Learning from the Observational Medical Outcomes Partnership (OMOP)

Patrick Ryan, Manager, Drug Development Sciences, GlaxoSmithKline Research and Development and Co-Investigator, Observational Medical Outcomes Partnership

Dr. David Madigan, Professor and Chair of Statistics, Columbia University and Research Investigator, Observational Medical Outcomes Partnership

July 1, 2010
Some Initial Housekeeping

- To minimize feedback, please confirm that the microphone on your telephone is muted.

- To mute your phone, press the mute button or ‘*6’. (To unmute, press ‘*7’ as well.)

- There will be several opportunities for questions and discussion throughout today’s session. Please use the Q&A tab at the top of your screen to submit your questions into the queue at any point and we will call upon you to state your question.

- We will open up the lines for questions from those participating only by phone at the end of each Q&A session.

- Call the Brookings IT Help Desk at 202-797-6193 with technical problems.
OMOP Overview and Methods Development Progress

Patrick Ryan, David Madigan
on behalf of OMOP Research Team
July 1, 2010
Outstanding questions for active surveillance

**Governance**
- What are the keys to a successful public-private partnership?

**Data**
- Which types of data? administrative claims, electronic health records
- Which sources? healthcare providers, insurers, data aggregators
- What are viable data access models:
  - centralized?
  - distributed?

**Performance**
- What are appropriate analyses for:
  - hypothesis generating?
  - hypothesis strengthening?

**Architecture**
- What is the appropriate infrastructure:
  - hardware?
  - software?
  - processes?
  - policies?

**Methods**
- How to maintain collaborations and engage research community?

**Technology**
- What are best practices for protecting data?
Breadth and diversity of OMOP research community

OMOP’s research community requires active participation from all key stakeholders, including government, academia, industry, health care organizations, and patient groups.

Governance
- 10 Executive Board members, chaired by FDA and managed by Foundation for NIH
- 21 Advisory Board members
- Led by 6 research investigators and PMO

Methods
- 17 methods collaborators

Data
- 6 distributed research partners
- 5 central databases included in the OMOP Research Lab

Technology
- 2 data access models, 7 different systems architectures

Over 100 collaborating partners
OMOP research experiment workflow

Drugs
- ACE Inhibitors
- Amphotericin B
- Antibiotics
- Antiepileptics
- Benzodiazepines
- Beta blockers
- Bisphosphonates
- Tricyclic antidepressants
- Typical antipsychotics
- Warfarin

Non-specified conditions
- All outcomes in condition terminology
- ‘Labeled events’ as reference
  - Warning
  - Precautions
  - Adverse Reactions
  - Postmarketing Experience

Health Outcomes of Interest
- Angioedema
- Aplastic Anemia
- Acute Liver Injury
- Bleeding
- GI Ulcer Hospitalization
- Hip Fracture
- Hospitalization
- Myocardial Infarction
- Mortality after MI
- Renal Failure
Methods Philosophy

• Many different analysts have considerable experience with different statistical and epidemiological approaches in different observational settings
• Massive claims/EHR databases present challenges where little theory exists to guide methodological choices
• Our approach is empirical
• Implement a broad swathe of methods and evaluate them against “ground truth.” Place code in the public domain.
• No magic bullet - but we expect to identify gross differences between different methods in different scenarios
Characterizing drug-outcome associations

- Active surveillance means different things to different people
  - Studying one drug-condition pair at a time
  - Broad-based screening across medical products and outcomes
- OMOP approach is to develop and test methods that may be appropriate anywhere on this continuum

Fundamental task: *Estimate* the strength of the drug-outcome relationship
What do the data look like?

Computational considerations require few passes through the data.
OMOP’s Methods Landscape

### Disproportionality Analysis

<table>
<thead>
<tr>
<th>Drug i = Yes</th>
<th>AE j = Yes</th>
<th>AE j = No</th>
</tr>
</thead>
<tbody>
<tr>
<td>a = 20</td>
<td>b = 100</td>
<td></td>
</tr>
<tr>
<td>c = 100</td>
<td>d = 1080</td>
<td></td>
</tr>
</tbody>
</table>

- Distinct Patients
- SRS
- Modified SRS
- MGPS
- BCPNN
- PRR
- Stratified
- Chi etc.
- Temporal Pattern Discovery (WHO)

### Sequential Methods

<table>
<thead>
<tr>
<th>Drug i = Yes</th>
<th>AE j = Yes</th>
<th>AE j = No</th>
</tr>
</thead>
<tbody>
<tr>
<td>a = 20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[\text{Compare to baseline Poisson}\]

- Maximized Sequential Probability Ratio Test (MaxSPRT)
- Conditional Sequential Sampling Procedure (CSSP)

### Exposure Based Methods

- Exposed → Case?
- Non-exposed → Case?

