US Initiatives
OMOP and Sentinel

Tobias Gerhard, BSPharm, PhD
Ernest Mario School of Pharmacy
Rutgers University
Disclaimer

• I am a Mini-Sentinel investigator and an OHDSI collaborator.

• I do not speak on behalf of any of the organizations involved.

• Many of the slides were graciously provided by Richard Platt (Mini-Sentinel) and Patrick Ryan (OMOP/OHDSI).
Overview and Timeline

• **FDA Amendments Act (2007)**
  – Mandate to perform active surveillance of the safety of approved drugs through use of routinely collected electronic health information from the care of at least 100 million people.

• **FDA’s Sentinel Initiative (2008-ongoing)**
  – Development and implementation of a proactive system that will complement existing systems that the Agency has in place to track reports of adverse events linked to the use of its regulated products.

• **Mini-Sentinel (2009-ongoing)**
  – Pilot program charged with developing the framework, data resources, analytic capabilities, policies and procedures to satisfy the FDAA mandate.

• **Observational Medical Outcomes Partnership (2008-2013)**
  – Public-private partnership (PhRMA, FDA, FNIH) established to inform the appropriate use of observational healthcare databases for studying the effects of medical products.
THE OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP (OMOP)
OMOP Objectives

• Conduct methodological research to empirically evaluate the performance of various analytical methods on their ability to identify true associations and avoid false findings.

• Develop tools and capabilities for transforming, characterizing, and analyzing disparate data sources across the health care delivery spectrum.

• Establish a shared resource so that the broader research community can collaboratively advance the science.
OMOP Objectives

• Conduct methodological research to empirically evaluate the performance of various analytical methods on their ability to identify true associations and avoid false findings.

  → OMOP Research Experiments

• Develop tools and capabilities for transforming, characterizing, and analyzing disparate data sources across the health care delivery spectrum.

  → OMOP Common Data Model and Vocabulary, etc

• Establish a shared resource so that the broader research community can collaboratively advance the science.

  → OMOP Research Laboratory
2010-2013 OMOP Research Experiments

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

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

Common Data Model

- Open-source
- Standards-based

Drug

<table>
<thead>
<tr>
<th>Outcome</th>
<th>ACE Inhibitors</th>
<th>Amphotericin B</th>
<th>Antibiotics: erythromycins, sulfonamides, tetracyclines</th>
<th>Antiepileptics: carbamazepine, phenytoin</th>
<th>Beta blockers</th>
<th>Bisphosphonates, alendronate</th>
<th>Tricyclic antidepressants</th>
<th>Typical antipsychotics</th>
<th>Warfarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angioedema</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aplastic Anemia</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute Liver Injury</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bleeding</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip Fracture</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalization</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality after MI</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal Failure</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI Ulcer Hospitalization</td>
<td>Red</td>
<td>Blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend
- True positive' benefit
- True positive' risk
- Negative control

| Total | 2 | 9 | 44 |
OMOP Finding 1: Database Heterogeneity
Holding analysis constant, different data may yield different estimates

• When applying a propensity score adjusted new user cohort design to 10 databases for 53 drug-outcome pairs:
  • 43% had substantial heterogeneity ($I^2 > 75\%$) where pooling would not be advisable
  • 21% of pairs had at least 1 source with significant positive effect and at least 1 source with significant negative effect

“Evaluating the Impact of Database Heterogeneity on Observational Study Results”
OMOP Finding 2: Parameter Sensitivity
Holding data constant, different analytic design choices may yield different estimates

Holding all parameters constant, except:
• Matching on age, sex and visit (within 30d) yields:
  \[ RR = 0.73 \, (0.65 - 0.81) \]
• Controls per case: up to 10 controls per case
• Required observation time prior to outcome: 180d
• Time-at-risk: 30d from exposure start
• Include index date in time-at-risk: No
• Case-control matching strategy: Age and sex
• Nesting within indicated population: No
• Exposures to include: First occurrence
• Metric: Odds ratio with Mantel Haenszel adjustment by age and gender

Sertaline-GI Bleed: \[ RR = 2.45 \, (2.06 - 2.92) \]

Madigan D, Ryan PB, Scheumie MJ, Therapeutic Advances in Drug Safety, 2013: “Does design matter? Systematic evaluation of the impact of analytical choices on effect estimates in observational studies”
OMOP Finding 3: Empirical Performance
Most observational methods do not have nominal statistical operating characteristics

- Applying the cohort design to MDCR against 34 negative controls for acute liver injury:
  - If 95% confidence interval was properly calibrated, then $95\% \times 34 = 32$ of the estimates should cover $RR = 1$
  - We observed 17 of negative controls did cover $RR=1$
  - Estimated coverage probability $= \frac{17}{34} = 50\%$
  - Estimates on both sides of null suggest high variability in the bias

OMOP Finding 4: Empirical calibration can help restore interpretation of study findings

- Negative controls can be used to estimate empirical null distribution: how much bias and variance exists when no effect should be observed.
- Empirical null can replace theoretical null to estimate calibrated p-value to test for statistical significance.

