OHDSI

SNOMED-CT Expo
19-October-2018
FDA Regulatory Action over Time

Number of FDA-caused Withdrawals

- 1960ies
- 1970ies
- 1980ies
- 1990ies
- 2000ies
FDAAA calls for establishing 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
OMOP Experiment 1 (2009-2010)

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

Common Data Model

- Open-source
- Standards-based

OMOP Methods Library

- 14 methods
- Epidemiology designs
- Statistical approaches adapted for longitudinal data

Drug

<table>
<thead>
<tr>
<th>Drug</th>
<th>ACE Inhibitors</th>
<th>Amphotericin B</th>
<th>Antibiotics: erythromycins, sulfonamides, tetracyclines</th>
<th>Antiepileptics: carbamazepine, phenytoin</th>
<th>Benodiazepines</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>Red</td>
<td>Red</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Aplastic Anemia</td>
<td>Red</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Acute Liver Injury</td>
<td>Red</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Bleeding</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Hip Fracture</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Hospitalization</td>
<td>Green</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Mortality after MI</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Renal Failure</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>GI Ulcer Hospitalization</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Blue</td>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
</tbody>
</table>
OMOP Experiment 2 (2011-2012)

**Observational Data**
- 4 claims databases
- 1 ambulatory EMR

**Methods**
- Case-Control
- New User Cohort
- Disproportionality methods
- ICTPD
- LGPS
- Self-Controlled Cohort
- SCCS

**Drug-outcome pairs**

<table>
<thead>
<tr>
<th></th>
<th>Positives</th>
<th>Negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></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>
European OMOP Experiment

Methods
- Case-Control
- New User Cohort
- Disproportionality methods
- ICTPD
- LGPS
- Self-Controlled Cohort
- SCCS

Drug-outcome pairs

<table>
<thead>
<tr>
<th></th>
<th>Positives</th>
<th>Negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></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>

Observational Data
- ARS
- IPCI
- HS
- PHARMO
Criteria for positive controls:
- Event listed in Boxed Warning or Warnings/Precautions section of active FDA structured product label
- Drug listed as ‘causative agent’ in Tisdale et al, 2010: Drug-Induced Diseases
- Literature review identified no powered studies with refuting evidence of effect

Criteria for negative controls:
- Event not listed anywhere in any section of active FDA structured product label
- Drug not listed as ‘causative agent’ in Tisdale et al, 2010: Drug-Induced Diseases
- Literature review identified no powered studies with evidence of potential positive association

<table>
<thead>
<tr>
<th></th>
<th>Positive controls</th>
<th>Negative controls</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute Liver Injury</strong></td>
<td>81</td>
<td>37</td>
<td>118</td>
</tr>
<tr>
<td><strong>Acute Myocardial Infarction</strong></td>
<td>36</td>
<td>66</td>
<td>102</td>
</tr>
<tr>
<td><strong>Acute Renal Failure</strong></td>
<td>24</td>
<td>64</td>
<td>88</td>
</tr>
<tr>
<td><strong>Upper Gastrointestinal Bleeding</strong></td>
<td>24</td>
<td>67</td>
<td>91</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>165</td>
<td>234</td>
<td>399</td>
</tr>
</tbody>
</table>
Results
Main findings in OMOP experiment

- Heterogeneity in estimates due to choice of database
- Heterogeneity in estimates due to analysis choices
- Except little heterogeneity due to outcome definitions
- Good performance (AUC > 0.7) in distinguishing positive from negative controls for optimal methods when stratifying by outcome and restricting to powered test cases
- Self controlled methods perform best for all outcomes
Observational Health Data Sciences and Informatics (OHDSI) Plans and Ambitions
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:

- Public, Open
- Not Pharma-funded
- International

http://ohdsi.org
OHDSI’s Mission & Vision

To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.

A world in which observational research produces a comprehensive understanding of health and disease.

Join us on the journey

http://ohdsi.org
OHDSI: a global community

OHDSI Collaborators:
• >220 researchers in academia, industry and government
• >21 countries

OHDSI Data Network:
• >114 databases from 19 countries
• 1.9 billion patients records (duplicates)
• ~222 million non-US patients
Observational Research has a Problem
Oral bisphosphonates and risk of cancer of oesophagus, stomach, and colorectum: case-control analysis within a UK primary care cohort

Jane Green, clinical epidemiologist,\textsuperscript{1} Gabriela Czanner, statistician,\textsuperscript{1} Gillian Reeves, statistical epidemiologist,\textsuperscript{1} Joanna Watson, epidemiologist,\textsuperscript{1} Lesley Wise, manager, Pharmacoepidemiology Research and Intelligence Unit,\textsuperscript{2} Valerie Beral, professor of cancer epidemiology\textsuperscript{1}
Exposure to Oral Bisphosphonates and Risk of Esophageal Cancer

Chris R. Cardwell, PhD
Christian C. Abnet, PhD
Marie M. Cantwell, PhD
Liam J. Murray, MD

Context  Use of oral bisphosphonates has increased dramatically in the United States and elsewhere. Esophagitis is a known adverse effect of bisphosphonate use, and recent reports suggest a link between bisphosphonate use and esophageal cancer, but this has not been robustly investigated.

Objective  To investigate the association between bisphosphonate use and esoph-
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?

BMJ May 2012: “The use of pioglitazone is associated with an increased risk of incident bladder cancer among people with type 2 diabetes.”

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.”
What is the quality of the current evidence from observational analyses?

