OMOP Common Data Model and Standardized Vocabularies

15- September-2019

Christian Reich, MD, PhD, Mui van Zandt
Erica A. Voss, MPH, PMP, Hamed Abedtash, PharmD, PhD
Dmitry Dymshyts, MD, Melanie Philofsky, RN, MS
Don Torok, MS
After the Tutorials, you will know...

1. History of OMOP, OHDSI
2. How the Standardized Vocabulary works
3. How to find codes and Concepts
4. How to navigate the concept hierarchy
5. The OMOP Common Data Model (CDM)
6. How to use the OMOP CDM
<table>
<thead>
<tr>
<th>Section</th>
<th>Speaker</th>
<th>Time</th>
<th>Item(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration</td>
<td>-</td>
<td>8:00 - 9:00</td>
<td>(1 hour)</td>
</tr>
<tr>
<td>Introduction</td>
<td>Christian</td>
<td>9:00 - 10:00</td>
<td>(1 hour)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Introductions and Ground Rules</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Foundational</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• History of OMOP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Why and How</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Birth of OHDSI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Introduction to OMOP Common Data Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OHDSI Community</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Example of Remote Study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VM Overview</td>
</tr>
<tr>
<td>Vocabulary – Part 1</td>
<td>Hamed</td>
<td>10:00 – 10:30</td>
<td>(30 min)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Basic Relationship</td>
</tr>
<tr>
<td>Break</td>
<td>-</td>
<td>10:30 - 10:45</td>
<td>(15 min)</td>
</tr>
<tr>
<td>Vocabulary – Part 2</td>
<td>Dima</td>
<td>10:45- 12:30</td>
<td>(1 hour &amp; 45 min)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ancestors &amp; Descendants</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How does it work for Drugs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SQL Examples</td>
</tr>
<tr>
<td>Section</td>
<td>Speaker</td>
<td>Time</td>
<td>Item(s)</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lunch</td>
<td>-</td>
<td>12:30 - 13:30</td>
<td>(1 hour)</td>
</tr>
<tr>
<td>Vocabulary – Part 3</td>
<td>Melanie</td>
<td>13:30 - 14:00</td>
<td>Continued</td>
</tr>
<tr>
<td>Common Data Model</td>
<td>Don</td>
<td>14:00 - 15:35</td>
<td>History of the model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1 hour &amp; 30 min)</td>
<td>In depth discussion of model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Era discussion</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Real World Scenario</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ETL Pitfalls</td>
</tr>
<tr>
<td>Break</td>
<td>-</td>
<td>15:30 - 15:45</td>
<td>(15 min)</td>
</tr>
<tr>
<td>CDM Examples</td>
<td>Mui / Erica</td>
<td>15:45 - 17:00</td>
<td>Leveraging OHDSI Tools (GitHub/Forums/Working Group)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1 hour &amp; 16 min)</td>
<td>Exercises</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OHDSI Community</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conclusion Game</td>
</tr>
</tbody>
</table>
# Instructors

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Christian Reich, MD, PhD</td>
<td>Mui van Zandt</td>
<td>Erica A. Voss, MPH, PMP</td>
<td>Hamed Abedtash, PharmD, PhD</td>
</tr>
<tr>
<td>Dmitry Dymshyts, MD</td>
<td>Melanie Philofsky, RN, MS</td>
<td>Don Torok, MS</td>
<td></td>
</tr>
</tbody>
</table>
Ground Rules

• We are recording today’s session.

• We may table some questions if they are too specific.

• If we cannot get the remote desktop working on your machine let’s try to buddy you up. Do not worry the presentation will still walk you through the content.
Foundational

What is OMOP/OHDSI?
OMOP Common Data Model (CDM) – Why and How
FDA Regulatory Action over Time

Number of FDA-caused Withdrawals

![Graph showing the number of FDA-caused withdrawals over decades.](image-url)
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.—
## OMOP Experiment 1 (2009-2010)

### OMOP Extended Consortium
- 10 data sources
- Claims and EHRs
- 200M+ lives

### OMOP Research Core

#### 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: sulfa, erythromycins, tetracyclines</th>
<th>Antiepileptics: carbamazepine, phenytoin</th>
<th>Benzo-diazepines</th>
<th>Beta blockers</th>
<th>Blephosphates: alendronate</th>
<th>Tricyclic antidepressants</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 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>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>

Observational Data

- ARS
- IPCI
- HS
- PHARMO

eu-adr
### 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>Condition</th>
<th>Positive controls</th>
<th>Negative controls</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute Liver Injury</td>
<td>81</td>
<td>37</td>
<td>118</td>
</tr>
<tr>
<td>Acute Myocardial Infarction</td>
<td>36</td>
<td>66</td>
<td>102</td>
</tr>
<tr>
<td>Acute Renal Failure</td>
<td>24</td>
<td>64</td>
<td>88</td>
</tr>
<tr>
<td>Upper Gastrointestinal Bleeding</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>

- isoniazid
- fluticasone
- indomethacin
- clindamycin
- ibuprofen
- loratadine
- sertraline
- pioglitazone
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
Fate of OMOP - 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
  – Public, Open
  – Not Pharma-funded
  – International

http://ohdssi.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
Current pace of evidence generation in healthcare

All health outcomes of interest

All drugs
“What's the adherence to my drug in the data assets I own?”

Current Approach: “One Study – One Script”

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

- Adherence
- Mortality
- Source of Business
- Safety Signals
- OHDSI Tools
- OMOP CDM

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
Virtual Machine
OHDSI in a Box

- PostgreSQL
  - cdm
  - webapi
- PGAdmin4
- EC2
- Atlas
- WebAPI
- Tomcat
- Methods Library
  - OHDSI R packages
  - Studio
- WhiteRabbit
- Raw Lauren
- CDM Lauren (EMPTY)
- Raw Synthea
- CDM Synthea
- CDM Synpuf (100K)
- Usagi
How to Sign into the Remote Desktop

From your command prompt, type %systemroot%/system32/mstsc.exe to launch Remote Desktop
How to Sign into the Remote Desktop

• Use the shortcut on the desktop named “Remote Desktop”

UPDATED LATER

• Pick one of the rows and put your name on the second column
How to Sign into the Remote Desktop

- Take Column A from spreadsheet and copy into the “Computer” field
How to Sign into the Remote Desktop

1. Pick ‘Use Another Account’
2. Copy username from Column C
3. Copy password from Column D
How to Sign into the Remote Desktop

• If you get this page, select “Yes”
OHDSI in a Box – Ready
CDM Database: pgAdmin III New Server

- Click on pgAdmin
CDM Database: Connect

- Password: ohdsi
CDM Database: Open SQL Sheet
CDM Database: Ready

set search_path to 'ohdsi';
Vocabulary

Basic Relationship, Ancestors, & Descendants
How does it work for Drugs
SQL Examples
OMOP Common Vocabulary Model

What it is
• **Standardized structure** to house existing vocabularies used in the public domain
• **Compiled standards** from disparate public and private sources and some OMOP-grown concepts

What it’s not
• **Static dataset** – the vocabulary updates regularly to keep up with the continual evolution of the sources
• **Finished product** – vocabulary maintenance and improvement is ongoing activity that requires community participation and support
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`
All vocabularies stacked up in one table.
Dozens of schemes, formats, rules

LOINC_248_MULTI-AXIAL_HIERARCHY.CSV

<table>
<thead>
<tr>
<th>PATH_TO_ROOT</th>
<th>SEQUENCE</th>
<th>IMMEDIATE_PARENT</th>
<th>CODE</th>
<th>CODE_TEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP31755-9</td>
<td>1</td>
<td>LP31755-9</td>
<td>LP31755-9</td>
<td>Microbiology</td>
</tr>
<tr>
<td>LP31755-9.1</td>
<td>1</td>
<td>LP14559-6</td>
<td>LP14559-6</td>
<td>Microorganism</td>
</tr>
<tr>
<td>LP31755-9.1</td>
<td>1</td>
<td>LP98185-9</td>
<td>LP98185-9</td>
<td>Bacteria</td>
</tr>
<tr>
<td>LP31755-9.1</td>
<td>1</td>
<td>LP14082-9</td>
<td>LP14082-9</td>
<td>Bacteria</td>
</tr>
</tbody>
</table>

LOINC_NL COMPONENT

<table>
<thead>
<tr>
<th>PROPERTY</th>
<th>TIME_ASPECT</th>
<th>SYSTEM</th>
<th>SCALE_TYP</th>
<th>METHOD_TYP</th>
<th>CLASS</th>
<th>SOURCE</th>
<th>DATE_LAST_CHANGED</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>10454-7</td>
<td>Xylose&lt;2H post 25 g xylose PO</td>
<td>MCnc</td>
<td>Pt</td>
<td>Set/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>19961220 ADD</td>
<td></td>
</tr>
<tr>
<td>10455-4</td>
<td>Xylose&lt;30M post 25 g xylose PO</td>
<td>MCnc</td>
<td>Pt</td>
<td>Set/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>19961220 ADD</td>
<td></td>
</tr>
<tr>
<td>1046-2</td>
<td>Xylose&lt;post 6H CFst</td>
<td>MCnc</td>
<td>Pt</td>
<td>Set/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>19961220 ADD</td>
<td></td>
</tr>
<tr>
<td>10457-0</td>
<td>Actin Ag</td>
<td>MCnc</td>
<td>Pt</td>
<td>Set/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>19961220 ADD</td>
<td></td>
</tr>
<tr>
<td>10458-8</td>
<td>Alkaline phosphatase placebo</td>
<td>MCnc</td>
<td>Pt</td>
<td>Set/Plas</td>
<td>Qn</td>
<td>CHAL</td>
<td>19961220 ADD</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 septicaemia</td>
<td>Salmonella septicaemia</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>Other localized salmonella infections</td>
</tr>
</tbody>
</table>
## What's in a Concept

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONCEPT_ID</td>
<td>313217</td>
</tr>
<tr>
<td>CONCEPT_NAME</td>
<td>Atrial fibrillation</td>
</tr>
<tr>
<td>DOMAIN_ID</td>
<td>Condition</td>
</tr>
<tr>
<td>VOCABULARY_ID</td>
<td>SNOMED</td>
</tr>
<tr>
<td>CONCEPT_CLASS_ID</td>
<td>Clinical Finding</td>
</tr>
<tr>
<td>STANDARD_CONCEPT</td>
<td>S</td>
</tr>
<tr>
<td>CONCEPT_CODE</td>
<td>49436004</td>
</tr>
<tr>
<td>VALID_START_DATE</td>
<td>01-Jan-1970</td>
</tr>
<tr>
<td>VALID_END_DATE</td>
<td>31-Dec-2099</td>
</tr>
<tr>
<td>INVALID_REASON</td>
<td></td>
</tr>
</tbody>
</table>

For use in CDM
- English description
- Domain
- Vocabulary
- Class in SNOMED
- Concept in data
- Code in SNOMED
- Valid during time interval
MiniSentinel in use: Dabigatran and bleeding

