Analyzing Large Drug Prescription Datasets

Principles, Tools and Examples

Olivier Bodenreider, MD, PhD, NLM
Vojtech Huser, MD, PhD, NLM
Christian Reich, MD, PhD, IQVIA
Disclosure

OB, VH: No relationships with commercial interest

The views and opinions expressed do not necessarily state or reflect those of the U.S. Government, and they may not be used for advertising or product endorsement purposes.

CR: IQVIA has many clients in the pharmaceutical industry, whose products are the subject of this Tutorial. However, all products are intended to only be mentioned for the purposes of this Tutorial and without any recommendation or judgment of their use.
Learning Objectives

To become familiar with large prescription datasets

To gain knowledge about tools available for analyzing prescription datasets

To gain knowledge about clinical data models, such as OMOP
Overview

Part 1 – Resources and use cases (OB)
- Prescription datasets
- RxNorm and NLM drug APIs; drug classification systems
- Common use cases

Part 2 – Drug data processing in practice (VH)
- Implementing use cases with RxMix and R

Part 3 – Experience with OHDSI (CR)
- Clinical data models
- Handling international drugs
Part 1 – Resources and use cases
Part 1 overview

Types of drug entities
- Drugs
- Drug classes

Prescription datasets
- Structure
- Sources

Drug data processing
- Mapping drugs to standards
- Aggregating drugs (by ingredient, by class)

Resources for processing drug datasets
- RxNorm
- Drug classification systems
- NLM drug APIs

Common use cases
- Pharmaco-epidemiology
- Assess exposure to drugs by ingredient or class
- Identify potentially inappropriate medications
Types of drug entities

Part 1 – Resources and use cases
Types of drug entities (drugs I)

Ingredient
- Azithromycin

Drug form
- Ingredient + dose form
  - Azithromycin Oral Tablet

Clinical drug (unit of prescription)
- Ingredient + dose form + strength
  - Azithromycin 250 MG Oral Tablet

Pack (packaging level)
- Collection of clinical drugs (unit of dispensation)
  - Z-PAK (6 (Azithromycin 250 MG Oral Tablet [Zithromax]))
- Collection of clinical drugs (bulk)
  - Manufacturer: Apotex Corp.; pack size: 500 in 1 BOTTLE
Types of drug entities (drugs II)

Generic vs. Brand
- **G**: Azithromycin 250 MG Oral Tablet
- **B**: Zithromax 250 MG Oral Tablet

Single vs. multi-ingredient drug
- **S**: Amoxicillin 250 MG Oral Capsule
- **M**: Amoxicillin 250 MG / Clavulanate 125 MG Oral Tablet

Systemic vs. topical drugs
- **S**: Azithromycin 250 MG Oral Tablet
- **T**: Azithromycin 10 MG/ML Ophthalmic Solution

Base vs. salt/ester (as basis of strength substance)
- **B**: Erythromycin 250 MG Oral Tablet
- **S**: Erythromycin Ethylsuccinate 400 MG Oral Tablet
Types of drug entities (drug classes)

**Atorvastatin**

**Therapeutic class**
- Anticholesteremic Agent

**Chemical structure**
- n/a

**Mechanism of action**
- HMG-CoA Reductase Inhibitor

**Physiologic effect**
- Decreased Cholesterol Synthesis
Types of drug entities (drug classes)

- **ANTIINFECTIVES FOR SYSTEMIC USE**
  - **ANTIBACTERIALS FOR SYSTEMIC USE**
    - MACROLIDES, LINCOSAMIDES AND STREPTOGRAMINS
      - Macrolides
        - azithromycin
    - **SENSORY ORGANS**
      - **OPHTHALMOLOGICALS**
    - **ANTIINFECTIVES**
      - Antibiotics
Prescription datasets

Part 1 – Resources and use cases
## Structure

### Transactions captured at dispensation time (pharmacy)
- Pharmaceutical claims data sent to payers

### Common elements
- **Prescription ID**: 999999
- **Patient identifier**: 123456
- **Date (prescription)**: 20161112
- **Product ID (NDC)**: 00071015623
- **Total quantity dispensed**: 30
- **Days supply**: 30
- **Cost data**: --
- **Drug name**: LIPITOR TAB 20MG
- **Strength**: 20
- **Generic name**: Atorvastatin
- **Prescriber ID**: 789

NDCs are not unique identifiers for clinical drugs.

Drug names are not standardized.
## Product Name from Medicaid Claims

<table>
<thead>
<tr>
<th>Product Name</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>COUMADIN</td>
<td>270</td>
</tr>
<tr>
<td>COUMADIN TAB</td>
<td>13</td>
</tr>
<tr>
<td>COUMADIN TABLET</td>
<td>215</td>
</tr>
<tr>
<td>JANTOVEN</td>
<td>130</td>
</tr>
<tr>
<td>JANTOVEN TAB</td>
<td>24</td>
</tr>
<tr>
<td>JANTOVEN TABLET</td>
<td>262</td>
</tr>
<tr>
<td>WARFARIN</td>
<td>1,093</td>
</tr>
<tr>
<td>WARFARIN TAB</td>
<td>763</td>
</tr>
<tr>
<td>WARFARIN SODIUM</td>
<td>8,717</td>
</tr>
<tr>
<td>WARFARIN SODIUM TA</td>
<td>95</td>
</tr>
<tr>
<td>WARFARIN SODIUM TAB</td>
<td>516</td>
</tr>
<tr>
<td>WARFARIN SODIUM TABL</td>
<td>587</td>
</tr>
<tr>
<td>WARFARIN SODIUM TABLET</td>
<td></td>
</tr>
</tbody>
</table>

Source: Dan Malone
Product Identification: NDCs

National Drug Codes

- Product identification system
- Three components
  - Manufacturer
  - Product
  - Packaging

Introduced in 1972 by FDA

Only format permitted by NCPDP

Mandated by HIPAA regulations for drug transactions

Source: Dan Malone
NDC Elements: 3 segments

[Diagram showing the NDC Elements: XXXXX-XXXX-XX with arrows pointing to Manufacturer, Product, and Packaging]
NDC Forms

Warfarin Sodium 1 MG Oral Tablet

XXX-XXXX-XX (4-4-2) → 0XXXX XXXX XX

0555-0831-02  (Teva Pharmaceuticals USA, Inc.; 100 in 1 BOTTLE) → 00555083102

XXXXX-XXX-XX (5-3-2) → XXXXX 0XXX XX

21695-672-30  (Rebel Distributors Corp; 30 in 1 BOTTLE) → 21695067230

XXXXX-XXXX-X (5-4-1) → XXXXX XXXX 0X

50090-1213-0  (A-S Medication Solutions; 30 in 1 BOTTLE) → 50090121300
NDC Characteristics

11 Digit code (leading zero for 4-4-2 format)

Hyphens between segments are missing in claims transmission (Field 407 in NCPDP claim format)

NDC codes set by the manufacturer/labeler

High turnover compared to other drug IDs

Product codes are unique to manufacturer – not to the chemical entity

Package codes are unique to the manufacturer and product – there is no standardization for packaging codes

Source: Dan Malone
Source of prescription datasets

Surescripts transactions

Data from payers
- Medicaid
- Medicare Part D

Commercial health analytics companies
- Truven (120M patients)
- PharMetrics Plus (100M patients)
- Ambulatory EMR (35M patients)
- Open Claims (250M patients)