Drug Safety

Studying the Science of Observational Research: Empirical Findings from the Observational Medical Outcomes Partnership

Guest Editor
Stephen J. W. Evans
Professor of Pharmacoepidemiology, London School of Hygiene and Tropical Medicine, London, UK

Peer Reviewer
Olaf H. Klungel
Associate Professor, Utrecht Institute for Pharmaceutical Sciences, Utrecht, The Netherlands

This supplement was sponsored by the Foundation for the National Institutes of Health and the Reagan-Udall Foundation for the Food and Drug Administration.
Current Status

• **IMEDS:** The OMOP Research Lab has transitioned to the IMEDS (Innovation in Medical Evidence Development and Surveillance) program of the Reagan Udall Foundation for the FDA in 2013. IMEDS serves to advance the science and tools necessary to support post-market evidence generation on regulated products and to facilitate utilization of a robust secondary electronic healthcare data platform for generating better evidence on regulated products in the post-market settings.
Announcing the 2015 IMEDS-Methods Research Agenda

The 2015 IMEDS-Methods Research Agenda was formulated to build on the successes achieved during 2014, and to reflect the goals and priorities articulated by all stakeholder groups. The 2015 IMEDS-Methods Research Agenda can be found here.

For more information, please review the full announcement here.

Be a Part of the IMEDS Research Lab

Do you want to know about research that is being completed in the lab? Find out about the IMEDS Research Laboratory, how to gain access, and the available datasets and tools to complete your research. Review the IMEDS Research Lab information.

Symposium on Health Care Data Analytics - Sept 28-30, 2014

Join biostatisticians and other scientists Sept. 28-30, 2014, in Seattle, Washington for the the 1st Seattle Symposium on Health Care Data Analytics featuring research on pragmatic clinical trial design, inference and prediction using EHR data, and drug and vaccine safety surveillance. Susan Gruber will be addressing "Gaps and opportunities: Methodologic challenges in post-market safety surveillance." View conference agenda [PDF].
Current Status

• **IMEDS:** The OMOP Research Lab has transitioned to the IMEDS (Innovation in Medical Evidence Development and Surveillance) program of the Reagan Udall Foundation for the FDA in 2013. IMEDS serves to advance the science and tools necessary to support post-market evidence generation on regulated products and to facilitate utilization of a robust secondary electronic healthcare data platform for generating better evidence on regulated products in the post-market settings.

• **OHDSI:** The Observational Health Data Sciences and Informatics (OHDSI) program is a multi-stakeholder, interdisciplinary collaborative that includes all of the members of the OMOP investigator team. Whereas OMOP was restricted to methodological research, OHDSI develops and applies methods to observational data to answer real-world clinical questions.
Welcome to OHDSI!

The Observational Health Data Sciences and Informatics (or OHDSI, pronounced "Odyssey") program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All our solutions are open-source.

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

Read more about us, about our goals, and how you can help support the OHDSI community.

Join the Journey

ACHILLES Released

OHDSI released its first open-source software application, ACHILLES, at the 2014 EDM Forum in San Diego, CA. Congratulations to the ACHILLES

OHDSI on YouTube

Welcome to OHDSI

Latest News

- OHDSI paper published in Drug Safety

http://ohdsi.org
Mini-Sentinel
Mini-Sentinel Partner Organizations

Lead – HPHC Institute

Data and scientific partners

Scientific partners
Post-Market Safety Surveillance

Signal Identification:
Potential safety concern identified

Signal Refinement:
Initial evaluation of safety concerns

Signal Evaluation:
Detailed assessment
Post-Market Safety Surveillance

Signal identification: Potential safety concern identified

Signal Refinement: Initial evaluation of safety concerns

Signal Evaluation: Detailed assessment

Data Mining

Summary Tables Modular Programs

Protocol-based Evaluations

PROMPT = Prospective Routine Observational Monitoring Program Tools
Impact / Dissemination