Dabigatran and Postmarketing Reports of Bleeding
Mary Ross Southworth, Pharm.D., Marsha E. Reichman, Ph.D., and Ellis F. Unger, M.D.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Dabigatran</th>
<th>Warfarin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Patients</td>
<td>No. of Events</td>
</tr>
<tr>
<td>Gastrointestinal hemorrhage</td>
<td>10,599</td>
<td>16</td>
</tr>
<tr>
<td>Analysis with required diagnosis of atrial fibrillation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity analysis without required diagnosis of atrial fibrillation</td>
<td>12,195</td>
<td>19</td>
</tr>
<tr>
<td>Intracranial hemorrhage</td>
<td>10,587</td>
<td>8</td>
</tr>
<tr>
<td>Analysis with required diagnosis of atrial fibrillation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity analysis without required diagnosis of atrial fibrillation</td>
<td>12,182</td>
<td>10</td>
</tr>
</tbody>
</table>
All Content in CDM is Coded as Concepts

• Concepts are referred to by concept_id

• All details are in the CONCEPT table:

```sql
SELECT * FROM concept WHERE concept_id = 313217
```

<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>
<th>valid_start_date</th>
<th>valid_end_date</th>
<th>invalid_reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>313217</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>49436004</td>
<td>1970-01-01</td>
<td>2099-12-31</td>
<td>NULL</td>
</tr>
</tbody>
</table>
Condition Concepts

- **Classification Concepts**
  - Standard Concepts
    - Top-level classification
    - Higher-level classifications
    - Low-level concepts
  - Source Codes
    - ICD10
    - ICD10CM
    - Read
    - SNOMED
    - Oxmis
    - Ciel
    - MeSH
    - ICD9CM

- **Source Concepts**
  - MedDRA
    - System organ class
    - High-level group terms
    - High-level terms
    - Preferred terms
    - Low-level terms
# Finding the Right Concept #1

## 1. ..if I know the ID

```
SELECT * FROM concept WHERE concept_id = 313217
```

<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>
<th>VALID_START_DATE</th>
<th>VALID_END_DATE</th>
<th>INVALID_REASON</th>
</tr>
</thead>
<tbody>
<tr>
<td>313217</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>49436004</td>
<td>01-Jan-1970</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
</tbody>
</table>

## 2. ..if I know the code

```
SELECT * FROM concept WHERE concept_code = '49436004'
```

<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>
<th>VALID_START_DATE</th>
<th>VALID_END_DATE</th>
<th>INVALID_REASON</th>
</tr>
</thead>
<tbody>
<tr>
<td>313217</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>49436004</td>
<td>01-Jan-1970</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
</tbody>
</table>
```sql
SELECT *
FROM concept
WHERE concept_code = '1001';
```

<table>
<thead>
<tr>
<th>Concept Name</th>
<th>Concept Class</th>
<th>Vocabulary ID</th>
<th>Concept Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antipyrine</td>
<td>Ingredient</td>
<td>RxNorm</td>
<td>1001</td>
</tr>
<tr>
<td>Aceprometazine maleate</td>
<td>Ingredient</td>
<td>BDPM</td>
<td>1001</td>
</tr>
<tr>
<td>Serum</td>
<td>Specimen</td>
<td>CIEL</td>
<td>1001</td>
</tr>
<tr>
<td>methixene hydrochloride</td>
<td>Ingredient</td>
<td>Multilex</td>
<td>1001</td>
</tr>
<tr>
<td>Brompheniramine Maleate, 10 mg/mL injectable solution</td>
<td>Ingredient</td>
<td>Multum</td>
<td>1001</td>
</tr>
<tr>
<td>ABBOTT COLD SORE BALM 4%/0.06% W/</td>
<td>Drug Product</td>
<td>LPD_Australia</td>
<td>1001</td>
</tr>
<tr>
<td>Residential Treatment - Psychiatric</td>
<td>Revenue Code</td>
<td>Revenue Code</td>
<td>1001</td>
</tr>
</tbody>
</table>
SELECT * FROM concept WHERE concept_name = 'Atrial fibrillation';

<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>313217</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>49436004</td>
</tr>
<tr>
<td>44821957</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>ICD9CM</td>
<td>5-digit billing</td>
<td>code</td>
<td>427.31</td>
</tr>
<tr>
<td>35204953</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
<td>10003658</td>
</tr>
<tr>
<td>45500085</td>
<td>Atrial fibrillation</td>
<td>Condition</td>
<td>Read</td>
<td>Read</td>
<td></td>
<td>G573000</td>
</tr>
<tr>
<td>45883018</td>
<td>Atrial fibrillation</td>
<td>Meas Value</td>
<td>LOINC</td>
<td>Answer</td>
<td>S</td>
<td>LA17084-7</td>
</tr>
</tbody>
</table>
Finding the Right Concept #3

1. if don't know any of this, but I know the code in another vocabulary

```sql
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>

```sql
SELECT * FROM concept_relationship WHERE concept_id_1 = 44821957;
```

<table>
<thead>
<tr>
<th>_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>21001551</td>
<td>ICD9CM - FDB Ind</td>
<td>01-Oct-13</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
<tr>
<td>44821957</td>
<td>35204953</td>
<td>ICD9CM - MedDRA</td>
<td>01-Jan-70</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
<tr>
<td>44821957</td>
<td>44824248</td>
<td>Is a</td>
<td>01-Oct-14</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
<tr>
<td>44821957</td>
<td>44834731</td>
<td>Is a</td>
<td>01-Oct-14</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
<tr>
<td>44821957</td>
<td>313217</td>
<td>Maps to</td>
<td>01-Jan-70</td>
<td>31-Dec-2099</td>
<td></td>
</tr>
</tbody>
</table>

ICD-9 is not a Standard Concept

Mapping to different vocabularies

Kind of relationship
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></td>
<td>Bulgarian            French            Maltese</td>
</tr>
<tr>
<td></td>
<td>Croatian              German              Polish</td>
</tr>
<tr>
<td></td>
<td>Czech                 Greek               Portuguese</td>
</tr>
<tr>
<td></td>
<td>Danish                Hungarian           Romanian</td>
</tr>
<tr>
<td></td>
<td>Dutch                 Irish               Slovak</td>
</tr>
<tr>
<td></td>
<td>English               Italian             Slovenian</td>
</tr>
<tr>
<td></td>
<td>Estonian              Latvian             Spanish</td>
</tr>
<tr>
<td></td>
<td>Finnish               Lithuanian          Swedish</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-digit 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;
```
Exercise: Find Standard Concept ID from Source Concept

ICD9: '427.31' : 313217
ICD10CM: 'I48.91' : 313217
ICD10: 'I48.0' : 4154290 'Paroxysmal Atrial Fibrillation'

Step 1. Lookup
SELECT * FROM concept WHERE concept_code = ...;

Step 2. Translate
SELECT * FROM concept_relationship WHERE concept_id_1 = ...
AND relationship_id = 'Maps to';

Step 3. Check out
SELECT * FROM concept WHERE concept_id = ...;
Break

Please return in 15 minutes
Reason #2: Disease Hierarchy

- Disease of the cardiovascular system
  - Heart disease
    - Cardiac arrhythmia
      - Supraventricular arrhythmia
        - Fibrillation
        - Atrial arrhythmia
          - Atrial fibrillation
            - Controlled atrial fibrillation
            - Persistent atrial fibrillation
            - Chronic atrial fibrillation
            - Paroxysmal atrial fibrillation
            - Rapid atrial fibrillation
            - Permanent atrial fibrillation
SELECT * FROM concept_relationship WHERE concept_id_1 = 313217
# Exploring Relationships

```sql
SELECT cr.relationship_id, c.*
FROM concept_relationship cr
JOIN concept c ON cr.concept_id_2 = c.concept_id
WHERE cr.concept_id_1 = 313217
```

## Ancestor Concepts

- **Asso finding of**
  - Unspecified atrial fibrillation
  - Atrial fibrillation and atrial fibrillation not documented
  - Fibrillation - atrial

## Descendant Concepts

- **Mapped from**
  - Patient with heart failure and atrial fibrillation documented to be on warfarin therapy
  - Unspecified atrial fibrillation
  - Atrial fibrillation

## Find out related concept

- **Focus of**
  - Insertion of pacemaker for control of atrial fibrillation

- **Due to of**
  - Transient cerebral ischemia due to atrial fibrillation

- **Is a**
  - Fibrillation
  - Atrial arrhythmia

- **Has finding site**
  - Atrial structure
Ancestry Relationships: Higher-Level Relationships

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

Concept Relationships:
- Ancestor
- Descendant

Ancestry Relationships:
- 5 levels of separation
- 2 levels of separation

- Controlled atrial fibrillation
- Persistent atrial fibrillation
- Chronic atrial fibrillation
- Paroxysmal atrial fibrillation
- Rapid atrial fibrillation
- Permanent atrial fibrillation
### Exploring Ancestors of a Concept

```sql
SELECT max_levels_of_separation, concept.*
FROM concept_ancestor
JOIN concept ON ancestor_concept_id = concept_id
WHERE descendant_concept_id = 313217 /* Atrial fibrillation */
ORDER BY max_levels_of_separation
```

<table>
<thead>
<tr>
<th>max_levels_of_separation</th>
<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>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>313217</td>
<td>Atrial fibrillation</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>35204953</td>
<td>Atrial fibrillation</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4226399</td>
<td>Fibrillation</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4068155</td>
<td>Atrial arrhythmia</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>35204969</td>
<td>Cardiac fibrillation</td>
<td></td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
</tr>
<tr>
<td>2</td>
<td>448028</td>
<td>Supraventricular arrhythmia</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>2</td>
<td>35204952</td>
<td>Arrhythmia supraventricular</td>
<td></td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
</tr>
<tr>
<td>2</td>
<td>35204934</td>
<td>Rate and rhythm disorders NEC</td>
<td></td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
</tr>
<tr>
<td>3</td>
<td>44784217</td>
<td>Cardiac arrhythmia</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>3</td>
<td>35202455</td>
<td>Supraventricular arrhythmias</td>
<td></td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>321588</td>
<td>Heart disease</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>4</td>
<td>35204989</td>
<td>Cardiac disorder</td>
<td></td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>35202050</td>
<td>Cardiac arrhythmias</td>
<td></td>
<td>MedDRA</td>
<td>HLGT</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>4103183</td>
<td>Cardiac finding</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>5</td>
<td>440142</td>
<td>Disorder of mediastinum</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>5</td>
<td>134057</td>
<td>Disorder of cardiovascular system</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>35204998</td>
<td>Cardiovascular disorder</td>
<td></td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>37219970</td>
<td>Mediastinal disorder</td>
<td></td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>37622411</td>
<td>Phlebosclerosis</td>
<td></td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>35202457</td>
<td>Cardiac disorders NEC</td>
<td></td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
</tr>
<tr>
<td>6</td>
<td>4115390</td>
<td>Mediastinal finding</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>6</td>
<td>4023995</td>
<td>Cardiovascular finding</td>
<td></td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
</tbody>
</table>
**Exploring Descendants of a Concept**

```sql
SELECT max_levels_of_separation, concept.*
FROM concept_ancestor
JOIN concept ON descendant_concept_id = concept_id
WHERE ancestor_concept_id = 44784217 /* cardiac arrythmia */
ORDER BY max_levels_of_separation
```