**Nov2012:** FDA released risk communication about the bleeding risk of dabigatran, based on unadjusted cohort analysis performed within Mini-Sentinel.

**Dec2013:** “This analysis shows that the RCTs and Mini-Sentinel Program show completely opposite results.”

**Aug2013:** “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.”
The current literature is severely biased

In reality, most results are not significant, but they are never published (missing)

85% of published studies are positive

Drug protects

Drug harms

29,982 estimates
11,758 papers
Current pace of evidence generation in healthcare

All health outcomes of interest

| Other Int. | Acid Prep. | Magnesium | Propulsil | Other Bl. | Other LA. | Enzymes | Vitamin D | Intermed. | Second. | Anti-HIV | Rota Virus | Other Al. | Anti-CD3 | Benznida | Benzimidazoles | Iodine | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. | Other Al. |
|------------|------------|-----------|-----------|-----------|-----------|---------|-----------|-----------|----------|---------|-----------|-----------|----------|----------|-----------|----------------|--------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|

All drugs
Current Approach: “One Study – One Script”

"What's the adherence to my drug in the data assets I own?"

Analytical method: Adherence to Drug

Application to data

Current solution:

One SAS or R script for each study

- Not scalable
- Not transparent
- Expensive
- Slow
- Prohibitive to non-expert routine use
Solution: Data Standardization Enables Systematic Research

OHDSI Tools

OMOP CDM

Source of Business

Mortality

Adherence

Safety Signals

North America
Southeast Asia
China
Europe
UK
Japan
India
So Africa
Switzerland
Italy
Israel

Standardized data
Analytics can be remote
Analytics can be behind firewall
Network Studies
Networks of networks

Another Network

Network

Coordinating Center

EMR
ISDN
University Medical Center
Inpatient Hospital
Outpatient Hospital

Claims Asset
EMR Asset
Claims Asset
EMR Asset
Claims Asset
EMR Asset
Claims Asset
EMR Asset
Claims Asset
EMR Asset
Claims Asset
EMR Asset
A) Incentives for the Node

• Enabling data for research
• ~Free Tools, Methods
  – Vocabulary browsing
  – Population characterization
  – Adjudication and validation
  – Population-based estimation
  – Patient-level prediction

• Quality benchmarks
• Scientific reputation
• Potentially money
b) Feeding the Network

- Foundational
  - CDM
  - Vocabulary, Mapping
  - Community
  - Training
- Trust
  - Open Source
  - Nodes keep control over data
- Methodology
- Technology, tools, automation
- Use cases, scientific impact
- Reciprocity, no autocracy
Tutorials

- OMOP CDM and Vocabulary
- Overview of the OHDSI Analysis
- OHDSI Tech Stack
- Data ETL
- Cohort Definition/Phenotyping
- Patient-Level Prediction
- Population-level Effect Estimation
- Data Quality
### Forum, Workgroups

<table>
<thead>
<tr>
<th>Category</th>
<th>Latest</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td>Welcome to OHDSI - Please introduce yourself 8h</td>
</tr>
<tr>
<td></td>
<td>Where can I find OHDSI policies on human subjects issues? new 12h</td>
</tr>
<tr>
<td></td>
<td>OHDSI Community Call 19Jun2018 1d</td>
</tr>
<tr>
<td><strong>Implementers</strong></td>
<td>Cerner ETL Workgroup new 4h</td>
</tr>
<tr>
<td></td>
<td>Tracking Source References when loading Data from Separated by Systems 2d</td>
</tr>
<tr>
<td></td>
<td>EpicUserWeb , Epic ETL documentation scripts 19d</td>
</tr>
<tr>
<td><strong>Developers</strong></td>
<td>Open Source Architecture Meeting Notes Jan’15</td>
</tr>
<tr>
<td></td>
<td>1K sample of simulated CMS SnpPUF data in CDV5 format available for download 2d</td>
</tr>
<tr>
<td></td>
<td>How to push Impala db v5.3 fixed code 2d</td>
</tr>
<tr>
<td><strong>Researchers</strong></td>
<td>Potential for CancerOncology Studies in OHDSI–Need your thoughts! 11h</td>
</tr>
<tr>
<td></td>
<td>Building a validated OHDSI outcome library 12h</td>
</tr>
<tr>
<td></td>
<td>Treatment Pathways: Combination drugs (Posting 1) 22h</td>
</tr>
<tr>
<td><strong>CDM Builders</strong></td>
<td>Multiple race solution5 1d</td>
</tr>
<tr>
<td></td>
<td>Geographical/Year-Based Multiplier 2d</td>
</tr>
<tr>
<td></td>
<td>Care sites and specialty, specialty code clean-up 3d</td>
</tr>
<tr>
<td><strong>Vocabulary Users</strong></td>
<td>The concepts in CONCEPT table is defined by OHDSI itself? 5h</td>
</tr>
<tr>
<td></td>
<td>CPT4 Place of Service Codes 5h</td>
</tr>
<tr>
<td></td>
<td>Sunscreen - Device or Drug new 7h</td>
</tr>
<tr>
<td><strong>Collaborators</strong></td>
<td>Sharing Research Results - Authentication Details Feb’15</td>
</tr>
<tr>
<td><strong>Leadership</strong></td>
<td>Tomorrow’s leadership call 15June2018 at 11:30am ET 7d</td>
</tr>
<tr>
<td></td>
<td>Tomorrow’s leadership call 15June2018 at 11:30am ET 14d</td>
</tr>
<tr>
<td></td>
<td>Tomorrow’s leadership call 15June2018 at 11:30am ET 19d</td>
</tr>
</tbody>
</table>
Atlas, Achilles, Athena
ARACHNE Research Collaboration Network and Workflow Suite

