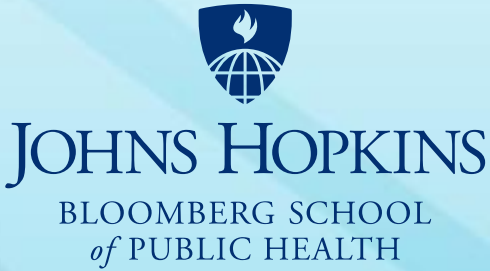
PCORI Methodology Standards: Academic Curriculum

Module 4: Planning a Nonexperimental Study

Category 8: Causal Inference Methods



Prepared by Elizabeth Stuart, PhD
Scott Zeger, PhD

Presented by Elizabeth Stuart, PhD

Planning a Nonexperimental Study

- When setting out to estimate a causal effect in a nonexperimental study, need to make many choices
  - Treatment and comparison conditions: What is the causal contrast of interest?
  - Population of interest
  - Outcomes and their timing
  - Dataset, study design

- Plan and pre-specify the same way it would be done in a randomized experiment

Specifying Treatment and Comparison Conditions

- First step is to state very clearly the treatment and comparison conditions
  - What is the relevant causal contrast?
- For example, in a study of a new case management program for individuals with severe mental illness, what is the comparison condition?
  - "Usual care"?
  - A less intense intervention?
- For example, in a study of the effect of statins on acute myocardial infarction rates:
  - What is the comparison condition? No statin? No treatment? One type of statin vs. another?
  - What is the "exposure" to statins required? One prescription? 60 days? 90 days?
- One is not necessarily right or wrong; they answer different scientific questions

Specifying the Population of Interest

- Need to be clear about who the individuals under study should be
  - Generally, those for whom the causal contrast is relevant
  - For example, individuals newly diagnosed with depression and trying to decide on treatment options
- Connect this population with the data to be examined
  - For example, don't include individuals for whom the treatment choice is not relevant
- In study reports, document clearly who is in the analysis sample

Example: CATIE Trial

- CATIE trial on the effectiveness of antipsychotic drugs among individuals with schizophrenia
  - Individuals randomized to one of five antipsychotics (first and second generation)

- Population of interest
  - 18–65 years old
  - Diagnosed with schizophrenia, as based on DSM-IV
  - Able to take oral antipsychotic medication
  - Excluded from the trial were those who had certain co-occurring conditions or history of treatment resistance

Source: Lieberman, J. A., et al. (2005, September 22). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med, 353*(12), 1209–1223. http://www.ncbi.nlm.nih.gov/pubmed/16172203

Use Only Covariates to Define a Population

- A population should be defined based on the characteristics measured at the time of treatment choice or before
  - Not on things that may be affected by the treatment, such as side effects, outcomes, or adherence to the intervention
  - This is true when exposure is time-varying as well

- In a prospective study, such as a randomized trial, use the time of study entry to define the study population
  - All individuals randomized should be included in the analysis

- In a retrospective study, use a defined time period prior to exposure to define the study population
  - Similar to inclusion/exclusion criteria in a randomized trial

Define Outcomes Clearly

- Specify in advance what the outcome(s) of interest is

- Clarify timing with respect to initiation and duration of the exposure
  - Must be measured after the exposure
  - For example, the number of missed days of work after 3, 6, and 12 months following the initiation of talk therapy or medication for depression

- CATIE trial:
  - Primary outcome: discontinuation of treatment for any reason
  - Secondary outcomes: reasons for discontinuation (e.g., weight gain), symptom scores
  - Measured at 1, 3, 6, 9, 12, 15, and 18 months after randomization

Develop Understanding of Why People Got Which Treatment

- In choosing a nonexperimental study design, it helps to understand why some people got one treatment vs. another
  - Confounding by indication?
  - Qualitative interviews can help this

- Especially, do we observe the characteristics based on which treatment choices were made?
  - Was a doctor choosing based on factors observed in the electronic health record?
  - Or was the patient choosing—potentially based on some factors we can't observe?

- Polsky et al. *(PharmacoEconomics,* 2009): value in including clinical baseline hemoglobin levels when comparing different erythropoiesis-stimulating agents
  - This is especially true if the treatment choice depends on the clinical measures

Source: Polsky, D., Eremina, D., Hess, G., et al. (2009, September). The importance of clinical variables in comparative analyses using propensity-score matching: The case of ESA costs for the treatment of chemotherapy-induced anaemia. *PharmacoEconomics, 27*(9), 755-765.