- 4 FDA drug safety communications
  - Tri-valent inactivated flu vaccine and febrile seizures (no increased risk)
  - RotaTeq, Rotarix and intussusception (label change for RotaTeq, no label change for Rotarix)
  - Dabigatran and bleeding (no increased risk)
  - Olmesartan and sprue-like enteropathy (label change)

- 26 Presentations by FDA
- 48 Methods reports / white papers
- 70 Peer-reviewed articles
- 137 Assessments of products, conditions, product-outcome pairs
Mini-Sentinel Journal Supplement

The U.S. Food and Drug Administration’s Mini-Sentinel Program
Edited by: Richard Platt and Ryan Carnahan
Three major domains

- Data
- Methods
- Active surveillance
Three major domains

- Data
- Methods
- Active surveillance
Mini-Sentinel Distributed Database*

- Populations with well-defined person-time for which most medically-attended events are known
- 178 million members**
- 358 million person-years of observation time
- 48 million people currently accruing new data
- 4 billion dispensings
- 4.1 billion unique encounters
  - 42 million acute inpatient stays
- 30 million people with ≥1 laboratory test result

*As of July 2014
** Double counting exists for individuals who change health plans
Mini-Sentinel’s Common Data Model

**Enrollment**
- Person ID
- Enrollment start & end dates
- Drug coverage
  - Medical coverage

**Demographic**
- Person ID
- Birth date
- Sex
- Race
- Etc.

**Dispensing**
- Person ID
- Dispensing date
- National drug code (NDC)
- Days supply
- Amount dispensed

**Lab Result**
- Person ID
- Dates of order, collection & result
- Test type, immediacy & location
- Provider seen
  - Type of encounter
- Facility
  - Etc.

**Tobacco use & type**
- National drug code (NDC)

**Diagnosis**
- Person ID
- Date
- Principal diagnosis flag
- Source
- Encounter type & provider
- Diagnosis code & type
  - Etc.

**Encounter**
- Person ID
- Dates of service
- Procedure code & type
- Provider seen
  - Type of encounter
- Facility
  - Etc.

**Procedure**
- Person ID
- Dates of service
- Procedure code & type
- Encounter type & provider
  - Etc.

**Vital Signs**
- Person ID
- Date & time of measurement
- Height
- Weight
- Diastolic & systolic BP
- Tobacco use & type
- BP type & position

**Death**
- Person ID
- Date of death
- Source
- Confidence
  - Etc.

**Cause of Death**
- Person ID
- Cause of death
- Diagnosis code & code type
- Source
- Confidence
  - Etc.

Also:
- Vaccine table
- Birth certificate table
- Blood components table

Data Quality Assurance review process

1. Perform Data Update
2. Execute data quality program package
3. Review output; identify and resolve issues
4. Deliver summary output to MSOC
5. Review #1 of data quality output
6. Prepare initial report of findings
7. Review #2 of data quality output
8. Annotate initial report of findings
9. Review and finalize report
10. Review report; resolve issues, respond to MSOC
11. Review Data Partner’s response to report; send additional questions if needed
12. Approve Data Update

Data Analyst

Data Quality Analyst

Data Quality Analyst 1

Data Quality Analyst 2

Data Manager

Data Partner

MSOC
Mini-Sentinel Distributed Analysis

1- User creates and submits query (a computer program)

2- Data partners retrieve query

3- Data partners review and run query against their local data

4- Data partners review results

5- Data partners return results via secure network

6 Results are aggregated
Three major domains

- Data
- **Methods**
- Active surveillance
## Domains of methods development / examples

<table>
<thead>
<tr>
<th>Data Fitness and Capacity</th>
<th>Evaluating Methods</th>
<th>Target Monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Integrity (validity, completeness)</td>
<td>• Validity, power/robustness, time-to-signal detection</td>
<td>• Preparedness Design</td>
</tr>
<tr>
<td>• Environments</td>
<td>• Empirical, simulation</td>
<td>– Systematic selection</td>
</tr>
<tr>
<td>– Claims, EHR, registries</td>
<td>• Heterogeneity across databases</td>
<td>– Self-controlled</td>
</tr>
<tr>
<td>• Anonymous linkage</td>
<td>• In collaboration with IMEDS</td>
<td>– Cohort methods</td>
</tr>
<tr>
<td>• Enriching the CDM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>– Lab results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Data sharing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Data mining (untargeted)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Sample size tools</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signal Generation</th>
<th>Signal Follow-up</th>
<th>Decision Making</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Data mining</td>
<td>• Data/code quality</td>
<td>• Decision analysis framework</td>
</tr>
<tr>
<td>• Sample size tools</td>
<td>• Sensitivity analyses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Timing of signals</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 2-phase designs</td>
<td></td>
</tr>
</tbody>
</table>
Health Outcome and Confounder Libraries

