

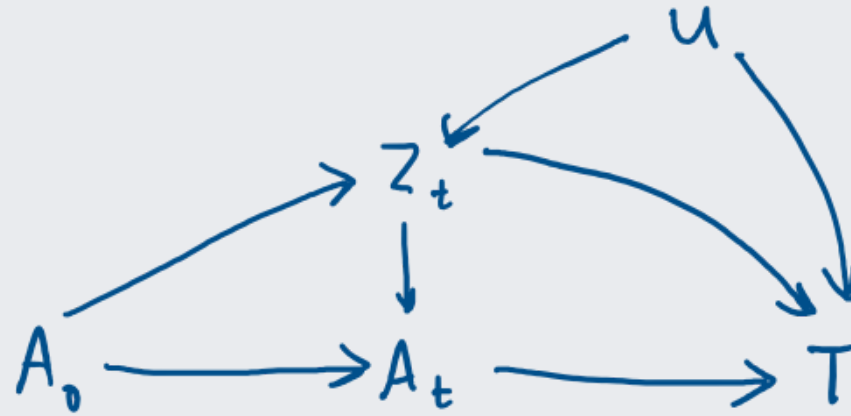
You perform a trial in which you recruit people with disease X and randomize them to no treatment or treatment with Drug A for the duration of illness.

Some people who you randomize to Drug A discontinue treatment during the year.

You want to compare the probability of recovery by 1 year.

Assume no loss to follow-up and no deaths or other competing events occurred during the course of the study.

Example DAG



Choosing a parameter

What do we want to estimate?

Choices

The "**intent-to-treat**" parameter: What would the probability of recovery have been if I had assigned everyone to take drug A vs if I had assigned everyone to no treatment?

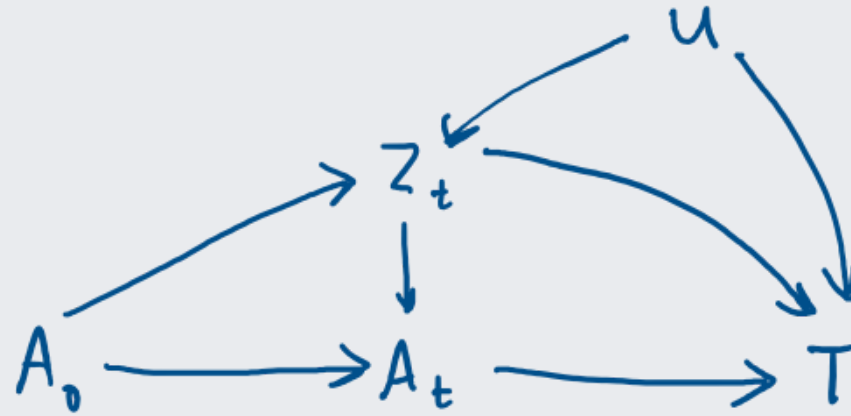
- We estimate risk in the treatment arm without regard to treatment discontinuation

The "**per protocol**" parameter: What would the probability of recovery have been if I had assigned everyone to take drug A and they had taken it throughout follow-up vs if I had assigned everyone to no treatment?

- We perform a second "intervention" to prevent treatment discontinuation

[*] "Intent-to-treat" and "per protocol" are just words. Do not expect everyone who uses these words to refer to the same parameters. If the parameter of interest is unclear, ask!

Example DAG, revisited



Statistical analyses

We will discuss how to conduct analyses to estimate the "intent-to-treat" and "per protocol" parameters.

