Using observational studies to determine RCT generalizability

Eloise Kaizar
The Ohio State University

Joel Greenhouse, Howard Seltman, Carnegie Mellon University
Kelly Kelleher, Bill Gardner, Jack Stevens Nationwide Children’s Hospital
Target Questions

- In the development setting (“efficacy”)
  - Treatment effect for the “best” individuals, those likely to:
    - experience the largest positive treatment effect
    - be cost-efficient
    - not be harmed by treatment

- In the policy setting (“effectiveness”)
  - Population average treatment effect (PATE)
  - Sub-population average treatment effect (SPATE)
  - Personal treatment effect, SPATE for n=1

Can we generalize studies designed to evaluate efficacy to inform effectiveness?
Populations

- **Target Population**
  - Population of all individuals for which treatment may be considered for its intended purpose.

- **Trial Population**
  - Theoretical population that consists of all individuals who would be eligible to enroll in the RCT.

- How is trial population related to target population?
Simple Random Sample

- RCT results are directly generalizable
- Distribution of baseline variables should be identical
  - Special attention paid to known/suspected treatment moderators
- Distribution of outcome variables should be logically related
  - Example: Suicidality in Antidepressant RCTs, compared with YRBS, Greenhouse, et al. (2008)
Weighted Sample

- RCT results can be weighted to generalize
- Reweighting schemes include:
  - Propensity-based standardization; Cole and Stuart (2010), Stuart, et al (2011)
Weighted Sample of Subpopulation

- Exclusion/inclusion criteria are known
- Studies consistently show that it is usual for RCTs to include half or less of the target population due to eligibility criteria
  - Asthma, <43% (Travers, et al., 2007)
  - Hypertension, Comorbidities <20% (Fortin et al., 2006)
  - Alcohol Treatment 30-94% (Humphreys and Weisner, 2000)
  - Antidepressants mean=34% (Zimmerman, et al., 2004)
- Generally unknown if these exclusions are important (modify treatment)
Extrapolating

- Sensitivity analysis
  - What must the effect size be in the excluded subpopulation for the inference for the target population to change?
- Compare RCT data to parallel observational study
  - Marcus (1997)
- Apply exclusion criteria to data representative of target population
  - Based on the Cross Design Synthesis framework, GAO (1992)
  - Kaizar (2011); Pressler and Kaizar (2013)
## Simple Case: Linear Bias

<table>
<thead>
<tr>
<th>Population</th>
<th>Randomized Data Estimand</th>
<th>Obs. Data Estimand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Represented by RCT</td>
<td>$D^R_{included}$</td>
<td>$D^O_{included}$</td>
</tr>
<tr>
<td>Not Represented by RCT</td>
<td>$= D^R_{included}$ + extrapolation</td>
<td>$D^O_{excluded}$</td>
</tr>
</tbody>
</table>

- **Data requirements:**
  - Treatment, outcome, exclusion recorded

- **Assumptions:**
  - No residual confounding within subpopulations, or
  - Residual confounding is “separate” from exclusion criteria
Example

Extensions

- Multiple RCTs and Observational data sets
  - Additional strata used for multiple inclusion criteria
- Multiple treatments
  - Doses
  - Comparative effectiveness
- “Fuzzy” population membership
References


References, cont.

References, cont.