- Observational screening
- HSIU
- Incident User Designs
- High-Dimensional Propensity Scoring
- Local control

OMOP Methods Library at: http://omop.fnih.org/MethodsLibrary
OMOP’s Methods Landscape

Case Based Methods

Exposed? ← Case ← Exposed?
Exposed? ← Non-case ← Exposed?

- Case control surveillance
- Multiset case control
- Self-controlled case series
- Case crossover

Other Methods

- Hi-Dimensional logistic regression
- Statistical relational learning

Future Methods

- Multivariate self-controlled case series
- Case-time control
- Lasso propensity scoring
- Online algorithms
- OMOP Cup (50+ submissions)

OMOP Methods Library at: http://omop.fnih.org/MethodsLibrary
OMOP Methods Library

- Standardized procedures are being developed to analyze *any* drug and *any* condition
- All programs being made publicly available to promote transparency and consistency in research
- Methods will be evaluated in OMOP research against specific test case drugs and Health Outcomes of Interest

http://omop.fnih.org/MethodsLibrary
Methodological considerations common across multiple approaches

• Exposure definition
  – Incident vs. prevalent exposure
  – Source of data capture

• Outcome definition
  – Incident vs. prevalent events
  – Diagnosis codes vs. HOI

• Defining temporal relationship
  – Time from exposure start
  – Time after exposure end

• Comparator selection

• Inclusion/exclusion criteria
  – Baseline history
  – Follow-up time

• Covariate selection and adjustment
  – Matching
  – Stratification
  – Multivariate modeling

• Output metric/statistic
  – Estimation vs. testing
  – Relative vs. attributable risk
  – Measure of uncertainty

*Each method has user input parameters that encode these choices*
Need for establishing ‘ground truth’ in methodological research

• Methodological research: goal is to measure performance of methods in their ability to identify ‘true’ relationships and discern from false positive findings

• Research requires ‘ground truth’, classification of test cases to evaluate methods against

• Challenge with real data: ‘truth’ is unknown or ill-defined

• Challenge with simulated data: data may not reflect complexities of real-world data
Analysis problems under study by OMOP

- **Monitoring of Health Outcomes of Interest (HOIs):**
  - Estimate the strength of the association between drug exposure and specific events (e.g. acute liver failure, bleeding, MI)
  - Modest in number so can customize analytic approach
  - Expert assessment of drug-HOI causal associations based on literature search

- **Identification of non-specified associations:**
  - More exploratory in nature
  - Same goal: estimate the strength of the association between drug exposure and conditions
  - Necessarily more generic analyses (e.g., adjust for age and sex)
  - Causality assessment relies on the product labels

- **Performance against simulated data**
  - Complement ‘real world’ experiments
  - Ground truth explicitly defined
Reference set for Monitoring Health Outcomes of Interest

Test cases to be used for evaluating method performance for 'Monitoring of Health Outcomes of Interest'

<table>
<thead>
<tr>
<th>Drug</th>
<th>Angioedema</th>
<th>Aplastic Anemia</th>
<th>Acute Liver Injury</th>
<th>Bleeding</th>
<th>GI Ulcer Hospitalization</th>
<th>Hip Fracture</th>
<th>Hospitalization</th>
<th>Myocardial Infarction</th>
<th>Mortality after MI</th>
<th>Renal Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE Inhibitors</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>B</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Amphotericin B</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Antibiotics</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Antiepileptics</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Benzodiazepines</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Beta blockers</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Bisphosphonates</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Tricyclic antidepressants</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>R</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Typical antipsychotics</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Warfarin</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>R</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

Legend

- B - 'True positive' benefit
- R - 'True positive' risk
- N - 'Negative control'

Total

2
9
44

Not selected due to correlation with HOI
# Measuring method performance

Drug-condition status  
Y – ‘true association’,  
N – ‘negative control’