<table>
<thead>
<tr>
<th>MAX_LEVELS_OF_SEPARATION</th>
<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>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>44784217</td>
<td>Cardiac arrhythmia</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>313224</td>
<td>Anomalous atrioventricular excitation</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>315643</td>
<td>Tachyarrhythm</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>316429</td>
<td>Premature beats</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>316999</td>
<td>Conduction disorder of the heart</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>321042</td>
<td>Cardiac arrest</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4030583</td>
<td>Pacemaker twiddler's syndrome</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4057008</td>
<td>Accelerated atrioventricular conduction</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4086313</td>
<td>Withdrawal arrhythmia</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4088507</td>
<td>Ventricular escape complex</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4088986</td>
<td>Atrial escape complex</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4091901</td>
<td>Aberrant premature complexes</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4092011</td>
<td>Aberrantly conducted complex</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4124704</td>
<td>Postoperative sinoatrial disease</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4143042</td>
<td>Ectopic beats</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4164083</td>
<td>Ectopic rhythm</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4172863</td>
<td>Fetal dysrhythm</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4173170</td>
<td>Neonatal dysrhythm</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4175473</td>
<td>Atrioventricular dissociation</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4185572</td>
<td>Ventricular arrhythm</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4217221</td>
<td>Nodal rhythm disorder</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4226399</td>
<td>Fibrillation</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
<tr>
<td>1</td>
<td>4228448</td>
<td>Bradycardia</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
</tr>
</tbody>
</table>
Let Us find Upper Gastrointestinal Bleeding

1. Find some initiation concept

   ```
   SELECT * FROM concept WHERE concept_name = 'Upper gastrointestinal bleeding'
   ```

<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>42891225</td>
<td>Upper gastrointestinal bleeding</td>
<td>Condition</td>
<td>MedDRA</td>
<td>LLT</td>
<td>C</td>
<td>10071910</td>
</tr>
</tbody>
</table>

2. Find standard concepts

   ```
   SELECT * FROM concept WHERE lower(concept_name) LIKE '%upper gastrointestinal%' 
   AND domain_id = 'Condition' AND standard_concept = 'S'
   ```

<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>4308202</td>
<td>Acute upper gastrointestinal hemorrhage</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>38938002</td>
</tr>
<tr>
<td>4291649</td>
<td>Upper gastrointestinal hemorrhage</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>37372002</td>
</tr>
<tr>
<td>4115581</td>
<td>Finding of upper gastrointestinal gas</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>300370006</td>
</tr>
<tr>
<td>4103011</td>
<td>Chronic upper gastrointestinal hemorrhage</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>25349007</td>
</tr>
<tr>
<td>4012503</td>
<td>Excessive upper gastrointestinal gas</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>162076009</td>
</tr>
<tr>
<td>4000609</td>
<td>Disorder of upper gastrointestinal tract</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>119291004</td>
</tr>
<tr>
<td>4332645</td>
<td>Upper gastrointestinal hemorrhage associated with hypercoagulability state</td>
<td>Condition</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>430349003</td>
</tr>
</tbody>
</table>
SELECT max_levels_of_separation, concept.* FROM concept_ancestor
JOIN concept ON ancestor_concept_id = concept_id
WHERE descendant_concept_id = 4332645 /* Upper gastrointestinal hemorrhage associated...*/
ORDER BY max_levels_of_separation

Going up the hierarchy:
Finding the right concept

<table>
<thead>
<tr>
<th>max_levels_of_separation</th>
<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>0</td>
<td>4332645</td>
<td>Upper gastrointestinal hemorrhage associated with hypercoag</td>
<td>S</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>430349003</td>
</tr>
<tr>
<td>1</td>
<td>35708054</td>
<td>Gastritis haemorrhagic</td>
<td>C</td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
<td>10017866</td>
</tr>
<tr>
<td>1</td>
<td>4291649</td>
<td>Upper gastrointestinal hemorrhage</td>
<td>C</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>37372002</td>
</tr>
<tr>
<td>1</td>
<td>35707871</td>
<td>Upper gastrointestinal haemorrhage</td>
<td>C</td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
<td>10046274</td>
</tr>
<tr>
<td>2</td>
<td>35707864</td>
<td>Gastrointestinal haemorrhage</td>
<td>S</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>119291004</td>
</tr>
<tr>
<td>2</td>
<td>4000609</td>
<td>Disorder of upper gastrointestinal tract</td>
<td>C</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>10059175</td>
</tr>
<tr>
<td>2</td>
<td>35707858</td>
<td>Intestinal haemorrhage</td>
<td>C</td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
<td>10017854</td>
</tr>
<tr>
<td>2</td>
<td>35702752</td>
<td>Gastritis (excl infective)</td>
<td>C</td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
<td>10017854</td>
</tr>
<tr>
<td>2</td>
<td>192671</td>
<td>Gastrointestinal hemorrhage</td>
<td>S</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>74474003</td>
</tr>
<tr>
<td>3</td>
<td>37604042</td>
<td>Gastrointestinal haemorrhages</td>
<td>C</td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
<td>10052742</td>
</tr>
<tr>
<td>3</td>
<td>37622518</td>
<td>Haemorrhage</td>
<td>C</td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
<td>10055798</td>
</tr>
<tr>
<td>3</td>
<td>437312</td>
<td>Bleeding</td>
<td>S</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>131148009</td>
</tr>
<tr>
<td>3</td>
<td>4198525</td>
<td>Disorder of upper digestive tract</td>
<td>S</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>50410009</td>
</tr>
<tr>
<td>3</td>
<td>37622515</td>
<td>Extravasation blood</td>
<td>C</td>
<td>MedDRA</td>
<td>PT</td>
<td>C</td>
<td>10015867</td>
</tr>
<tr>
<td>3</td>
<td>4000610</td>
<td>Disorder of gastrointestinal tract</td>
<td>S</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>119292006</td>
</tr>
<tr>
<td>3</td>
<td>35702116</td>
<td>Gastrointestinal inflammatory conditions</td>
<td>C</td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
<td>10017969</td>
</tr>
<tr>
<td>3</td>
<td>35702743</td>
<td>Intestinal haemorrhages</td>
<td>C</td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
<td>10022653</td>
</tr>
<tr>
<td>3</td>
<td>35702744</td>
<td>Non-site specific gastrointestinal haemorrhages</td>
<td>C</td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
<td>10017958</td>
</tr>
<tr>
<td>4</td>
<td>35702114</td>
<td>Gastrointestinal haemorrhages NEC</td>
<td>C</td>
<td>MedDRA</td>
<td>HLGT</td>
<td>C</td>
<td>10017959</td>
</tr>
<tr>
<td>4</td>
<td>4304916</td>
<td>Gastrointestinal tract finding</td>
<td>C</td>
<td>SNOMED</td>
<td>Clinical Finding</td>
<td>S</td>
<td>386618008</td>
</tr>
<tr>
<td>4</td>
<td>35702767</td>
<td>Nausea and vomiting symptoms</td>
<td>C</td>
<td>MedDRA</td>
<td>HLT</td>
<td>C</td>
<td>10028817</td>
</tr>
</tbody>
</table>

Hold the descendant
SELECT max_levels_of_separation, concept.*
FROM concept_ancestor
JOIN concept ON descendant_concept_id = concept_id
WHERE ancestor_concept_id = 4291649 /* Upper gastrointestinal hemorrhage */
ORDER BY max_levels_of_separation

<table>
<thead>
<tr>
<th>max_levels_of_separation</th>
<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>0</td>
<td>4291649</td>
<td>Upper gastrointestinal hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>37372002</td>
</tr>
<tr>
<td>1</td>
<td>4318535</td>
<td>Duodenal hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>64121000</td>
</tr>
<tr>
<td>1</td>
<td>23245</td>
<td>Esophageal bleeding</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>25349007</td>
</tr>
<tr>
<td>1</td>
<td>4308202</td>
<td>Acute upper gastrointestinal hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>8765009</td>
</tr>
<tr>
<td>1</td>
<td>4271696</td>
<td>Peptic ulcer with hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>430349003</td>
</tr>
<tr>
<td>1</td>
<td>4103011</td>
<td>Chronic upper gastrointestinal hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>61401005</td>
</tr>
<tr>
<td>1</td>
<td>26727</td>
<td>Hematemesis</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>27281001</td>
</tr>
<tr>
<td>1</td>
<td>4332645</td>
<td>Upper gastrointestinal hemorrhage associated with hypercoagulability</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>1.5072E+14</td>
</tr>
<tr>
<td>1</td>
<td>193250</td>
<td>Gastric hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>64121000</td>
</tr>
<tr>
<td>2</td>
<td>4131525</td>
<td>Hemorrhagic gastropathy</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>308904008</td>
</tr>
<tr>
<td>2</td>
<td>4204041</td>
<td>Hematemesis - cause unknown</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>413212000</td>
</tr>
<tr>
<td>2</td>
<td>4134808</td>
<td>Hemorrhagic duodenopathy</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>409506009</td>
</tr>
<tr>
<td>2</td>
<td>4260059</td>
<td>Hemorrhagic gastroenteritis</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>27281001</td>
</tr>
<tr>
<td>2</td>
<td>4099014</td>
<td>Duodenal ulcer with hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>1.5072E+14</td>
</tr>
<tr>
<td>2</td>
<td>46270145</td>
<td>Gastric hemorrhage due to atrophic gastritis</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>262843005</td>
</tr>
<tr>
<td>2</td>
<td>4096032</td>
<td>Duodenal hematoma</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>49232000</td>
</tr>
<tr>
<td>2</td>
<td>4174044</td>
<td>Chronic peptic ulcer with hemorrhage</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>262790002</td>
</tr>
<tr>
<td>2</td>
<td>4095555</td>
<td>Esophageal hematoma</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>1.0866E+15</td>
</tr>
<tr>
<td>2</td>
<td>46269904</td>
<td>Hemorrhage of duodenum co-occurent and due to diverticul</td>
<td></td>
<td></td>
<td>Clinical Finding</td>
<td>S</td>
<td>7.071E+12</td>
</tr>
</tbody>
</table>
Exercise: Find Standard Concept ID for Conditions

• Asthma
  317009

• Plague
  434271

• Ingrown toenail
  4065236  4290993

• Your favorite condition here
<table>
<thead>
<tr>
<th>ID</th>
<th>CODE</th>
<th>NAME</th>
<th>CLASS</th>
<th>CONCEPT</th>
<th>VALIDITY</th>
<th>DOMAIN</th>
<th>VOCAB</th>
</tr>
</thead>
<tbody>
<tr>
<td>2744127</td>
<td>0BD70ZZ</td>
<td>Excision of Left Main Bronchus, Open Approach</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744128</td>
<td>0BBF3ZX</td>
<td>Excision of Left Main Bronchus, Percutaneous Approach, Diagnostic</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744129</td>
<td>0BB73ZZ</td>
<td>Excision of Left Main Bronchus, Percutaneous Approach</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744130</td>
<td>0BB74ZX</td>
<td>Excision of Left Main Bronchus, Percutaneous Endoscopic Approach, Diagnostic</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744131</td>
<td>0BB74ZX</td>
<td>Excision of Left Main Bronchus, Percutaneous Endoscopic Approach</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744132</td>
<td>0BD77ZX</td>
<td>Excision of Left Main Bronchus, Via Natural or Artificial Opening, Diagnostic</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744133</td>
<td>0BD77ZX</td>
<td>Excision of Left Main Bronchus, Via Natural or Artificial Opening</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744134</td>
<td>0BD70ZX</td>
<td>Excision of Left Main Bronchus, Via Natural or Artificial Opening Endoscopic, Diagnostic</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744135</td>
<td>0BD70ZX</td>
<td>Excision of Left Main Bronchus, Via Natural or Artificial Opening Endoscopic</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744136</td>
<td>0BB80ZX</td>
<td>Excision of Left Upper Lobe Bronchus, Open Approach, Diagnostic</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744137</td>
<td>0BB80ZX</td>
<td>Excision of Left Upper Lobe Bronchus, Open Approach</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
<tr>
<td>2744138</td>
<td>0BB83ZX</td>
<td>Excision of Left Upper Lobe Bronchus, Percutaneous Approach, Diagnostic</td>
<td>ICD10PCS</td>
<td>Standard</td>
<td>Valid</td>
<td>Procedure</td>
<td>ICD10PCS</td>
</tr>
</tbody>
</table>
Lunch

Please return in 1 hour
Does it Work that Way with Drugs?

