Observational Health Data Sciences and Informatics (OHDSI) Overview

Patrick Ryan, PhD
Janssen Research and Development
on behalf of OHDSI collaborative
15 May 2014
FDAAA call for establishing the Risk Identification and Analysis System

SEC. 905. ACTIVE POSTMARKET RISK IDENTIFICATION AND ANALYSIS.

(a) IN GENERAL.—Subsection (k) of section 505 of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 355) is amended by adding at the end the following:

“(3) ACTIVE POSTMARKET RISK IDENTIFICATION.—

(A) DEFINITION.—In this paragraph, the term ‘data’ refers to information with respect to a drug approved under this section or under section 351 of the Public Health Service Act, including claims data, patient survey data, standardized analytic files that allow for the pooling and analysis of data from disparate data environments, and any other data deemed appropriate by the Secretary.

(B) DEVELOPMENT OF POSTMARKET RISK IDENTIFICATION AND ANALYSIS METHODS.—The Secretary shall, not later than 2 years after the date of the enactment of the Food and Drug Administration Amendments Act of 2007, in collaboration with public, academic, and private entities—

(i) develop methods to obtain access to disparate data sources including the data sources specified in subparagraph (C);

(ii) develop validated methods for the establishment of a postmarket risk identification and analysis system to link and analyze safety data from multiple sources, with the goals of including, in aggregate—

(I) at least 25,000,000 patients by July 1, 2010; and

(II) at least 100,000,000 patients by July 1, 2012; and

(iii) convene a committee of experts, including individuals who are recognized in the field of protecting data privacy and security, to make recommendations to the Secretary on the development of tools and methods for the ethical and scientific uses for, and communication of, postmarketing data specified under subparagraph (C), including recommendations on the development of effective research methods for the study of drug safety questions.

(C) ESTABLISHMENT OF THE POSTMARKET RISK IDENTIFICATION AND ANALYSIS SYSTEM.—

Risk Identification and Analysis System:

a systematic and reproducible process to efficiently generate evidence to support the characterization of the potential effects of medical products from across a network of disparate observational healthcare data sources
Risk identification and analysis system: One additional piece of evidence to inform medical decision-making

- Pre-clinical toxicology
- Pharmacology
- Clinical trials
- Spontaneous case reports
- Perspectives in literature from medical experts
- Pharmacoepidemiology evaluation studies
- Risk identification and analysis system

Evidence to support safety assessment
- Evidence about the benefits of the product
- Evidence about alternative treatments

Decision-making about appropriate use
141 patients exposed in pivotal study for metformin
>10,000 patients exposed across canagliflozin clinical development program
>1,000,000 new users of metformin in one administrative claims database
Patient profiles from observational data
What is the quality of the current evidence from observational analyses?

August 2010: “Among patients in the UK General Practice Research Database, the use of oral bisphosphonates was not significantly associated with incident esophageal or gastric cancer.”

Sept 2010: “In this large nested case-control study within a UK cohort [General Practice Research Database], we found a significantly increased risk of oesophageal cancer in people with previous prescriptions for oral bisphosphonates.”
What is the quality of the current evidence from observational analyses?

April 2012: “Patients taking oral fluoroquinolones were at a higher risk of developing a retinal detachment”

Dec 2013: “Oral fluoroquinolone use was not associated with increased risk of retinal detachment”
What is the quality of the current evidence from observational analyses?

BJCP May 2012: “In this study population, pioglitazone does not appear to be significantly associated with an increased risk of bladder cancer in patients with type 2 diabetes.”

BMJ May 2012: “The use of pioglitazone is associated with an increased risk of incident bladder cancer among people with type 2 diabetes.”
What is the quality of the current evidence from observational analyses?

Nov 2012: FDA released risk communication about the bleeding risk of dabigatran, based on unadjusted cohort analysis performed within Mini-Sentinel.

