Discussion

Elizabeth Stuart
Professor
Departments of Mental Health, Biostatistics,
and Health Policy and Management
Johns Hopkins Bloomberg School of Public Health

estuart@jhu.edu
www.biostat.jhsph.edu/∼estuart

June 8, 2016

Thanks to Joe and co-authors!

- Very interesting work
- Nice to see the connections, and distinctions, between mathematical modeling and statistical modeling
- Have often thought that there should be more work connecting the two
- Also a nice example of bridging important public health/policy questions and statistical research
  - HIV Care Cascade makes a lot of sense intuitively and from a public health perspective; how do we then formalize that from a statistical perspective, to study it?

Targeting analyses to parameters of interest

- One challenge is that large-scale mathematical models often try to model everything all together and thereby need a lot more assumptions
  - e.g., model of how covariates relate to each other
- Another strategy is to tailor analyses and frameworks for the *particular* questions of interest
  - e.g., to understand the effects of interventions to increase adherence to treatment, can focus on parameters related to that and maybe don't have to worry about the whole cascade
  - When interest is in just a few pieces of the larger model maybe it makes sense to focus on those and treat the rest as nuisance parameters
  - This is what Joe did when looking at two treatment policies (treat immediately vs. treat when CD4 below 350) and detailing the specific estimand of interest

Connection with Structural Equation Models and estimating causal effects

VanderWeele (2012; *American Journal of Epidemiology*) has a nice discussion of this. From the abstract:

> *"SEMs estimate more types of effects than do these other techniques [regression, causal diagrams, marginal structural models, etc.], but this comes at the price of additional assumptions. Many of these assumptions have often been ignored and not carefully evaluated when SEMs have been used in practice. In light of the strong assumptions employed by SEMs, the author argues that they should be used principally for the purposes of exploratory analysis and hypothesis generation when a broad range of effects are potentially of interest."*

So how do we deal with those assumptions?

- Assess sensitivity to them
  - Consider ranges of values for parameters, especially those without good empirical estimates
- Fully acknowledge that they exist (e.g., implications of missing arrows, potential unobserved factors)
- Consider this particularly in relation to the quantities of ultimate interest
  - e.g., maybe some pieces of a large model, but NOT the quantities of actual interest, are sensitive to particular parameter values

The role of potential outcomes

- Another challenge when estimating (and drawing) these sorts of models is how to formalize the potential outcomes
- When trying to estimate causal effects of investigate causal relationships important to consider the potential outcomes
  - Outcome we would see if a given person gets the "treatment" of interest (Y(1))
  - Outcome we would see if a given person gets the "control" condition of interest (Y(0))

- Potential outcomes almost always absent from SEM and graphical models
- But not thinking about these potential outcomes as two separate random variables can cause problems
  - e.g., relationship between some predictor of interest and Y(1) may differ from the relationship of that variable with Y(0)
  - Particularly a problem when thinking about mediation (a particular value of a mediator under one intervention condition implies something very different than under the other; e.g., healthy marriage intervention)
- Joe gave an example of how to do this when defining $P_{\bar{a}}(S_j)$ and $P_{\bar{a}*}(S_j)$ and is related to his points about prediction vs. causation

Pros and cons of study designs

- This problem also gets us thinking about standards of evidence and weighing pros and cons of different study designs
- An analogy is the debate about experiments "vs." non-experimental studies
- Conventional wisdom is that experiments are always better than non-experimental studies
- But when estimating a population treatment effect (effect in a well-defined target population) the trade-offs are less clear
- Imai, King, & Stuart (2008; JRSS-A) decomposes overall bias into that due to internal and external validity
  - When estimating a population effect, a small non-representative randomized trial may yield more bias than a large non-experimental study in a representative sample
  - (Has potential relevance when thinking about large-scale EHR data used by Joe et al.)

Formalizing trade-offs of different designs (Imai, King, & Stuart, 2008)

- Interested in the effect of some treatment, T, on an outcome, Y, in some target population
  - Population average treatment effect (PATE)

$$\text{PATE} = \frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0))$$

- Simple estimate of PATE is just a difference in means of the outcome between the observed treated and control groups

$$D = \frac{1}{n_1}\sum_{i=1}^{n_1} Y_i(1) - \frac{1}{n_0}\sum_{i=1}^{n_0} Y_i(0)$$

- Bias in estimated treatment effect: Δ=PATE - D
- How do different design elements affect the size of Δ?

Decomposition of Δ

Different study designs entail trade-offs between sources of bias

$$\Delta = (\Delta_{SX} + \Delta_{SU}) + (\Delta_{TX} + \Delta_{TU})$$

- Decompose Δ into 4 parts, due to:
  - Sample selection bias (S), Treatment selection bias (T)
  - Observed variables (X), Unobserved variables (U)

Examples of trade-offs of different designs

Study design	Sample Selection Bias		Treatment Selection Bias	
	Δ_{SX}	Δ_{SU}	Δ_{TX}	Δ_{TU}
Ideal experiment	0	0	0	0
Typical experiment	Big	Big	0	0
Typical non-exp study	Small	Small	Big	Big
Well-done non-exp study	Small	Small	Small	?

Need for more investigation . . .

- How to assess which parameters in a large model are the ones that may matter most for estimand(s) of interest
- Formalizing sensitivity analyses, especially when there are many different assumptions and parameters that are of potential concern
- Methods for bringing in parameter estimates from varying data sources
  - Accounting for uncertainty and for potential differences in populations across the sources
- How to calibrate/validate when interest is in causal effects
  - Never observe the true effects; is prediction of the potential outcomes sufficient? What about (non)-overlap between treatment groups (extrapolation)?