Reagan-Udall Foundation for the FDA
- IMEDS Research Lab (temporarily suspended)
Drug data processing

Part 1 – Resources and use cases
Mapping and aggregation

Mapping to standard resources (e.g., RxNorm)

- Standard names and codes
- Standard set of relations among drug entities
- Link to drug classification systems
- E.g., NDC → clinical drug

Aggregation

- “roll up” to the appropriate level of granularity for analytics (use case-dependent)
- E.g., branded drug → clinical drug → ingredient → drug class

Temporal aggregation

- Aggregate individual prescriptions into longer spans (“drug eras” in OHDSI)
### Esomeprazole (283742)

**Esomeprazole 40 MG Delayed Release Oral Capsule (606730)**

**Esomeprazole 40 MG Delayed Release Oral Capsule [Nexium] (606731)**

---

**ATC code** | **Name**      | **DDD** | **U** | **Adm.R.** | **Note**
--- | --- | --- | --- | --- | ---
A02BC05 | esomeprazole | 30 mg | O | | 
A02BC05 | esomeprazole | 30 mg | P | |
Resources for processing drug datasets – RxNorm
Part 1 – Resources and use cases
RxNorm

Terminology integration system

- Structured Product Labels, First DataBank, Micromedex, Multum, MeSH, SNOMED CT, MED-RT, ATC, ...

Scope

- Drug names and codes
- Drugs available on the U.S. market

Developer: National Library of Medicine

Publicly available*

Monthly updates

Size: > 10k ingredients; 19k clinical drugs

Uses: e-prescription, information exchange, analytics

https://www.nlm.nih.gov/research/umls/rxnorm/
### Normalization Lexical level

<table>
<thead>
<tr>
<th>Name</th>
<th>Code</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>WARFARIN (COUMADIN) NA 1MG TAB</td>
<td>4005203</td>
<td>VANDF</td>
</tr>
<tr>
<td>warfarin 1 mg oral tablet</td>
<td>3617</td>
<td>MMSL</td>
</tr>
<tr>
<td>WARFARIN NA 1MG TAB,UD</td>
<td>4014039</td>
<td>VANDF</td>
</tr>
<tr>
<td>WARFARIN NA 1MG TAB,UD [VA Product]</td>
<td>N0000161787</td>
<td>NDFRT</td>
</tr>
<tr>
<td>WARFARIN SODIUM 1 mg ORAL TABLET</td>
<td>14198</td>
<td>NDDF</td>
</tr>
<tr>
<td>WARFARIN SODIUM 1 mg ORAL TABLET</td>
<td>60429-784</td>
<td>MTHSPL</td>
</tr>
<tr>
<td>Warfarin Sodium 1 MG Oral Tablet</td>
<td>104045</td>
<td>MMX</td>
</tr>
<tr>
<td>WARFARIN SODIUM 1 mg ORAL TABLET</td>
<td>63629-4017</td>
<td>MTHSPL</td>
</tr>
<tr>
<td>WARFARIN SODIUM 1 mg ORAL TABLET [Warfarin Sodium]</td>
<td>53808-0985</td>
<td>MTHSPL</td>
</tr>
<tr>
<td>Warfarin Sodium 1 MILLIGRAM In 1 TABLET ORAL TABLET</td>
<td>15330-100</td>
<td>MTHSPL</td>
</tr>
<tr>
<td>WARFARIN SODIUM 1.09 MG ORAL TABLET</td>
<td>281572</td>
<td>MTHFDA</td>
</tr>
<tr>
<td>Warfarin Sodium 1mg Oral tablet</td>
<td>933</td>
<td>GS</td>
</tr>
<tr>
<td>Warfarin sodium 1mg tablet (product)</td>
<td>319733000</td>
<td>SNOMEDCT_US</td>
</tr>
<tr>
<td>Warfarin Sodium Tab 1 MG</td>
<td>6749</td>
<td>MDDB</td>
</tr>
<tr>
<td>Warfarin Sodium, 1 mg oral tablet</td>
<td>3617</td>
<td>MMSL</td>
</tr>
<tr>
<td>WARFARIN SODIUM@1 mg@ORAL@TABLET</td>
<td>14198</td>
<td>NDDF</td>
</tr>
</tbody>
</table>

---

**Warfarin Sodium 1 MG Oral Tablet (855288)**
**Warfarin Sodium 1 MG**

- **Dose form**: Oral Tablet
- **Clinical drug form**:
  - Warfarin Sodium 1 MG (855287)
  - Warfarin Oral Tablet (374319)
  - Warfarin Sodium 1 MG Oral Tablet (855288)
Relations among drug entities

- **Ingredient**: Azithromycin
- **Brand Name**: Zithromax
- **C. Drug Comp.**: Azithromycin 250 MG
- **C. Drug Form**: Azithromycin Oral Tablet
- **B. Drug Comp.**: Azithromycin 250 MG [Zithromax]
- **B. Drug Form**: Azithromycin Oral Tablet [Zithromax]
- **C. Drug**: Azithromycin 250 MG Oral Tablet
- **B. Pack**: Zithromax 250 MG Oral Tablet
- **G. Pack**: {6 (Azithromycin 250 MG Oral Tablet)} Pack
- **B. Pack**: Z-PACK
RxNav – RxNorm browser

https://mor.nlm.nih.gov/RxNav/
Resources for processing drug datasets – Drug classification systems

Part 1 – Resources and use cases
ATC/DDD Index

Origin

- World Health Organization (WHO) Collaborating Centre for Drug Statistics Methodology (Norway)

Purpose

- For drug utilization research / pharmaco-epidemiology

~1300 classes (1-4)

Organization

- Drug classification on 4 levels
  - Anatomical
  - Therapeutic
  - Chemical

- Drugs (5th level)

<table>
<thead>
<tr>
<th>ATC code</th>
<th>Name</th>
<th>DDD</th>
<th>U</th>
<th>Adm.R</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>J01CA04</td>
<td>amoxicillin</td>
<td>1 g</td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 g</td>
<td>P</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Established Pharmacologic Classes (EPCs)

Origin
- Veterans Health Administration’s Medication Reference Terminology (MED-RT)
  - For use by the U.S. Food and Drug Administration (FDA)

Purpose
- For drug classification in the Structured Product Labels

~600 classes

No hierarchical organization

Examples
- Macrolide antibacterial (e.g., Azithromycin)
Mechanism of action (MoA)  
Physiologic effect (PE)  
Chemical structure (Chem)

Origin

- Veterans Administration’s Medication Reference Terminology (MED-RT)

Purpose

- For drug classification in the Structured Product Labels
- For drug classification at the VA

Number of classes

- MoA: ~600; PE: ~1900; Chem: ~10,000

Hierarchical organization

Examples

- MoA: HMG-CoA Reductase Inhibitor (e.g., atorvastatin)
- PE: Decreased Blood Pressure (e.g., enalapril)
- Chem: Penicillins (e.g., amoxicillin)
VA Classes

Origin
- Veterans Administration’s National Drug File

Purpose
- For drug classification in the VA formulary

~500 classes

Shallow hierarchical organization (3 levels)

Examples
- **L1:** ANTIMICROBIALS
- **L2:** PENICILLINS AND BETA-LACTAM ANTIMICROBIALS
- **L3:** QUINOLONES (e.g., Ofloxacin 200 MG Oral Tablet)

