Causal Inference in HIV Research

Sherri Rose
Johns Hopkins University
&
Elise Riley
University of California, San Francisco

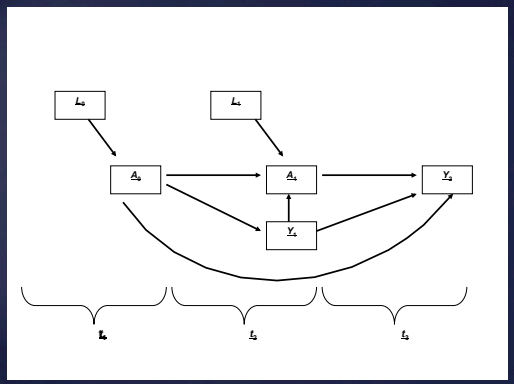
We will start by motivating the methods presentation with challenging problems found in the study of HIV.

Study Objective

To determine the impact of multiple risks
on the overall health status of
HIV+ homeless and unstably housed adults

Potential Challenges

- HIV is increasingly characterized as a chronic condition that can be managed through adherence to a healthy lifestyle, complex drug regimens and treatment monitoring;
  - however, social and structural factors can be significant determinants of an individual's ability to meet these requirements and achieve better overall health status.
- The effects of exposures that change over time and influence one another, such as drug use and housing status, now have the opportunity to influence a longer disease course.



Outcome of Interest

- Overall Physical Health Status (SF-36)
- Overall Mental Health Status (SF-36)

Exposures of Interest

- Age, race, education
- Employment, income
- Subsistence needs (housing, food, clothing, hygiene needs)
- Incarceration
- Drug use, alcohol use
- Victimization, social support
- Insurance status
- Adherence to antiretroviral therapy
- CD4 cell count, viral load

Physical Health Status (N=288)

Main Effect	Adjusted Population Effect	Adjusted 95% Confidence Interval	Adjusted p-value
Unmet subsistence needs	-3.83	(-5.27,-1.6)	<.0001
Caucasian race/ethnicity	-3.71	(-6.03,-1.29)	.0012
No source of instrumental support	-1.56	(-2.88,-0.21)	.0220
Viral load	-0.000018	(-0.000038 , - 0.000003)	.0410

Mental Health Status (N=288)

Main Effect	Adjusted Population Effect	Adjusted 95% Confidence Interval	Adjusted p-value
Unmet subsistence needs	-3.51	(-5.08, -1.29)	.000036
Close friend/ Confidant	3.19	(1.64, 4.72)	.000045
Any drug use	-3,67	(-5.53, -1.8)	.0002
No Sources of Instrumental Support	-2.2	(-3.62, -0.89)	.0012
>90% ART Adherence	1.66	(0.07,3.27)	.0043

Cautionary Note

- Analyses do not necessarily indicate the highest health priorities for specific individuals; instead they indicate exposures with the largest population-level effects on the health of unstably housed HIV-positive adults and that the biggest population-wide impact on health would be made by focusing on these issues.

Causal Modeling

Let's Take Step Back…

What is the Goal of the Study?

When we ask scientific questions, we frequently collect data in an attempt to answer these questions.

We are often, *but not always*, interested in causal effects. We prefer not to merely conclude that there is an association or correlation between two variables. Instead, we want to know that *A* causes *Y*.

Causal Modeling

Key Take Home Message:

Causal assumptions allow us to interpret the parameter of interest as a causal effect.

- These additional assumptions are untestable; we cannot use the data to verify their accuracy
- The causal modeling assumptions are separate from the estimation procedure.
- If we choose not to make causal assumptions, perhaps because we know they do not hold, the parameter has a statistical interpretation, just not a causal one.

Causal Modeling

Multiple Frameworks

- The Neyman-Rubin Causal Model assumes:
  - No unmeasured confounders
  - Consistency
  - No interference (the counterfactual outcome of one subject should not be affected by the treatment assignment of other subjects)

Causal Modeling

Multiple Frameworks

- Structural Causal Models of Pearl:
  - Describes each endogenous variable X_j as a deterministic function of other endogenous variables and an exogenous error. (The errors are never observed.)
  - For each X_j, the deterministic function depends on the other endogenous variables only through the parents of X_j.
  - The exogenous variables have a particular joint distribution.

Causal Modeling

We could specify the following SCM:

$$\begin{aligned} W &= f_W(U_W), \\ A &= f_A(W, U_A), \\ Y &= f_Y(W, A, U_Y), \end{aligned}$$

Recall that we assume for the full data:

1. for each X_j, $X_j = f_j(Pa(X_j), U_{X_j})$ depends on the other endogenous variables only through the parents $Pa(X_j)$,
2. the exogenous variables have a particular joint distribution P_U.

