Propensity Score-Based Methods for Causal Inference

Module 2: Developing the Research Question



**I. Module Objectives**

Once we have defined what we mean by a causal effect, we can define the research question in terms of a causal effect. One approach to doing so is to frame the question as a hypothetical randomized target trial, as recommended by Hernan and Robins (2016). This module describes their approach and provides other guidance in terms for the PICOTS framework.

By the end of this module, you will be able to:

1. Frame your research question as a randomized target trial (and/or modify the question to do so)
2. Specify the population, intervention and comparator (or more generally exposure), outcomes, timing and setting (PICOTS) for your question

**II. Module Assignments**

**Required Assignments: (~ 40 minutes - 7 pages to read + 26 minutes over 2 videos)**

One of the most important aspects of asking a research question in the context of causal inference using observation data is to be able to frame the question as if it were a randomized trial. Without doing so, one cannot properly specify the causal effect. For an introduction to this concept, read the following article: Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American journal of epidemiology*, *183*(8), 758-764.

Another key aspect of asking a research question is to specify key components of that question, including the population and the exposure of interest, the comparator to that exposure (e.g. no exposure or some baseline exposure level), the outcome of interest, and the timing of exposures and outcomes, as well as the setting. A related framework (which is again more specific to comparative effectiveness research) is PICOTS (for population, intervention, comparator, outcome, timing and setting).

The PCORI Methodology Standards Academic Curriculum (although somewhat specifically focused on patient-centered outcomes research) also provides some relevant description of PICOTS in Module 8 (**~4 minutes**) of [Category 1. Standards for Formulating Research Questions](https://www.pcori.org/research-results/about-our-research/research-methodology/methodology-standards-academic-curriculum-0). They also discuss issues in designing an observational study consistent with a given causal question and target trial in Module 4 (**~18 minutes**) of [Category 8. Standards for Causal Inference](https://www.pcori.org/research-results/about-our-research/research-methodology/methodology-standards-academic-curriculum-5).

**Optional Assignments: (~11 pages to read)**

Read the following article for further explanation of the target trial approach using an example in cancer screening.

García-Albéniz, X., Hsu, J. and Hernán, M.A., 2017. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *European journal of epidemiology*, *32*(6), pp.495-500.

For further discussion of the target trial approach, and some underlying probabilities and assumptions, see Chapter 3, Section 3.6, pages 36-40, of Hernan and Robins (in draft form, and freely available at <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/> as of June, 2019): Hernán MA, Robins JM (2019). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming.

Additional resources have been developed specific to comparative effectiveness of two or more treatments; see <http://www.cerbot.org/> and <http://www.cerbot.org/resources>. These resources, however, focus more specifically on treatments and interventions (rather than just exposures), and expand coverage of methods into more complex time-varying scenarios (beyond just propensity score-based methods).

**III. Project Exercises**

Create a copy of this Google Doc or download the Module onto your computer and review the material offered above under Module Assignments before beginning these workbook exercises.

Thinking about what you learned in this module so far, continue developing the analysis plan for your project by answering the following questions:

1. Describe your research question as a hypothetical randomized controlled trial. In answering this question, including explanations for each of the following components (as outlined in the Hernan and Robins manuscript, but slightly modified to be more general to exposures rather than just treatments):
   1. Eligibility criteria
   2. Exposure patterns of interest, or treatment or intervention strategies being compared
   3. Assignment procedures
   4. Follow-up period
   5. Outcome
   6. Causal contrasts of interest
   7. Overarching analysis plan
2. To reinforce these concepts, describe your study using the PICOTS framework, but where I can stand for an intervention or, more generally some exposure.
3. Begin researching the possible observational data sets that you could use to emulate the randomized trial. List some possible data sets with a brief (2-3 sentence) description of each. If you plan to collect new data for this research, write a brief description of your plans for data collection (as this will be covered in more detail in the next module).
4. For each data set, consider your responses to questions #1 and #2 above. For each data set, describe any potential strengths or weaknesses for each specific point. For instance, does the data set mimic your eligibility criteria of interest, and/or does it have sufficient follow-up time, collect the outcomes of interest, etc.?

[Link to go back to the Course Overview Document](https://docs.google.com/document/d/1UDTkp3rbhqdun7jvSvktaZmTtoUWOz_VUDQw3HIsElg/edit?usp=sharing)