1. How are the data organized?
2. What estimators are used?

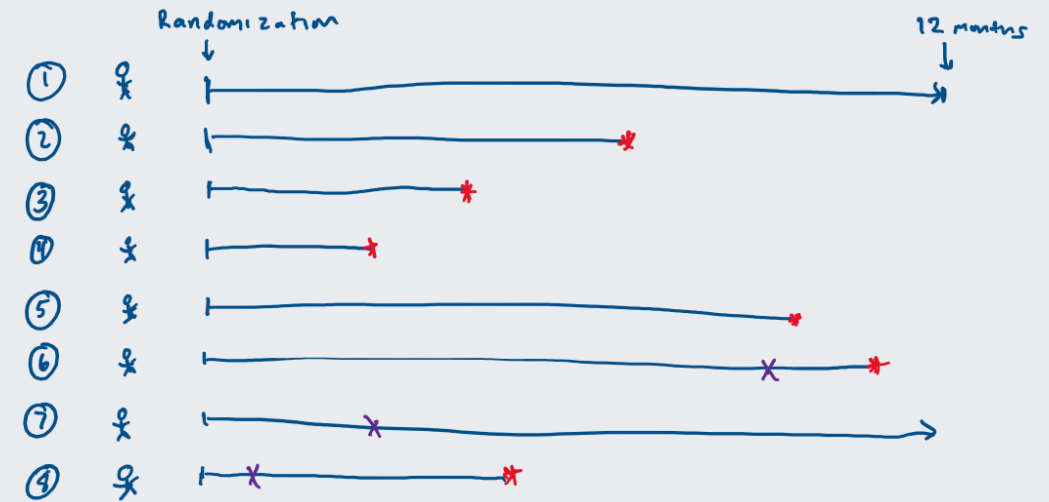
Example data records

ID	<i>A</i>	<i>T_D</i>	<i>Z₀</i>	<i>T_Y</i>	<i>Y</i>
1	0	.	0	12.0	0
2	0	.	1	6.5	1
3	0	.	0	4.0	1
4	0	.	1	3.0	1
5	1	.	0	9.5	1
6	1	9.0	1	11.0	1
7	1	3.0	0	12.0	0
8	1	1.0	1	4.5	1

"Intent-to-treat" analysis

Data set-up

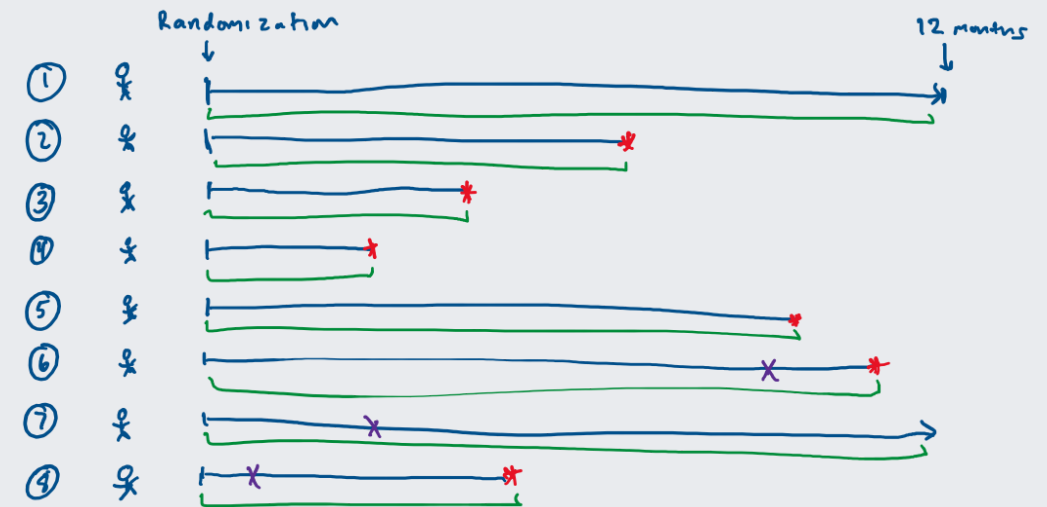
ID	A	T_D	Z_0	T_Y	Y
1	0	.	0	12.0	0
2	0	.	1	6.5	1
3	0	.	0	4.0	1
4	0	.	1	3.0	1
5	1	.	0	9.5	1
6	1	9.0	1	11.0	1
7	1	3.0	0	12.0	0
8	1	1.0	1	4.5	1