NB: links to clinical drugs rather than ingredients
RxClass – Drug class browser

Class Browser
- Anatomical Therapeutic Chemical (ATC1-4)
- Established Pharmacologic Classes (EPC) [from DailyMed]
  - Names from 4-H - alp
  - Names from alp - Ant
  - Names from Ant - Asp
  - Names from Asp - Bio
  - Names from Bio - CD1
  - Names from CD2 - Col
  - Names from Com - Dih
  - Names from Dih - Ery
  - Names from Est - Gon
  - Names from Gro - Hum
  - Names from Hum - Ina
  - Names from Ina - Int
  - Names from Int - Lin
  - Names from Lip - Low
  - Names from Lym - Mus
    - Lymphocyte Growth Factor (1)
    - Macrolide (1)
    - Macrolide Antibacterial (1)
    - Macrolide Antimicrobial (3)
    - Mast Cell Stabilizer (3)
    - Megakaryocyte Growth Factor (1)
    - Melanin Synthesis Inhibitor (1)
    - Melatonin Receptor Agonist (2)
    - Metabolic Alkalizer (9)
    - Metal Chelator (2)
    - Methylated Sulfonamides Antibacterial (1)
    - Methylation Agent (2)
    - Methylxanthine (2)
    - Microsomal Triglyceride Transfer Protein Inhibitor (1)

RxClass
Exploring drug classes and their RxNorm drug members

azithromycin

- by class name/id
- by RxNorm drug name/id
- ingredient drug only
- Edit Drug Sources

Macrolide Antimicrobial

class: Macrolide Antimicrobial / id: N0000175935 / class type: EPC / show context

3 RxNorm generic drugs for has_EPC in DailyMed / similar classes

<table>
<thead>
<tr>
<th>Type</th>
<th>RXCUI</th>
<th>RxNorm Name</th>
<th>Relation</th>
<th>All classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>IN</td>
<td>19631</td>
<td>Azithromycin</td>
<td>DIRECT</td>
<td>Show</td>
</tr>
<tr>
<td>IN</td>
<td>21212</td>
<td>Clarithromycin</td>
<td>DIRECT</td>
<td>Show</td>
</tr>
<tr>
<td>IN</td>
<td>4053</td>
<td>Erythromycin</td>
<td>DIRECT</td>
<td>Show</td>
</tr>
</tbody>
</table>
Resources for processing drug datasets – NLM drug APIs

Part 1 – Resources and use cases
NLM drug APIs

Expose the content of RxNorm, RxTerms and MED-RT (and other resources)

• Logical structure, not storage format
• Up-to-date information (monthly updates of RxNorm)
• Additional features
  • Normalized and approximate matching; spelling correction
  • Drug-drug interactions checking (from DrugBank)
  • Link to drug classes (from ATC, DailyMed, MeSH, MED-RT)
  • Archive of NDCs since 2007
• Optimized graph traversal (pre-computed)

For use in applications

• Web services
• SOAP, REST (XML, JSON)
• Independent of any programming language
API documentation and examples

RxNorm API

The RxNorm API is a web service for accessing the current RxNorm data set. With one exception, no license is needed to use the RxNorm API. This is because the data returned from the API is from the RxNorm vocabulary, a non-proprietary vocabulary developed by the National Library of Medicine.

The API can be accessed by clients in two different ways:
- RxNorm RESTful web services
- RxNorm SOAP web services

Please check RxNorm API changes for the current updates.

Functions and Resources

In the table, the base URI (https://rxnav.nlm.nih.gov/REST/) for the REST resources has been omitted to improve readability.

<table>
<thead>
<tr>
<th>SOAP Function</th>
<th>REST resource</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>filterByProperty</td>
<td>/rxcu/{rxcu}/filter</td>
<td>Filter by property</td>
</tr>
<tr>
<td>findRxcuById</td>
<td>/rxcu/{idtype}</td>
<td>Search by identifier to find RxNorm concepts</td>
</tr>
<tr>
<td>findRxcuByString</td>
<td>/rxcu/{name}</td>
<td>Search by name to find RxNorm concepts</td>
</tr>
<tr>
<td>getRxCpuClasses</td>
<td>/classes</td>
<td>Return all RxNorm Concept classes</td>
</tr>
<tr>
<td>getRxCpuConceptsByTTY</td>
<td>/allconcepts</td>
<td>Return the RxNorm Concepts for the specified term types</td>
</tr>
<tr>
<td>getRxCpuHistoricalNDCs</td>
<td>/rxcu/{rxcu}/allhistoricalncds</td>
<td>Return all National Drug Codes (NDC) for a concept</td>
</tr>
<tr>
<td>getRxCpuNDCs</td>
<td>/rxcu/{rxcu}/allNDCs</td>
<td>TO BE DEPRECATED. Use getRxCpuHistoricalNDCs or /rxcu/{rxcu}/allhistoricalncds instead.</td>
</tr>
<tr>
<td>getRxCpuProperties</td>
<td>/rxcu/{rxcu}/allproperties</td>
<td>Return all properties for a concept</td>
</tr>
<tr>
<td>getRxCpuRelatedInfo</td>
<td>/rxcu/{rxcu}/allrelated</td>
<td>Return all related concept information</td>
</tr>
<tr>
<td>ATC code</td>
<td>Name</td>
<td>DDD</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------</td>
<td>-----</td>
</tr>
<tr>
<td>A02BC05</td>
<td>esomeprazole</td>
<td>30 mg</td>
</tr>
<tr>
<td>A02BC05</td>
<td>esomeprazole</td>
<td>30 mg</td>
</tr>
</tbody>
</table>

**Esomeprazole (283742)**

**Esomeprazole 40 MG Delayed Release Oral Capsule (606730)**

**Esomeprazole 40 MG Delayed Release Oral Capsule [Nexium] (606731)**

```plaintext
rxnorm:findRxcuiById("NDC", "0186-5040-31", 0) → 606731
```
Esomeprazole (A02BC05)

Esomeprazole 40 MG Delayed Release Oral Capsule (606731)
Esomeprazole 40 MG Delayed Release Oral Capsule [Nexium] (606731)

rxclass:getClassByRxNormDrugId ("283742", "ATC", "ALL") → A02BC, Proton pump inhibitors
Esomeprazole (A02BC05)

<table>
<thead>
<tr>
<th>ATC Code</th>
<th>Name</th>
<th>DDD</th>
<th>U</th>
<th>Adm.</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>A02BC</td>
<td>esomeprazole</td>
<td>30</td>
<td>mg</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>A02BC</td>
<td>esomeprazole</td>
<td>30</td>
<td>mg</td>
<td>P</td>
<td></td>
</tr>
</tbody>
</table>

**rxclass:getClassGraph( “A02BC”)**

- A02B, DRUGS FOR PEPTIC ULCER AND GASTRO-OESOPHAGEAL REFLUX DISEASE (GORD);
- A02, DRUGS FOR ACID RELATED DISORDERS;
- A, ALIMENTARY TRACT AND METABOLISM
RxMix

Graphical interface to the drug APIs

- RxNorm, MED-RT, RxTerms, RxImage, Interactions, RxClass, MedEx, DailyMed

Handles interoperability between functions

Helps users compose complex queries

- Find all the NDC codes for a given allergy class (e.g., barbiturates)

Supports batch execution

https://mor.nlm.nih.gov/RxMix/
Common use cases

Part 1 – Resources and use cases
Common use cases

Pharmaco-epidemiology
- Assess exposure to drugs (by ingredient or class)
- Assess prescribed daily dose