Create study write up
Study Notebook

Build a research team
Expert Finder

Find relevant patient data
Clinical Data Catalog

Perform data analysis
Analysis Execution

Create research paper
Study Publisher

Publish into Insights Library

Knowledge

Information

Traditional approach

ARACHNE Approach

Insights & Decisions

Data

Effort = Cost

CDM

Insights & Decisions

CDM

CDM

CDM

CDM

CDM

CDM
Common Data Model

- 10 data sources

OMOP Extended Consortium
- OMOP Research Core
- Distributed partners

OMOP Research Lab

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

Common Data Model

- 14 methods

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

Data sources and methods

34
CDM Version 6 Key Domains

Standardized clinical data:
- Person
  - Observation_period
  - Visit_occurrence
  - Visit_detail
  - Condition_occurrence
  - Drug_exposure
  - Procedure_occurrence
  - Device_exposure
  - Measurement
  - Note
  - Note_NLP
  - Survey_conduct
  - Observation
  - Specimen
  - Fact_relationship

Standardized health system data:
- Location
  - Location_history
  - Care_site
  - Provider

Standardized derived elements:
- Condition_era
- Drug_era
- Dose_era

Results Schema:
- Cohort
  - Cohort_definition

Standardized health economics:
- Cost
  - Payer_plan_period

Standardized metadata:
- CDM_source
- Metadata

Standardized vocabularies:
- Concept
- Vocabulary
- Domain
- Concept_class
- Concept_relationship
- Relationship
- Concept_synonym
- Concept_ancestor
- Source_to_concept_map
- Drug_strength
Structure of OMOP Vocabulary

All content: concepts in `concept`

Direct relationships between concepts in `concept_relationship`

Multi-step hierarchical relationships pre-processed into `concept_ancestor`
Single Concept Reference Table

All vocabularies stacked up in one table

Vocabulary ID
Dozens of schemes, formats, rules

### LOINC_248_MULTI-AXIAL_HIERARCHY.CSV

<table>
<thead>
<tr>
<th>PATH_TO_ROOT</th>
<th>SEQUENC IMMEDIATE_PARENT</th>
<th>CODE</th>
<th>CODE_TEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP31755-9</td>
<td></td>
<td>LP31755-9</td>
<td>Microbiology</td>
</tr>
<tr>
<td>LP31755-9</td>
<td>1 LP31755-9</td>
<td>LP14559-6</td>
<td>Microorganism</td>
</tr>
<tr>
<td>LP14559-6</td>
<td>1 LP98185-9</td>
<td>LP98185-9</td>
<td>Bacteria</td>
</tr>
<tr>
<td>LP98185-9</td>
<td></td>
<td>LP14082-9</td>
<td>Bacteria</td>
</tr>
</tbody>
</table>

### loinc.csv

<table>
<thead>
<tr>
<th>LOINC_NL_COMPONENT</th>
<th>PROPERTY</th>
<th>TIME ASPCT</th>
<th>SYSTEM</th>
<th>SCALE_TYP</th>
<th>METHOD_TYP</th>
<th>CLASS</th>
<th>SOURCE</th>
<th>DATE_LAST_CHANGED</th>
<th>TC/COMM</th>
<th>DATE_LAST_CHANGED</th>
<th>TC/COMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>10454-7</td>
<td>Xylose*2H post 25 g xylose PO</td>
<td>MCnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>SH</td>
<td>19961220 ADD</td>
<td>PATH</td>
<td>SH/TL-M</td>
<td>PATH</td>
</tr>
<tr>
<td>10455-4</td>
<td>Xylose*30M post 25 g xylose PO</td>
<td>MCnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>SH</td>
<td>19961220 ADD</td>
<td>PATH</td>
<td>SH/TL-M</td>
<td>PATH</td>
</tr>
<tr>
<td>10456-2</td>
<td>Xylose*post 6H Grst</td>
<td>MCnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>SH</td>
<td>19961220 ADD</td>
<td>PATH</td>
<td>SH/TL-M</td>
<td>PATH</td>
</tr>
<tr>
<td>10457-0</td>
<td>Actin Ag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10458-8</td>
<td>Alkaline phosphatase,place</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10459-6</td>
<td>Alpha-1-Fetoprotein Ag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10460-4</td>
<td>Lactalbumin alpha Ag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10461-2</td>
<td>Alpha-1-Antichymotrypsin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10462-0</td>
<td>Alpha 1 antitrypsin Ag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10463-8</td>
<td>Amyloid A component Ag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10464-5</td>
<td>Amyloid P component Ag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10465-3</td>
<td>Amyloid:prealbumin Ag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10466-1</td>
<td>Anion gap 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### CMS32_DESC_LONG_SHORT_DX.xlsx