<table>
<thead>
<tr>
<th>Y</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>True positives</td>
<td>False positives</td>
</tr>
<tr>
<td>False negatives</td>
<td>True negatives</td>
</tr>
</tbody>
</table>

**Positive predictive value**  
\[ \text{Precision} = \frac{TP}{TP+FP} \]

**Negative predictive value**  
\[ \text{Specificity} = \frac{TN}{FN+TN} \]

**Sensitivity**  
\[ \text{Recall} = \frac{TP}{TP+FN} \]

**Specificity**  
\[ \text{Specificity} = \frac{TN}{FP+TN} \]

**Method prediction:**  
Drug-condition pair met a defined threshold

Page 17
Accuracy measures from different perspectives

- **Mean average precision (MAP)** – on average, what proportion of predictions are true at different thresholds?
- **Area under ROC curve (AUC)** – what is the composite tradeoff between sensitivity and specificity at all possible threshold values?
- **p@k** - among the top k (e.g. 100) scores, what percentage are true?
- **recall@FPR** - if we can tolerate a particular false positive rate (e.g. 5%), what fraction of the true positives will we identify?
- **Average false positive rate** – what is the average false positive rate observed for each of the true positives?
## Mean Average Precision

<table>
<thead>
<tr>
<th>Drug</th>
<th>Condition</th>
<th>Original Values</th>
<th>Sorted Values</th>
<th>( P^{(K)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( z_i )</td>
<td>( y_i )</td>
<td>( z_{(i)} )</td>
<td>( y_{(i)} )</td>
</tr>
<tr>
<td>D1</td>
<td>C1</td>
<td>5</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>0</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>8</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>D2</td>
<td>C1</td>
<td>8</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D3</td>
<td>C1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Total Score**

\[
\frac{(1+1+0.75+0.8+0.55)}{5} = 0.82
\]

E.g. also consider, for example, \( P^{(100)} \), “Precision at 100”
Threshold-free Methods

Gives substantial weight to unrealistically high false positive rates
Visualizing performance of alternative methods across a network of databases

*Data is for illustrative purposes only*
OMOP Cup: Methods Competition

- **Simulated data**
  - **Observation Period**
  - **Persons**
  - **Drug Exposure**
  - **Condition Occurrence**

10 years of observations
10m persons
4000 drugs
5000 conditions
<5% of drug-condition pairs are true associations

- **Validation**
  - Drug ID
  - Condition ID
  - Relationship

- **OSIM**

- **Methods Submissions**
  - Description
  - Source code
  - Results dataset

- **Prediction**
  - Drug ID
  - Condition ID
  - Relationship

- **TP**
- **FP**
- **FN**
- **TN**

- **Summary**
  - Method
  - Challenge
  - MAP Score

- **Two competitions: http://omopcup.orwik.com**
  - Challenge 1: Identifying drug-condition associations within an entire observational dataset
  - Challenge 2: Identifying drug-condition associations as data accumulates over time

- **Evaluation criteria: Weighted Mean Average Precision**

- **Winning entries were given cash prize and methods and invited to be further tested against OMOP data community**
OMOP Cup: Interactive Website

Welcome to the 2009/2010 Cup!

Organized by the Observational Medical Outcomes Partnership

Rules – Background – Methods Problem – Data – Challenges – Register

The Cup Grand Prize competition has ended. Stay tuned for official results!

Ensuring the safety of prescription drugs has emerged as a major global public health challenge. Several recent events have highlighted the need for new data sources and new algorithms to respond to the challenge. The new data sources have now become available but algorithmic progress has lagged. Here is your chance to develop a groundbreaking approach that can help protect the lives of millions of patients worldwide.
## Challenge 1 Leaderboard