• Codes
  – NDC, GPI, Multilex, HCPCS, etc.

• Concepts
  – Drug products (Generic and Brand)
  – Drug ingredients
  – Drug Classes

• Relationships

• Ancestry
Let us find Warfarin

1. Find active compound Warfarin by keyword

```sql
SELECT * FROM concept WHERE lower(concept_name) = 'warfarin'
```

<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>21253542</td>
<td>Warfarin</td>
<td>Drug</td>
<td>dm+d</td>
<td>VTM</td>
<td>NULL</td>
<td>48603004</td>
</tr>
<tr>
<td>40772858</td>
<td>Warfarin</td>
<td>Measurement</td>
<td>LOINC</td>
<td>LOINC Hierarchy</td>
<td>C</td>
<td>LP16309-4</td>
</tr>
<tr>
<td>1847834</td>
<td>WARFARIN</td>
<td>Drug</td>
<td>DA_France</td>
<td>Ingredient</td>
<td>NULL</td>
<td>OMOP13547</td>
</tr>
<tr>
<td>4293218</td>
<td>WARFARIN</td>
<td>Drug</td>
<td>NDFRT</td>
<td>Pharma Preparation</td>
<td>NULL</td>
<td>N0000148057</td>
</tr>
<tr>
<td>35715598</td>
<td>WARFARIN</td>
<td>Drug</td>
<td>LPD_Australia</td>
<td>Ingredient</td>
<td>NULL</td>
<td>OMOP582219</td>
</tr>
<tr>
<td>43081820</td>
<td>warfarin</td>
<td>Drug</td>
<td>Multlex</td>
<td>Ingredient</td>
<td>NULL</td>
<td>2849</td>
</tr>
<tr>
<td>4187015</td>
<td>Warfarin</td>
<td>Drug</td>
<td>SNOMED</td>
<td>Substance</td>
<td>NULL</td>
<td>372758006</td>
</tr>
<tr>
<td>43343324</td>
<td>Warfarin</td>
<td>Drug</td>
<td>AMT</td>
<td>AU Substance</td>
<td>NULL</td>
<td>2714011000038109</td>
</tr>
<tr>
<td>4174989</td>
<td>Warfarin</td>
<td>Drug</td>
<td>SNOMED</td>
<td>Pharma/Biol Product</td>
<td>NULL</td>
<td>48603004</td>
</tr>
<tr>
<td>4325514</td>
<td>Warfarin</td>
<td>Drug</td>
<td>NDFRT</td>
<td>Chemical Structure</td>
<td>C</td>
<td>N0000006403</td>
</tr>
<tr>
<td>1310149</td>
<td>Warfarin</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
<td>S</td>
<td>11289</td>
</tr>
<tr>
<td>21600965</td>
<td>warfarin</td>
<td>Drug</td>
<td>ATC</td>
<td>ATC 5th</td>
<td>C</td>
<td>E01AA03</td>
</tr>
<tr>
<td>45618204</td>
<td>Warfarin</td>
<td>Drug</td>
<td>MeSH</td>
<td>Main Heading</td>
<td>NULL</td>
<td>D014859</td>
</tr>
</tbody>
</table>
Let us find Clopidogrel

1. Find drug product containing Clopidogrel by NDC code:
   Bristol Meyer Squibb's Plavix 75mg capsules: NDC 67544050474

\[
\text{SELECT} \ * \ \text{FROM} \ \text{concept} \ \text{WHERE} \ \text{concept}_\text{code} = '67544050474'
\]

<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>
<th>valid_start_date</th>
<th>valid_end_date</th>
<th>invalid_reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>67544050474</td>
<td>clopidogrel 75 MG Oral Tablet [Plavix]</td>
<td>Drug</td>
<td>NDC</td>
<td>11-digit NDC</td>
<td>NULL</td>
<td>67544050474</td>
<td>2014-07-01</td>
<td>2099-12-31</td>
<td>NULL</td>
</tr>
</tbody>
</table>

\[
\text{SELECT} \ * \ \text{FROM} \ \text{concept}_\text{relationship} \ \text{WHERE} \ \text{concept}_\text{id}_1 = 45867731 \\
\text{AND} \ \text{relationship}_\text{id} = 'Maps to'
\]

<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>45867731</td>
<td>1322185</td>
<td>Maps to</td>
<td>2015-01-29</td>
<td>2099-12-31</td>
<td>NULL</td>
</tr>
</tbody>
</table>

\[
\text{SELECT} \ * \ \text{FROM} \ \text{concept} \ \text{WHERE} \ \text{concept}_\text{id} = 1322185
\]

<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>
<th>valid_start_date</th>
<th>valid_end_date</th>
<th>invalid_reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>1322185</td>
<td>clopidogrel 75 MG Oral Tablet [Plavix]</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Branded Drug</td>
<td>S</td>
<td>213169</td>
<td>1970-01-01</td>
<td>2099-12-31</td>
<td>NULL</td>
</tr>
</tbody>
</table>
2. Find ingredient Clopidogrel as Ancestor of drug product

```sql
SELECT max_levels_of_separation, concept.*
FROM concept_ancestor
JOIN concept ON ancestor_concept_id = concept_id
WHERE descendant_concept_id = 1322185 /* clopidogrel 75 MG Oral Tablet [Plavix]*/
ORDER BY max_levels_of_separation
```
3. Check Descendants (other drug products containing Warfarin and Dabigatran)

```sql
SELECT max_levels_of_separation, concept.*
FROM concept_ancestor
JOIN concept ON descendant_concept_id = concept_id
WHERE ancestor_concept_id = 1310149 /* Warfarin or 1322185 Clopidogrel*/
ORDER BY max_levels_of_separation
```

<table>
<thead>
<tr>
<th>concept_id</th>
<th>concept_name</th>
<th>vocabulary_id</th>
<th>concept_class_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1310149</td>
<td>Warfarin</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>36221229</td>
<td>Jantoven Pill</td>
<td>RxNorm</td>
<td>Branded Dose Group</td>
</tr>
<tr>
<td>40163559</td>
<td>Warfarin Sodium 6 MG</td>
<td>RxNorm</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>40163544</td>
<td>Warfarin Sodium 3 MG [Jantoven]</td>
<td>RxNorm</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>21134746</td>
<td>Warfarin 0.2 MG/ML</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>21105414</td>
<td>Warfarin 5 MG/ML</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>36221228</td>
<td>Jantoven Oral Product</td>
<td>RxNorm</td>
<td>Branded Dose Group</td>
</tr>
<tr>
<td>40163565</td>
<td>Warfarin Sodium 7.5 MG</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>21115236</td>
<td>Warfarin 0.3 MG/ML</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>40163509</td>
<td>Warfarin Sodium 1 MG</td>
<td>RxNorm</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>21156284</td>
<td>1 ML Warfarin 0.02 MG/ML Oral Solution</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>21095537</td>
<td>Warfarin 0.3 MG/ML Oral Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21105427</td>
<td>Warfarin 0.4 MG/ML Oral Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21046557</td>
<td>Warfarin 1 MG/ML Oral Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>40093133</td>
<td>Warfarin Oral Tablet [Coumadin]</td>
<td>RxNorm</td>
<td>Branded Drug Form</td>
</tr>
<tr>
<td>40093134</td>
<td>Warfarin Oral Tablet [Jantoven]</td>
<td>RxNorm</td>
<td>Branded Drug Form</td>
</tr>
<tr>
<td>21077698</td>
<td>1 ML Warfarin 1 MG/ML Oral Solution</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>40163534</td>
<td>Warfarin Sodium 2.5 MG Oral Tablet</td>
<td>RxNorm</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>40163530</td>
<td>Warfarin Sodium 2 MG/ML Injectable Solution</td>
<td>RxNorm</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21066136</td>
<td>Warfarin 5 MG Oral Tablet [Marevan]</td>
<td>RxNorm Extension</td>
<td>Branded Drug</td>
</tr>
<tr>
<td>40163542</td>
<td>Warfarin Sodium 3 MG Oral Tablet [Jantoven]</td>
<td>RxNorm</td>
<td>Branded Drug Form</td>
</tr>
<tr>
<td>21116822</td>
<td>1 ML Warfarin 0.6 MG/ML Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>21175784</td>
<td>1 ML Warfarin 0.1 MG/ML Oral Solution</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>21175783</td>
<td>1 ML Warfarin 0.832 MG/ML Oral Solution</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>1322184</td>
<td>clopidogrel</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>21043471</td>
<td>clopidogrel Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Form</td>
</tr>
<tr>
<td>36229332</td>
<td>Plavix Pill</td>
<td>RxNorm</td>
<td>Branded Dose Group</td>
</tr>
<tr>
<td>21043470</td>
<td>clopidogrel Oral Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Form</td>
</tr>
<tr>
<td>21023802</td>
<td>clopidogrel Injectable Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Form</td>
</tr>
<tr>
<td>21023806</td>
<td>clopidogrel 5 MG</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>1322187</td>
<td>clopidogrel 75 MG</td>
<td>RxNorm</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>21141600</td>
<td>clopidogrel 1 MG/ML</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>36222254</td>
<td>clopidogrel Oral Product</td>
<td>RxNorm</td>
<td>Clinical Dose Group</td>
</tr>
<tr>
<td>21092477</td>
<td>clopidogrel 5 MG/ML</td>
<td>RxNorm Extension</td>
<td>Clinical Drug Comp</td>
</tr>
<tr>
<td>21177192</td>
<td>100 ML clopidogrel 1 MG/ML Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>21047899</td>
<td>1 ML clopidogrel 5 MG/ML Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>21121870</td>
<td>clopidogrel 5 MG/ML Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21063106</td>
<td>clopidogrel 75 MG Oral Tablet [Grepid]</td>
<td>RxNorm Extension</td>
<td>Branded Drug</td>
</tr>
<tr>
<td>1322190</td>
<td>clopidogrel 300 MG Oral Tablet [Plavix]</td>
<td>RxNorm</td>
<td>Branded Drug Form</td>
</tr>
<tr>
<td>21121869</td>
<td>clopidogrel 75 MG Injectable Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21053280</td>
<td>clopidogrel 6 MG Injectable Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21023810</td>
<td>clopidogrel 4 MG Injectable Solution</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21106783</td>
<td>1 ML clopidogrel 1 MG/ML Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
<tr>
<td>19075601</td>
<td>clopidogrel 75 MG Oral Tablet</td>
<td>RxNorm</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>21102364</td>
<td>clopidogrel 1 MG/ML Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Clinical Drug</td>
</tr>
<tr>
<td>40095879</td>
<td>clopidogrel Oral Tablet [Plavix]</td>
<td>RxNorm</td>
<td>Branded Drug Form</td>
</tr>
<tr>
<td>40095878</td>
<td>clopidogrel Oral Tablet</td>
<td>RxNorm</td>
<td>Clinical Drug Form</td>
</tr>
<tr>
<td>21088717</td>
<td>100 ML clopidogrel 15 MG/ML Oral Suspension</td>
<td>RxNorm Extension</td>
<td>Quant Clinical Drug</td>
</tr>
</tbody>
</table>
Find members of Drug Classes

