Real World Assessment and Research of Drug performance (REWARD)
Software demonstration

James Gilbert PhD
OHDSI Community Call 8th December 2020

Co-authors: Rachel Teneralli, Chris Knoll, Dave Kern, Soledad Cepeda, Patrick Ryan
Introduction

REWARD Enables the discovery of unknown effects of existing medications

This is currently achieved through:

- Exploiting OMOP CDM compliant databases using standardized vocabularies
- Self-controlled cohort designs
- Empirical calibration to minimize systematic error
- Data Visualization and Exploration in Shiny Dashboards
The problem

Where does knowledge about drugs come from?

- Mechanistic understanding and compound discovery
- *In silico* simulation and *in vitro* lab studies
- *In vivo* animal model testing
- Clinical trials
- **Still have only a limited idea of what happens when you expose a population at large**
Real World Assessment and Research of Drug performance (REWARD)
To strengthen and expand our existing observational health data analytics system that proactively provides real-world evidence through the population-level effect estimation to enable exploration of performance of existing medical products.
Population level estimation provides evidence on what happens to real people exposed to real medicines.

This can be exploited to generate new hypotheses about any existing medications.

How can we apply this at the scale of all exposures by all outcomes?
## Claims Databases currently used for REWARD

<table>
<thead>
<tr>
<th>Source Name</th>
<th>Persons</th>
<th>Average Person-years</th>
<th>Years covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optum SES</td>
<td>84,310,086</td>
<td>2.97</td>
<td>May 1, 2000 to Dec 31, 2019</td>
</tr>
<tr>
<td>IBM CCAE</td>
<td>152,963,555</td>
<td>2.73</td>
<td>Jan 1, 2000 to Feb 29, 2020</td>
</tr>
<tr>
<td>IBM MDCD</td>
<td>28,917,265</td>
<td>3.35</td>
<td>Jan 1, 2006 to June 30, 2019</td>
</tr>
<tr>
<td>IBM MDCR</td>
<td>10,115,120</td>
<td>3.93</td>
<td>Jan 1, 2000 to Jan 31, 2020</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>276,306,026</strong></td>
<td><strong>3.25</strong></td>
<td><strong>Jan 1, 2000 to Feb 29, 2020</strong></td>
</tr>
</tbody>
</table>
Exposure Cohorts:
• Rx Norm Ingredient
• New user cohort with 365 day washout period
• Custom cohorts defined in ATLAS e.g. drug classes
• ~2500 ingredient exposures (prescribed drugs)

Outcome cohorts:
• SNOMED standard concepts
• First occurrence of condition record
• Custom cohorts defined in ATLAS, e.g. phenotype library
• ~13,000 outcomes

Resulting combination of exposures and outcomes:
– 32,500,000 exposure-outcome combinations
– Sparse, heavy tailed distribution – most exposures have no outcomes.
Self-controlled cohort studies

Time prior to treatment
Equal in length to time on treatment

Time on treatment

First diagnosis of condition: can occur prior to treatment, during treatment, or neither

- Individual serves as own control
- Identify occurrence of condition for all patients with drug exposure
- Post-exposure and pre-exposure incidence rates are calculated, rate ratio is calculated
- Note: Drugs indicated for the condition would look protective
  - Condition occurs before exposure
Empirical calibration

Adjusting for systematic bias

- Systematic bias occurs because any study design is inherently imperfect

- Controlling for bias is achieved through using negative controls i.e. medications and diseases that do not interact.

- Estimate an empirical null distribution

- Allows adjustment of p-values and confidence intervals

- Achieved using EmpiricalCalibration package in R

https://ohdsi.github.io/EmpiricalCalibration/
Automated negative control selection

- Adapted from Voss et al
- Scaled of all diseases by all treatments
- Use Common Evidence Model to construct map of drug-disease associations
- Based on Spontaneous Reports, PubMed MESH literature terms and Product Labels
- Mapping drugs to conditions in data-set is straightforward
- Mapping conditions to drugs requires looking up the concept hierarchy
- Initial validation performed on Depression data (Teneralli et al)
Examples - False positives

Lactulose for outcome of impaired cognition

Null distribution mean is IRR = 0.697 SD = 1.454.

Metformin for bipolar disorder

Null distribution mean is IRR = 0.689 SD = 1.282

Calibration plot shows that the null distribution of negative controls is heavily biased to the left. Blue dots show negative controls. Yellow dots show uncalibrated IRR effect estimates for exposure, outcome pairs. Orange shaded area indicates significant results after calibration.
System overview

- Central Postgresql database for cohort references
- Integration with ATLAS – cohorts and concept sets
- Aims to be platform independent\(^1\)
- Cohort references exported in a zip file
- Based on OHDSI R packages
- Shiny Application for visualization and exploration of results based on study areas

\(^1\) Currently built on MS RDW, Redshift under development.
Shiny Dashboard

Skin cells at 20x magnification
1. Or flat JSON files, e.g., from the Phenotype Library.
Conclusions

- REWARD enables the discovery of unknown effects of existing medications
- One use case, providing clinical insights for drug development
- This software builds on the OMOP CDM and OHDSI Open-Source software tools
- Provides convenient mechanism for exploring results
- Set up to support multiple CDM compliant data sources
- Allows a quick test for clinical questions
Future work

- To include the phenotype library for well validated cohorts
- Self-controlled case series studies in addition to self-controlled cohort
- True all-by all interface for fast hypothesis checking
- "one click" cohort characterization through integration with CohortDiagnostics
- End to end unit and integration testing
References


