Leveraging Randomized Clinical Trials to Generate Real-World Evidence for Regulatory Purposes

Real-World Designs and Implications for Causal Inference

Vincent Willey, PharmD, BCACP
Principal Scientist
HealthCore
July 11, 2019
AIRWISE Trial Methodology

Study design

• Planned enrollment of 3200 COPD patients
  – Anthem and non-Anthem health plan members

• Randomized (1:1)
  – Dual bronchodilator therapy
  – Triple therapy

• Open label

• 12 month follow-up

• Community-based physician sites
  – Use of the administrative claims data to identify practices with the highest density of potential study patients

• Analysis
  – Non-inferiority
  – Primary endpoint: Time to first moderate or severe COPD exacerbation
Work With Existing Data To...

Develop the protocol
• Gaps in care
• Current treatment patterns
• Inform inclusion/exclusion criteria, sample size calculation and endpoints

Perform protocol feasibility
• Sites
• Patients

Recruit sites and patients
### “Preparatory” Retrospective Cohort Study

<table>
<thead>
<tr>
<th>260,750</th>
<th>patients with ≥1 claim for COPD (index date), aged ≥40 years, and ≥12-month pre-index enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>53,484</td>
<td>newly diagnosed patients with COPD with ≥24-month post-index enrollment with ≥1 more COPD claim</td>
</tr>
<tr>
<td>14,293</td>
<td>patients (27%) had ≥1 spirometry claim in 12 months before or 12 months after index date</td>
</tr>
<tr>
<td>4130</td>
<td>screened medical record population with confirmed spirometry results</td>
</tr>
<tr>
<td>1505</td>
<td>patients with spirometry-confirmed COPD FEV₁ / FVC &lt;0.7</td>
</tr>
</tbody>
</table>


COPD = Chronic Obstructive Pulmonary Disease; FEV₁ = Forced Expiratory Volume in 1 second; FVC = Forced Vital Capacity
“Preparatory” Retrospective Cohort Study


GOLD = Global Initiative for Chronic Obstructive Lung Disease; ICS = Inhaled Corticosteroids
LAMA = Long-Acting Muscarinic Antagonists; LABA = Long-Acting Beta Agonists
AIRWISE Inclusion/Exclusion Criteria

Inclusion Criteria

• COPD diagnosis as defined by the study physician  
  – no spirometry required
• Currently on one of the following COPD medication therapies  
  – LAMA monotherapy
  – LABA monotherapy
  – ICS/LABA
• Physician determination that patient is not controlled on current medication therapy
• Age ≥ 40 years

Exclusion Criteria

• Currently on LAMA/LABA or triple therapy (ICS/LAMA/LABA)
• Contraindications to any of the study drugs
• Documented diagnosis of current asthma
• Pregnancy or nursing
Real-World Designs and Implications for Causal Inference
Leveraging Randomized Clinical Trials to Generate Real-World Evidence for Regulatory Purposes: Real World Designs and Implications for Causal Inference

11 July 2019

Mark Levenson, Ph.D.
Center for Drug Evaluation and Research
Food and Drug Administration
By now we heard:

A randomized trial becomes an observational study on Day 2.
Figure 1. Cumulative Hazard Rates for the Primary Outcome of Stroke or Systemic Embolism, According to Treatment Group.

Connolly, et al. 2009, Dabigatran versus Warfarin in Patients with Atrial Fibrillation
<table>
<thead>
<tr>
<th>Variable</th>
<th>Dabigatran, 110 mg (N=6015)</th>
<th>Dabigatran, 150 mg (N=6076)</th>
<th>Warfarin (N=6022)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study-drug discontinuation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued at 1 yr↑</td>
<td>862 (15)</td>
<td>935 (16)</td>
<td>608 (10)</td>
</tr>
<tr>
<td>Discontinued at 2 yr↑</td>
<td>1161 (21)</td>
<td>1211 (21)</td>
<td>902 (17)</td>
</tr>
<tr>
<td><strong>Reason for discontinuation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient's decision</td>
<td>440 (7.3)</td>
<td>474 (7.8)</td>
<td>375 (6.2)</td>
</tr>
<tr>
<td>Outcome event</td>
<td>192 (3.2)</td>
<td>164 (2.7)</td>
<td>130 (2.2)</td>
</tr>
<tr>
<td>Serious adverse event‡</td>
<td>163 (2.7)</td>
<td>166 (2.7)</td>
<td>105 (1.7)</td>
</tr>
<tr>
<td>Gastrointestinal symptoms¶</td>
<td>134 (2.2)</td>
<td>130 (2.1)</td>
<td>38 (0.6)</td>
</tr>
<tr>
<td>Gastrointestinal bleeding</td>
<td>58 (1.0)</td>
<td>80 (1.3)</td>
<td>54 (0.9)</td>
</tr>
</tbody>
</table>
Graham et al. 2019, Comparative Stroke, Bleeding, and Mortality Risks in Older Medicare Patients Treated with Oral Anticoagulants for Nonvalvular Atrial Fibrillation.
Estimands (ICH E9R1 Draft)

