Head-to-head Comparisons using Real-World Data

Design and data source considerations from pilot investigations in CVD

July 28, 2020
Speakers

Alind Gupta  
Cytel

Miguel Hernán  
Harvard University

Judson Schneider  
Nashville Biosciences
This is a part of a webinar series

Webinar 1
Introduction and overview of pilot investigations

Webinar 2
Target trial design and data considerations for CVD pilot

Webinar 3
Pilot of pancreatic cancer, overview and methods

Webinars focusing on practical application and results from pilot projects will be hosted in the summer and early fall 2020.
Introduction: Target trial emulation
Randomized clinical trials

• Randomized trials are the standard for comparing effectiveness of interventions
  • Randomization reduces selection bias and confounding
  • Provide causal estimates of effects

• Decision-making and health policy need to be informed by causal knowledge about comparative effectiveness and safety

Treat with drug A, B or C?
Treat with drug B or C if no response to A? Treat now or later?
Which individuals to treat?
Randomized trials are not always feasible

• Possible reasons:
  • Costly
  • Untimely
  • Unethical
  • Impractical

• Trials may be limited in scope
  • E.g. eligibility criteria, outcomes measured
  • Real-world generalizability

• Decisions need to be made even in the absence of a randomized trial to address them – maintaining status quo is also a decision
Plan B: Emulate a randomized trial using RWD

- Try to emulate a hypothetical randomized clinical trial using observational data
- Limitations:
  - Can not emulate placebo, blinding, or force adherence
  - Limited by data available (e.g. only approved therapies)
Target trial framework

Step 1: Ask a causal question
• Equivalent to specifying the protocol of the analogous randomized trial explicitly

Step 2: Answer the causal question
• Identify a suitable RWD source
• Try and emulate the randomized trial using RWD

• Usually, it is necessary to cycle between steps 1 and 2 iteratively to tune them due to practical limitations, e.g. by restricting eligibility criteria
Problems with naïve analysis of RWD

• Asking the “wrong” or a poorly-defined causal question
  • Explicit representation of the causal question as a target trial prevents this

• Answering the causal question incorrectly
  • Not accounting for various biases that may arise
  • Not adjusting for relevant confounders/adjusting where we shouldn’t
  • Not using the appropriate methods

• Result: biased effect estimates, wrong conclusions
Example: post-menopausal hormone therapy and CHD

- Effect of post-menopausal hormone therapy on risk for coronary heart disease?

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Outcome Comparison</th>
<th>Hazard Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trial</td>
<td>“Initiators vs non-initiators”</td>
<td>1.24</td>
</tr>
<tr>
<td>Observational analysis (naïve analysis)</td>
<td>“Users vs non-users”</td>
<td>0.68</td>
</tr>
<tr>
<td>Observational analysis (target trial emulation)</td>
<td>“Initiators vs non-initiators”</td>
<td>1.05</td>
</tr>
</tbody>
</table>

- Manson et al. NEJM (2003)
- Grodstein et al. J. Women’s Health (2006)
Use cases for target trial emulation

- Generation of efficacy or safety evidence for conditional regulatory approval or post-market assessment
- Providing a comparison when network meta-analysis is not possible
- Expanding the scope (e.g. eligibility criteria) of a randomized trial
- Refining aspects of an existing treatment protocol
- Enabling a comparison to identify optimal treatment regimen
Pilot project: Head-to-head comparison of second-line therapies for type 2 diabetes

Considerations for protocol design and data source
Overview

• Patients with type 2 diabetes on metformin often fail to achieve/maintain glycemic control and require second-line therapy

• >6 classes of anti-diabetic drugs have been approved for use
  • A head-to-head comparison of effectiveness, safety is lacking
  • Choice of second-line agent varies across clinical practices

• There is a need for comparative effectiveness evidence to guide choice of second-line agent
  • To reduce detrimental effects, improve cost-effectiveness
  • Guide personalized therapy
Efficacy is the main driver in prescription practices globally

Reason for choice of second-line therapy by investigators for type 2 diabetes

Step 1 – asking the causal question

• Framing the causal question explicitly in the form of a target trial protocol

• Considerations for protocol design:
  • Who needs to be involved?
  • What are the roles for those involved?
  • How do you define the data/variables necessary for emulation?
  • How do you refine the protocol given practical limitations (e.g. data)?
GRADE trial

- Ongoing open-label randomized trial for comparative effectiveness of second-line therapy in type 2 diabetes
  - Started in 2013
  - Est. primary completion in July 2021

- Outcomes
  - HbA1C (primary)
  - Cardiovascular, microvascular side effects, adherence-tolerability

- No SGLT-2 inhibitor arm

Data considerations for target trial emulation

- Data from combination of EMR/EHR, claims data, clinician notes

- Brief checklist:
  - “Big” data – sufficiently large sample size, large number of variables
  - Representative sample covered
  - Bias/irregularities in capture of study variables
  - Long follow up
  - Database expertise
  - Access to clinical practitioners familiar with data
  - Data normalization, harmonization, de-identification
  - Cohort generation – minimizing bias, false positives/negatives