- Need standardized operational definitions for health outcomes and confounding conditions
- Summarize literature sources
- Document definitions used in protocol-based assessments
Taxonomy

Structured decision table to facilitate methods selection for particular active medical product monitoring scenarios

<table>
<thead>
<tr>
<th>Monitoring scenario characteristics with implication for design choice(^a)</th>
<th>Characteristics of the (potential) exposure-HOI link</th>
<th>Monitoring scenario characteristics with implication for analytic choice(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure persistence (transient, sustained)</td>
<td>Onset of exposure risk window (Immediate, delayed)</td>
<td>Duration of exposure risk window (short, long)</td>
</tr>
<tr>
<td>Transient (e.g. vaccine, initiation of a drug; including episodic drug use [e.g., triptans] to the extent that the question pertains to its transient nature)</td>
<td>Immediate</td>
<td>Short</td>
</tr>
<tr>
<td>Negligible Needs to be addressed</td>
<td>Abrupt</td>
<td>Needs to be addressed</td>
</tr>
<tr>
<td>Negligible Needs to be addressed</td>
<td>Abrupt</td>
<td>Needs to be addressed</td>
</tr>
<tr>
<td>Negligible Needs to be addressed</td>
<td>Abrupt</td>
<td>Needs to be addressed</td>
</tr>
</tbody>
</table>

Exposure-outcome scenarios linked to design strategies

Analytic choice

1. Infrequent
2. Infrequent
3. Infrequent
4. Infrequent
5. Infrequent
6. Infrequent
7. Infrequent
8. Infrequent
9. Rare
10. Rare
11. Rare
12. Rare
13. Rare
14. Rare
15. Rare
16. Rare
17. Rare
Tool development steps

1. Review needs & existing tools
   - Example deliverable: Findings from literature review

2. Methods development / enhancement
   - Example deliverable: Proof-of-concept paper / report with simulated data

3. Methods evaluation using known associations
   - Example deliverable: Proof-of-concept paper / report with actual data

4. Prototype development
   - Example deliverable: Analytic code that runs against Sentinel Distributed Data

5. Tool development
   - Example deliverable: Fully QC-ed code, documented, with input forms

6. Tool enhancement
   - Example deliverable: Fully QC-ed code with documentation and input forms
Reusable Rapid Query Tools

Cohort Identification and Descriptive Analysis

Analytic Adjustment

Self Controlled Risk Interval

Cohort matching / stratification

Sequential Analysis and Signaling

Binomial maxSPRT
Maximized Sequential Probability Ratio Testing

General Estimating Equations Regression

Inverse Probability of Treatment Weighting Regression

Group Sequential GEE Signaling

Group Sequential IPTW Signaling
Online First

Comparative Risk for Angioedema Associated With the Use of Drugs That Target the Renin-Angiotensin-Aldosterone System

Sengwee Toh, ScD; Marsha E. Reichman, PhD; Monika Houstoun, PharmD; Mary Ross Southworth, PharmD; Xiao Ding, PhD; Adrian F. Hernandez, MD; Mark Levenson, PhD; Lingling Li, PhD; Carolyn McCloskey, MD, MPH; Azadeh Shoaibi, MS, MHS; Eileen Wu, PharmD; Gwen Zornberg, MD, MS, ScD; Sean Hennessy, PharmD, PhD

- Used data for 3.9 million new users of anti-hypertensives in 18 organizations
- Propensity score matched stratified analysis
- No person-level data was shared
- Five months and $250,000 required for programming and analysis – compared to 1-2 years and $2 million without analysis-ready distributed dataset

Replication of the same study with the new rapid query tools: ACEI vs β-blocker: $HR = 3.1$ (95% CI, 2.9-3.4)

Toh et al findings: $HR = 3.0$ (95% CI, 2.8-3.3)

Time and cost requirements: Weeks and 10s of thousands of $s$
Three major domains

- Data
- Methods
- Active surveillance
Query Fulfillment

- Year 5 Activities
  - 48 Summary Table Requests
  - 63 Modular Program Requests
    - Over 2000 “scenarios”
    - Over 90 reports to FDA

- To Date
  - ~350 Summary Table Requests
  - ~175 Modular Program Requests
Selected Protocol Based Assessments Planned or Under Way

