The ICH E9 addendum from an academic causal inference perspective and feedback on talks

Fabrizia Mealli

University of Florence

Dept. of Statistics, Computer Science, Applications

Florence Center for Data Science

fabrizia.mealli@unifi.it

RCTs meeting causal inference: principal stratum strategy and beyond

BBS Webinar

September 7, 2020

Bridging the IV-PS Causal Literature with the Addendum

- Principal Stratification (PS, Frangakis, Rubin, 2002) has its roots in Instrumental Variable (IV) literature but MORE general

- Angrist, Imbens, Rubin (1996) introduced the notions of Complier (C), Never-Taker (NT), Always-Taker (AT), Defier (D)

- The global *ITT* may be written as the weighted average of the *ITT* effects across the four subpopulations:

$$ITT = \pi_C ITT_C + \pi_{NT} ITT_{NT} + \pi_{AT} ITT_{AT} + \pi_D ITT_D$$

- **ITT** ignores treatment actually receipt like the **Treatment Policy Strategy**

Bridging the IV-PS Causal Literature with the Addendum

- **ITT$_C$** (or CACE or LATE) is about treatment effect heterogeneity w.r.t. a posttreatment variable, like the **Principal Stratum Strategy**
  - ✓ Depending on the subgroup, treatment effects may be have different meaning and can be attributed to different interventions
  - ✓ It is like a subgroup analysis, where groups are defined by post-treatment variables
  - ✓ This is why is crucial to characterize the subgroups based on the covariates
- **Average Treatment Effect (ATE)**, that is, the effect if everybody, contrary to fact, were forced to take or not take the treatment, is like a **Hypothetical Strategy**, often requiring assumptions of a different type

What can we learn from the PS literature and the talks today

- RCTs with intercurrent events should be designed and analyzed through the lenses of observational studies
- What makes observational studies credible?
  - ✓ Plausibility of assumptions
  - ✓ Covariates (to assess and describe populations)
  - ✓ Sensitivity analysis to deviations from such assumptions
- The plus of having initial randomization on our side!
  - ✓ some assumption hold by design because of randomization (e.g., random assignment=random instrument, monotonicity)
  - ✓ randomization has useful implications for identification and estimation (e.g., PS distribution is the same across treatment arms)

What are the implications for study protocol

- Anticipate intercurrent events so that estimands are defined a priori and possibly incoporated into treatment regimes (different complications imply focus on different subgroups/principal strata)
- Think about plausible assumptions and collect relevant baseline covariates or plan to collect secondary outcomes that may help identification in different ways
- Identify and collect baseline covariates that you anticipate being associated with PS membership that may help characterize/describe the principal stratum (as with missing data/dropout)
- Do not forget other principal strata! (size and characteristics); include it as a secondary analysis, it may help extrapolate effects to other subgroups or to whole population.
- Plan analysis (Bayesian) and sensitivity analysis

How can RCTs and causal inference literatures connect

- A lot can be learned from RCTs even with intercurrent events!
- Tools for estimation and sensitivity analysis are possible
- Develop tools that are flexible enough to accomodate several types of intercurrent events and covariates

Some Related References

Baccini M., Mattei A., Mealli F. (2017) Bayesian inference for causal mechanisms with application to a randomized study for postoperative pain control, *Biostatistics*, 18, 605-617.

Feller A., Mealli F., Miratrix L. (2017) Principal Score Methods: Assumptions and Extensions, *Journal of Educational and Behavioral Statistics*, 42, 726-758.

Mealli F., Pacini B., Stanghellini E. (2016) Identification of Principal Causal Effects Using Additional Outcomes in Concentration Graphs, *Journal of Educational and Behavioral Statistics*, 41, 463-480.

Forastiere L., Mealli F., VanderWeele T. (2016) Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs: Disentangling Bed Nets using Bayesian Principal Stratification, *Journal of the American Statistical Association*, 111, 510-525.

Mattei A., Li F., Mealli F. (2013) Exploiting multiple outcomes in Bayesian inference for causal effects with intermediate variables, *The Annals of Applied Statistics*,7, 2336-2360.

Mealli F., Pacini B. (2013) Using secondary outcomes to sharpen inference in randomized experiments with noncompliance, *Journal of the American Statistical Association*, 108, 1120-1131.

Frumento P., Mealli F., Pacini B., and Rubin D.B. (2012) Evaluating the Effect of Training on Wages in the Presence of Noncompliance, Nonemployment, and Missing Outcome Data, *Journal of the American Statistical Association*, 107, 450-466.

Mattei A., Mealli F. (2012) A Refreshing Account of Principal Stratification, *Inter Journal of Biostatistics*, 1, 1-37.

Rubin D.B. (1974). Estimating Causal Effects of Treatments in Randomized and nonrandomized studies. *Journal of Educational Psychology*, 66: 688-701.

Rubin D.B. (1978). Bayesian Inference for Causal Effects. *The Annals of Statistics* 6, 34-58.