Gender-Based Analysis and Directed Acyclic Graphs

Developed in Women, Gender and Health 207:
Advanced Topics of Women, Gender, and Health,
Harvard School of Public Health, Spring 2014

Course Instructed by Jerel P. Calzo and Sabra L. Katz-Wise

Teaching Example Authored by Matthew Hamill and Emily R. Smith

Appropriate HSPH Core Courses:
The HSPH core course(s) in which the teaching example could be used are EPI201/202 or EPI 500. The teaching example (a gender analysis problem within a problem set) should be given after students have learned about confounding and DAGs.

Brief Background:
Students are expected to have read Chapter 6 of the Causal Inference textbook. Additionally they will have attended lectures covering sections 1.1 and 1.2 of the course material. As such, students will know that Directed Acyclic Graphs (DAGs) are part of a graph theory that is very useful for identifying sources of bias (e.g. confounding), problems with study design (e.g. selection bias), and can be used to develop appropriate data analysis plans.

Additionally, students are expected to acquire "Subject-Matter Knowledge" through research as part of the homework assignment. The basis for immune function are CD4 positive T lymphocytes, these are a proxy measure of a person's ability to fight off infections. HIV attaches to, infects, and destroys CD4 cells thereby leading to immune compromise and related illnesses. Antiretroviral therapy (ART) is used in combination to disrupt HIV viral replication at several points in its lifecycle. When a patient commences ART there is usually a rapid fall in HIV 'viral load' because of this impaired reproductive capacity. The fall in HIV viral load results in a rise in CD4 positive cells and usually indicates a reconstitution of immune function. CD4 counts tend to increase to a normal level in proportion to the duration of ART exposure. It is common that a rapid rise in CD4 count is seen within 6 months of starting ART and that the rate of increase slows over the following 12-24 months.

Learning Objectives for Students:
The learning goals for students include:

- To create a DAG prior to analysis in order to identify relevant confounders and develop a technically and conceptually appropriate model.
- To understand the difference between and importance of including sex and gender in model related to HIV disease progression.
- To further appreciate that gender can be further divided into categories such as gender identity and gender expression which may be distinct from sexual orientation. Each component of gender may occupy a different position in the DAG as each may be related to different 'risk' behaviors for delayed commencement of ART.

Teaching Methods:

The assignment below asks students to use recent knowledge about confounding, DAGs, and model building in order to explain why it is advisable to consider gender and sex separately when analyzing the homework dataset.

Assignment

The HIVPOINT dataset, available at EPI202's web site includes information on 6452 HIV-infected individuals in Europe and the U.S. You will use these data to estimate the effect of antiretroviral treatment A (1: yes, 0; no) at baseline on CD4 cell count after 3 years of follow-up. CD4 count, measured in cells/µl, is a marker of immunosuppression; the higher the CD4 cell count, the better. The goal is to estimate the average causal effect $E[Y^{a=1,c=0}] - E[Y^{a=0,c=0}]$ of treatment A on outcome Y had nobody been lost to follow-up (C=0). Provide your SAS code and output (relevant output only, please) at the end of each question. A code book for the HIVPOINT dataset can be found at the end of this assignment.

1. Consider the following list of variables measured at baseline: age, sex, **gender**, geographic origin, calendar year, mode of HIV transmission, ART adherence, health seeking behavior, CD4 cell count, and viral load (concentration of viral RNA in blood). Use your subject-matter knowledge to justify why adjusting for the variable gender and sex is advisable. Use a causal DAG to support your argument.

Possible Answer:

Adjusting for sex is necessary because it likely causes confounding between the effect of ART and 3-year CD4 cell count. A criteria used to start ART (thus a cause of the exposure) is also linked to long term health for people with HIV (thus a cause of the outcome). Thus, failure to account for baseline sex cell count may result in a spurious relationship between treatment and outcome. Additionally, gender or gender identity is important to consider in the model because it is highly correlated with health seeking behavior (e.g. some evidence suggests transgender people are less likely to seek care), and gender is related to ART adherence, which is an important predictor of 3-year CD4 cell count. The DAG below illustrates how both sex and gender are potential confounders and must be considered in an analysis in order to "close the backdoor pathways" between ART and 3-year CD4 cell count.