Identify potentially inappropriate medications
- Elderly (Beers)
- Pregnant women (Briggs)
Use case #1

*Pharmaco-epidemiology*

[Bodenreider, AMIA, 2014]
Prescribed vs. defined daily dose

Dataset
- Surescripts feed
- All prescriptions to ER patients
- For 3 months in 2011 in a Bethesda hospital

Reference for defined daily dose: ATC

Methods
- RxNorm clinical drug → RxNorm ingredient ↔ ATC ingredient → ATC defined daily dose ↔ prescribed daily dose
- Restricted to systemic drugs (based on dose form)

Findings
- Confirmed feasibility
- 25% of the prescriptions exactly match the ATC DDD
- 50% of the prescriptions within 66-150% of the ATC DDD
- 75% of the prescriptions within 50-200% of the ATC DDD
Methods

Example

RxNorm

Amoxicillin 500 MG Oral Capsule (308191)
Amoxicillin (723)
Oral Capsule

ATC/DDD Index

amoxicillin (J01CA04)
O
1 g

Surescripts

Amoxicillin 500 MG Oral Capsule (308191)
40 capsules
10 days
40 x 500 mg / 10 = 2 g

J ANTIINFECTIVES FOR SYSTEMIC USE
J01 ANTIBACTERIALS FOR SYSTEMIC USE
J01C BETA-LACTAM ANTIBACTERIALS, PENICILLINS
J01CA Penicillins with extended spectrum
Results  Prescription classification

Frequency of drugs by level-1 ATC group in the Surescripts prescription dataset

N=86,578

ALIMENTARY TRACT AND METABOLISM (A)
BLOOD AND BLOOD FORMING ORGANS (B)
CARDIOVASCULAR SYSTEM (C)
DERMATOLOGICALS (D)
GENITO URINARY SYSTEM AND SEX HORMONES (G)
SYSTEMIC HORMONAL PREP., EXCL. SEX HORMONES AND INSULINS (H)
ANTIINFECTIVES FOR SYSTEMIC USE (J)
ANTEOPOPLASTIC AND IMMUNOMODULATING AGENTS (L)
MUSCULO-SKELETAL SYSTEM (M)
NERVOUS SYSTEM (N)
ANTIPARASITIC PRODUCTS, INSECTICIDES AND REPELLENTS (P)
RESPIRATORY SYSTEM (R)
SENSORY ORGANS (S)
VARIOUS (V)

Atorvastatin  Simvastatin  Lisinopril  Metoprolol  Amlodipine  Furosemide  Atenolol  Hydrochlorothiazide
Zolpidem  Sertraline  Escitalopram  Alprazolam  Clonazepam  Gabapentin  Quetiapine  Oxycodone  Fluoxetine  Duloxetine
Deviation of the prescribed daily dose (PDD) in Surescripts from the defined daily dose (DDD) in ATC for 68,462 oral solid dose form prescriptions

- 86.1% of the prescriptions are within 33%-300% of the ATC DDD
- 76.1% of the prescriptions are within 50%-200% of the ATC DDD
- 49.5% of the prescriptions are within 66%-150% of the ATC DDD
- 28.6% of the prescriptions exactly match the ATC DDD
- 10.4% < 33% of the ATC DDD
- 3.5% > 300% of the ATC DDD
Use case #2

Identifying potentially inappropriate medications for elderly patients

[Mundkur, AMIA, 2016]
PIMs for elderly patients

Dataset
- Medicare Part D
- 1M beneficiaries ≥ 65
- All prescriptions for one year (2009)

Reference list of PIMs: Beers list

Methods
- NDC → RxNorm clinical drug → ingredient ↔ Beers
- Restricted to systemic drugs (based on dose form)

Findings
- 47% of all beneficiaries were prescribed at least 1 PIM
- Top PIMs: zolpidem (6.3%), nitrofurantoin (4.5%)
Methods  Example

RxNorm

55111047901  
Zolpidem tartrate 10 MG Oral Tablet (854873)
Oral Tablet

zolpidem

Beers

zolpidem

Oral Pill

DFG filter

Medicare

55111047901  
Demographic data
470,523 prescriptions
Use case #3

Identifying potential risk in drug prescriptions during pregnancy

[Dhombres, AMIA, 2016]
Potential risk during pregnancy

Dataset
• Large prescription dataset from private insurer (150M patients)
• 3.7M pregnant women; 19M prescriptions (2003-2014)
• OMOP common data model

Reference list for risk during pregnancy: Briggs textbook

Methods
• RxNorm clinical drug → ingredient ↔ Briggs drug → fetal risk
• Restricted to systemic drugs (based on dose form)

Findings
• 41.2% compatible with pregnancy or probably compatible
• 55.6% potential risk
• 3.29% high risk or contraindicated
Specific challenge

Obsolete identifiers

- NDC = drug + manufacturer + packaging information
  - ~250,000 active NDCs
  - ~300,000 obsolete NDCs in the past 10 years
  - ~220,000 “alien” NDCs (not curated by RxNorm)

- Obsolete NDCs
  - Removed from RxNorm (e-prescribing use case)
  - Needed for analytics (longitudinal datasets)

- RxNorm API provides access to obsolete NDCs
  - Mapping obsolete NDCs to active drugs
    - rxnorm:getNDCStatus( ndc, startDate, endDate, option )
  - List of all NDCs – active or obsolete – for a given drug
    - rxnorm:getAllHistoricalNDCs( rxcui, history )
Other challenges

Reuse of identifiers
  • NDCs (time-indexed)

Insufficient coverage in RxNorm
  • International drugs
  • Over-the-counter drugs

Granularity of knowledge
  • Ingredient-class vs. clinical drug-class

Heterogeneity of drug classification
  • Different use cases
Impact assessment

1B prescriptions from Medicare analyzed

- Over a 10-year period (2005-2014)

Vast majority of NDCs can be resolved with the RxNorm API functions

- Minor issues
  - Start/End date do not match prescription date
  - Ambiguous mapping (multiple RxCUIs; often clinically insignificant – generic vs. brand)
- <5% unmapped NDCs (mostly supplies; OTCs)
Part 2 – Drug data processing in practice

Olivier Bodenreider, MD, PhD, NLM
Vojtech Huser, MD, PhD, NLM
Christian Reich, MD, PhD, IQVIA
Part 2 overview

Detailed look at the API

Try it yourself (Follow-along examples with RxMix)
  • and get help (if you hit a problem)
  • 5-10 min

R code examples
Links

https://github.com/lhnbc/r-snippets-bmi/tree/master/rxnorm

https://github.com/mpancia/RxNormR (not used in this session)
Detailed look at the API

Part 2 – Drug data processing in practice
API and RxMix web-tool

https://rxnav.nlm.nih.gov/

https://mor.nlm.nih.gov/RxMix/
<table>
<thead>
<tr>
<th>ATC code</th>
<th>Name</th>
<th>DDD</th>
<th>U</th>
<th>Adm. R</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>A02BC05</td>
<td>esomeprazole</td>
<td>30</td>
<td>mg</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>mg</td>
<td>P</td>
<td></td>
</tr>
</tbody>
</table>

**Esomeprazole**

(283742)

Esomeprazole 40 MG Delayed Release Oral Capsule (606730)

Esomeprazole 40 MG Delayed Release Oral Capsule [Nexium] (606731)

00186504

031

0186-5040-31

**Nexium** (esomeprazole magnesium) 40 mg

Rx only

Dispense the accompanying Medication Guide to each patient.
Functions

rxnorm:findRxcuiById

- Parameters
  - id_string: NDC
  - AllSourcesFlag: 0
- Input: 00186504031 or 0186-5040-31
- Output: 606731

rxnorm:getRelatedByType

- Parameters:
  - term_type: IN
- Input: 606731
- Output: Esomeprazole  283742
Try it yourself (Follow-along examples with RxMix)

Part 2 – Drug data processing in practice
RxMix

Creating Applications from NLM Drug APIs

Introduction

RxMix is an interface for building applications using RxNorm, RxTerms, NDF-RT, RxClass, Interact, and Interact run either interactively or in batch mode.

