Generalize Randomized Clinical Trial (RCT) treatment effects to Real World target populations AND calibrate missingness in Real World data with RCT or other data sources

Sanofi Lead: Robert LoCasale (Robert.LoCasale@sanofi.com)
Department: Real World Evidence & Clinical Outcomes
Endorsement: Javier Jimenez, VP RWE & CO
Project/Drug: Any drug with Ph 3 data available for access
Functions Involved: RWE, Biostatistics, Clinical Operations, Medical Affairs

OVERVIEW

- Objectives:
  o Compare target population estimates of RCT treatment effects based on Propensity Score (PS) and Disease Risk Score (DRS) weighting.
  o Explore estimation of treatment effect in populations that are restricted in the RCT or in the target population when weighting based on PS or DRS.
  o Utilize propensity score approach in previous objectives with implementing Pressler’s approach to generalizability bias
- Disease focus: Any with Ph 3 data available for access
- Data sources: Ph 3 RCT and licensed electronic healthcare database (e.g. Optum/Humedica, PHEI); sourced through Sanofi or partnering institution
- Aim: Establish a best practice approach and extend the exploration of methods and challenges with generalizing RCTs to a target population identified in observational data
- Potential Impact: Clinical trial design, inform real world data needs, generalizing treatment effects and prioritizing future RWE studies.

DATA CHALLENGE

Common Ground: RCTs satisfy standard conditions to identify average causal effects of treatment assignment and treatment use (assuming negligible loss to follow-up and non-adherence)

Problem: RCTs are often conducted in populations that differ from real world and treatment effects in the real world are likely to differ from those estimated in RCT for a variety of issues, including (but not limited to) treatment effect heterogeneity. Under the assumption of non-uniform (heterogeneous) treatment effects, RCT results strictly apply to populations with the same (joint) distribution of effect modifiers than the one present in the RCT. We can make mistakes in estimating the causal effect when these differences are for factors that express modest (read non-significant) treatment-effect heterogeneity.

Solution: “Standardize” RCT results by adapting the distribution of the effect modifier(s) in the RCT population to the target population of interest. However, standardization is only straightforward when based on small number of categorical effect modifiers and requires point estimates for all combinations of factors leading to treatment effect heterogeneity. A solution
could be to standardize RCT results to a specific, user-defined target or ‘real world’ population by extending historical standardization to model-based standardization using inverse-probability weighting. Cole and Stuart proposed this weighting approach based on predicting the individual probability of being in the trial versus the target population as a function of multiple factors and to reweight the RCT participants to be representative of the target population which was identified in observational study data. The challenges of standardization with this method due to missing data in the target population were further explored by Hong et al and described during an oral presentation at the Annual International Conference for Pharmacoepidemiology in 2016.

**Data Challenge:** Extend the exploration of methods and challenges with generalizing RCTs to a target population identified in observational data. Some challenging areas could be as follows:

*Use of a disease risk score (DRS) versus propensity scores:* In clinical practice, treatment decisions are routinely based on clinical estimation of disease risk rather than propensity for treatment. The objective would be to compare target population estimates of treatment effects based on PS and DRS weighting.

**Positivity:** Weighting methods assume that everyone has a non-zero probability of being in each group compared. This is likely violated given that the target population may not be fully represented in the RCT and vice-versa. The objective would be to explore estimation of treatment effect in populations that are restricted in the RCT or in the target population when weighting based on PS or DRS.

**Applying methods from Pressler et al 2013:**
With suitable registry or secondary data and the use of inclusion/exclusion criteria for the RCT, we estimate the treatment effect D(I) on patients that could be included in the RCT and the treatment effect D(E) among patients who could not be included in the RCT. The generalizability bias is then defined as p(E)(D(I)-D(E)) where p(E) is the fraction non eligible registry patients. This method does not depend on having access to the RCT data, but it assumes that all patients that could be included in the RCT has the same probability of being so. The idea would be to use the propensity score approach to correct for the variability in selection probability when we estimate D(I). This would essentially estimate the D(I) we would see if the RCT eligible registry patient has the same distribution of patient characteristics as patients in the RCT. We could also do the same for the RCT patients with the same distribution of patient characteristics as the patients in the registry. The benefit is that we’d capture generalizability bias both from not covering all patients through inclusion/exclusion criteria and from selected patients in a controlled way that alters the distribution of patient characteristics, e.g. risk enrichment or use of primarily specialist clinics.

Separately, explore whether RCT or other clinical trial data can be used as a form of external adjustment or calibration for missing data in observational studies. Can RCTs be used to calibrate unmeasured confounding in a comparative observational analysis of
trial eligible patients? Can we account for unmeasured confounding by calibrating from multiple external sources into 1 study? For example, RCTs for calibrating one subset of patients for a select group of confounders and a national survey to calibrate for a different set of patients and perhaps different confounders thus incrementally improving the measurement of variables in the main study.

**Added Value to Sanofi:** Understanding RCT generalizability of patient population and treatment effects could:

1. Inform trial design through adjustments of inclusion/exclusion criteria or adding requirements for collection of specific data elements.
2. Inform future RWE data needs by identifying data elements necessary to compare with RCT.
3. Provide insights into how treatment effects may be generalized and help decide whether future RWE studies will be needed for regulatory and payer interactions and/or whether future RWE studies should focus on particular patient segments.

**References:**