4. Check Ingredient Descendants of Drug Class Anticoagulants

```
SELECT max_levels_of_separation, concept.*
FROM concept_ancestor
JOIN concept ON descendant_concept_id = concept_id
WHERE ancestor_concept_id = 21600961 /* ATC Antithromboic Agent */
  AND concept_class_id = 'Ingredient'
ORDER BY max_levels_of_separation
```

<table>
<thead>
<tr>
<th>concept_id</th>
<th>concept_name</th>
<th>domain_id</th>
<th>vocabulary_id</th>
<th>concept_class_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>46275677</td>
<td>cangrelor</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>45892847</td>
<td>edoxaban</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>1322184</td>
<td>clopidogrel</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>44818499</td>
<td>vorapaxar</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>43013024</td>
<td>apixaban</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>42898933</td>
<td>defibrotide</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>42801108</td>
<td>Protein C</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>40241331</td>
<td>rivaroxaban</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>1310149</td>
<td>Warfarin</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>40241186</td>
<td>Ticagrelor</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>40228152</td>
<td>dabigatran etexilate</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>40163718</td>
<td>prasugrel</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>35604848</td>
<td>selexipag</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>19136187</td>
<td>Streptokinase</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>19136187</td>
<td>Penicillin</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
<tr>
<td>19136187</td>
<td>Penicillin</td>
<td>Drug</td>
<td>RxNorm</td>
<td>Ingredient</td>
</tr>
</tbody>
</table>
Exercise:
Find Standard Concept ID

Metformin  1503297

Tolazamide  1502809

Telmisartan  1317640

Your favorite ingredient here
Exercise:
Find Standard Concept ID

Your favorite drug here
Common Data Model

In depth discussion of model & era discussion
CDM Version 6 Key Domains
OMOP CDM Principles

• Patient centric

• Vocabulary and Data Model are blended

• Domain-oriented concepts

• Accommodates data from various sources

• Preserves data provenance

• Extendable & Evolving

• Database Platform Independent
## OMOP CDM Standard Domain Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description &amp; Purpose</th>
<th>Field Name Convention</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient centric</td>
<td>Every domain table has <strong>patient identifier</strong>. Patient data can be retrieved independently from other domains.</td>
<td><code>person_id</code></td>
<td><code>person_id</code> 123</td>
</tr>
<tr>
<td>Unique domain identifiers</td>
<td>Every domain table has a unique primary key to identify domain entities.</td>
<td><code>&lt;entity&gt;_id</code></td>
<td><code>condition_occurrence_id</code> 470985</td>
</tr>
<tr>
<td>Standard concept from a respective vocabulary domain</td>
<td>Integration with the Vocabulary. Foreign key into the Standard Vocabulary for <strong>Standard Concept</strong>.</td>
<td><code>&lt;entity&gt;_concept_id</code></td>
<td><code>condition_concept_id</code> 313217 (SNOMED “Atrial Fibrillation”)</td>
</tr>
<tr>
<td>Source value</td>
<td>Provenance. Verbatim information from the source data, <strong>not to be used</strong> by any standard analytics.</td>
<td><code>&lt;entity&gt;_source_value</code></td>
<td><code>condition_source_value</code> 427.31 (ICD9CM “Atrial Fibrillation”)</td>
</tr>
<tr>
<td>Source concept from a respective vocabulary domain</td>
<td>Provenance. Foreign key into Standard Vocabulary for <strong>Source Concept</strong>.</td>
<td><code>&lt;entity&gt;_source_concept_id</code></td>
<td><code>condition_source_concept_id</code> 44821957 (ICD9CM “Atrial Fibrillation”)</td>
</tr>
<tr>
<td>Source type</td>
<td>Provenance. Foreign key into Vocabulary for the <strong>origin of the data</strong>.</td>
<td><code>&lt;entity&gt;_type_concept_id</code></td>
<td><code>condition_type_concept_id</code> 38000199 (“Inpatient header – primary”)</td>
</tr>
</tbody>
</table>
A Patient’s Story: Lauren

Lauren’s story

“Every step of this painful journey I’ve had to convince everyone how much pain I was in.”

“My first surgery taught me that I had to be very patient with my recovery and very patient with myself in general.”

https://www.endometriosis-uk.org/laurens-story
What data do we have?

• Guided Exercise:
  – Where and how do we think Lauren’s data is generated?
  – Where do we think Lauren’s data could go into the CDM?
What data do we have?

Lauren’s Timeline

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

-3 Years
-2 Years
-1 Years

-2 Weeks
-3 Days
Day 0

-3 Years -2 Years -1 Years

Lauren’s Timeline

-3 Years -2 Years -1 Years

-2 Weeks
-3 Days
Day 0
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.
PERSON

- Need to create one unique record per person
- No history of location/demographics: need to select latest available
- Year of birth required...day/month optional
- Foreign key to the LOCATION, PROVIDER, and CARE_SITE table that contains one record
Lauren’s Timeline

-3 Years  -2 Years  -1 Years

/ / /

-2 Weeks

/ / /

-3 Days

Day 0

Endometriosis

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

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

What data do we have?

What data do we have?
## PERSON

### Sample of Table’s Columns

<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</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</td>
</tr>
<tr>
<td>ethnicity_concept_id</td>
<td>0</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>
<tr>
<td>ethnicity_source_value</td>
<td>NULL</td>
</tr>
</tbody>
</table>
OBSERVATION_PERIOD

• Spans of time where data source has capture of data

• One person may have multiple periods if there is interruption in data capture

• Required to run analytical methods

• Challenge: determine observation periods based on the source data
Lauren’s Timeline

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

What data do we have?

- dysmenorrhea
- abdominal pain
- missed work
- acetaminophen
- acetaminophen
- acetaminophen

GP visit
ultrasound
cyst of ovary

Hospital Visit

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

Endometriosis
# OBSERVATION_PERIOD

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>observation_period_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>observation_periods_start_date</td>
<td>2000-01-01</td>
</tr>
<tr>
<td>observation_periods_start_date</td>
<td>2010-12-31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>observation_period_id</td>
<td>2</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>observation_periods_start_date</td>
<td>2012-01-01</td>
</tr>
<tr>
<td>observation_periods_start_date</td>
<td>2013-12-31</td>
</tr>
</tbody>
</table>
VISIT_OCCURRENCE

- Visits are ‘Encounters’
- Contains spans of time where a person receives medical services

Visit Types
- Emergency room
- Inpatient
- Inpatient/Emergency
- Outpatient
- Long-term care
What data do we have?

Lauren's Timeline

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

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

Endometriosis

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

Dysmenorrhea
Abdominal pain
Missed work
Missed work
Acetaminophen
Acetaminophen
Acetaminophen

GP visit
Ultrasound
Pelvic exam
Cyst of ovary

Hospital Visit
Severe pain
Temp 103°F
Ultrasound
Ambulance
Bloated abdomen
Ascites
Surgery
Endometrioma
## 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>

**Lauren’s ID**  
**Outpatient Visit**

<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>

**Lauren’s ID**  
**Inpatient Visit**

(sample of table’s columns)
• Records suggesting the presence of a disease or medical condition stated as a diagnosis, a sign or a symptom

• Examples:
  – Billing diagnosis
  – Problem list
Lauren’s Timeline

**-3 Years**
- abdominal pain
- missed work
- acetaminophen

**-2 Years**
- acetaminophen
- acetaminophen

**-1 Years**
- acetaminophen
- acetaminophen

**-2 Weeks**
- ulcerative colitis
- pelvic exam
- ultrasound
- cyst of ovary

**-3 Days**
- ulcerative colitis
- ultrasound
- severe pain
- ambulance
- ascites
- surgery

**Day 0**
- ulcerative colitis

**Hospital Visit**
- temp 103°F
- ultrasound
- severe pain
- ambulance
- Bloated abdomen
- ascites
- surgery
- endometrioma

What data do we have?

**Endometriosis and Dysmenorrhea**

Endometriosis:
- Cyst of ovary
- Severe pain

Dysmenorrhea:
- Abdominal pain
- Acetaminophen
- Missed work

Lauren’s Timeline:
- 3 years
- 2 years
- 1 year

Medical History:
- GP visit
- Ultrasound
- Pelvic exam
- Cyst of ovary

Symptoms:
- Abdominal pain
- Ulcerative colitis
- Bloated abdomen
- Ascites

Diagnosis:
- Endometriosis

 Corona's diagnosis

 Surgical intervention

 Emergency room intervention
## 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</td>
</tr>
<tr>
<td>condition_concept_id</td>
<td>433527</td>
</tr>
<tr>
<td>condition_start_date</td>
<td>2008-04-24</td>
</tr>
<tr>
<td>condition_start_datetime</td>
<td>2008-04-24 00:00:00</td>
</tr>
<tr>
<td>condition_type_concept_id</td>
<td>38000183</td>
</tr>
<tr>
<td>condition_status_concept_id</td>
<td>4230359</td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>2</td>
</tr>
<tr>
<td>condition_source_value</td>
<td>6171</td>
</tr>
<tr>
<td>condition_source_concept_id</td>
<td>44832501</td>
</tr>
</tbody>
</table>

Lauren’s ID

Endometriosis

Inpatient detail - primary

ICD9, missing decimal

Endometriosis of ovary
DRUG_EXPOSURE

• Records about the utilization of a drug when ingested or otherwise introduced into the body

• Data sources:
  – Pharmacy dispensing
  – Prescriptions written
  – Medication history

• If drug is represented as a procedure, the OMOP Vocabulary realigns as drug
Lauren's Timeline

-3 Years
-2 Years
-1 Years

-2 Weeks
-3 Days
Day 0

What data do we have?

- dysmenorrhea
- abdominal pain
- acetaminophen
- missed work
- missed work
- acetaminophen
- acetaminophen

GP visit
pelvic exam
ultrasound
cyst of ovary

Hospital Visit

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

Endometriosis
## DRUG_EXPOSURE

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug_exposure_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>person_id</td>
<td>Lauren’s ID</td>
</tr>
<tr>
<td>drug_concept_id</td>
<td>40162494</td>
</tr>
<tr>
<td>drug_exposure_start_date</td>
<td>2007-02-01</td>
</tr>
<tr>
<td>drug_exposure_end_date</td>
<td>2007-02-08</td>
</tr>
<tr>
<td>verbatim_end_date</td>
<td>NULL</td>
</tr>
<tr>
<td>drug_type_concept_id</td>
<td>38000183</td>
</tr>
<tr>
<td>refills</td>
<td>0</td>
</tr>
<tr>
<td>quantity</td>
<td>14</td>
</tr>
<tr>
<td>days_supply</td>
<td>7</td>
</tr>
<tr>
<td>drug_source_value</td>
<td>54348001301</td>
</tr>
<tr>
<td>drug_source_concept_id</td>
<td>45904353</td>
</tr>
</tbody>
</table>

Table example:
- **Prescription dispensed in pharmacy**
- **NDC 11-digit code**
- **Drug_exposure_start_date + days_supply**
- **Sample of table’s columns**
PROCEDURE_OCCURRENCE

- Contains records of activities or processes ordered by, or carried out by, a healthcare provider on the patient to have a diagnostic or therapeutic purpose.