- **CDER**
  - Mirabegron and several outcomes (prospective monitoring)
  - Rivaroxaban and several outcomes (prospective monitoring)
  - Dabigatran and several outcomes
  - Metabolic effects of 2\textsuperscript{nd} generation antipsychotics in youth
  - Diabetes drugs and acute myocardial infarction
  - IV Iron and anaphylaxis

- **CBER**
  - IV Immune Globulin and thromboembolic events
  - Gardasil and venous thromboembolism
  - Influenza vaccines and pregnancy outcomes
  - Gardasil 9 and Pregnancy Outcomes
  - Prevnar 13 and Kawasaki disease
  - Blood components and Transfusion-Related Lung Injury (TRALI)
Current Status

- The program in the process of transitioning from its pilot stage to the full-fledged program
- The Sentinel Contract was awarded to Harvard Pilgrim Health Care Institute (PI: Richard Platt) in October 2014
- Within FDA’s Center for Drug Evaluation and Research (CDER) the Sentinel Program is moving from the Office of Medical Policy to the Office of Surveillance and Epidemiology (OSE)
Plans for Sentinel: New Populations

- Part of Sentinel contract
  - BCBS Massachusetts
  - Hospital Corporation of America
  - PCORnet Clinical Data Research Networks

- Potential Future Populations
  - CMS data (Medicare, Medicaid)
  - Veterans Health Administration
  - Department of Defense
Plans for Sentinel: Methods Priorities

- Data linkage: National death index (NDI)
- Method evaluation: Comprehensive evaluation of Sentinel programs’ operational and statistical performance
- Targeted prospective surveillance (enhancing PROMPT)
  - Historical comparison groups (vaccines, rare outcomes)
  - More flexible survival data estimation/signaling methods
  - Improving sequential design selection processes
  - Prospective temporal scans in self-controlled & cohort designs
- Signal follow-up from prospective surveillance
  - Practical guidance for follow-up of safety signal
  - Electronic claims profile retrieval tool to review HOIs
- Signal generation: extending tree scan data mining
Plans for Sentinel: External Engagements

- Clinical Trials Transformation Initiative
- PCORnet – Nat’l Patient Centered Research Network
- NIH Health Care System Collaboratory
- Reagan Udall Foundation – IMEDS
- ONC Standards & Interoperability Framework (Query Health)
- IOM Roundtable on Value & Science-Driven Health Care
- Academy Health EDM Forum
Welcome to Mini-Sentinel

Mini-Sentinel is a pilot project sponsored by the U.S. Food and Drug Administration (FDA) to create an active surveillance system - the Sentinel System - to monitor the safety of FDA-regulated medical products. Mini-Sentinel uses pre-existing electronic healthcare data from multiple sources. Collaborating institutions provide access to data as well as scientific and organizational expertise. Mini-Sentinel is part of the FDA’s Sentinel Initiative, which is exploring a variety of approaches for improving the Agency’s ability to quickly identify and assess safety issues.

Most Mini-Sentinel activities focus on assessments, methods, or data. Visit the following links to learn more about each type of activity:

- **Assessments** - Medical product exposures, health outcomes, and links between them
- **Methods** - Techniques for identifying, validating, and linking medical product exposures and health outcomes
- **Data** - Mini-Sentinel Distributed Dataset and tools used to access the data

**Spotlight**
- Brookings Seventh Annual Sentinel Initiative Public Workshop (February 5, 2015 from 9am--4pm - registration required)
- Employment Opportunities
- FDA Sentinel Contract Awarded to Harvard Pilgrim Health Care Institute

**Latest Postings**

**Ongoing Projects**
- Decision Analysis for Surveillance and Health - Pandemic influenza (PRISM)
- Quantifying Uncertainty in Protocol Based Approaches (PRISM)
Summary

- “OMOP efforts have drawn important and cautionary attention to issues of design, data quality, and replicability of observational studies” (Psaty et al., *NEJM*, 2014)

- Methods work (including methods evaluation) is ongoing within Sentinel, IMEDS, and OHDSI. “We are not there yet, with the solution to problems of drug safety, but we are moving in the right direction” (Evans, *Drug Saf*, 2013)

- Mini-Sentinel is in the process of transitioning from its pilot stage to the full-fledged program. It is intended to complement the existing FDA post-marketing resources, not to replace existing activities and systems. Sentinel can be expected to continue to evolve and to play an integral role in the future of FDA’s postmarketing activities.
Thank you!

Contact: tgerhard@rci.rutgers.edu