Interoperability, Innovation, and Strategy using Real World Evidence

Taylor Pressler, Director- Healthcare Consulting
Barry D. Reese, PgD-MSPA

Outline for Today's Talk

- Interoperability between Randomized Clinical Trials (RCT) and Real World Data (RWD)
- Finding patterns in the RWD using propensity scores and linear models to support advances in healthcare
- Components of Real World Evidence (RWE) and strategy for growth and development

Clinical Trials versus Observational Studies

- RCTs are can make causal inferences, but are expensive to run.
- Results are not typically generalizable to a population.
- Observational studies are relatively inexpensive, but causality can be confounded.
- Self selection of treatment means exclusion criteria not needed to minimize risk.

Interoperability of RCT and RWD

- Interoperability allows practitioners to combine benefits of both approaches:
  - We still want strong efficacious design for clinical trials
  - Extrapolate results of RCT to excluded subpopulation for increased external validity
  - Interpolate RWE into design of future RCT with increased external validity
- Generalizability Bias:
  - Difference in Population Average Treatment Effect (PATE) between the included and excluded population.

Patterns in RWD

- Linear models measure treatment effect in included population from RWD and compare against RCT:
  - RCTs assume stronger internal validity due to selection criteria and documented focus or purpose.
- Propensity scores predict likelihood of treatment selection:
  - Observational studies assumer stronger external validity due to relaxed exclusion criteria and patient choice.

Linear Models and Propensity Scores

- Linear models effective in estimating causal treatment effects.
- Propensity scores effective in predicting likelihood of treatment selection.
- Sample Selection Error (SSE) measures difference in treatment effect across subpopulations and weights linear models and propensity scores according to optimum performance

Components of a RWE Use Case

- PATE of included population
- PATE of excluded population
- Propensity score for patient treatment choice
- Sample selection error and generalizability bias
- Interoperability of RCT and RWD
- Strategy for growth and development

Method for Estimating SSE

- Assuming a non-constant treatment effect across Subpopulations I_i, we can use observational data to build an estimator for SSE using two elements:
  - Linear model to estimate the treatment effect on a particular individual
  - Propensity score model to estimate the likelihood that the individual would select the treatment of interests

Sample Selection Error

$$
\begin{aligned}
\gamma &= PATE - PATE^I \\
&= \frac{1}{N}\sum_{i=1}^{N} TE_i - \frac{1}{N_I}\sum_{i=1}^{N} TE_i(1 - I_i) \\
\text{where } I_i &= \begin{cases} 0 & \text{if exclusion criteria are satisfied for subject i;} \\ 1 & \text{otherwise.} \end{cases} \\
\text{and } N_I &= \sum_{i=1}^{N} I_i.
\end{aligned}
$$

Method for Estimating SSE

- Linear response model with related covariates and an interaction term for the treatment:

$$E[Y_i|(\mathbf{X}, \hat{\beta}, \mathbf{T}, \delta)] = \beta_0 + \beta_1 X_{1i} + \beta_2 X_2 + \beta_3 X_{3i} + \beta_4 X_{1i} t + \delta t_i$$

- Note: The model for the response should include all variables that could have an influence on the treatment response. Non-measurement may lead to bias.

Method for Estimating SSE

- Logistic model to estimate propensity of an individual selecting the treatment of interest, used as a weighting mechanism:

$$e(\mathbf{X}) = P(T = 1|\mathbf{X}, \alpha) = \frac{e^{\alpha_0+\alpha_1 X_1+\alpha_2 X_2+\alpha_3 X_3}}{1 + e^{\alpha_0+\alpha_1 X_1+\alpha_2 X_2+\alpha_3 X_3}}, \text{ where}$$

$$T_i \sim Bernoulli(e(X_i))$$

- Note: The model for the propensity scores should include all variables that could have an influence on the treatment response. Non-measurement may lead to bias.