Confounding by Indication

- A particular concern in this regard is confounding by indication
- Those given a treatment of interest may be those individuals who are the most sick/the most at risk for the outcome
- For example, blood pressure medication may look like it increases risk of heart attacks, but only because it is given to people at higher risk for heart attacks

Document Potential Confounders and Their Timing

- Measure confounders before the start of the exposure
  - Also true in a setting with time-varying exposure
- Ensure not affected by exposure
- Report data on the confounders with the study results
  - For example, similarities or differences between groups

Danger in Adjusting for Post-exposure Variables

- Adjusting for variables that may be affected by the exposure (e.g., are measured after exposure) can cause bias

- Example: conditioning on birth weight when estimating association between exposure during pregnancy and some perinatal outcome can cause bias*
  - For example, when looking at smoking and infant mortality, conditioning on birth weight can cause the appearance of a protective effect from smoking
  - Smoking appears protective when looking at lowest birth weight babies

*Source: VanderWeele, T. J., Mumford, S. L., and Schisterman, E. F. (2012, January). Conditioning on intermediates in perinatal epidemiology. *Epidemiology, 23*(1), 1–9. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3240847/

Graphical Model from VanderWeele et al. (2012)

- Essentially, low birth weight babies from mothers who did and didn't smoke are likely very different from one another in lots of ways

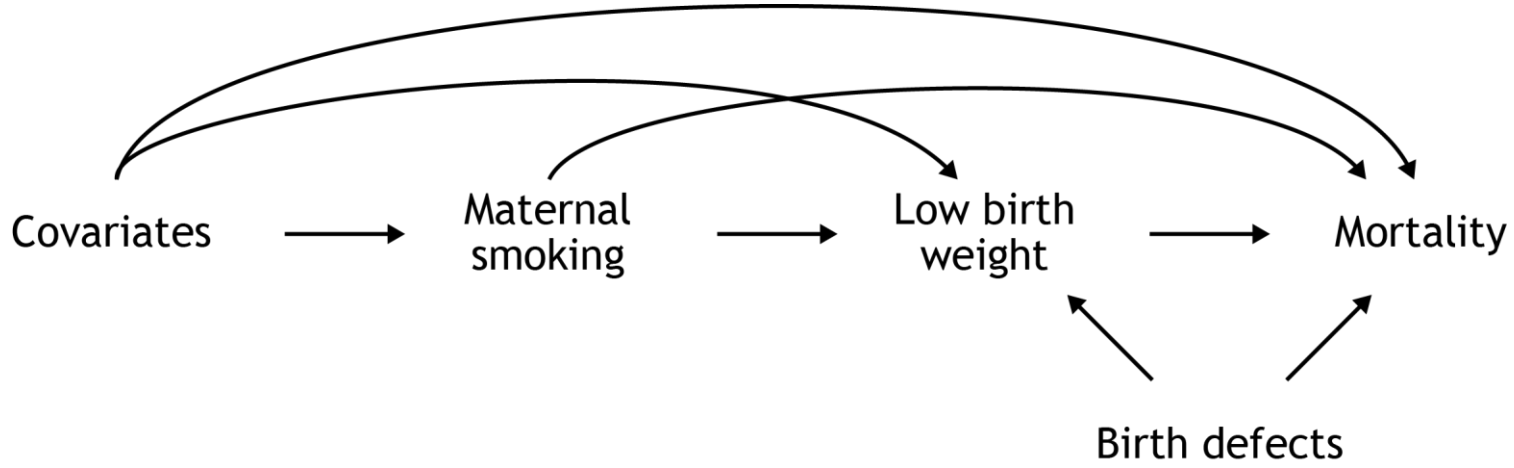


Image source: VanderWeele, T. J., Mumford, S. L., and Schisterman, E. F. (2012, January). Conditioning on intermediates in perinatal epidemiology. *Epidemiology*, *23*(1), 1–9.

Choice of Study Designs and Datasets

- An understanding of the extent of confounding can help in the selection of a study design and dataset

- Do we observe most of the ways in which we think individuals in different treatment conditions differ? Or could there be strong unobserved differences?
  - Are some datasets more comprehensive in terms of having potential confounders observed?
  - If a set of observed confounders is relatively weak, methods such as instrumental variables may be more appealing

- What sorts of study designs for a nonexperimental study may be appropriate to enhance causal inferences?

Conclusions

- When estimating causal effects in experimental or nonexperimental studies, it is important to clearly specify the population, the causal contrast of interest, the outcomes
  - Make sure that the population is not defined on the basis of post-treatment variables

- In nonexperimental studies, it may be useful to think about how the study would be designed if it were a randomized experiment