Dec 2013: “This analysis shows that the RCTs and Mini-Sentinel Program show completely opposite results”

Aug 2013: “However, the absence of any adjustment for possible confounding and the paucity of actual data made the analysis unsuitable for informing the care of patients”
2010-2011 OMOP Research Experiment

- Open-source
- Standards-based

Common Data Model

OMOP Methods Library
- Inception cohort
- Case control
- Logistic regression

- 10 data sources
- Claims and EHRs
- 200M+ lives

Drug

Outcome

<table>
<thead>
<tr>
<th></th>
<th>ACE inhibitors</th>
<th>Amphotericin B</th>
<th>Antibiotics: erythromycin, sulfonamides, tetracyclines</th>
<th>Antiepileptics: carbamazepine, phenytoin</th>
<th>Benzodiazepines</th>
<th>Beta blockers</th>
<th>Bisphosphonates: alendronate</th>
<th>Thiazide diuretics</th>
<th>Typical antipsychotics</th>
<th>Warfarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angioedema</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aplastic Anemia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute Liver Injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bleeding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip Fracture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality after MI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal Failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI Ulcer Hospitalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
OMOP 2011/2012 Research Agenda

Drug-outcome pairs

<table>
<thead>
<tr>
<th>Drug-outcome pairs</th>
<th>Positives</th>
<th>Negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>165</td>
<td>234</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>36</td>
<td>66</td>
</tr>
<tr>
<td>Upper GI Bleed</td>
<td>24</td>
<td>67</td>
</tr>
<tr>
<td>Acute Liver Injury</td>
<td>81</td>
<td>37</td>
</tr>
<tr>
<td>Acute Renal Failure</td>
<td>24</td>
<td>64</td>
</tr>
</tbody>
</table>

+ EU-ADR replication

- Improve HOI definitions
- Explore false positives

Methods development

- Evaluate study design decisions (EDDIE)
- Methods enhancements
  - Multivariate self-controlled case series
  - Increased parameterization
  - Case-control, new user cohort designs
  - Application of existing tools
  - ICTPD, OS, LGPS, DP

- Expand CDM for additional use cases

Observational data

Real-world performance:

- Thomson MarketScan
- GE

+ OMOP Distributed Partners
+ EU-ADR network

Simulated data:

- OSIM2
  - Strength (RR)
  - Type (timing)

All results and presentations available at:

http://omop.org
Takeaways from research toward risk identification

- **Performance of different methods**
  - Self-controlled designs appear to consistently perform well.

- **Evaluating alternative HOI definitions**
  - Broader definitions have better coverage and comparable performance to more specific definitions.

- **Performance across different signal sizes**
  - A risk identification system should confidently discriminate positive effects with RR>2 from negative controls.

- **Data source heterogeneity**
  - Substantial variation in estimates across sources suggest replication has value but may result in conflicting results.

- **Method parameter sensitivity**
  - Each method has parameters that are expected to be more sensitive than others, but all parameters can substantially shift some drug-outcome estimates.

---

**Journal References**

- Ryan PB et al, Statistics in Medicine, 2012: “Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational Medical Outcomes Partnership”
- Madigan D et al, American Journal of Epidemiology, 2013: “Evaluating the Impact of Database Heterogeneity on Observational Study Results”
- Madigan D et al, Therapeutic Advances in Drug Safety, 2013: “Does design matter? Systematic evaluation of the impact of analytical choices on effect estimates in observational studies”
Why large-scale analysis is needed in healthcare

All health outcomes of interest
Introducing OHDSI

• The Observational Health Data Sciences and Informatics (OHDSI) program is a multi-stakeholder, interdisciplinary collaborative to create open-source solutions that bring out the value of observational health data through large-scale analytics

• OHDSI has established an international network of researchers and observational health databases with a central coordinating center housed at Columbia University

http://ohdsi.org
OHDSI’s mission

To transform medical decision making by creating reliable scientific evidence about disease natural history, healthcare delivery, and the effects of medical interventions through large-scale analysis of observational health databases for population-level estimation and patient-level predictions.

http://ohdsi.org
OHDSI’s vision

OHDSI collaborators access a network of 1 billion patients to generate evidence about all aspects of healthcare. Patients and clinicians and other decision-makers around the world use OHDSI tools and evidence every day.