"Intent-to-treat" analysis

Data set-up

ID	A	T_D	Z_0	T_Y	Y
1	0	.	0	12.0	0
2	0	.	1	6.5	1
3	0	.	0	4.0	1
4	0	.	1	3.0	1
5	1	.	0	9.5	1
6	1	9.0	1	11.0	1
7	1	3.0	0	12.0	0
8	1	1.0	1	4.5	1



"Intent-to-treat" analysis

ID	A	T_D	Z_0	T_Y	Y
1	0	.	0	12.0	0
2	0	.	1	6.5	1
3	0	.	0	4.0	1
4	0	.	1	3.0	1
5	1	.	0	9.5	1
6	1	9.0	1	11.0	1
7	1	3.0	0	12.0	0
8	1	1.0	1	4.5	1

Arm 1: No treatment. Kaplan-Meier estimator of the risk of recovery, limited to the no treatment arm

Arm 2: Assign Drug A. Kaplan-Meier estimator of the risk of recovery, limited to those assigned drug A

-- Or, estimate HR using

$$h^{itt}(t) = h_0(t) \exp(\beta A_0)$$

"Intent-to-treat" analysis

Why can we use crude comparisons of unweighted risk functions in this example?

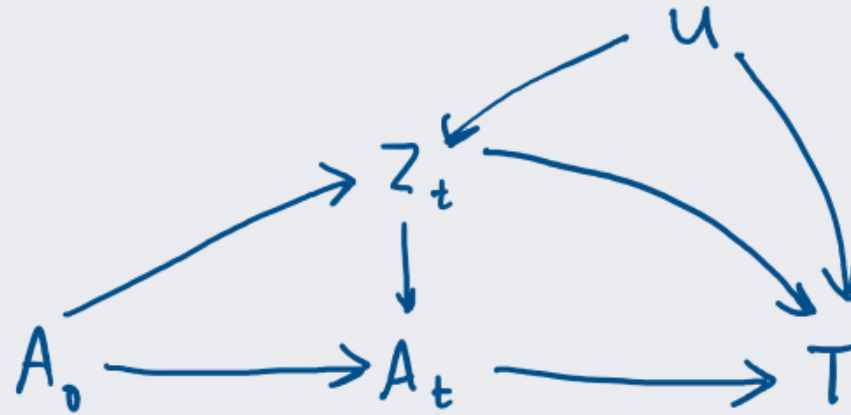
"Per protocol" analysis

Challenge: We want to know what would have happened had everyone in the treatment arm remained on treatment throughout follow-up, **but some participants discontinued treatment.**

Solution: Censor people when they stop following the "protocol" (i.e., at treatment discontinuation).

Account for this *informative censoring* using IPW.

Recall, DAG



Data setup

Step 1: Reorganize data to capture time-updated covariates, if necessary

ID	A	T_D	Z_0	T_Y	Y
4	0	.	1	3.0	1

ID	in	out	A	D_{in}	D_{out}	Z_{in}	Y
4	0	1	0	0	0	1	0
4	1	2	0	0	0	1	0
4	2	3	0	0	0	0	1

ID	A	T_D	Z_0	T_Y	Y
7	1	3.0	0	12.0	0

ID	in	out	A	D_{in}	D_{out}	Z_{in}	Y
7	0	1	1	0	0	1	0
7	1	2	1	0	0	0	0
7	2	3	1	0	1	0	0
7	3	4	1	1	1	1	0
7	4	5	1	1	1	0	0

Data setup

Step 2: Censor at treatment discontinuation

ID	A	T_D	Z_0	T_Y	Y
4	0	.	1	3.0	1

ID	in	out	A	D_{in}	D_{out}	Z_{in}	Y
4	0	1	0	0	0	1	0
4	1	2	0	0	0	1	0
4	2	3	0	0	0	0	1

ID	A	T_D	Z_0	T_Y	Y
7	1	3.0	0	12.0	0

ID	in	out	A	D_{in}	D_{out}	Z_{in}	Y
7	0	1	1	0	0	1	0
7	1	2	1	0	0	0	0
7	2	3	1	0	1	0	0
7	3	4	1	1	1	1	0
7	4	5	1	1	1	0	0

"Per protocol" analysis with IPW

Step 3: Use IPW to account for informative censoring induced in the previous step.