<table>
<thead>
<tr>
<th>DIAGNOSIS_CODE</th>
<th>LONG DESCRIPTION</th>
<th>SHORT DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>0010</td>
<td>Cholera due to vibrio cholerae</td>
<td>Cholera d/t vib cholerae</td>
</tr>
<tr>
<td>0011</td>
<td>Cholera due to vibrio cholerae el tor</td>
<td>Cholera d/t vib el tor</td>
</tr>
<tr>
<td>0019</td>
<td>Cholera, unspecified</td>
<td>Cholera NOS</td>
</tr>
<tr>
<td>0020</td>
<td>Typhoid fever</td>
<td>Typhoid fever</td>
</tr>
<tr>
<td>0021</td>
<td>Paratyphoid fever A</td>
<td>Paratyphoid fever a</td>
</tr>
<tr>
<td>0022</td>
<td>Paratyphoid fever B</td>
<td>Paratyphoid fever b</td>
</tr>
<tr>
<td>0023</td>
<td>Paratyphoid fever C</td>
<td>Paratyphoid fever c</td>
</tr>
<tr>
<td>0029</td>
<td>Paratyphoid fever, unspecified</td>
<td>Paratyphoid fever NOS</td>
</tr>
<tr>
<td>0030</td>
<td>Salmonella gastroenteritis</td>
<td>Salmonella enteritis</td>
</tr>
<tr>
<td>0031</td>
<td>Salmonella septicemia</td>
<td>Salmonella septicemia</td>
</tr>
<tr>
<td>00320</td>
<td>Localized salmonella infection, unspecified</td>
<td>Local salmonella inf NOS</td>
</tr>
<tr>
<td>00321</td>
<td>Salmonella meningitis</td>
<td>Salmonella meningitis</td>
</tr>
<tr>
<td>00322</td>
<td>Salmonella pneumonia</td>
<td>Salmonella pneumonia</td>
</tr>
<tr>
<td>00323</td>
<td>Salmonella arthritis</td>
<td>Salmonella arthritis</td>
</tr>
<tr>
<td>00324</td>
<td>Salmonella osteomyelitis</td>
<td>Salmonella osteomyelitis</td>
</tr>
<tr>
<td>00329</td>
<td>Other localized salmonella infections</td>
<td>Local salmonella inf NEC</td>
</tr>
</tbody>
</table>
Vocabulary Goals

✓ **Domains:** Every Standard Concept belongs to the right Domain

• **No duplicates:** For every entity exists one Standard Concept

• **Comprehensive:** For every Domain exists a complete finite set of Concepts covering all possible entities in this domain

• **Hierarchy:** All Concepts are connected through a comprehensive hierarchy

• **Mapping:** For every existing code in a vocabulary there is a map to a Standard Concept or a map to 0
Condition Concepts

Source Concepts

Classification Concepts

Standard Concepts

Top-level classification

Higher-level classifications

Low-level concepts

Source codes

ICD10  ICD10CM  Read  SNOMED  Oxmis  Ciel  MeSH  ICD9CM

MedDRA

System organ class

High-level group terms

High-level terms

Preferred terms

Low-level terms

SNOMED-CT

SNOMED-CT

SNOMED-CT

Condition Concepts

SNOMED-CT

ICD10

ICD10CM

Read

SNOMED

Oxmis

Ciel

MeSH

ICD9CM
Why are we mapping?

Official languages of the EU

<table>
<thead>
<tr>
<th>What is it?</th>
<th>The European Union has 24 official and working languages. They are:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulgarian</td>
<td>French</td>
</tr>
<tr>
<td>Croatian</td>
<td>German</td>
</tr>
<tr>
<td>Czech</td>
<td>Greek</td>
</tr>
<tr>
<td>Danish</td>
<td>Hungarian</td>
</tr>
<tr>
<td>Dutch</td>
<td>Irish</td>
</tr>
<tr>
<td>English</td>
<td>Italian</td>
</tr>
<tr>
<td>Estonian</td>
<td>Latvian</td>
</tr>
<tr>
<td>Finnish</td>
<td>Lithuanian</td>
</tr>
</tbody>
</table>

What is the Commission doing?

With a permanent staff of 1,750 linguists and 600 support staff, the Commission has one of the largest translation services in the world, bolstered by a further 600 full-time and 3,000 freelance interpreters.
How many different ways do you express one meaning?

Cheers
Mapping = Translating

Step 1. Lookup the Source Concept

```
SELECT * FROM concept WHERE concept_code = '427.31';
```

<table>
<thead>
<tr>
<th>CONCEPT_ID</th>
<th>CONCEPT_NAME</th>
<th>DOMAIN_ID</th>
<th>VOCABULARY_ID</th>
<th>CONCEPT_CLASS_ID</th>
<th>STANDARD_CONCEPT</th>
<th>CONCEPT_CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>44821957</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>ICD9CM</td>
<td>5-dig billing code</td>
<td></td>
<td>427.31</td>
</tr>
</tbody>
</table>

Step 2. Translate to Standard

```
SELECT * FROM concept_relationship WHERE concept_id_1 = 44821957 AND relationship_id = 'Maps to';
```

<table>
<thead>
<tr>
<th>CONCEPT_ID_1</th>
<th>CONCEPT_ID_2</th>
<th>RELATIONSHIP_ID</th>
<th>VALID_START_DATE</th>
<th>VALID_END_DATE</th>
<th>INVALID_REASON</th>
</tr>
</thead>
<tbody>
<tr>
<td>44821957</td>
<td>313217</td>
<td>Maps to</td>
<td>01-Jan-1970</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
</tbody>
</table>

Step 3. Check out the translated Concept

```
SELECT * FROM concept WHERE concept_id = 313217;
```
Ancestry Relationships: Higher-Level Relationships

1. Ancestor
   - Disease of the cardiovascular system
     - Heart disease
       - Cardiac arrhythmia
         - Supraventricular arrhythmia
           - Atrial arrhythmia
             - Fibrillation
               - Atrial fibrillation