<table>
<thead>
<tr>
<th>User</th>
<th>Best Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martijn J Schuemie</td>
<td>0.2662359</td>
</tr>
<tr>
<td>David S Vogel</td>
<td>0.2570616</td>
</tr>
<tr>
<td>Hawkeye DORP</td>
<td>0.2569417</td>
</tr>
<tr>
<td>Mohammad Khoshneshin</td>
<td>0.2569404</td>
</tr>
<tr>
<td>Nick Street</td>
<td>0.2568678</td>
</tr>
<tr>
<td>Craig G Carmichael</td>
<td>0.2483813</td>
</tr>
<tr>
<td>Harris T Lin</td>
<td>0.2483137</td>
</tr>
<tr>
<td>girishkumar ramesh sabhnani</td>
<td>0.2358521</td>
</tr>
<tr>
<td>Liang Xiong</td>
<td>0.2317307</td>
</tr>
<tr>
<td>Christophe G Giraud-Carrier</td>
<td>0.2310854</td>
</tr>
<tr>
<td>Vladimir N Nikulin</td>
<td>0.2309664</td>
</tr>
<tr>
<td>Bin Liu</td>
<td>0.2303903</td>
</tr>
<tr>
<td>Andrew J Zitzelberger</td>
<td>0.2301842</td>
</tr>
<tr>
<td>Derrall Heath</td>
<td>0.2301842</td>
</tr>
<tr>
<td>Nathaniel Gustafson</td>
<td>0.2298511</td>
</tr>
<tr>
<td>Sam Ogden</td>
<td>0.2298361</td>
</tr>
<tr>
<td>David L Wilcox</td>
<td>0.2293719</td>
</tr>
<tr>
<td>Rob Smith</td>
<td>0.2292806</td>
</tr>
<tr>
<td>Michael T Roscheck</td>
<td>0.2263714</td>
</tr>
<tr>
<td>Yisong Guo</td>
<td>0.2262477</td>
</tr>
<tr>
<td>Robin Sabhnani</td>
<td>0.2256460</td>
</tr>
<tr>
<td>Benchmark (BLR)</td>
<td>0.2244814</td>
</tr>
<tr>
<td>Benchmark (BCPNN-M)</td>
<td>0.2241564</td>
</tr>
</tbody>
</table>
Challenge 2 : Identifying associations over time

<table>
<thead>
<tr>
<th>User</th>
<th>Best Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vladimir N Nikulin</td>
<td>0.2376259</td>
</tr>
<tr>
<td>Martijn J Schuemie</td>
<td>0.2132523</td>
</tr>
<tr>
<td>Hawkeye DORP</td>
<td>0.2032173</td>
</tr>
<tr>
<td>David S Vogel</td>
<td>0.2000853</td>
</tr>
<tr>
<td>Harris T Lin</td>
<td>0.1949657</td>
</tr>
<tr>
<td>Peng Liu</td>
<td>0.1859670</td>
</tr>
<tr>
<td>Ed Ramsden</td>
<td>0.1719514</td>
</tr>
<tr>
<td>Robin Sabhnani</td>
<td>0.1037598</td>
</tr>
<tr>
<td>Lisa D Friedland</td>
<td>0.1032716</td>
</tr>
<tr>
<td>Benchmark (Random)</td>
<td>0.0156415</td>
</tr>
</tbody>
</table>
OMOP Cup: Top Four Performers

- Binary prediction model
- Drug-condition pair is unit of analysis
- Construct a feature vector per pair
- Random forest

- Two stage disproportionality (IC)
  - Stage 1: find signals
  - Stage 2: remove signals confounded by indication

- Ensemble of two disproportionality methods plus Poisson method using exposure times

- Disproportionality
- Ensembling
- Matrix factorization
- Calibration
OMOP Cup: Summary

• Two-fold purpose: Build community and generate new ideas
• 60+ competitors with over 600 submissions
• Community-building requires continuity so we need to plan for OMOP Cup 2
• Some methods did beat benchmarks, but no method dramatically outperformed current approaches
• Improving simulation capability may provide opportunities for further methods innovation
OMOP Current Status

• OMOP data community established to conduct methodological research
• Open-source methods implementation largely complete
• Methods evaluation underway
• OMOP Cup provides another avenue to stimulate methods innovation...watch out for OMOP Cup 2!
Contact Information

Patrick Ryan
Research Investigator
Patrick.b.ryan@gsk.com

David Madigan
OMOP Methods Lead
madigan@stat.columbia.edu

Thomas Scarnecchia
Executive Director
tscarnecchia@fnih.org

Emily Welebob
Senior Program Manager, Research
ewlebob@fnih.org

OMOP website: http://omop.fnih.org
OMOP Cup website: http://omopcup.orwik.com
Roundtable Discussion and Questions

View this and past Active Medical Product Surveillance webinars at:
http://www.brookings.edu/health/Events.aspx