- Vocabularies include CPT-4, HCPCS, ICD-9 Procedures, ICD-10 Procedures, LOINC, SNOMED.

- Procedures have the least standardized vocabularies that causes some redundancy.
Lauren’s Timeline

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?

Endometriosis

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

Lauren’s Timeline

What data do we have?
## 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</td>
</tr>
<tr>
<td>procedure_concept_id</td>
<td>2211740</td>
</tr>
<tr>
<td>procedure_date</td>
<td>2008-04-08</td>
</tr>
<tr>
<td>procedure_type_concept_id</td>
<td>38000267</td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>1</td>
</tr>
<tr>
<td>procedure_source_value</td>
<td>76700</td>
</tr>
<tr>
<td>procedure_source_concept_id</td>
<td>2211740</td>
</tr>
</tbody>
</table>

- **Lauren’s ID**: Ultrasound, abdominal, real time with image documentation; complete
- **CPT4**: Ultrasound, abdominal, real time with image documentation; complete

*sample of table’s columns*
### MEASUREMENT

- Contains records of Measurement, i.e. structured values (numerical or categorical) obtained through systematic and standardized examination or testing of a Person or Person's sample.

- Data sources: structured, quantitative measures, such as laboratory tests.

- Measures have associated units.

```
<table>
<thead>
<tr>
<th>Measurement Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>person_id</td>
</tr>
<tr>
<td>measurement_concept_id</td>
</tr>
<tr>
<td>measurement_date</td>
</tr>
<tr>
<td>measurement_datetime</td>
</tr>
<tr>
<td>measurement_type_concept_id</td>
</tr>
<tr>
<td>operator_concept_id</td>
</tr>
<tr>
<td>value_as_number</td>
</tr>
<tr>
<td>value_as_concept_id</td>
</tr>
<tr>
<td>unit_concept_id</td>
</tr>
<tr>
<td>range_low</td>
</tr>
<tr>
<td>range_high</td>
</tr>
<tr>
<td>provider_id</td>
</tr>
<tr>
<td>visit_occurrence_id</td>
</tr>
<tr>
<td>measurement_source_value</td>
</tr>
<tr>
<td>measurement_source_concept_id</td>
</tr>
<tr>
<td>unit_source_value</td>
</tr>
<tr>
<td>value_source_value</td>
</tr>
</tbody>
</table>
```
What data do we have?

Lauren’s Timeline

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

-3 Years

-2 Years

-1 Years

-2 Weeks

-3 Days

Day 0

Endometriosis

dysmenorrhea

abdominal pain

missed work

acetaminophen

acetaminophen

acetaminophen

GP visit

pelvic exam

ultrasound

cyst of ovary

Hospital Visit

severe pain

ultrasound

ambulance

Bloated abdomen

ascites

surgery

endometrioma

temp 103°F

103°F

107

What data do we have?

-3 Years

-2 Years

-1 Years

-2 Weeks

-3 Days

Day 0

Endometriosis
## 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*
OBSERVATION

- Captures clinical facts about a Person obtained in the context of examination, questioning or a procedure.

- Any data that cannot be represented by any other domains, such as social and lifestyle facts, medical history, family history, etc. are recorded here.

- Instrument for CDM extension, playpen.
Lauren’s Timeline

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

What data do we have?

Endometriosis

- dysmenorrhea
- abdominal pain
- acetaminophen
- missed work
- acetaminophen
- acetaminophen
- GP visit
- pelvic exam
- ultrasound
- cyst of ovary
- ultrasound
- endometrioma
- ambulance
- surgery
- ascites
- bloated abdomen
- temp 103°F
- severe pain
- missed work
- missed work
- abdominal pain
- bloated abdomen
- ascites
- ultrasound
# Observation

## Example of Table's Columns

<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>person_id</td>
<td><strong>Lauren's ID</strong></td>
</tr>
<tr>
<td>observation_concept_id</td>
<td>0</td>
</tr>
<tr>
<td>observation_concept_id</td>
<td><strong>No matching concept</strong></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>observation_type_concept_id</td>
<td><strong>Patient reported</strong></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>
<tr>
<td>observation_source_concept_id</td>
<td><strong>No matching concept</strong></td>
</tr>
</tbody>
</table>
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

- **Note**
  - NLP
DRUG_ERA

- Standardized inference of length of exposure to product for all active ingredients
- Derived from records in DRUG_EXPOSURE under certain rules to produce continuous Drug Eras
### DRUG_EXPOSURE

**Acetaminophen 500 MG / Hydrocodone Bitartrate 5 MG Oral Tablet**

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug_exposure_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>drug_concept_id</td>
<td>40162494</td>
</tr>
<tr>
<td>drug_exposure_start_date</td>
<td>2007-02-01</td>
</tr>
<tr>
<td>drug_exposure_end_date</td>
<td>2007-02-08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug_exposure_id</td>
<td>2</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>drug_concept_id</td>
<td>40162494</td>
</tr>
<tr>
<td>drug_exposure_start_date</td>
<td>2007-02-10</td>
</tr>
<tr>
<td>drug_exposure_end_date</td>
<td>2007-02-17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug_era_id</td>
<td>1</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>drug_concept_id</td>
<td>1125315</td>
</tr>
<tr>
<td>drug_era_start_date</td>
<td>2007-02-01</td>
</tr>
<tr>
<td>drug_era_end_date</td>
<td>2007-02-17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>drug_era_id</td>
<td>2</td>
</tr>
<tr>
<td>person_id</td>
<td>123456</td>
</tr>
<tr>
<td>drug_concept_id</td>
<td>1174888</td>
</tr>
<tr>
<td>drug_era_start_date</td>
<td>2007-02-01</td>
</tr>
<tr>
<td>drug_era_end_date</td>
<td>2007-02-17</td>
</tr>
</tbody>
</table>

**Sample of table’s columns**

*Acetaminophen*

*Hydrocodone*
Illustrating inferences needed within longitudinal pharmacy claims data for one patient

Person Timeline

NDC: 00179198801
Lisinopril 5 MG Oral Tablet

How do we handle reversals?

How do we handle NDC change?

How do we handle overlap?

NDC: 00310013010
Zestril 5 MG TABLET

How do we handle discontinuation?

30d gap

NDC: 00038013134
Lisinopril 10 MG Oral Tablet [Zestril]

NDC: 00038013210
Lisinopril 20 MG Oral Tablet [Zestril]

How do we handle change in dose?

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

How do we handle gaps?

How do we handle combination products?

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](https://ohdsi.github.io/TheBookOfOhdsi/)

- Please contact Clair Blacketer ([mblacke@its.jnj.com](mailto:mblacke@its.jnj.com)) for more information
Break

Please return in 15 minutes
CDM Examples

Leveraging OHDSI Tools
(GitHub /Forums/
Working Group)
Exercises
ETL: Real world scenario

**PharMetrics Plus**

<table>
<thead>
<tr>
<th>CLAIMS</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pat_id</td>
<td>claimno</td>
<td>from_dt</td>
<td>to_dt</td>
<td>diagprc_ind</td>
<td>Diag_admit</td>
<td>diag1</td>
</tr>
<tr>
<td>05917921689</td>
<td>IPA333393946</td>
<td>1/5/2006</td>
<td>1/5/2006</td>
<td>1</td>
<td>41071</td>
<td>41071</td>
</tr>
</tbody>
</table>

**LRx/Dx**

<table>
<thead>
<tr>
<th>MEDICAL_CLAIMS</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>md_clm_id</td>
<td>ims_pat_nbr</td>
<td>dt_of_service</td>
<td>rxer_id</td>
<td>diag_cd</td>
<td></td>
</tr>
<tr>
<td>95963982102</td>
<td>80445908</td>
<td>8/1/2012 0:00</td>
<td>680488</td>
<td>41071</td>
<td></td>
</tr>
</tbody>
</table>

**German DA**

**Problem Events**

<table>
<thead>
<tr>
<th>db_country</th>
<th>international_practice_num</th>
<th>international_doctor_num</th>
<th>international_patient_num</th>
<th>age</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>GE6326</td>
<td>GE8784</td>
<td>GE46478747</td>
<td></td>
</tr>
</tbody>
</table>

**Diagnosis**

<table>
<thead>
<tr>
<th>db_country</th>
<th>international_diagnosis_num</th>
<th>diagnosis_num</th>
<th>icd10_4_code</th>
<th>icd10_3_text</th>
<th>diagnosis_confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>GE2397573</td>
<td>2397573</td>
<td>l21.4</td>
<td>Non-ST elevation (NSTEMI) myocardial infarction</td>
<td>Confirmed</td>
</tr>
</tbody>
</table>

**Ambulatory EMR**

**Problem**

<table>
<thead>
<tr>
<th>Patient_id_synth</th>
<th>Diag_dt</th>
<th>icd10_cd</th>
</tr>
</thead>
<tbody>
<tr>
<td>271138</td>
<td>4/11/2013</td>
<td>214</td>
</tr>
</tbody>
</table>

4 real observational databases, all containing an inpatient admission for a patient with a diagnosis of ‘acute subendocardial infarction’

- Not a single table name the same...
- Not a single variable name the same....
- Different table structures (rows vs. columns)
- Different conventions (with and without decimal points)
- Different coding schemes (ICD9 vs. ICD10)
What does it mean to ETL to OMOP CDM?
Standardize **structure** and **content**

<table>
<thead>
<tr>
<th>pat_id</th>
<th>claimno</th>
<th>from_dt</th>
<th>to_dt</th>
<th>diagprc_ind</th>
<th>Diag_admit</th>
</tr>
</thead>
<tbody>
<tr>
<td>05917921689</td>
<td>IPA333393946</td>
<td>1/5/2006</td>
<td>1/5/2006</td>
<td>1</td>
<td>41071</td>
</tr>
</tbody>
</table>

Structure optimized for large-scale analysis for clinical characterization, population-level estimation, and patient-level prediction.

<table>
<thead>
<tr>
<th>PERSON_ID</th>
<th>CONDITION_START_DATE</th>
<th>CONDITION_SOURCE_VALUE</th>
<th>CONDITION_TYPE_CONCEPT_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>05917921689</td>
<td>1/5/2006</td>
<td>41071</td>
<td>Inpatient claims - primary position</td>
</tr>
</tbody>
</table>

Content using international vocabulary standards that can be applied to any data source.