Sample RxMix configurations

- Find drug interaction brands for Morphine (RxNorm)
- Find allergy drugs for Proton Pump Inhibitors

APIs

- RxNorm
- NDF-RT
- RxTerms
- RxImageAccess
- Interaction
- RxClass
- DailyMed
- MedEx
Library of pre-built workflows

Find brand and ingredients of an ATC drug class
This workflow finds the RxNorm ingredients and brands associated with an ATC drug class. Sample input: N03CF.

Find pill image information for a brand or an ingredient
This workflow retrieves pill image information for an ingredient or brand name. If an ingredient is specified, then all pill image information for the ingredient and brands containing the ingredient are returned. A brand name input will return the pill information only for the brand. Sample input: Cymbalta.

Find brand names containing an ingredient
This simple workflow gets a string and finds the RxNorm concept identifier, then finds the related brands. It can be used to enter ingredients and find the brand names which contain that ingredient. Sample input: warfarin.

Find the National Drug Codes (NDCs) for clinical drugs of an ingredient
This workflow finds the RxNorm Identifier (RxCID) for an ingredient name and retrieves all the clinical drugs for that name. The final step retrieves the NDCs for all the clinical drugs. Sample input: simvastatin.

Find the clinical drugs of an ingredient class
This workflow finds the drug class for an ingredient name (ex: hydantoin) and gets all the ingredients of the class. Then it gets the clinical drugs associated with ingredients. Sample input: hydantoin.
Lyrica 150 MG Oral Capsule
NCD: 00071101668

Pfizer
Lyrica®
(pregabalin) capsules
150 mg

ALWAYS DISPENSE WITH MEDICATION GUIDE
Rx only

90 Capsules
MADE IN SINGAPORE

DOSAGE AND USE
See accompanying prescribing information.
Each capsule contains 150 mg pregabalin.
Distributed by
Parke-Davis
Division of Pfizer Inc, NY, NY 10017

Store at 25°C (77°F); excursions permitted to 15-30°C (59-86°F)
[see USP Controlled Room Temperature].
Dispense in tight (USP), child-resistant containers.

PAA089715

NIH NLM
Resolving NDC 00071101668 in RxNav
Getting properties for NDC 00071101668
Getting all RxNorm entities related to Lyrica 150 MG Oral Capsule

<table>
<thead>
<tr>
<th>UMLSCUI</th>
<th>term_type</th>
<th>name</th>
<th>RXCUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1570032</td>
<td>EN</td>
<td>Lyrica</td>
<td>598441</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DF**

<table>
<thead>
<tr>
<th>UMLSCUI</th>
<th>term_type</th>
<th>name</th>
<th>RXCUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C099123</td>
<td>DF</td>
<td>Oral Capsule</td>
<td>216985</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SBDC**

<table>
<thead>
<tr>
<th>UMLSCUI</th>
<th>term_type</th>
<th>name</th>
<th>RXCUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1649612</td>
<td>SBDC</td>
<td>pregabalin 150 MG [Lyrica]</td>
<td>607019</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SBDF**

<table>
<thead>
<tr>
<th>UMLSCUI</th>
<th>term_type</th>
<th>name</th>
<th>RXCUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1649611</td>
<td>SBDF</td>
<td>pregabalin Oral Capsule [Lyrica]</td>
<td>697017</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Selecting parameters
Getting ATC classes for Lyrica 150 MG Oral Capsule
Getting drug-drug interactions for Lyrica 150 MG Oral Capsule
Nonsensical query returning no results
{
    "function": {
        "input": "raloxifene",
        "level": 0,
        "name": "findRxucuiByString",
        "outputs": {
            "output": {
                "RXCUI": 72143
            }
        },
        "function": {
            "input": 72143,
            "level": 1,
            "name": "getRelatedByType",
            "outputs": {
                "output_type": "BN",
                "output": {
                    "RXCUI": 217010,
                    "name": "Evista",
                    "term_type": "BN"
                }
            }
        }
    }
}
Saving your workflow
R code examples

Part 2 – Drug data processing in practice
#url %>% URLencode() %>% fromJSON()

# wrapping a call to API into R function
findRxCuiByString <- function(input) {
  url2 <- URLencode(url)
  j <- jsonlite::fromJSON(url2)
  # result in variable j, we can traverse JSON as traversing
table <- data.frame(rxnormId = as.integer(j$idGroup$rxcui$rxcuiId))
  # output is just a list of strings, we will extend it with more
  if (nrow(table) > 0) {
    out$input = input; out$match = 1: nrow(table)
  } else {
    out$input = input; out$match = NULL
  }

  out
}

example = 'lyrica'
findRxCuiByString(example)
out

1 593441 lyrica 1

example = 'atenolol'
example = 'nexitum'
example = 'Esomeprazone'
example = 'esomeprazole'
callingAPI.R

Example = 'lyrica'
findRxCuibyString(example)

#JSON parsing
input = '283742' # esomeprazole
whatFunction = 'allrelated'

j2 <- jsonlite::fromJSON(URLencode(url))
print(j2)
#display highest level
str(j2$max.level = 1)

#traverse the tree
str(j2$allRelatedGroup)

62:1 (Top Level)  R Script
# traverse the tree
str(j2$allRelatedGroup)

# using JSON parsing shortcut
j3 <- jsonlite::fromJSON(URLEncode(url), flatten = T)
str(j3$allRelatedGroup$conceptGroup$Properties)