Arm 1: No treatment. Limit dataset to those randomized to no treatment. No discontinuation, so everyone has a weight of 1 for entire period.

Arm 2: Continuous treatment. Limit dataset to those randomized to treatment, censor at treatment discontinuation, and apply IPCW.

Let T_D represent the time of treatment Discontinuation.

$$\pi_D = \frac{P(T_D > t | A_0)}{P(T_D > t | Z_t, A_0)}$$

Let $D(t)$ be an indicator of treatment discontinuation at time t

$$\pi_D = \prod_{out=1}^{\lceil t \rceil} \frac{1 - P(D_{out} = 1 | D_{in} = 0, A_0)}{1 - P(D_{out} = 1 | \bar{D}_{in} = 0, Z_{in}, A_0)}$$

We can estimate $P(D_{out} = 1 | D_{in} = 0, Z_{in}, A_0)$ using pooled logistic regression.

"Per protocol" analysis with IPW

Step 3: Use IPW to account for informative censoring induced in the previous step.

Arm 1: No treatment. Limit dataset to those randomized to no treatment. No discontinuation, so everyone has a weight of 1 for entire period.

Arm 2: Continuous treatment. Limit dataset to those randomized to treatment, censor at treatment discontinuation, and apply IPCW.

ID	in	out	A	D	Z_{in}	Y	π_D
4	0	1	0	0	1	0	1
4	1	2	0	0	1	0	1
4	2	3	0	0	0	1	1

ID	in	out	A	D_{in}	D_{out}	Z_{in}	Y	π_D
7	0	1	1	0	0	1	0	1
7	1	2	1	0	0	0	0	1.03
7	2	3	1	0	1	0	0	1.10
7	3	4	1	1	1	1	0	
7	4	5	1	1	1	0	0	

"Per protocol" analysis with IPW

Step 3: Use IPW to account for informative censoring induced in the previous step.

Arm 1: No treatment. Limit dataset to those randomized to no treatment. No discontinuation, so everyone has a weight of 1 for entire period.

Arm 2: Continuous treatment. Limit dataset to those randomized to treatment, censor at treatment discontinuation, and apply IPCW.

"Per protocol" analysis with IPW

Step 4: Apply IPW

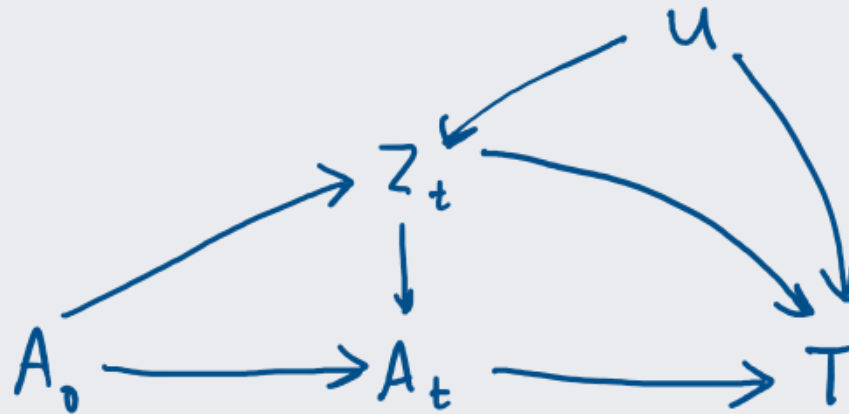
Arm 1: No treatment. Kaplan-Meier estimator of the risk of recovery, limited to the no treatment arm.