Causal Modeling

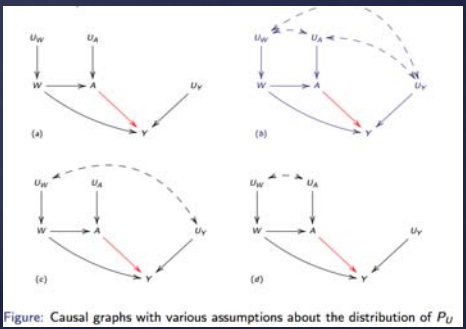


Figure: Causal graphs with various assumptions about the distribution of P_U

What is a Causal Effect?

- How would outcomes change in the population under different exposures/treatments?
- In the SCM framework, we want to know what happens when we intervene on the system to, for example, set $A=a$.

Recall Elise's Cautionary Note….

Cautionary Note

- Analyses do not necessarily indicate the highest health priorities for specific individuals; instead they indicate exposures with the largest population-level effects on the health of unstably housed HIV-positive adults and that the biggest population-wide impact on health would be made by focusing on these issues.

Why Not Randomized Studies?

- Ethical issues
- Time
- Cost
- Randomization may not occur perfectly

Data

- Our study is an experiment where we draw a random variable from our population n times.
- The data we observe are realizations of these n random variables, and the random variables have an underlying probability distribution.

Data

- Formally, the data consists of n i.i.d. copies of random variable $O \sim P$, where P is the true underlying probability distribution for O.
- We'll start with the same simple case used for the SCM, where W is a vector of baseline variables, A is an intervention, and Y is an outcome.
- $O=(W,A,Y) \sim P$.

Statistical Model

- A statistical model represents the set of possible probability distributions of the data.
- You are likely familiar with parametric statistical models, where one assumes that the probability distribution underlying the data is known (up to a certain number of parameters).
- You can also assume nonparametric and semiparametric models.

Parameters

The target parameter of interest will depend on your scientific question.

One simple parameter is the risk difference.

$$\begin{aligned}\psi_{RD} = \Psi(P) &= E[E(Y \mid A=1, W) - E(Y \mid A=0, W)] \\ &= E(Y_1) - E(Y_0) \\ &= P(Y_1 = 1) - P(Y_0 = 1)\end{aligned}$$

Parameters

In HIV research, we frequently address complicated research questions that require more complex parameters.

Marginal Stuctural Models

Marginal structural models (MSMs) are a useful tool to describe additional parameters.

MSMs are not estimators.

MSMs are simply a way to define parameters.

Marginal Stuctural Models

Example (Effect Modification):

One may be interested in the treatment-specific mean of an outcome conditional on a particular baseline covariate. Now we have a treatment effect that is a function of a baseline covariate.

We could use an MSM to define such a parameter:

$E(Y_a \mid V)=\beta_0+\beta_1 a+\beta_2 V+\beta_3 aV,$

with effect modifier V and a continuous Y.

Reference: Rosenblum (2011). Marginal structural models. In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Marginal Stuctural Models

Consider effect modification examples from Rosenblum (2011):

- What is the effect of an antidepressant medication on Hamilton Depression Rating Scale (HAM-D) score for those who enter a study with severe depression vs. moderate depression?
- What is the effect of a cancer therapy for those who test positive for over-expression of a particular gene and for those who test negative?

Reference: Rosenblum (2011). Marginal structural models. In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Marginal Stuctural Models

Example (High-Dimensional Treatment):

What if we are interested in the effect of a continuous treatment?

Reference: Rosenblum (2011). Marginal structural models. In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Parameters

Recall our risk difference parameter…

$$\begin{aligned}\psi_{RD} = \Psi(P) &= E[E(Y \mid A=1, W) - E(Y \mid A=0, W)] \\ &= E(Y_1) - E(Y_0) \\ &= P(Y_1 = 1) - P(Y_0 = 1)\end{aligned}$$

Marginal Stuctural Models

Example (High-Dimensional Treatment):

What if we are interested in the effect of a continuous treatment?

We could use an MSM to define such a parameter:

$E(Y_a)=\beta_0+\beta_1 a$,

with continuous outcome Y.

Reference: Rosenblum (2011). Marginal structural models. In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Marginal Stuctural Models

Example (Dynamic Treatment Regimes):

What if we are interested in the effect of a particular "rule" for assigning the intervention in response to baseline or intermediate variables?

We could use an MSM to define such a parameter.

Reference: Rosenblum (2011). Marginal structural models. In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Marginal Stuctural Models

Examples (Dynamic Treatment Regimes):

- Understanding adherence to combined antiretroviral therapy in HIV treatment

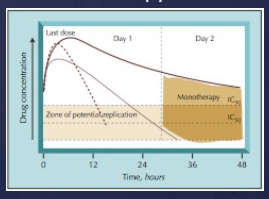
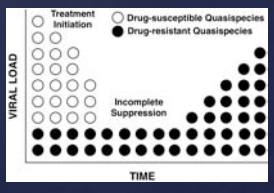


Image credit: Bangsberg D et al. (2007). Adherence-resistance relationships to combination HIV antiretroviral therapy. *Curr HIV/AIDS Rep*,4(2):65-72.