2. Descendant
   - Controlled atrial fibrillation
   - Persistent atrial fibrillation
   - Chronic atrial fibrillation
   - Paroxysmal atrial fibrillation
   - Rapid atrial fibrillation
   - Permanent atrial fibrillation

Concept Relationships

- 5 levels of separation
- 2 levels of separation
Lauren's Timeline

-3 Years  -2 Years  -1 Years  //  -2 Weeks  //  -3 Days  //  Day 0

Endometriosis

What data do we have?

dysmenorrhea
abdominal pain
missed work
acetaminophen
acetaminophen
acetaminophen
GP visit
pelvic exam
ultrasound
cyst of ovary

Hospital Visit

ultrasound
severe pain
temp 103°F
ambulance
Bloated abdomen
ascites
surgery
endometrioma

3 Years
2 Years
1 Year
-2 Weeks
-3 Days
Day 0

acacetaminophen
 acetaminophen
 acetaminophen
Examples of how Researchers get Lauren’s data?

• Health Insurance Claim Form (HCFA-1500)

• Universal Billing form (UB-92)
Examples of how Researchers get Lauren’s data?

• Health Insurance Claim Form (HCFA-1500)

• Universal Billing form (UB-92)

• Prescriptions
Examples of how Researchers get Lauren’s data?

- Health Insurance Claim Form (HCFA-1500)
- Universal Billing form (UB-92)
- Prescriptions
- Doctors notes

Patient: Lauren
Date of Procedure: 12-March
Surgeon: Dr. Patrick Ryan
Assistant: Dr. Erica Voss
Procedure: Endometrial biopsy
Operative Summary: Endometrial biopsy performed with sterile technique. Adequate sample.
Presence of endometrial tissues outside the uterus.
CDM Version 6 Key Domains

Standardized clinical data
- Person
  - Observation_period
  - Visit_occurrence
  - Visit_detail
  - Condition_occurrence
  - Drug_exposure
  - Procedure_occurrence
  - Device_exposure
  - Measurement
  - Note
  - Note_NLP
  - Survey_conduct
  - Observation
  - Specimen
  - Fact_relationship

Standardized health system data
- Location
- Location_history
- Care_site
- Provider

Standardized derived elements
- Condition_era
- Drug_era
- Dose_era

Results Schema
- Cohort
- Cohort_definition

Standardized health economics
- Cost
- Payer_plan_period

Standardized metadata
- CDM_source
- Metadata

Standardized vocabularies
- Concept
- Vocabulary
- Domain
- Concept_class
- Concept_relationship
- Relationship
- Concept_ancestor
- Concept_synonym
- Source_to_concept_map
- Drug_strength
<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>gender_concept_id</td>
<td>8532 Female</td>
</tr>
<tr>
<td>year_of_birth</td>
<td>1982</td>
</tr>
<tr>
<td>month_of_birth</td>
<td>NULL</td>
</tr>
<tr>
<td>day_of_birth</td>
<td>NULL</td>
</tr>
<tr>
<td>race_concept_id</td>
<td>8527 White</td>
</tr>
<tr>
<td>person_source_value</td>
<td>123456</td>
</tr>
<tr>
<td>gender_source_value</td>
<td>F</td>
</tr>
<tr>
<td>race_source_value</td>
<td>W</td>
</tr>
</tbody>
</table>

Sample of table's columns
<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>observation_period_id</td>
<td>1</td>
</tr>
</tbody>
</table>
| person_id                          | 123456  
-shared by Lauren’s ID          |
| observation_period_start_date      | 2000-01-01            |
| observation_period_end_date        | 2010-12-31            |

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>observation_period_id</td>
<td>2</td>
</tr>
</tbody>
</table>
| person_id                          | 123456  
-shared by Lauren’s ID          |
| observation_periods_start_date     | 2012-01-01            |
| observation_periods_end_date       | 2013-12-31            |
## VISIT_OCCURRENCE

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>visit_occurrence_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>visit_start_date</td>
<td>2008-04-07</td>
</tr>
<tr>
<td>visit_end_date</td>
<td>2008-04-07</td>
</tr>
<tr>
<td>visit_concept_id</td>
<td>9202</td>
</tr>
<tr>
<td>visit_source_value</td>
<td>OP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>visit_occurrence_id</td>
<td>2</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>visit_start_date</td>
<td>2008-04-21</td>
</tr>
<tr>
<td>visit_end_date</td>
<td>2008-04-26</td>
</tr>
<tr>
<td>visit_concept_id</td>
<td>9201</td>
</tr>
<tr>
<td>visit_source_value</td>
<td>IP</td>
</tr>
</tbody>
</table>
## CONDITION_OCCURRENCE