<table>
<thead>
<tr>
<th>PERSON_ID</th>
<th>CONDITION_START_DATE</th>
<th>CONDITION_SOURCE_VALUE</th>
<th>CONDITION_TYPE_CONCEPT_ID</th>
<th>CONDITION_SOURCE_CONCEPT_ID</th>
<th>CONDITION_CONCEPT_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>05917921689</td>
<td>1/5/2006</td>
<td>41071</td>
<td>Inpatient claims - primary position</td>
<td>44825429</td>
<td>444406</td>
</tr>
</tbody>
</table>
OMOP CDM = Standardized structure: same tables, same fields, same datatypes, same conventions across disparate sources

- Consistent structure optimized for large-scale analysis
- Structure preserves all source content and provenance

PharMetrics Plus: CONDITION_OCCURRENCE

<table>
<thead>
<tr>
<th>PERSON_ID</th>
<th>CONDITION_START_DATE</th>
<th>CONDITION_SOURCE_VALUE</th>
<th>CONDITION_TYPE_CONCEPT_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>157033702</td>
<td>1/5/2006</td>
<td>41071</td>
<td>Inpatient claims - primary position</td>
</tr>
</tbody>
</table>

LRX/DX: CONDITION_OCCURRENCE

<table>
<thead>
<tr>
<th>PERSON_ID</th>
<th>CONDITION_START_DATE</th>
<th>CONDITION_SOURCE_VALUE</th>
<th>CONDITION_TYPE_CONCEPT_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>80445908</td>
<td>8/1/2012</td>
<td>41071</td>
<td>Primary Condition</td>
</tr>
</tbody>
</table>

German DA: CONDITION_OCCURRENCE

<table>
<thead>
<tr>
<th>PERSON_ID</th>
<th>CONDITION_START_DATE</th>
<th>CONDITION_SOURCE_VALUE</th>
<th>CONDITION_TYPE_CONCEPT_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>46478747</td>
<td>11/19/2014</td>
<td>I21.4</td>
<td>EHR problem list entry</td>
</tr>
</tbody>
</table>

Ambulatory EMR: CONDITION_OCCURRENCE

<table>
<thead>
<tr>
<th>PERSON_ID</th>
<th>CONDITION_START_DATE</th>
<th>CONDITION_SOURCE_VALUE</th>
<th>CONDITION_TYPE_CONCEPT_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>271138</td>
<td>4/11/2013</td>
<td>I214</td>
<td>Primary Condition</td>
</tr>
</tbody>
</table>
OMOP CDM = Standardized content: common vocabularies across disparate sources

- Standardize across vocabularies to a common referent standard (ICD9/10→SNOMED)
- Source codes mapped into each domain standard so that now you can talk across different languages
- Standardize source codes to be uniquely defined across all vocabularies
- No more worries about formatting or code overlap

<table>
<thead>
<tr>
<th>PharMetrics Plus: CONDITION_OCCURRENCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON_ID</td>
<td>CONDITION_START_DATE</td>
</tr>
<tr>
<td>05917921689</td>
<td>1/5/2006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LRx/Dx: CONDITION_OCCURRENCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON_ID</td>
<td>CONDITION_START_DATE</td>
</tr>
<tr>
<td>80445908</td>
<td>8/1/2012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>German DA : CONDITION_OCCURRENCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON_ID</td>
<td>CONDITION_START_DATE</td>
</tr>
<tr>
<td>6478747</td>
<td>11/19/2014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ambulatory EMR : CONDITION_OCCURRENCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON_ID</td>
<td>CONDITION_START_DATE</td>
</tr>
<tr>
<td>271138</td>
<td>4/11/2013</td>
</tr>
</tbody>
</table>
Data Used for Demonstration

• Medicare Claims Synthetic Public Use Files (SynPUFs)
  – synthetic US Medicare insurance claims database
  – Medicare is a government based insurance program for primarily 65 and older but also individuals with disabilities
  – SynPUF not for research but rather demonstration/development purposes
  – Has been converted to the Common Data Model

Data Used for Demonstration

- Five types of data:

<table>
<thead>
<tr>
<th>DE-SynPUF</th>
<th>Unit of record</th>
<th>Number of Records 2008</th>
<th>Number of Records 2009</th>
<th>Number of Records 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beneficiary Summary</strong></td>
<td>Beneficiary</td>
<td>2,326,856</td>
<td>2,291,320</td>
<td>2,255,098</td>
</tr>
<tr>
<td><strong>Inpatient Claims</strong></td>
<td>claim</td>
<td>547,800</td>
<td>504,941</td>
<td>280,081</td>
</tr>
<tr>
<td><strong>Outpatient Claims</strong></td>
<td>claim</td>
<td>5,673,808</td>
<td>6,519,340</td>
<td>3,633,839</td>
</tr>
<tr>
<td><strong>Carrier Claims</strong></td>
<td>claim</td>
<td>34,276,324</td>
<td>37,304,993</td>
<td>23,282,135</td>
</tr>
<tr>
<td><strong>Prescription Drug Events (PDE)</strong></td>
<td>event</td>
<td>39,927,827</td>
<td>43,379,293</td>
<td>27,778,849</td>
</tr>
</tbody>
</table>

SynPUF High Level Diagram

- Beneficiary Summary
  - Inpatient Claims
  - Outpatient Claims
  - Carrier Claims
  - Prescription Drug Events (PDE)
Mapping SynPUF to CDM

SYNPUF DIAGRAM

- **Beneficiary Summary**
  - Inpatient Claims
  - Outpatient Claims
  - Carrier Claims
  - Prescription Drug Events (PDE)

SYNPUF

CDM

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
OHDSI in a Box
CDM Database: pgAdmin III New Server

- Click on pgAdmin
CDM Database: Connect

- Password: ohdsi
CDM Database: Open SQL Sheet
CDM Database: Ready

set search_path to 'public', 'ohdsi';
Some Example Questions

Finding Warfarin

New Users of Warfarin

New Users of Warfarin who are >=65?

New Users of Warfarin with prior Atrial Fibrillation?
Warfarin Exposure

• Warfarin is a blood thinner that is used to treat/prevent blood clots.
  
  – Where do you find drug data in the CDM?
  
  – What codes do I use to define drugs?
Where are Drug Exposures in the CDM?

Captures records about the utilization of a drug when ingested or otherwise introduced into the body.
How do I define Warfarin?

• When raw data is transformed into the CDM raw source codes are transformed into standard OMOP Vocabulary concepts

• In the CDM, we no longer care what source codes existed in the raw data, we just need to use concept identifiers

• We can use the OMOP Vocabulary to identify all concepts that contain the ingredient warfarin
How do I define Warfarin?

• Writing SQL Statement

• OHDSI Tool ATLAS
Some Example Questions

Ex 1
Finding Warfarin

Ex 2
New Users of Warfarin

Ex 3
New Users of Warfarin who are >= 65?

Ex 4
New Users of Warfarin with prior Atrial Fibrillation?
Finding Warfarin

/* (Exercise 0) Finding Warfarin */

/* Just looking for the ingredient concept */
SELECT COUNT(DISTINCT de.PERSON_ID)
FROM DRUG_EXPOSURE de
WHERE DRUG_CONCEPT_ID = 1310149 /* warfarin */;

/* Looking for drugs associated with the ingredient */
SELECT COUNT(DISTINCT de.PERSON_ID)
FROM DRUG_EXPOSURE de
WHERE de.DRUG_CONCEPT_ID IN ( 
    SELECT DESCENDANT_CONCEPT_ID
    FROM CONCEPT>Ancestor
    WHERE ANCESTOR_CONCEPT_ID = 1310149 /* warfarin */
); 

/* Looking for anticoagulants, a class of drugs warfarin belongs */
SELECT COUNT(DISTINCT de.PERSON_ID)
FROM DRUG_EXPOSURE de
WHERE de.DRUG_CONCEPT_ID IN ( 
    SELECT DESCENDANT_CONCEPT_ID
    FROM CONCEPT>Ancestor
    WHERE ANCESTOR_CONCEPT_ID = 4283987 /* ANTICOAGULANTS (VA Class) */
);
Finding Warfarin

0 individuals

22,247 individuals
Finding Warfarin

Ex 0

0 individuals

22,247 individuals

32,955 individuals
How do I define new users of a drug?

Someone who has recently started taking the drug, typically with a 6 or 12 month wash out
How do I define new users of a drug?

Someone who has recently started taking the drug, typically with a 6 or 12 month wash out.
What is Needed in the CDM?

• OMOP Vocabulary
to find the concepts

• CDM Table DRUG_EXPOSURE
to find individuals with exposure

• CDM Table OBSERVATION_PERIOD
to know people’s time within the database
New Users of Warfarin

```sql
WITH CTE_DRUG_INDEX AS (
    SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE
    FROM DRUG_EXPOSURE de
    WHERE de.DRUG_CONCEPT_ID IN (
        SELECT DESCENDANT_CONCEPT_ID
        FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/
    )
    GROUP BY de.PERSON_ID
)

SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE,
    (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX
FROM CTE_DRUG_INDEX i
JOIN OBSERVATION_PERIOD op
    ON op.PERSON_ID = i.PERSON_ID
    AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180
ORDER BY i.PERSON_ID
```
Step 1: Get the codes you need

```sql
WITH CTE_DRUG_INDEX AS (  
    SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE  
    FROM DRUG_EXPOSURE de  
    WHERE de.DRUG_CONCEPT_ID IN (  
        SELECT DESCENDANT_CONCEPT_ID  
        FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/  
    )  
    GROUP BY de.PERSON_ID
  )  
SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE,  
  (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX  
FROM CTE_DRUG_INDEX i  
JOIN OBSERVATION_PERIOD op  
  ON op.PERSON_ID = i.PERSON_ID  
  AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE  
WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180  
ORDER BY i.PERSON_ID
```
Step 2: Find Drug Exposures

/* Exercise 1: Warfarin New Users */

WITH CTE_DRUG_INDEX AS (  
    SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE  
    FROM DRUG_EXPOSURE de  
    WHERE de.DRUG_CONCEPT_ID IN (  
        SELECT DESCENDANT_CONCEPT_ID  
        FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/  
    )  
    GROUP BY de.PERSON_ID  
)

SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE,  
    (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX  
FROM CTE_DRUG_INDEX i  
JOIN OBSERVATION_PERIOD op  
    ON op.PERSON_ID = i.PERSON_ID  
    AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE  
WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180  
ORDER BY i.PERSON_ID
Step 3: Find New Users

WITH CTE_DRUG_INDEX AS (  
  SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE  
  FROM DRUG_EXPOSURE de  
  WHERE de.DRUG_CONCEPT_ID IN (  
    SELECT DESCENDANT_CONCEPT_ID  
    FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/  
  )  
  GROUP BY de.PERSON_ID
)

SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE, (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX
FROM CTE_DRUG_INDEX i
  JOIN OBSERVATION_PERIOD op
    ON op.PERSON_ID = i.PERSON_ID
    AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180
ORDER BY i.PERSON_ID
New Users of Warfarin

WITH CTE_DRUG_INDEX AS (
    SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE
    FROM DRUG_EXPOSURE de
    WHERE de.DRUG_CONCEPT_ID IN (
        SELECT DESCENDANT_CONCEPT_ID
        FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/
    )
    GROUP BY de.PERSON_ID
)

SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE, (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX
FROM CTE_DRUG_INDEX i
JOIN OBSERVATION_PERIOD op
    ON op.PERSON_ID = i.PERSON_ID
    AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180
ORDER BY i.PERSON_ID
New Users of Warfarin

Try running this on your own!

How many people do you get?

15,685 individuals
Some Example Questions

**Ex 1**
Finding Warfarin

**Ex 2**
New Users of Warfarin

**Ex 3**
New Users of Warfarin who are >=65?

**Ex 4**
New Users of Warfarin with prior Atrial Fibrillation?
How do I define new users of warfarin who are >=65?

Someone who has recently started taking the drug, typically with a 6 or 12 month wash out period.

>=65 years old

6 months

index drug

time in database

Ex 3
What is Needed in the CDM?