# making it one large table
oneBigTable <- plyr::bind.fill(j3$allRelatedGroup$conceptGroup)

# A tibble: 68 x 7
  rxcui name synonym tty language suppress umlsCui
  <chr> <chr> <chr>  <chr>  <chr>    <chr>     <chr>
1 593441 Lyrica  ""  BN  ENG    N   C15702~
2 316945 Extende~ ""  DF  ENG    N   C09915~
3 316965 Oral Ca~ ""  DF  ENG    N   C09915~
4 316968 Oral So~ ""  DF  ENG    N   C09915~
5 187832 pregaba~ ""  IN  ENG    N   C06579~
6 607018 pregaba~ Lyrica 1~ SBD  ENG   N   C16367~
7 607020 pregaba~ Lyrica 1~ SBD  ENG   N   C16383~
8 607022 pregaba~ Lyrica 2~ SBD  ENG   N   C16373~
9 607024 pregaba~ Lyrica 2~ SBD  ENG   N   C16367~
<table>
<thead>
<tr>
<th>rxcui</th>
<th>name</th>
<th>synonym</th>
<th>tty</th>
<th>language</th>
<th>suppress</th>
<th>umlscode</th>
</tr>
</thead>
<tbody>
<tr>
<td>593441</td>
<td>Lyrica</td>
<td></td>
<td>BN</td>
<td>ENG</td>
<td>N</td>
<td>C1570232</td>
</tr>
<tr>
<td>316945</td>
<td>Extended Release Oral Tablet</td>
<td></td>
<td>DF</td>
<td>ENG</td>
<td>N</td>
<td>C0991507</td>
</tr>
<tr>
<td>316965</td>
<td>Oral Capsule</td>
<td></td>
<td>DF</td>
<td>ENG</td>
<td>N</td>
<td>C0991533</td>
</tr>
<tr>
<td>316968</td>
<td>Oral Solution</td>
<td></td>
<td>DF</td>
<td>ENG</td>
<td>N</td>
<td>C0991536</td>
</tr>
<tr>
<td>187832</td>
<td>pregabalin</td>
<td></td>
<td>IN</td>
<td>ENG</td>
<td>N</td>
<td>C0657912</td>
</tr>
<tr>
<td>607018</td>
<td>pregabalin 100 MG Oral Capsule</td>
<td>Lyrica 100 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1636761</td>
</tr>
<tr>
<td>607020</td>
<td>pregabalin 150 MG Oral Capsule</td>
<td>Lyrica 150 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1638398</td>
</tr>
<tr>
<td>607022</td>
<td>pregabalin 200 MG Oral Capsule</td>
<td>Lyrica 200 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1637375</td>
</tr>
<tr>
<td>607024</td>
<td>pregabalin 225 MG Oral Capsule</td>
<td>Lyrica 225 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1636762</td>
</tr>
<tr>
<td>607026</td>
<td>pregabalin 25 MG Oral Capsule</td>
<td>Lyrica 25 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1637376</td>
</tr>
<tr>
<td>607028</td>
<td>pregabalin 300 MG Oral Capsule</td>
<td>Lyrica 300 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1638300</td>
</tr>
<tr>
<td>607033</td>
<td>pregabalin 50 MG Oral Capsule</td>
<td>Lyrica 50 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1638399</td>
</tr>
<tr>
<td>607038</td>
<td>pregabalin 75 MG Oral Capsule</td>
<td>Lyrica 75 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1639484</td>
</tr>
<tr>
<td>rxcui</td>
<td>name</td>
<td>synonym</td>
<td>tty</td>
<td>language</td>
<td>suppress</td>
<td>umls cui</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------</td>
<td>--------------------------------------</td>
<td>-----</td>
<td>----------</td>
<td>----------</td>
<td>--------------</td>
</tr>
<tr>
<td>13</td>
<td>pregabalin 75 MG Oral Caps...</td>
<td>Lyrica 75 MG Oral Capsule</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1639484</td>
</tr>
<tr>
<td>14</td>
<td>pregabalin 20 MG/ML Oral S...</td>
<td>Lyrica 20 MG/ML Oral Solut...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C2740404</td>
</tr>
<tr>
<td>15</td>
<td>24 HR pregabalin 165 MG Ext...</td>
<td>24 HR Lyrica 165 MG Extens...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C4541401</td>
</tr>
<tr>
<td>16</td>
<td>24 HR pregabalin 330 MG Ext...</td>
<td>24 HR Lyrica 330 MG Ext...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C4541405</td>
</tr>
<tr>
<td>17</td>
<td>24 HR pregabalin 82.5 MG Ext...</td>
<td>24 HR Lyrica 82.5 MG Ext...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C4541408</td>
</tr>
<tr>
<td>18</td>
<td>pregabalin 100 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1648911</td>
</tr>
<tr>
<td>19</td>
<td>pregabalin 150 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1648912</td>
</tr>
<tr>
<td>20</td>
<td>pregabalin 200 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1648913</td>
</tr>
<tr>
<td>21</td>
<td>pregabalin 225 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1649612</td>
</tr>
<tr>
<td>22</td>
<td>pregabalin 25 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1646070</td>
</tr>
<tr>
<td>23</td>
<td>pregabalin 300 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1649613</td>
</tr>
<tr>
<td>24</td>
<td>pregabalin 50 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1646071</td>
</tr>
<tr>
<td>25</td>
<td>pregabalin 75 MG [Lyrica]</td>
<td></td>
<td>SBDC</td>
<td>ENG</td>
<td>N</td>
<td>C1646072</td>
</tr>
</tbody>
</table>

Showing 12 to 26 of 68 entries
<table>
<thead>
<tr>
<th>rxcui</th>
<th>name</th>
<th>synonym</th>
<th>tty</th>
<th>language</th>
<th>suppress</th>
<th>umlsncui</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Esomeprazole Strontium</td>
<td></td>
<td>PIN</td>
<td>ENG</td>
<td>N</td>
<td>C3540733</td>
</tr>
<tr>
<td>13</td>
<td>esomeprazole magnesium dihydr...</td>
<td></td>
<td>PIN</td>
<td>ENG</td>
<td>N</td>
<td>C3883231</td>
</tr>
<tr>
<td>14</td>
<td>esomeprazole sodium</td>
<td></td>
<td>PIN</td>
<td>ENG</td>
<td>N</td>
<td>C1337241</td>
</tr>
<tr>
<td>15</td>
<td>Esomeprazole 20 MG Injection [N...</td>
<td>Nexium I.V. 20 MG I...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1631632</td>
</tr>
<tr>
<td>16</td>
<td>Esomeprazole 40 MG Injection [N...</td>
<td>Nexium I.V. 40 MG I...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1630443</td>
</tr>
<tr>
<td>17</td>
<td>Esomeprazole 20 MG Delayed Release [N...</td>
<td>Nexium 20 MG Delayed Release [N...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1637815</td>
</tr>
<tr>
<td>18</td>
<td>Esomeprazole 40 MG Delayed Release [N...</td>
<td>Nexium 40 MG Delayed Release [N...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1638385</td>
</tr>
<tr>
<td>19</td>
<td>Esomeprazole 10 MG Granules for...</td>
<td>Nexium 10 MG Granules for...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C2343719</td>
</tr>
<tr>
<td>20</td>
<td>Esomeprazole 20 MG Granules for...</td>
<td>Nexium 20 MG Granules for...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1876474</td>
</tr>
<tr>
<td>21</td>
<td>Esomeprazole 40 MG Granules for...</td>
<td>Nexium 40 MG Granules for...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C1876475</td>
</tr>
<tr>
<td>22</td>
<td>Esomeprazole 20 MG Delayed Release [N...</td>
<td>Nexium 20 MG Delayed Release [N...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C4084362</td>
</tr>
<tr>
<td>23</td>
<td>Esomeprazole 2.5 MG Granules for...</td>
<td>Nexium 2.5 MG Granules for...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C3472929</td>
</tr>
<tr>
<td>24</td>
<td>Esomeprazole 5 MG Granules for...</td>
<td>Nexium 5 MG Granules for...</td>
<td>SBD</td>
<td>ENG</td>
<td>N</td>
<td>C3472968</td>
</tr>
</tbody>
</table>
Part 3 – Experience with OHDSI

Olivier Bodenreider, MD, PhD, NLM
Vojtech Huser, MD, PhD, NLM
Christian Reich, MD, PhD, IQVIA
Part 3 overview

Clinical data models

Handling international drugs
Clinical data models

Part 3 – Experience with OHDSI
FDA Regulatory Action over Time

Number of FDA-caused Withdrawals


0 5 10 15 20 25 30
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—

“(I) at least 25
2010; and

“(II) at least 10
2012; and

“(iii) convene a committee of individuals who are recognized for their data privacy and security expertise to the Secretary on the methods for the ethical communication of, posts in subparagraph (C), including the development of effective methods of drug safety questions.