Arm 2: Continuous treatment. Kaplan-Meier estimator of the risk of recovery, limited to the treatment arm in the time prior to discontinuation, and weighted by π_D

Or, estimate weighted HR using

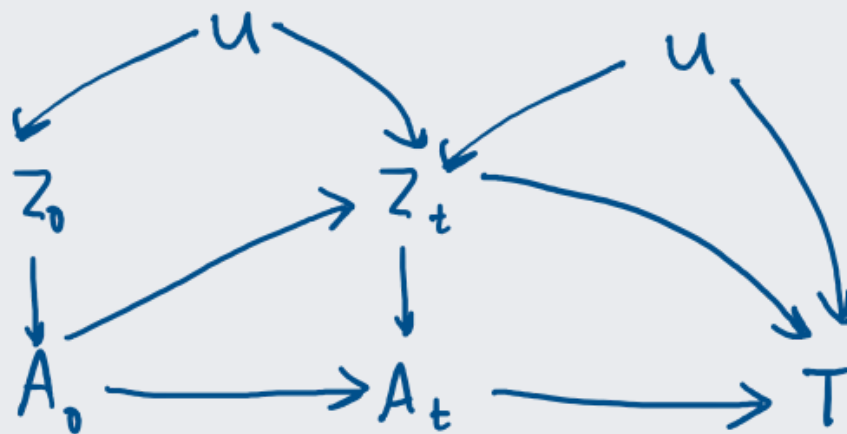
$$h^a(t) = h_0(t) \exp(\beta a)$$

Why not use regression adjustment to account for confounding?



Extend to settings with baseline confounding

What if A_0 were not randomized?



We would need to account for confounding of both A_0 and A_t .

Solution: Set $\pi(t) = \pi_A \times \pi_D(t)$

Extend to settings with baseline confounding

Step 1: Fit time-fixed inverse probability of treatment weight.

$$\pi_A = \frac{P(A_0 = a)}{P(A_0 = a | Z_0)}$$

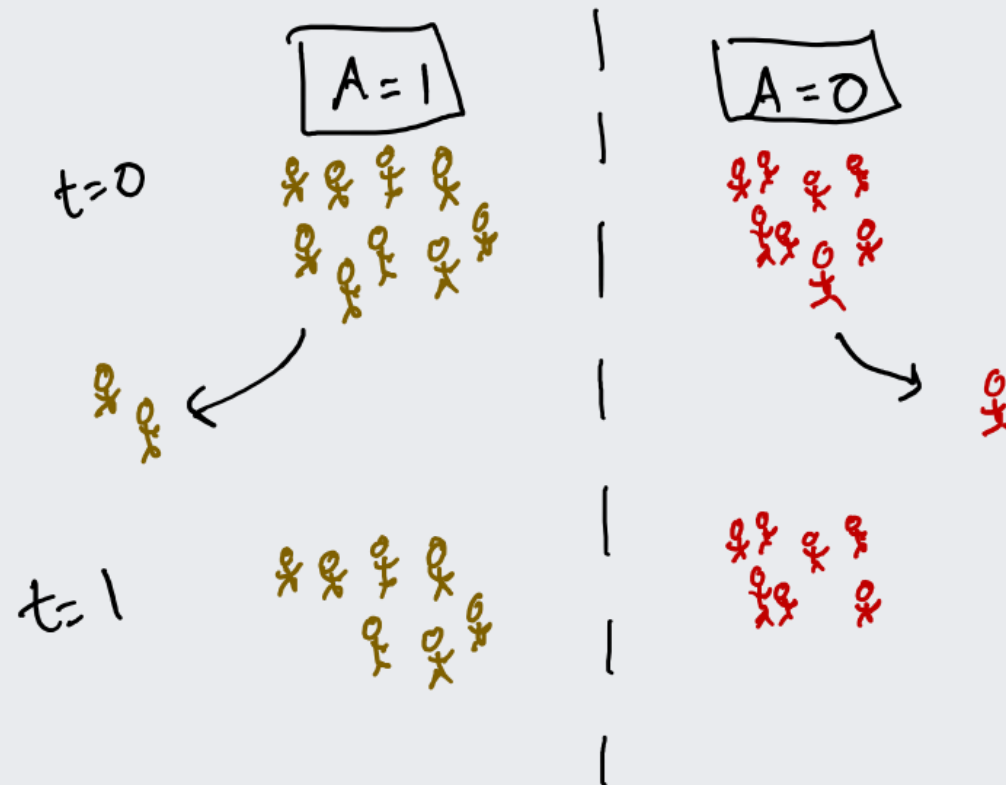
Step 2: Fit time-varying inverse probability of treatment discontinuation

$$\pi_D = \frac{P(T_D > t | A_0)}{P(T_D > t | A_0, Z_0, \bar{Z}_t)}$$

Step 3: Apply $\pi(t)$ to KM estimators to obtain risk estimates or Cox model to obtain HR