Method for Estimating SSE

$$\hat{\Delta}_{dr} = \left[\frac{1}{n}\sum_{i=1}^{n} m_{1i}(\mathbf{X}_i, \hat{\beta}_1) + \sum_{i=1}^{n} \frac{T_i\left(Y_i - m_{1i}(\mathbf{X}_i, \hat{\beta}_1)\right)}{\hat{e}_i}\right] - \left[\frac{1}{n}\sum_{i=1}^{n} m_{0i}(\mathbf{X}_i, \hat{\beta}_0) + \sum_{i=1}^{n} \frac{(1-T_i)\left(Y_i - m_{0i}(\mathbf{X}_i, \hat{\beta}_0)\right)}{(1-\hat{e}_i)}\right].$$

- When the response model is correct, the error of the expected value of the response goes to zero. If incorrect, the second term provides a correction and is weighted by the propensity score.

Method for Estimating SSE

- Having already defined SSE as the following:

$$\gamma = PATE - PATE^I$$

- The following estimator can be used:

$$\hat{\gamma} = \hat{\pi}_E \left(\hat{\Delta}^E - \hat{\Delta}^I \right)$$

$$\hat{\Sigma}_\gamma = \frac{(1-\hat{\pi}_E)}{n\hat{\pi}_E}\hat{\gamma}^2 + \left(\frac{\hat{\pi}_E(1-\hat{\pi}_E)}{n} + \hat{\pi}_E^2 \right) \left[\hat{Var}(\hat{\Delta}^E) + \hat{Var}(\hat{\Delta}^I) \right]$$

Full Scientific Paper

Pressler TR and EE Kaizar. The use of propensity scores and observational data to estimate randomized controlled trial generalizability bias. *Stat Med.* 32(20): 3552-68, 2013.

Example Use Case

- Looking at the use of forceps or vacuum in an assisted delivery:
  - In the original RCT, mothers under the age of 18 were excluded from the study.
  - Number of previous births (gravida) was not considered and was an additional research question to understand if prima gravida had an impact on outcomes.
- Looks at degree of laceration, length of stay

Example Use Case

- Created response and propensity models using demographic and clinical information on 56,300 women to see if the choice of instrument had an effect on outcomes during an assisted delivery.
- Results: Use of vacuum in an assisted delivery led to 13% fewer severe lacerations.
  - In the original RCT, the difference was found to be insignificant at less than 1% difference

Why This is Useful

- Examine segments of the population where expected effect of treatment may be higher/lower
  - Analysis of adverse effects potentially caused by multiple co-morbidities or genetic factors
    - Are women with the BRCA2 gene more likely to have a lower response to a particular course of chemotherapy?
  - Develop new scientific hypotheses
    - Is there a genetic factor that might be able to explain why a certain subset of the population is having less of a response to a new blood pressure medication?
- Segment Populations according to Strategic Objectives
  - Drug Development and identify gaps/ overlaps
  - Drug Efficacies and linked diseases
  - Reach new patients
  - “Beyond the pill” correlation

Strategy for Growth and Development

- Use results of modeling to estimate which patients will select treatment and measure efficacy of treatment
- Segment patient populations according to strategic objectives
- Patient Safety
- Hypothesis generation leading to new drug development

References

[1] Imai, K., King, G. and E. Stuart. Misunderstandings between experimentalists and observationalists about casual inference. *J.R. Statist. Soc.* 171:481-502, 2008.

[2] Kang, J. and J. Shafer. Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data. *Stat. Sci.* 22(4):523-539.

[3] Lunceford, J. and M. Davidian. Stratification and weighting via the propensity score in estimation of casual treatment effects: a comparative study. *Statist Med.* 23:2937-2960, 2004.

[4] P. Rothwell. External validity of randomised controlled trials: To whom do these results of this trial apply? *Lancet.* 365:82-93, 2005.

[5] Stevens, J., Kelleher, K., Greenhouse, J., Chen, G., Xiang, H., Kaizar, E., Jensen, P. and L.E. Arnold. Empirical evaluation of the generalizability of the sample from the multimodal treatment study for ADHD. *Adm Policy Ment Health and Ment Health Serv Res.* 34:221-232, 2007.

[6] Wiltsey, S. and DeRubeis, R. and Crits-Christoph, P. and Brody, P. Are samples in randomized controlled trials of psychotherapy representative of community outpatients? A new methodology and initial findings. *J Consult and Clin Psychol.* 71(6):963-972, 2003.

[7] Zimmerman, M. and Mattia, J. and Posternak, M. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *Am J Psychiatry.* 159(3):469-573, 2002.