“(C) ESTABLISHMENT OF RISK IDENTIFICATION AND ANALYSIS SYSTEM:

Risk Identification and Analysis System:
a systematic and reproducible process to efficiently generate evidence to support the characterization of the potential effects of medical products from across a network of disparate observational healthcare data sources
OMOP Experiment 1 (2009-2010)

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

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: erythromycin, sulfonamides, tetracyclines</th>
<th>Antiepileptics: carbamazepine, phenytoin</th>
<th>Benzodiazepines</th>
<th>Beta Blockers</th>
<th>Bisphosphonates: alendronate</th>
<th>Tricyclic antidepressants</th>
<th>Typical antipsychotics</th>
<th>Warfarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angioedema</td>
<td>Red</td>
<td></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)

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

Observational data
- 4 claims databases
- 1 ambulatory EMR

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

Observational data

- Aarhus
  - ARS
  - HS

- Pedianet
  - IPCI
  - PHARMO

Drug-outcome pairs

<table>
<thead>
<tr>
<th>Outcome</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>

Methods

- Case-Control
- New User Cohort
- Disproportionality methods
- ICTPD
- LGPS
- Self-Controlled Cohort
- SCCS
<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aarhus</td>
<td>Danish national health registry, covering the Aarhus region. Includes inhabitant registry, drug dispensations, hospital claims, lab values, and death registry.</td>
<td>🇩🇰2 M</td>
</tr>
<tr>
<td>ARS</td>
<td>Italian record linkage system covering the Tuscany region, including inhabitant registry, drug dispensations, hospital claims, and death registry</td>
<td>🇮🇹4 M</td>
</tr>
<tr>
<td>Health-Search</td>
<td>Italian general practice database (no children)</td>
<td>🇮🇹1 M</td>
</tr>
<tr>
<td>IPCI</td>
<td>Dutch general practice database</td>
<td>🇳🇱0.75 M</td>
</tr>
<tr>
<td>Pedianet</td>
<td>Italian general practice pediatric database</td>
<td>🇮🇹0.14 M</td>
</tr>
<tr>
<td>PHARMO</td>
<td>Dutch record linkage system. Includes inhabitant registry, drug dispensations, hospital claims, and lab values.</td>
<td>🇳🇱1.28 M</td>
</tr>
</tbody>
</table>
Results

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP
• 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
OMOP: ended in 2013 with Symposium

IMEDS: Program at Reagan-Udall Foundation of the FDA
  • Methodological research to inform Industry and Agency
  • Research Lab

OHDSI: Open Research Collaborative started by OMOP Pis and coordinated through Columbia University
  • Multiple stakeholders: academia, government, industry
  • Multiple geographies: US, Europe, Asia-Pacific
  • Multiple disciplines: Statistics, epidemiology, informatics, clinical sciences
  • Maintains OMOP CDM and Vocabularies
OHDSI’s vision

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

Join us on the journey

http://ohdsi.org
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
A caricature of the patient journey
Questions asked across the patient journey

- Which treatment did patients choose after diagnosis?
- Which patients chose which treatments?
- How many patients experienced the outcome after treatment?
- Does one treatment cause the outcome more than an alternative?
- Does treatment cause the outcome?
- What is the probability I will develop the disease?
- What is the probability I will experience the outcome?
Classifying questions across the patient journey

Clinical characterization: What happened to them?
• What treatment did they choose after diagnosis?
• Which patients chose which treatments?
• How many patients experienced the outcome after treatment?

Patient-level prediction: What will happen to me?
• What is the probability that I will develop the disease?
• What is the probability that I will experience the outcome?

Population-level effect estimation: What are the causal effects?
• Does treatment cause outcome?
• Does one treatment cause the outcome more than an alternative?
OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP
OMOP Common Data Model

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

Standardized health system data
- Location
- Care_site
- Provider
- Payer_plan_period
- Visit_cost
- Procedure_cost
- Drug_cost
- Device_cost

Standardized meta-data
- Concept
- Vocabulary
- Concept_relationship
- Relationship
- Concept_synonym
- Concept_ancestor
- Source_to_concept_map
- Drug_strength
- Cohort_definition

Standardized derived elements
- Cohort
- Drug_era
- Dose_era
- Condition_era

Standardized health economics

Standardized vocabulary
Drug Hierarchy

Classifications
- VA Class
- CVX
- NDFRT
- NDFRT Ind
- ATC
- FDB Ind
- ETC
- SPL
- SNOMED

Drugs
- Source codes:
  - CIEL
  - NDC
  - GPI
  - VA-Product
  - Gemscript
  - EU Product
  - DPD
  - HCPCS
  - MeSH
  - Multum
  - Oxmis
  - Read
  - Genseqno
  - dm+d
  - AMIS
  - BDPM
  - CPT4

Ingredients
- Drug Forms and Components
- Drug products
- RxNorm
- RxNorm Extension

Drug Classes
- Standard Drug Vocabulary:
  - Procedure Drugs
  - Source codes

Drug Codes
Simple Use Case

Give me all patients who take

Plavix 75 mg Tablets
OMOP Vocabulary

Branded Drug

clopidogrel 75 MG Oral Tablet [Plavix]
OMOP Vocabulary

Branded Drug

clopidogrel 75 MG Oral Tablet [Plavix] → Plavix
Clinical Drug

clopidogrel 75 MG Oral Tablet

clopidogrel 75 MG Oral Tablet [Plavix]
OMOP Vocabulary Relationships

Drug Component

clopidogrel 75 MG

clopidogrel 75 MG Oral Tablet

clopidogrel 75 MG Oral Tablet [Plavix]

clopidogrel 75 MG [Plavix]

75 MG
Relationships
Hierarchy

Platelet Aggregation Alteration

Ancestor

Concepts

Concept Relationships

clopidogrel 75 MG

Descendant

clopidogrel 75 MG Oral Tablet [Plavix]
OMOP Vocabulary Common Data Model

1. All concepts in the concept table
2. Direct relationships between concepts listed in the concept_relationship table
3. Multi-step hierarchical relationships pre-processed in the concept_ancestor table
4. Local codes mapped to concepts through the source_to_concept_map table
CONCEPT – Single standardized Table

All vocabularies stacked up in one table

Vocabulary ID
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 = 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>
</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>5</td>
<td>213169</td>
</tr>
</tbody>
</table>
### Drugs by Ingredients

3. Check Descendants (other drug products containing Warfarin and Dabigatran)

```sql
SELECT max_levels_of_separation, descendant.*
FROM concept_ancestor a, concept_descendant
WHERE a.ancestor_concept_id = 1310149 /* Warfarin or 1322185 Clopidogrel*/
AND descendant.concept_id = descendant_concept_id;
ORDER BY max_levels_of_separation;
```
SELECT max_levels_of_separation, descendant.
FROM concept_ancestor a, concept descendant
WHERE a.ancestor_concept_id = 21600961 /* ATC Antithromboic Agent */
AND a.descendant_concept_id = descendant.concept_id
AND descendant.concept_class = 'Ingredient'
ORDER BY max_levels_of_separation;
Fulfilling the Use Case