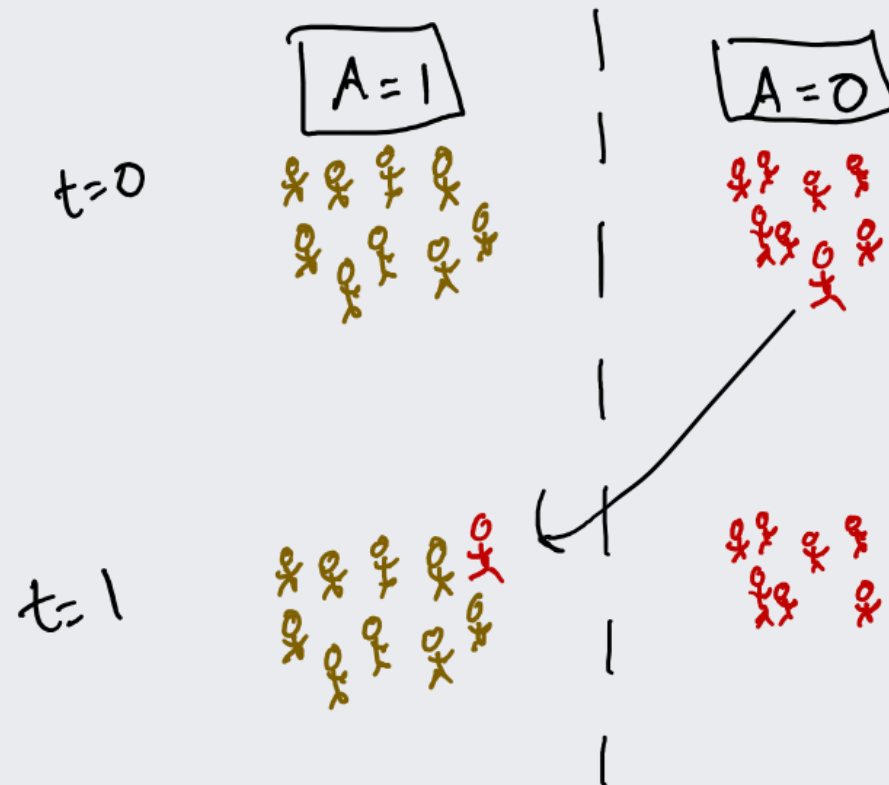
Other types of time-varying exposures

What we discussed today:



Other types of time-varying exposures

An alternative: allow people to "late enter" treatment plans



Other types of time-varying exposures

For an example, see

Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000 Sep 1:561-70.

Other types of time-varying exposures

Another alternative: allow people to contribute to multiple plans at the same time.

Example: Estimate the effect of starting treatment within 3 months of diagnosis compared to not starting within the first 3 months

Other types of time-varying exposures

Another alternative: allow people to contribute to multiple plans at the same time.



Other types of time-varying exposures

Another alternative: allow people to contribute to multiple plans at the same time.

Example::

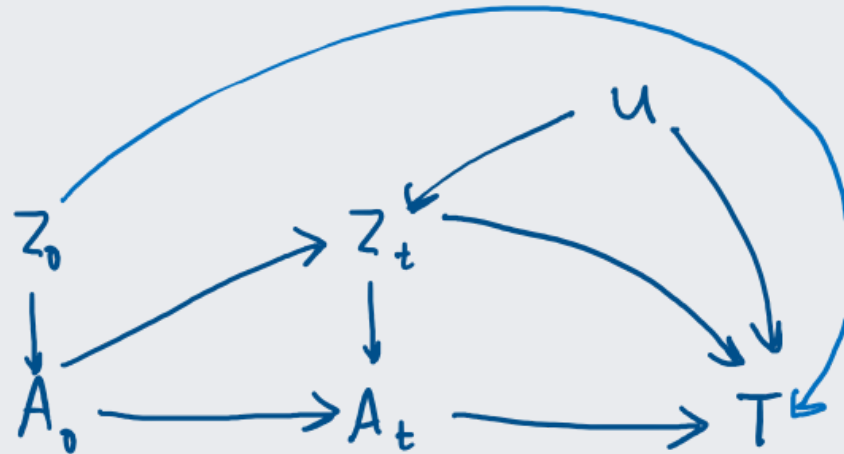
Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernán MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. The international journal of biostatistics. 2010 Apr 13;6(2).