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>condition_occurrence_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456 Lauren’s ID</td>
</tr>
<tr>
<td>condition_concept_id</td>
<td>433527 Endometriosis</td>
</tr>
<tr>
<td>condition_start_date</td>
<td>2008-04-24 Inpatient detail - primary</td>
</tr>
<tr>
<td>condition_type_concept_id</td>
<td>38000183 ICD9, missing decimal</td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>2</td>
</tr>
<tr>
<td>condition_source_value</td>
<td>6171 Endometriosis of ovary</td>
</tr>
<tr>
<td>condition_source_concept_id</td>
<td>44832501</td>
</tr>
<tr>
<td>COLUMN</td>
<td>EXAMPLE</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td><code>drug_exposure_id</code></td>
<td>1</td>
</tr>
<tr>
<td><code>person_id</code></td>
<td>123456</td>
</tr>
<tr>
<td><code>drug_concept_id</code></td>
<td>40162494</td>
</tr>
<tr>
<td><code>drug_exposure_start_date</code></td>
<td>2007-02-01</td>
</tr>
<tr>
<td><code>drug_exposure_end_date</code></td>
<td>2007-02-08</td>
</tr>
<tr>
<td><code>verbatim_end_date</code></td>
<td>NULL</td>
</tr>
<tr>
<td><code>drug_type_concept_id</code></td>
<td>38000183</td>
</tr>
<tr>
<td><code>refills</code></td>
<td>0</td>
</tr>
<tr>
<td><code>quantity</code></td>
<td>14</td>
</tr>
<tr>
<td><code>days_supply</code></td>
<td>7</td>
</tr>
<tr>
<td><code>drug_source_value</code></td>
<td>54348001301</td>
</tr>
<tr>
<td><code>drug_source_concept_id</code></td>
<td>45904353</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lauren’s ID</th>
<th>Acetaminophen 500 MG / Hydrocodone Bitartrate 5 MG Oral Tablet</th>
</tr>
</thead>
<tbody>
<tr>
<td>prescription dispensed in pharmacy</td>
<td>Drug_exposure_start_date + days_supply</td>
</tr>
<tr>
<td>NDC 11-digit code</td>
<td>Acetaminophen 500 MG / Hydrocodone Bitartrate 5 MG Oral Tablet</td>
</tr>
</tbody>
</table>
## PROCEDURE_OCCURRENCE

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>procedure_occurrence_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456  \textit{Lauren's ID}</td>
</tr>
<tr>
<td>procedure_concept_id</td>
<td>2211740  \textit{Ultrasound, abdominal, real time with image documentation; complete}</td>
</tr>
<tr>
<td>procedure_date</td>
<td>2008-04-08  \textit{Outpatient detail - 1st position}</td>
</tr>
<tr>
<td>procedure_type_concept_id</td>
<td>38000267  \textit{CPT4}</td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>1</td>
</tr>
<tr>
<td>procedure_source_value</td>
<td>76700  \textit{Ultrasound, abdominal, real time with image documentation; complete}</td>
</tr>
<tr>
<td>procedure_source_concept_id</td>
<td>2211740  \textit{Ultrasound, abdominal, real time with image documentation; complete}</td>
</tr>
</tbody>
</table>
### MEASUREMENT

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>measurement_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>measurement_concept_id</td>
<td>3020891</td>
</tr>
<tr>
<td>measurement_date</td>
<td>2008-04-21</td>
</tr>
<tr>
<td>measurement_type_concept_id</td>
<td>44818701</td>
</tr>
<tr>
<td>value_as_number</td>
<td>103</td>
</tr>
<tr>
<td>unit_concept_id</td>
<td>9289</td>
</tr>
<tr>
<td>measurement_source_value</td>
<td>8310-5</td>
</tr>
<tr>
<td>measurement_source_concept_id</td>
<td>3020891</td>
</tr>
</tbody>
</table>

*Lauren’s ID*

*Body temperature*

*From physical examination*

*Degree Fahrenheit*

*LOINC*

*Body temperature*
<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>observation_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>observation_concept_id</td>
<td>0</td>
</tr>
<tr>
<td>observation_date</td>
<td>2006-01-20</td>
</tr>
<tr>
<td>observation_type_concept_id</td>
<td>44814721</td>
</tr>
<tr>
<td>value_as_number</td>
<td>8</td>
</tr>
<tr>
<td>value_as_string</td>
<td>Work Hours Missed</td>
</tr>
<tr>
<td>observation_source_value</td>
<td>Work Hours Missed</td>
</tr>
<tr>
<td>observation_source_concept_id</td>
<td>0</td>
</tr>
</tbody>
</table>
Illustrating inferences needed within longitudinal pharmacy claims data for one patient

Person Timeline

Lisinopril era 1

How do we handle reversals?

NDC: 00179198801
Lisinopril 5 MG Oral Tablet

NDC: 00038013134
Lisinopril 10 MG Oral Tablet [Zestril]

NDC: 00038013210
Lisinopril 20 MG Oral Tablet [Zestril]

NDC: 58016078020
Hydrochlorothiazide 12.5 MG / Lisinopril 20 MG Oral Tablet [Zestoretic]

Era 2

How do we handle NDC change?

How do we handle overlap?

How do we handle change in dose?

How do we handle gaps?

How do we handle combination products?

How do we infer discontinuation?

How do we handle overlap?

Prescription dispensing
(Fill date + days supply)
CDM Tables Not Covered in Detail

- VISIT_DETAIL
- SPECIMEN
- DEATH
- DEVICE_EXPOSURE
- NOTE
- NOTE_NLP
- FACT_RELATIONSHIP
- LOCATION
- CARE_SITE

- PROVIDER
- PAYER_PLAN_PERIOD
- COST
- COHORT
- COHORT_ATTRIBUTES
- CONDITION_ERA
- DOSE_ERA
- CDM_SOURCE
Standards

• Patients without transaction

• Cleaning dirty data
  – Patient IDs reused
  – Bogus code records (e.g. ‘000’)

