ISPOR 21ST ANNUAL INTERNATIONAL MEETING
May 21 - 25, 2016
Washington Hilton, Washington, DC, USA

ISPOR SHORT COURSES

Informatics and Interoperability: Speaking the Same Language

Presented by:

Scott D. Nelson, PharmD, MS, Assistant Professor, Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA

Olivier Bodenreider, MD, PhD, Branch Chief, Cognitive Science Branch
U.S. National Library of Medicine, Bethesda, MD, USA

Daniel C. Malone, PhD, RPh, MS, Professor, Department of Pharmacy Practice & Science, University of Arizona, Tucson, AZ, USA

Richard D. Boyce, PhD, Assistant Professor, Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA

Saturday, May 21, 2016

Objectives

- Describe three levels of information interoperability and explain why they are important to pharmacoeconomic research.
- Explain the strengths and limitations of different methods for identifying medications within a clinical data set.
- Access and utilize medication information contained in RxNorm.
- List several potential strengths and limitations of using a common data model and standard vocabulary for pharmacoeconomic and pharmacoepidemiologic research.

Meet and greet!

- Communication
  - Meet your neighbor
    - Name
    - Where you're from
    - Why you're here
    - Something interesting

What is interoperability?

"the ability of two or more systems or components to exchange information and to use the information that has been exchanged."

- Institute for Electrical and Electronics Engineering (IEEE)

Example

- Send a message to your neighbor
  - Pt. X (1234) has a blood pressure of 120/80
  - Pt. Y (9876) is taking 81 mg of aspirin by mouth once a day
  - Pt. Z (4567) is blood type O negative
  - Must be parse-able
    - Special character for separating fields
      - .,|\@_-*

Example

- Send a message to your neighbor
  - What was the structure or sequence of things?
  - Was there a coding scheme?
    - Or was it free-text?
  - Using the rules for the message you created, could you understand the message?

Levels of interoperability

Foundational

- Ability to communicate
- Send and receive messages

Structural

- Functional understanding
- Can understand data fields, but not contents

Semantic

- Shared understanding
- Recognize and interpret data fields and contents

Foundational Interoperability

Foundational Interoperability

Hospital A	Patient taking APAP, and has serum creatinine of 1.2
Hospital B	50580-112-06|1|PO|Q4h OBX|1|CE|A028^S. CRT
Hospital C	D000082_1_PO_MESHPA 1.2_scrt_1245_pt

Structural Interoperability

CONSULADO GENERAL DE PANAMA

SOLICITUD DE VISAS AUTORIZADAS

FECHA: _____

NOMBRE COMPLETO: _____

PASAPORTE No._____ EXPIRACION:_____

NACIONALIDAD:_____ OCUPACION:_____
OCCUPATION:

FECHA DE NACIMIENTO:_____ PAIS DE RESIDENCIA_____

DEPENDIENTES:_____

VISA QUE SOLICITA:_____ TIEMPO DE ESTADIA:_____

MOTIVO DE VIAJE:_____

VISAS DE OTROS PAISES:_____

Structural Interoperability

ANSI

ISO

NCPDP™

DICOM℠
Digital Imaging and Communications in Medicine

ASC X12
The Accredited Standards Committee

HL7 INTERNATIONAL

Structural Interoperability

- HL7 V2.x
- HL7 V3.x
- Clinical Document Architecture (CDA)

Helps structure the information being sent

Structural Interoperability

Hospital A	OBX||1|CE|2730-72^Creatinine
Hospital B	OBX||1|CE|A028^S. CRT
Hospital C	OBX||1|CE|35^Serum Creatinine

HL7 2.x

Example modifications

- Make the following modifications to you message:
  - Blood pressure in the right arm
  - Blood pressure laying down
  - History of aspirin ¼ of a 325 mg tablet
  - Blood type of fetus
  - Blood type of donor
  - Negation and uncertainty

Pre- vs post-coordination

- warfarin 5 mg tablet

Post-coordinated

Medication

Ingredient: Warfarin

Dose form: tablet

Strength: 5 mg

Medication

SCDF: Warfarin tab

Strength: 5 mg

Pre-coordinated

Medication

SCD: Warfarin 5mg tab

Structural Interoperability

The/dog/eats/red/meat

Article/subject noun/verb/adjective/direct object noun

The/cat/sings/blue/trees

Structural Interoperability

When is a cold just a cold?

- Cold – sickness (patient has a cold)
- Cold – temperature (object is cold – give pt. cold packs)
- Cold – feeling (pt. complains of feeling cold in here)
- Cold – symbolism (gave me the cold shoulder)
- COLD - Computer Output to Laser Disk (term for CD or DVD)

Semantic Interoperability

Form **1040** Department of the Treasury—Internal Revenue Service (99)
U.S. Individual Income Tax Return **2014** OMB No. 1545-0074 IRS Use Only—Do not write or staple in this space.

For the year Jan. 1–Dec. 31, 2014, or other tax year beginning , 2014, ending , 20 See separate instructions.

Your first name and initial | Last name | Your social security number

If a joint return, spouse's first name and initial | Last name | Spouse's social security number

Home address (number and street). If you have a P.O. box, see instructions. | Apt. no.

▲ Make sure the SSN(s) above and on line 6c are correct.

City, town or post office, state, and ZIP code. If you have a foreign address, also complete spaces below (see instructions).

Presidential Election Campaign
Check here if you, or your spouse if filing jointly, want $3 to go to this fund. Checking a box below will not change your tax or refund. ☐ You ☐ Spouse

Foreign country name | Foreign province/state/county | Foreign postal code

Filing Status
Check only one box.