Marginal Stuctural Models

Examples (Dynamic Treatment Regimes):

- When to initiate combined antiretroviral treatment in therapy-naïve HIV-infected person



MSMs: Dynamic Regimes

- Consider data structure $O=(L_0,A_0,L_1,A_1,L_2=Y)$
- Let $D=(d_1,\ldots,d_K)$ be the set of dynamic regimes we consider. These dynamic regimes define a set of rules for guiding intervention $A(t)$ at each time point based on previous covariates and prior intervention.
- Thus, each rule d takes as input previous covariates and interventions to assign $a(t)$.

MSMs: Dynamic Regimes

- One could estimate the regime-specific mean – the population mean of Y had everyone followed regime d, and then do this for each of the k regimes.

MSMs: Dynamic Regimes

- However, suppose we have a large number of regimes, or that we have a small number of subjects. In these cases, we may wish to smooth across regimes to obtain a summary measure.

MSMs: Dynamic Regimes

- We could define an MSM m_β, which is a known function of our parameter $\beta=(\beta_0, \beta_1,\ldots, \beta_{K-1})$.
  - For example:

$$m_\beta=\text{logit}^{-1}(\beta_0+\beta_1 1_{d=d_1}+\ldots+\beta_{K-1}1_{d=d_{K-1}})$$

We can also consider nonsaturated MSMs, working MSMs, as well as MSMs that include baseline covariates

Marginal Stuctural Models

A few technical distinctions…

Marginal Stuctural Models

Approach #1:

- Assume a model m_β for the parameter $\Psi(a)$, for example:

$$\text{logit } \Psi(a)=\beta(a)$$

Here, we focus on estimating β, but have forced ourselves to make restrictive modeling assumptions that may not be true.

Reference: Rosenblum (2011). Marginal structural models. In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Marginal Stuctural Models

Approach #2:

- Define the parameter as a summary measure of the parameters $\{\Psi(a):a\}$.
- This is called a "working" marginal structural model.
- If helps us define a parameter that allows for smoothing, but does not represent an additional statistical or causal assumption.

Reference: Rosenblum (2011). Marginal structural models. In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Estimation Approaches

- Maximum-likelihood-based estimators
  - G-computation
    - requires estimate of outcome regression at each t
  - Targeted maximum likelihood estimation (TMLE)
    - requires estimate of **outcome regression** and **treatment mechanism** at each t
- Estimating equation Estimators
  - Inverse probability weighted estimators (IPW)
    - requires estimate of **treatment mechanism** at each t
  - Augmented inverse probability weighted estimators (A-IPW)
    - requires estimate of **outcome regression** and **treatment mechanism** at each t

Reference: Rose & van der Laan (2011). Why TMLE? In Targeted Learning: Causal Inference for Observational and Experimental Data, van der Laan & Rose, Springer: New York.

Study Objective

To determine the impact of multiple risks
on the overall health status of
HIV+ homeless and unstably housed adults

Study Objective

Used marginal structural models to define

$E(Y_{a(t)})=m_\beta=\beta_0+\beta_1 a(t\text{-}1)$

for their continuous outcome variables *Y*.

The authors used targeted maximum likelihood estimation to estimate the effect of each exposure, which led to a ranked list of "variable importance measures."

References

- Riley et al (2011) Population-level effects of uninterrupted health insurance on services use among HIV-positive unstably housed adults. AIDS Care, 23(7): 822-830.
- Riley et al (2011) Basic subsistence needs and overall health among human immunodeficiency virus-infected homeless and unstably housed women. AJE, 174(5):515-522.
- Riley et al. (2012) Social, structural and behavioral determinants of overall health status in a cohort of homeless and unstably housed HIV-infected men. PloS One, 7(4):e35207.
- van der Laan & Rose (2011) Targeted Learning: Causal Inference for Observational and Experimental Data. Springer: New York.
  - Rose & van der Laan, Chapter 2. (*For introduction to causal assumptions)*
  - Rose & van der Laan, Chapter 6. (*For comparison of different estimators)*
  - Rosenblum, Chapter 9. (*For chapter on MSMs we reference in this talk)*
- Snowden et al. (2011) Implementation of G-Computation on a simulated data set: demonstration of a causal inference technique. Am J Epid, 173(7): 731-738. (*For an introduction to g-computation for MSMs)*

Please feel free to email (srose@jhsph.edu and eriley@epi-center.ucsf.edu) for PDFs of papers or book chapters.