- **Population**: patients targeted by the scientific question
- **Variable (endpoint)**: obtained for each patient, that is required to address the scientific question
- **Intercurrent events**: Events that occur after randomization
- **Summary**: population-level summary for a comparison between treatment conditions
Addressing Intercurrent Events

- Treatment policy strategy (think ITT)
- While on treatment strategy (think per protocol)
- Composite strategy

- Hypothetical strategy
- Principal stratum strategy

- Others approaches
  - Fixed follow-up time.
Real-World Designs and Implications for Causal Inference
Leveraging Randomized Clinical Trials to Generate Real-World Evidence for Regulatory Purposes: Real World Designs and Implications for Causal Inference

Jesse A. Berlin, ScD
VP and Global Head of Epidemiology
Johson & Johnson
11 July 2019
General comments

• How do you know whether treatment effects vary across subgroups unless you include the subgroups
  • Those using the medication in practice (beyond the limited populations typically included in registration studies)
  • FDA draft guidance on increasing diversity in clinical trials

• What about a kind of compromise?
  • Do the larger study, with broad entry criteria, e.g., include elderly, those with comorbidities, etc.
  • Plan the PRIMARY analysis to focus on the narrower cohort, but now you have randomized evidence on the other populations

• Consider sub-studies
  • Validation (and applying “correction factors”)

Causal inference methods to assess safety upper bounds in randomized trials with noncompliance

Yiting Wang†, Jesse A Berlin‡, Jose Pinheiro§ and Marsha A Wilcox

Clinical Trials
1–11
© The Author(s) 2015 Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav DOI: 10.1177/1740774515572352
cjt.sagepub.com

Abstract
Background: Premature discontinuation and other forms of noncompliance with treatment assignment can complicate causal inference of treatment effects in randomized trials. The intent-to-treat analysis gives unbiased estimates for causal effects of treatment assignment on outcome, but may understate potential benefit or harm of actual treatment. The corresponding upper confidence limit can also be underestimated.

Purpose: To compare estimates of the hazard ratio and upper bound of the two-sided 95% confidence interval from causal inference methods that account for noncompliance with those from the intent-to-treat analysis.

Methods: We used simulations with parameters chosen to reflect cardiovascular safety trials of diabetes drugs, with a focus on upper bound estimates relative to 1.3, based on regulatory guidelines. A total of 1000 simulations were run under each parameter combination for a hypothetical trial of 10,000 total subjects randomly assigned to active treatment or control at 1:1 ratio. Noncompliance was considered in the form of treatment discontinuation and cross-over at specified proportions, with an assumed true hazard ratio of 0.9, 1, and 1.3, respectively. Various levels of risk associated with being a non-complier (independent of treatment status) were evaluated. Hazard ratio and upper bound estimates from causal survival analysis and intent-to-treat were obtained from each simulation and summarized under each parameter setting.

Results: Causal analysis estimated the true hazard ratio with little bias in almost all settings examined. Intent-to-treat was unbiased only when the true hazard ratio = 1; otherwise it underestimated both benefit and harm. When upper bound estimates from intent-to-treat were ≥1.3, corresponding estimates from causal analysis were also ≥1.3 in almost 100% of the simulations, regardless of the true hazard ratio. When upper bound estimates from intent-to-treat were ≥1.3 and the true hazard ratio = 1, corresponding upper bound estimates from causal analysis were ≥1.3 in up to 66% of the simulations under some settings.

Limitations: Simulations cannot cover all scenarios for noncompliance in real randomized trials.

Conclusion: Causal survival analysis was superior to intent-to-treat in estimating the true hazard ratio with respect to bias in the presence of noncompliance. However, its large variance should be considered for safety upper bound estimation especially when the true hazard ratio = 1. Our simulations provided a broad reference for practical considerations of bias–variance trade-off in dealing with noncompliance in cardiovascular safety trials of diabetes drugs. Further research is warranted for the development and application of causal inference methods in the evaluation of safety upper bounds.
Causal inference

• Various approaches to controlling confounding and/or addressing non-compliance

• Instrumental variables (from economics)
  • Related to exposure but NOT to outcome, except through exposure

• **KEY POINT:** Randomization is the perfect instrumental variable
Wang and colleagues methods

• Simulations to reflect *randomized* cardiovascular safety trials of diabetes drugs (could be done in EHR system!)
• Focus on upper bound estimates relative to 1.3, based on regulatory guidelines (*non-inferiority!*)
• Hypothetical trial of 10,000 total subjects *randomly* assigned to active treatment or control
• Treatment discontinuation and cross-over
• Considered compliers, always takers, never takers, assuming *increased risk of CV outcome with non-compliance*