Example: Infant flu shots

Infants require 2 doses of the flu shot, 1 month apart.

Consider a hypothetical example in which 2 "brands" of flu shot are available. You want to compare the risk of flu over 1 year if all infants received shot A vs if all infants received shot B.

When the first dose is given, parents receive a notice to return for the second shot.



Example: outline

We will estimate the

- "Intent to treat" (ITT) parameter: What if all infants were given a first dose of flu shot A vs what if all infants received a first dose of flu shot B?
- "Per protocol" (PP) parameter: What if all infants were given both doses of flu shot A vs what if all infants received both doses of flu shot B?

We will estimate risk functions under each plan and compute an HR each analysis above (ITT and PP).

Example: data

Shot A: 4971 participants elected shot A.

- 41% had covariate $Z_0 = 1$ at baseline
- of infants receiving shot A, 25% had noticeable side effects of the first dose
- 65% followed-up with a second dose

Shot B: 5029 participants elected shot B.

- 58% had covariate $Z_0 = 1$ at baseline
- of infants receiving shot B, 57% had noticeable side effects of the first dose
- 55% followed-up with a second dose.

ITT (crude)

- Stratified by A_0
- Estimated risk in each group

RR: 1.85 (95% CI: 1.58, 2.15)

HR: 1.89 (95% CI: 1.61, 2.21)

ITT (weighted)

- Stratified by A_0
- Estimated inverse probability of treatment weights $\frac{P(A_0=a)}{P(A_0=a|Z_0)}$
- Estimated risk in each group using weighted KM
- Estimated HR using weighted Cox model

RR: 1.54 (95% CI: 1.32, 1.8)

HR: 1.56 (95% CI: 1.33, 1.83)

Per protocol analysis (crude)

- Reshaped dataset to have 1 record per person-period; augmented with time varying covariates
- Stratified by A_0
- Censored after 1 month if infant did not receive second dose (removed 38678 rows)
- Estimated risk in each group (after censoring at discontinuation) using KM
- Estimated HR using Cox model

RR: 2.3 (95% CI: 1.91, 2.76)

HR: 2.29 (95% CI: 1.9, 2.77)

Per protocol analysis (weighted)

- Stratified by A_0
- Censored after 1 month if infant did not receive second dose (removed 38678 rows)
- Fit logistic regression model for probability of discontinuation given A_0 , Z_0 , and Z_1 (denominator model)
- Fit logistic regression model for probability of discontinuation given A_0 (numerator model)
- Estimated inverse probability of discontinuation weights $\pi_D = \prod_{k=1}^{\lceil t \rceil} \frac{1 - P(D=1|A_0)}{1 - P(D=1|A_0, Z_0, Z_1)}$
 - In this case, $\pi_D = 1$ in month 1, otherwise it is the inverse probability of receiving the second dose

PP (weighted)

- Final weight is $\pi = \pi_A \times \pi_D$
- Estimated risk in each group (after censoring at discontinuation) using KM weighted by π
- Estimated HR using Cox model weighted by π .

RR: 1.73 (95% CI: 1.43, 2.09)

HR: 1.71 (95% CI: 1.41, 2.07)

Discussion points

"Protocols" can be as simple or as complex as you wish

- In the reading, some types of treatment switch were not counted as protocol deviations. Why?

Just because you *can* run a per protocol analysis doesn't mean you want to.

- When might the ITT be useful?

The reading mentioned this idea of emulating a randomized trial using observational data. Why might this be useful?