Partnership with Nashville Biosciences for RWD

• Subsidiary of Vanderbilt University Medical Center (VUMC)
• Leverages a large de-identified EMR, biospecimen repository BioVU®
  • E.g. 163,000 EMRs for type 2 diabetes
• Also provide scientific, clinical expertise
Source of data

- 6100+ providers and 68 hospital locations accepting ~2.5 million patients annually
Target trial protocol – checklist

- Eligibility criteria
- Treatment strategies
- Causal contrasts of interest
- Confounders for assignment to treatment strategies
  - Other confounding variables will need to be defined depending on treatment strategies and causal contrasts (e.g. for adherence)
- Outcomes
- Follow-up
  - Start and end time
  - Defining time zero
- Analysis plan (will be discussed in next webinar)
# Eligibility criteria

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Available in NB data?</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2DM diagnosis</td>
<td>Yes</td>
<td>ICD, ICD+meds+labs</td>
</tr>
<tr>
<td>Age (&gt;45 years)</td>
<td>Yes</td>
<td>DOB</td>
</tr>
<tr>
<td>Antihyperglycemic monoTx with metformin for &gt;3 months in 2014/later</td>
<td>Yes</td>
<td>Rx records, self-reported drug use</td>
</tr>
<tr>
<td>Suboptimal glycemic control after metformin initiation (&gt;7% HbA1c)</td>
<td>Yes</td>
<td>Labs</td>
</tr>
<tr>
<td>No suspected T1D or other reasons for hyperglycemia (rare disorder, pancreatic surgery)</td>
<td>Yes (?) – depends on start of care at VUMC relative to treatment start date</td>
<td>ICD, CPT</td>
</tr>
<tr>
<td>Not pregnant</td>
<td>Yes</td>
<td>ICD, CPT</td>
</tr>
</tbody>
</table>
Treatment strategies

• Initiation of second-line therapy within 12 months of persistent inadequate HbA1c control with metformin use from one of:
  • GLP-1 RA
  • SGLT-2 inhibitor
  • DPP-4
  • Sulfonylureas
  • (TZD, insulin)

• Other strategies depend on data:
  • Sustained treatment with, addition of a third-line therapy

• Limitations:
  • Limited to drugs prescribed at/dispensed at VUMC pharmacy
Causal contrasts

• Intention-to-treat effect
  • Effect of being assigned to treatment strategy at baseline
  • Of interest in the GRADE trial
  • Requires adjustment for non-random allocation at baseline + non-random loss to follow-up

• Per-protocol effect
  • Effect of adhering to assigned treatment strategy
  • Also requires adjustment for non-adherence (pre + post-baseline factors, e.g. treatment discontinuation)
  • May be feasible depending on availability of sufficient relevant data on adherence
### Confounders for treatment assignment at baseline*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Available in NB data?</th>
<th>Source/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of medical insurance</td>
<td>Yes</td>
<td>For most</td>
</tr>
<tr>
<td>History of CVD</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Physician comfort level/perception of drug</td>
<td>Unknown</td>
<td>May need help of treating physician or their Rx histories</td>
</tr>
<tr>
<td>HbA1c levels</td>
<td>Yes</td>
<td>Labs</td>
</tr>
<tr>
<td>Contraindications</td>
<td>Yes</td>
<td>Rx records</td>
</tr>
<tr>
<td>BMI</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Frailty/tolerability</td>
<td>(?) - Depends on how we define this, may involve practitioners</td>
<td></td>
</tr>
</tbody>
</table>

*How can we tell if we have adjusted sufficiently, appropriately?
Outcomes

• Metabolic
  • HbA1C levels (may not be a patient-important outcome)
  • BMI

• Macrovascular
  • MACE
  • Heart failure

• Microvascular
  • Glomerular filtration rate (eGFR)
  • Retinopathy

Follow up

• Defining time zero
  • Time of treatment prescription

• Patients are followed from time zero until diagnosis of outcome or loss to follow-up due to various factors
  • E.g. due to discontinuation of pharmacy benefits, medical insurance etc.
  • Ideally, minimize loss to follow-up
  • Possible ways include only including patients with a history of regular visits
Next steps

• Cohort extraction and sample size estimation
• Iterative tuning of protocol given data limitations
• Definition of final protocol and analysis plan
• Trial emulation
  • Results – survival curves, relative risk estimates
  • Subgroup analyses
• Sensitivity analysis
• Benchmarking against results from GRADE if possible
Conclusion

• In the absence of a randomized clinical trial, we can try to emulate a hypothetical randomized trial using RWD
  • Can also complement existing trials
  • Fraction of time and cost
• Considerations discussed today:
  • Trial protocol design
  • Suitable high-quality RWD for emulation
• Trial design may need to be refined iteratively due to practical constraints
• Multiple parties are needed to successfully define protocol and interact with big data (e.g. for data validation, database expertise)
Next webinar

Webinar 1
Introduction and overview of pilot investigations
July 7, 2020

Webinar 2
Target trial design and data considerations for CVD pilot
July 28, 2020

Webinar 3
Pilot of pancreatic cancer, overview and methods
August (TBD)

Webinars focusing on practical application and results from pilot projects will be hosted in the summer and early fall 2020.
Questions?