- **OMOP Vocabulary**
  to find the concepts

- **DRUG_EXPOSURE**
  to find individuals with exposure

- **OBSERVATION_PERIOD**
  to know people’s time within the database

- **PERSON**
  to know year of birth
Step 1: Start with the previous query

```sql
WITH CTE_DRUG_INDEX AS (
    SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE
    FROM DRUG_EXPOSURE de
    WHERE de.DRUG_CONCEPT_ID IN (
        SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/
    )
    GROUP BY de.PERSON_ID
)

SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OB

FROM CTE_DRUG_INDEX i
JOIN OBSERVATION_PERIOD op
    ON op.PERSON_ID = i.PERSON_ID
    AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
JOIN PERSON p
    ON p.PERSON_ID = i.PERSON_ID
WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180
AND EXTRACT(YEAR FROM i.INDEX_DATE)-p.YEAR_OF_BIRTH >= 65
ORDER BY i.PERSON_ID
```
Step 2: Add the Person Table to calculate age

```sql
WITH CTE_DRUG_INDEX AS (  
  SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE  
  FROM DRUG EXPOSURE de  
  WHERE de.DRUG_CONCEPT_ID IN (  
    SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/  
  )  
  GROUP BY de.PERSON_ID  
)  
SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE,  
  (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX,  
  EXTRACT(YEAR FROM i.INDEX_DATE)-p.YEAR_OF_BIRTH AS AGE_AT_INDEX  
FROM CTE_DRUG_INDEX i  
JOIN OBSERVATION_PERIOD op  
  ON op.PERSON_ID = i.PERSON_ID  
AND 1.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE  
JOIN PERSON p  
  ON p.PERSON_ID = i.PERSON_ID  
WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180  
AND EXTRACT(YEAR FROM i.INDEX_DATE)-p.YEAR_OF_BIRTH >= 65  
ORDER BY i.PERSON_ID
```
New Users of Warfarin >= 65 years of age

Try running this on your own!

```sql
/* (Exercise 2) Warfarin New Users 65 or Older at Index */

WITH CTE_DRUG_INDEX AS (
  SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE
  FROM DRUG_EXPOSURE de
  WHERE de.DRUG_CONCEPT_ID IN (
    SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /* warfarin */
  )
  GROUP BY de.PERSON_ID
)

SELECT i.PERSON_ID, i_INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE,
  (i_INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX,
  EXTRACT(YEAR FROM i_INDEX_DATE)-p.YEAR_OF_BIRTH AS AGE_AT_INDEX
FROM CTE_DRUG_INDEX i
JOIN OBSERVATION_PERIOD op
  ON op.PERSON_ID = i.PERSON_ID
  AND i_INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
JOIN PERSON p
  ON p.PERSON_ID = i.PERSON_ID
WHERE (i_INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180
  AND EXTRACT(YEAR FROM i_INDEX_DATE)-p.YEAR_OF_BIRTH >= 65
ORDER BY i.PERSON_ID
```

How many people do you get?
Some Example Questions

- Finding Warfarin
- New Users of Warfarin
- New Users of Warfarin who are >=65?
- New Users of Warfarin with prior Atrial Fibrillation?
How do I define new users of Warfarin with prior Atrial Fibrillation?
What is Needed in the CDM?

- **OMOP Vocabulary**
  to find the concepts

- **DRUG_EXPOSURE**
  to find individuals with exposure

- **OBSERVATION_PERIOD**
  to know people’s time within the database

- **PERSON**
  to know year of birth

- **CONDITION_OCCURRENCE**
  to find presence of a disease
Step 1: Start with the Ex 1 query

```sql
/* (Exercise 3) Warfarin New Users With Prior AFIB
*******************************************************************************************/

WITH CTE_DRUG_INDEX AS (
  SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE
  FROM DRUG_EXPOSURE de
  WHERE de.DRUG_CONCEPT_ID IN (SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /* warfarin */
  )
  GROUP BY de.PERSON_ID
),

CTE_DRUG_NEW_USERS AS (
  SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE,
  (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX
  FROM CTE_DRUG_INDEX i
  JOIN OBSERVATION_PERIOD op
    ON op.PERSON_ID = i.PERSON_ID
    AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
  WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180
)

SELECT nu.*, MIN(nu.INDEX_DATE-co.CONDITION_START_DATE) AS DAYS_OF_CLOSEST_AFIB_PRIOR_TO_INDEX
FROM CTE_DRUG_NEW_USERS nu
  JOIN CONDITION_OCCURRENCE co
    ON co.PERSON_ID = nu.PERSON_ID
    AND co.CONDITION_START_DATE BETWEEN nu.OBSERVATION_PERIOD_START_DATE AND nu.OBSERVATION_PERIOD_END_DATE
WHERE co.CONDITION_CONCEPT_ID IN (SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 313217 /* Atrial fibrillation */
  )
  AND co.CONDITION_START_DATE < nu.INDEX_DATE
GROUP BY nu.PERSON_ID, nu.INDEX_DATE, nu.OBSERVATION_PERIOD_START_DATE, nu.OBSERVATION_PERIOD_END_DATE, nu.DAYS_BEFORE_INDEX
ORDER BY nu.PERSON_ID
```
Step 2: Define Atrial Fibrillation

```sql
/* (Exercise 3) Warfarin New Users With Prior AFIB */

WITH CTE_DRUG_INDEX AS (
    SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE
    FROM DRUG_EXPOSURE de
    WHERE de.DRUG_CONCEPT_ID IN (/*warfarin*/
        SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149
    )
    GROUP BY de.PERSON_ID
),

CTE_DRUG_NEW_USERS AS (
    SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE,
            (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) AS DAYS_BEFORE_INDEX
    FROM CTE_DRUG_INDEX i
    JOIN OBSERVATION_PERIOD op
        ON op.PERSON_ID = i.PERSON_ID
        AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
    WHERE (i.INDEX_DATE-op.OBSERVATION_PERIOD_START_DATE) >= 180
)

SELECT nu.*, MIN(nu.INDEX_DATE-co CONDITION_START_DATE) AS DAYS_OF_CLOSEST_AFIB_PRIOR_TO_INDEX
FROM CTE_DRUG_NEW_USERS nu
    JOIN CONDITION_OCCURRENCE co
        ON co.PERSON_ID = nu.PERSON_ID
        AND co.CONDITION_START_DATE BETWEEN nu.OBSERVATION.PERIOD_START_DATE AND nu.OBSERVATION.PERIOD_END_DATE
    WHERE co.CONDITION_CONCEPT_ID IN (/*Atrial fibrillation*/
        SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 313217
    )
    AND co.CONDITION_START_DATE < nu.INDEX_DATE
GROUP BY nu.PERSON_ID, nu.INDEX_DATE, nu.OBSERVATION_PERIOD_START_DATE, nu.OBSERVATION_PERIOD_END_DATE, nu.DAYS_BEFORE_INDEX
ORDER BY nu.PERSON_ID
```
Step 3: Prior Atrial Fibrillation

Keeps condition within the same observable time, exclude if you want all time prior.
How do I define new users of Warfarin with prior Atrial Fibrillation?
Try running this on your own!

---

```sql
WITH CTE_DRUG_INDEX AS (
    SELECT de.PERSON_ID, MIN(de.DRUG_EXPOSURE_START_DATE) AS INDEX_DATE
    FROM DRUG_EXPOSURE de
    WHERE de.DRUG_CONCEPT_ID IN (SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 1310149 /*warfarin*/)
    GROUP BY de.PERSON_ID
),

CTE_DRUG_NEW_USERS AS (
    SELECT i.PERSON_ID, i.INDEX_DATE, op.OBSERVATION_PERIOD_START_DATE, op.OBSERVATION_PERIOD_END_DATE, DATEDIFF(DAY, op.OBSERVATION_PERIOD_START_DATE, i.INDEX_DATE) AS DAYS_BEFORE_INDEX
    FROM CTE_DRUG_INDEX i
    JOIN OBSERVATION_PERIOD op
    ON op.PERSON_ID = i.PERSON_ID
    AND i.INDEX_DATE BETWEEN op.OBSERVATION_PERIOD_START_DATE AND op.OBSERVATION_PERIOD_END_DATE
    WHERE DATEDIFF(DAY, op.OBSERVATION_PERIOD_START_DATE, i.INDEX_DATE) >= 180

    SELECT nu.*, MAX(DATEDIFF(DAY, co_CONDITION_START_DATE, nu.INDEX_DATE)) AS DAYS_OF_CLOSEST_AFIBPRIOR_TO_INDEX
    FROM CTE_DRUG_NEW_USERS nu
    JOIN CONDITION OCCURRENCE co
    ON co.PERSON_ID = nu.PERSON_ID
    AND co_CONDITION_START_DATE BETWEEN nu.OBSERVATION_PERIOD_START_DATE AND nu.OBSERVATION_PERIOD_END_DATE
    WHERE co_CONDITION_CONCEPT_ID IN (SELECT DESCENDANT_CONCEPT_ID FROM CONCEPT_ANCESTOR WHERE ANCESTOR_CONCEPT_ID = 313217 /*Atrial fibrillation*/)
    AND co_CONDITION_START_DATE < nu.INDEX_DATE
    GROUP BY nu.PERSON_ID, nu.INDEX_DATE, nu.OBSERVATION_PERIOD_START_DATE, nu.OBSERVATION_PERIOD_END_DATE, nu.DAYS_BEFORE_INDEX
    ORDER BY nu.PERSON_ID
)"
Try on your own!

• Warfarin New Users 65 or Older at Index with Prior Atrial Fibrillation
  
  7,067 individuals

• Bonus: Clopidogrel New Users 65 or Older at Index with Prior Atrial Fibrillation
  
  2,683 individuals
Queries Can Be Automated

• Open up Google Chrome

• Open up ATLAS

• Example cohort under “Cohort Definitions”: “Warfarin New Users 65 or Older at Index with Prior Atrial Fibrillation”
**Cohort definition:** A cohort is defined as the set of persons satisfying one or more inclusion criteria for a duration of time. Criteria and cohort exit criteria. Cohort entry criteria involve selecting one or more initial events, which determine the start date of the entry record to determine the end date when the person’s episode no longer qualifies for the cohort.
Conclusions
OMOP CDM standardizes the structure
OMOP Vocabulary standardizes the terminology
Concept IDs link CDM and Vocabulary
Source data still preserved in the OMOP CDM
Concept domains decide what table each piece of data lands on
OMOP CDM can be used for many types of data (e.g. claims, EHR, survey, labs, etc.)
OMOP CDM development is Open Source, Community driven
OMOP CDM is patient centric
OMOP Vocabulary

- Is used to **standardize terminology**

- **Compiles standards** from disparate public and private sources and some OMOP-grown concepts

- Has **one uniform structure** to house multiple vocabularies used in the public domain

- Is designed to **facilitate efficient queries**

- Is **regularly updated, maintained, and improved**
OMOP CDM

- Is used to **standardize structure** and **queries**
- **Integrated with Controlled Vocabulary**
- **Consolidates data from heterogeneous data sources**: EMR, claims, registries
- **Patient centric**
- **Domain (subject area) based**: concepts decide what table each piece of data lands on
- **Preserves data provenance**
- **Database platform independent**
What Makes OMOP CDM Unique

- **Supports collaborative research** across data sources both within and outside of US

- Developed **based on analytic use cases** by community of collaborators

- **Specialized**: reflective of clinical domain, granular, well structured

- **Integrated with Vocabulary** that is uniformly structured and well curated

- **Extendable**: new concepts and attributes can be added

- **Supported by Community** of interdisciplinary developers and researchers