• How to handle tobacco information

https://github.com/OHDSI/CommonDataModel/wiki
CDM Version Control

• Working group meets once a month to discuss proposed changes to the CDM

• All CDM documentation, versions, and proposals located on GitHub
  – [https://github.com/OHDSI/CommonDataModel](https://github.com/OHDSI/CommonDataModel)
  – Proposals tracked and discussed as GitHub issues

• Meeting information can be found on the working group [wiki page](#)

• Please contact Clair Blacketer [(mblacke@its.jnj.com)](mailto:mblacke@its.jnj.com) for more information
OHDSI generates Evidence
Characterizing treatment pathways at scale using the OHDSI network

George Hripcsak\textsuperscript{a,b,c,1}, Patrick B. Ryan\textsuperscript{c,d}, Jon D. Duke\textsuperscript{e}, Nigam H. Shah\textsuperscript{c,f}, Rae Woong Park\textsuperscript{c,g}, Vojtech Huser\textsuperscript{c,h}, Marc A. Suchard\textsuperscript{c,i,j,k}, Martijn J. Schuemie\textsuperscript{c,d}, Frank J. DeFalco\textsuperscript{c,d}, Adler Perotte\textsuperscript{a,c}, Juan M. Banda\textsuperscript{c,f}, Christian G. Reich\textsuperscript{c,l}, Lisa M. Schilling\textsuperscript{c,m}, Michael E. Matheny\textsuperscript{c,n,o}, Daniella Meeker\textsuperscript{c,p,q}, Nicole Pratt\textsuperscript{c,r}, and David Madigan\textsuperscript{c,s}

\textsuperscript{a}Department of Biomedical Informatics, Columbia University Medical Center, New York, NY 10032; \textsuperscript{b}Medical Informatics Services, NewYork-Presbyterian Hospital, New York, NY 10032; \textsuperscript{c}Observational Health Data Sciences and Informatics, New York, NY 10032; \textsuperscript{d}Epidemiology Analytics, Janssen Research and Development, Titusville, NJ 08560; \textsuperscript{e}Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN 46205; \textsuperscript{f}Center for Biomedical Informatics Research, Stanford University, CA 94305; \textsuperscript{g}Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea, 443-380; \textsuperscript{h}Lister Hill National Center for Biomedical Communications (National Library of Medicine), National Institutes of Health, Bethesda, MD 20894; \textsuperscript{i}Department of Biostatistics, University of California, Los Angeles, CA 90095; \textsuperscript{j}Department of Human Genetics, University of California, Los Angeles, CA 90095; \textsuperscript{k}Real World Evidence Solutions, IMS Health, Burlington, MA 01809; \textsuperscript{l}Department of Medicine, University of Colorado School of Medicine, Aurora, CO 80045; \textsuperscript{m}Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN 37212; \textsuperscript{n}Geriatric Research, Education and Clinical Center, VA Tennessee Valley Healthcare System, Nashville, TN 37212; \textsuperscript{o}Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90089; \textsuperscript{p}Department of Pediatrics, University of Southern California, Los Angeles, CA 90089; \textsuperscript{q}Division of Health Sciences, University of South Australia, Adelaide, SA, Australia 5001; and \textsuperscript{r}Department of Statistics, Columbia University, New York, NY 10027

Edited by Richard M. Shiffrin, Indiana University, Bloomington, IN, and approved April 5, 2016 (received for review June 14, 2015)

Observational research promises to complement experimental research by providing large, diverse populations that would be infeasible for an experiment. Observational research can test its own clinical hypotheses, and observational studies also can contribute to the design of experiments and inform the generalizability of experimental research. Understanding the diversity of populations and the underlying risk factors involved in health and disease is critical.

Without sufficiently broad databases available in the first stage, randomized trials are designed without explicit knowledge of actual disease status and treatment practice. Literature reviews are restricted to the population choices of previous investigations, and pilot studies usually are limited in scope. By exploiting the ClinicalTrials.gov national trial registry (9) and electronic health record databases (2), we have been able to identify observations and outcomes, which have been then used to inform the design of future studies.
Treatment pathways for diabetes

T2DM: All databases

First drug

Second drug

Only drug
Heterogeneity in treatments

Japan differs in use of Metformin (due to genetics)

No agreement on depression
The current literature is severely biased

In reality, most results are not significant, but they are never published (missing)

85% of published studies are positive

Drug protects

Drug harms

29,982 estimates
11,758 papers
1. Address confounding that is measured
   • Propensity stratification

2. Address unmeasured confounding
   • Negative controls

3. Multiple databases, locations, practice types
   • Exploit international OHDSI network

4. Open: publish all
5. Run 17,000 studies at once

Duloxetine vs. Sertraline for these 22 outcomes:

<table>
<thead>
<tr>
<th>Acute liver injury</th>
<th>Hypotension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute myocardial infarction</td>
<td>Hypothyroidism</td>
</tr>
<tr>
<td>Alopecia</td>
<td>Insomnia</td>
</tr>
<tr>
<td>Constipation</td>
<td>Nausea</td>
</tr>
<tr>
<td>Decreased libido</td>
<td>Open-angle glaucoma</td>
</tr>
<tr>
<td>Delirium</td>
<td>Seizure</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>Stroke</td>
</tr>
<tr>
<td>Fracture</td>
<td>Suicide and suicidal ideation</td>
</tr>
<tr>
<td>Gastrointestinal hemorrhage</td>
<td>Tinnitus</td>
</tr>
<tr>
<td>Hyperprolactinemia</td>
<td>Ventricular arrhythmia and sudden cardiac death</td>
</tr>
<tr>
<td>Hyponatremia</td>
<td>Vertigo</td>
</tr>
</tbody>
</table>
Many treatments at once