Example queries for the Drug Era table

```sql
/*
Find all periods of exposure for patients exposed to warfarin
*/

CREATE TEMPORARY TABLE warfarin_all_exposures AS
SELECT
  person_id,
  drug_concept_id,
  drug_era_start_date,
  drug_era_end_date
FROM
  drug_era
WHERE
  drug_concept_id IN (1310149 /* warfarin */) 
```

WHERE
```
drug_concept_id IN (1310149 /* warfarin */) 
```
Fulfilling Another Use Case

Example queries for the Drug Era table

```sql
/* Find all periods of exposure for patients exposed to warfarin */

CREATE TEMPORARY TABLE warfarin_all_exposures AS
SELECT person_id, drug_concept_id, drug_era_start_date, drug_era_end_date
FROM drug_era
WHERE drug_concept_id IN (1310149 /* warfarin */)
```

```sql
WHERE
  drug_concept_id IN (1310149 /* warfarin */)
```

```sql
WHERE
  drug_concept_id IN (
    SELECT descendant.concept_id
    FROM concept_ancestor a, concept descendant
    WHERE a.ancestor_concept_id = 21500803 /* ETC Anticoagulants */
    AND a.descendant_concept_id = descendant.concept_id
  )
```
Many Other Use Cases

OMOP Vocabulary Queries

Drug Queries

The following drug domain queries are available:

- D01: Find drug concept by concept ID
- D02: Find drug or class by keyword
- D03: Find ingredients of a drug
- D04: Find drugs by ingredient
- D05: Find generic drugs by ingredient
- D06: Find branded drugs by ingredient
- D07: Find single ingredient drugs by ingredient
- D08: Find drug classes for a drug or ingredient
- D09: Find drugs by drug class
- D10: Find ingredient by drug class
- D11: Find source codes by drug class
- D12: Find indications for a drug
- D13: Find indications as condition concepts for a drug
- D14: Find drugs for an indication
- D15: Find drugs for an indication provided as condition concepts
- D16: Find drugs for an indication by indication type
- D17: Find ingredients for an indication

http://vocabqueries.omop.org/
Handling international drugs ("RxNorm extension")

Part 3 – Experience with OHDSI
Germany: AMIS

Ergebnisse

Suchschnitt

Sortierte Suchschnitte sind mit einem * markiert
» alle Suchschnitte anzeigen

☐ 2. Arzneimittelname: clopidogrel?

1. AJ29
Stand: 11.11.2016 07:32:00

» und » » oder » » nicht »
» sortieren »
» abbrechen » "Hilfe » » top »

Arzneimittel

Suchschnitt: Arzneimittelname: clopidogrel?

Ergebnis 1-10 von 338

<table>
<thead>
<tr>
<th>Art. Name</th>
<th>Dar.-form</th>
<th>Anmelder</th>
<th>Freie Infos</th>
<th>€</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clopidogrel 1 A Pharma 75mg Filmtabletten</td>
<td></td>
<td>1 A Pharma GmbH</td>
<td></td>
<td>1,82 EUR</td>
</tr>
<tr>
<td>Clopidogrel 1A Pharma 75 mg Filmtabletten - OP100</td>
<td></td>
<td>Adino Pharma GmbH (BS 2) ...</td>
<td></td>
<td>1,82 EUR</td>
</tr>
<tr>
<td>Clopidogrel 1A Pharma 75 mg Filmtabletten - OP14</td>
<td></td>
<td>Adino Pharma GmbH (BS 2) ...</td>
<td></td>
<td>1,82 EUR</td>
</tr>
<tr>
<td>Clopidogrel 1A Pharma 75 mg Filmtabletten - OP28</td>
<td></td>
<td>Adino Pharma GmbH (BS 2) ...</td>
<td></td>
<td>1,82 EUR</td>
</tr>
<tr>
<td>Clopidogrel 1A Pharma 75 mg Filmtabletten - OP30</td>
<td></td>
<td>Adino Pharma GmbH (BS 2) ...</td>
<td></td>
<td>1,82 EUR</td>
</tr>
<tr>
<td>Clopidogrel 1A Pharma 75 mg Filmtabletten - OP31</td>
<td></td>
<td>Adino Pharma GmbH (BS 2) ...</td>
<td></td>
<td>1,82 EUR</td>
</tr>
</tbody>
</table>
France

Format de données
Le lien suivant vous permet de télécharger la base de données publique des médicaments en français, ainsi que les fichiers dédiés à la transparence.

Liste des fichiers
La base de données publique des médicaments est mise à disposition des usagers au moyen des fichiers offerts au téléchargement libre et gratuit.

Termes du contrat de licence
Conformément à l'article L. 161-40-1 du code de la sécurité sociale dans sa rédaction issue de l'article 8 de la loi n° 2011-2012 du 29/12/2011, l'information contenue dans la base de données publique des médicaments (BDPM) est mise à disposition des usagers au moyen des fichiers offerts au téléchargement libre et gratuit.

Vous êtes libres de :
- Reproduire, diffuser, redistribuer et exploiter les données

Fichier des compositions (date de mise à jour : 28/10/2016, 2086 Ko)
Fichier des avis SMR de la HAS (date de mise à jour : 28/10/2016, 2386 Ko)
Fichier des avis ASMR de la HAS (date de mise à jour : 28/10/2016, 1561 Ko)
Fichier des liens vers les avis de la commission de la transparence de la HAS (date de mise à jour : 28/10/2016, 347 Ko)
Fichier des groupes génériques (date de mise à jour : 28/10/2016, 937 Ko)
Fichier des conditions de prescription et de délivrance (date de mise à jour : 28/10/2016, 716 Ko)
Fichier des informations importantes (génération en direct)
Drug Product Database - All Files

The Drug Product Database (DPD) system captures information on Canadian human, veterinary and disinfectant products approved for use by Health Canada. This extract contains both marketed (active) and discontinued (inactivated) products in separate files. This data extract contains information about Drug Product - All Files.

Publisher - Current Organization Name: Health Canada

Licence: Open Government Licence - Canada

Resources

<table>
<thead>
<tr>
<th>Resource Name</th>
<th>Resource Type</th>
<th>Format</th>
<th>Language</th>
<th>Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Products</td>
<td>Dataset</td>
<td>CSV</td>
<td>English</td>
<td>Access</td>
</tr>
<tr>
<td>Inactive Products</td>
<td>Dataset</td>
<td>CSV</td>
<td>English</td>
<td>Access</td>
</tr>
</tbody>
</table>
Plavix Prepackaged Product
Resources

Part 3 – Experience with OHDSI
RxNorm Extension Composition

- **BDMD**
  - Ingredient Drug Products: 2081, 243
  - Count: 574, 13,238

- **NDC**
  - Ingredient Drug Products: 2,745
  - Count: 19,139

- **dm+d**
  - Ingredient Drug Products: 2,240, 171
  - Count: 6,443, 25,203

Legend:
- **Blue**: RxNorm
- **Red**: RxNorm Extension
Drug Forms Internationally

142
Resources

1. Download
http://athena.ohdsi.org

2. Rebuild (not for the faint of heart)
https://github.com/OHDSI/Vocabulary-v5.0

3. Documentation (incomplete still)
Questions

Olivier Bodenreider, MD, PhD, NLM
olivier.bodenreider@nih.gov

Vojtech Huser, MD, PhD, NLM
vojtech.huser@nih.gov

Christian Reich, MD, PhD, IQVIA
reich@ohdsi.org