http://ohdsi.org
OHDSI’s objectives

- To establish a research community for observational health data sciences that enables active engagement across multiple disciplines (including statistics, computer science, epidemiology, physics, informatics, and biomedical sciences) spanning multiple stakeholder groups (including academia, healthcare providers, payers, medical product manufacturers, and government agencies)
- To develop and evaluate analytical methods that use observational health data to study the effects of medical interventions and predict health outcomes for patients, and to generate the empirical evidence base necessary to establish best practices in observational analysis
- To apply scientific best practices in the design and implementation of open-source systems for observational analysis to enable medical product risk identification, comparative effectiveness research, patient-level predictions, and healthcare improvement
- To generate evidence about disease natural history, healthcare delivery, and the effects of medical interventions, supporting medical decision making in a way that is credible, consistent, transparent, and personalized to the patient’s and provider’s specific needs
- To establish educational opportunities to train students, practitioners, and consumers about the foundational science of observational health data analysis
OHDSI’s guiding principles

- **Evidence-based**: OHDSI’s scientific research and development will be driven by objective, empirical evidence to ensure accuracy and reliability in everything we do.
- **Practical**: OHDSI will go beyond methodological research, developing applied solutions and generating clinical evidence.
- **Comprehensive**: OHDSI aims to generate reliable scientific evidence for all interventions and all outcomes.
- **Transparent**: All work products within OHDSI will be open source and publicly available, including source code, analysis results, and other evidence generated in all our activities. Best practices for large-scale open source collaboration will guide development activities.
- **Inclusive**: OHDSI encourages active participation from all stakeholders – patients, providers, payers, government, industry, academia – in all phases of research and development.
- **Secure**: OHDSI will protect patient privacy and respect data holder interests at all times in our work.
OHDSI’s path forward: systematic evaluation and calibration
To go forward, we must go back

“What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?”

- Strength
- Consistency
- Temporality
- Plausibility
- Experiment
- Coherence
- Biological gradient
- Specificity
- Analogy

http://omop.org/2013Symposium

Introducing HOMER

- Health Outcomes and Medical Effectiveness Research (HOMER) system

- Live, interactive evidence exploration system with fully functional implementations of all of the components of Sir Bradford Hill’s viewpoints for risk identification and assessment, plus some additional components designed by the OMOP team

http://omop.org/2013Symposium
HOMER implementation of Hill’s viewpoints

http://omop.org/2013Symposium
Join the journey

Interested in OHDSI?
Questions or comments?

Contact:

ryan@ohdsi.org
OMOP Common Data Model
v5

Patrick Ryan, PhD
Janssen Research and Development
Christian Reich, MD PhD
AstraZeneca
on behalf of OHDSI team
15 May 2014
OMOP CDM v5: what’s new?
OMOP CDM v5: what are the key domains?

- Person
  - Observation_period
  - Person_relationship
  - Visit_occurrence
  - Procedure_occurrence
  - Drug_exposure
  - Device_exposure
  - Condition_occurrence
  - Note
  - Observation
  - Lab_result
  - Death

- Standardized health system data
  - Location
  - Care_site
  - Care_site_relationship
  - Provider
  - Payer_plan_period
  - Visit_cost
  - Procedure_cost
  - Drug_cost
  - Device_cost
  - Standardized health economics
  - Observation_relationship
  - Lab_result_relationship

- CDM_source
  - Concept
  - Vocabulary
  - Concept_relationship
  - Relationship
  - Concept_synonym
  - Concept_ancestor
  - Source_to_concept_map
  - Drug_strength
  - Cohort_definition

- Standardized vocabulary
  - Standardized derived elements
    - Condition_era
    - Cohort
    - Drug_era
    - Dose_era
OMOP CDMv5 specification next steps

• Please review two draft documents open for public comment at http://omop.org/CDM:
  – OMOP CDMv5 specification
  – OMOP CDMv4 to CDMv5 conversion

• Email all comments to:
  – Patrick Ryan, ryan@omop.org
  – Christian Reich, reich@omop.org

• All comments due by 30 June 2014