<table>
<thead>
<tr>
<th>Type</th>
<th>Class</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug</td>
<td>Atypical</td>
<td>Bupropion</td>
</tr>
<tr>
<td>Drug</td>
<td>Atypical</td>
<td>Mirtazapine</td>
</tr>
<tr>
<td>Procedure ECT</td>
<td></td>
<td>Electroconvulsive therapy</td>
</tr>
<tr>
<td>Procedure y</td>
<td></td>
<td>Psychotherapy</td>
</tr>
<tr>
<td>Drug</td>
<td>SARI</td>
<td>Trazodone</td>
</tr>
<tr>
<td>Drug</td>
<td>SNRI</td>
<td>Desvenlafaxine</td>
</tr>
<tr>
<td>Drug</td>
<td>SNRI</td>
<td>duloxetine</td>
</tr>
<tr>
<td>Drug</td>
<td>SNRI</td>
<td>venlafaxine</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Citalopram</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Escitalopram</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Fluoxetine</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Paroxetine</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>Sertraline</td>
</tr>
<tr>
<td>Drug</td>
<td>SSRI</td>
<td>vilazodone</td>
</tr>
<tr>
<td>Drug</td>
<td>TCA</td>
<td>Amitriptyline</td>
</tr>
<tr>
<td>Drug</td>
<td>TCA</td>
<td>Doxepin</td>
</tr>
</tbody>
</table>
OHDSI’s results: less bias

11% of exposure-outcome pairs have calibrated p < 0.05
### Benefit and Harm of 2nd-generation Antidepressants

<table>
<thead>
<tr>
<th>Outcome</th>
<th>SSRI</th>
<th>SNRI</th>
<th>TCA</th>
<th>Other Drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute liver injury</td>
<td>Advers</td>
<td>Overd</td>
<td>Advers</td>
<td>Reacti</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>Advers</td>
<td>Reacti</td>
<td>Advers</td>
<td>Reacti</td>
</tr>
<tr>
<td>Alopecia</td>
<td>Advers</td>
<td>Reacti</td>
<td>Advers</td>
<td>Reacti</td>
</tr>
<tr>
<td>Constipation</td>
<td>Advers</td>
<td>Reacti</td>
<td>Advers</td>
<td>Reacti</td>
</tr>
<tr>
<td>Decreased libido</td>
<td>Advers</td>
<td>Reacti</td>
<td>Advers</td>
<td>Reacti</td>
</tr>
<tr>
<td>Delirium</td>
<td>Advers</td>
<td>Overd</td>
<td>Overd</td>
<td>Reacti</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>Advers</td>
<td>Reacti</td>
<td>Advers</td>
<td>Reacti</td>
</tr>
<tr>
<td>Fracture</td>
<td>Precau</td>
<td>Warning</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Gastrointestinal hemorrhage</td>
<td>Advers</td>
<td>Warning</td>
<td>Precau</td>
<td>Reacti</td>
</tr>
<tr>
<td>Hyperprolactinemia</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Hypoanorexia</td>
<td>Precau</td>
<td>Warning</td>
<td>Precau</td>
<td>Reacti</td>
</tr>
<tr>
<td>Hypotension</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Insomnia</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Nausea</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Open-angle glaucoma</td>
<td>Warning</td>
<td>Warning</td>
<td>Precau</td>
<td>Warnin</td>
</tr>
<tr>
<td>Seizure</td>
<td>Precau</td>
<td>Warning</td>
<td>Precau</td>
<td>Reacti</td>
</tr>
<tr>
<td>Stroke</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Suicide and suicidal thoughts</td>
<td>Boxed</td>
<td>Boxed</td>
<td>Boxed</td>
<td>Boxed</td>
</tr>
<tr>
<td>Tinnitus</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Ventricular arrhythmia</td>
<td>Warning</td>
<td>Advers</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
<tr>
<td>Vertigo</td>
<td>Advers</td>
<td>Overd</td>
<td>Reacti</td>
<td>Reacti</td>
</tr>
</tbody>
</table>

71
Comparing RCT and observational results for effects of sertraline on diarrhea
Comparative effectiveness hypotheses from Gartleher et al

- Venlafaxine has higher risk of nausea than SSRI
- No difference in nausea between duloxetine and paroxetine or fluoxetine
- Sertraline has higher risk of diarrhea than comparators
- Paroxetine has higher rate of sexual dysfunction than fluoxetine and sertraline.
- Bupropion has lower incidence of sexual dysfunction than fluoxetine, paroxetine, and sertraline.
- Trazodone increased risk of somnolence
FDA BEST Program

62 studies:
- Simple: Rapid queries
- Medium complexity: cohort characterization
- High complexity: safety, pharmacoepi

AE Reports

UCLA
- Develop new methods

Add unstructured data through NLP

Develop Studies/Reports
- Run Studies/Reports

Create MedWatch submission module

BEST Network

Report back

Hospital Charge Master
Ambulatory EMR
LRxDx Provider-based Claims

BEST Network
Summary

- OHDSI is a public world-wide collaborative. Everybody can participate. It's free. There is no catch.
- You don't have to give the data away, but you need to standardize the data. The standard is strict. No shortcuts!
- When you do that, you get tools, methods, and a lot of new colleagues. People in the OHDSI community are nice and competent.
- You can do meaningful and scientifically high-quality network research
Join the Journey

http://ohdsi.org