1 ☐ Single
2 ☐ Married filing jointly (even if only one had income)
3 ☐ Married filing separately. Enter spouse's SSN above and full name here. ▶
4 ☐ Head of household (with qualifying person). (See instructions.) If the qualifying person is a child but not your dependent, enter this child's name here. ▶
5 ☐ Qualifying widow(er) with dependent child

Exemptions

6a ☐ **Yourself.** If someone can claim you as a dependent, **do not** check box 6a . . . . .
b ☐ **Spouse** . . . . . .

Boxes checked on 6a and 6b

c **Dependents:**

(1) First name Last name	(2) Dependent's social security number	(3) Dependent's relationship to you	(4) ✓ if child under age 17 qualifying for child tax credit (see instructions)
			☐
			☐
			☐
			☐

No. of children on 6c who:
• lived with you
• did not live with you due to divorce or separation (see instructions)

Dependents on 6c not entered above

If more than four dependents, see instructions and check here ▶ ☐

d Total number of exemptions claimed

Add numbers on lines above ▶

Structural vs Semantic

Structural

Source	Lab	Value	Units
Hospital A	2730-72	1.6	mg/dl
Hospital B	A028	0.014	g/L
Hospital C	35	0.9	mg/dl

Semantic

Source	Lab	Value	Units
Hospital A	2160-0	1.6	mg/dl
Hospital B	2160-0	1.4	mg/dl
Hospital C	2160-0	0.9	mg/dl

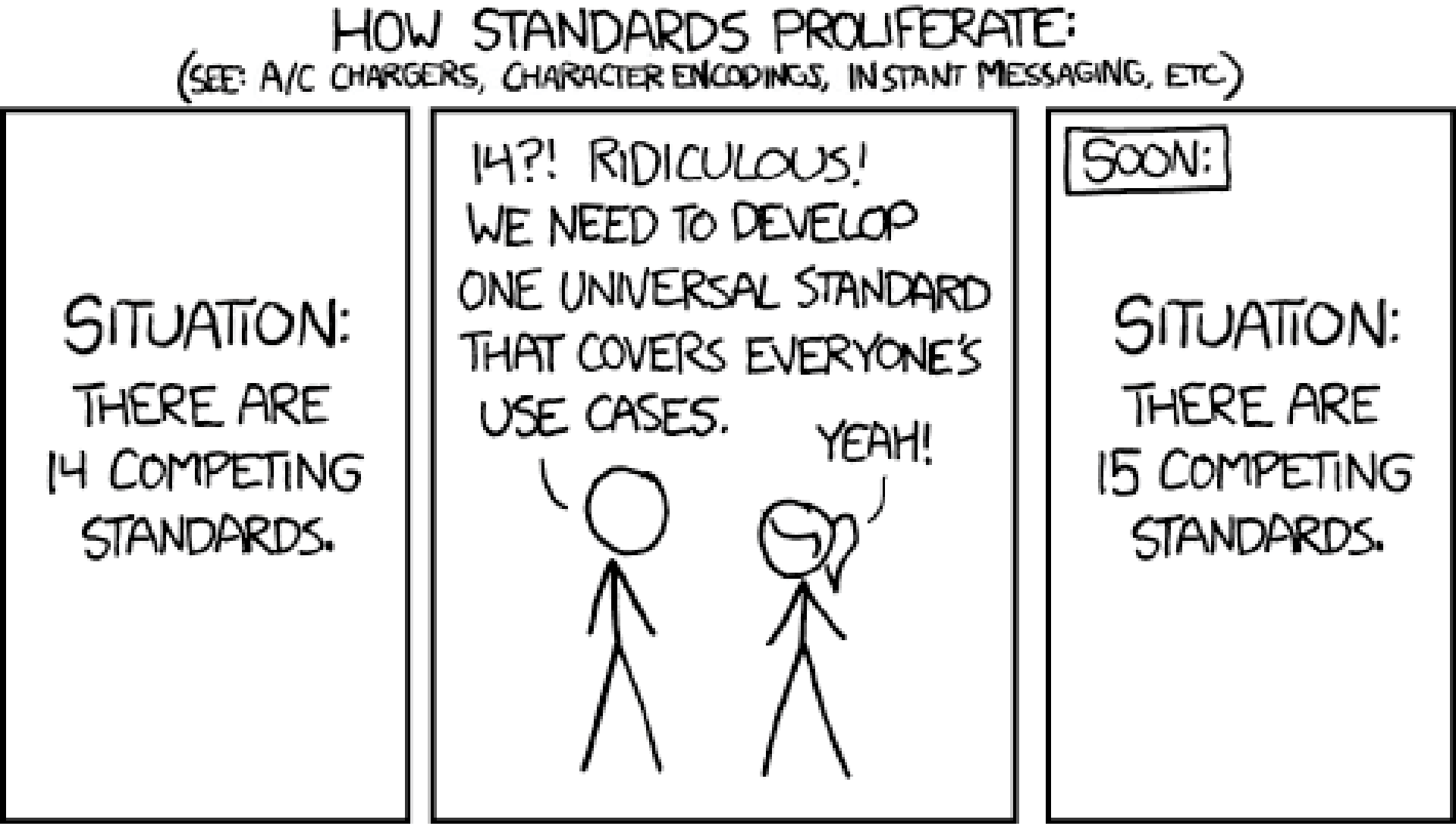
Semantic Interoperability

Requires standard format and standard terminology

ihealthtran.com

Standard Vocabularies

Domain	Vocabulary
Labs	LOINC, SNOMED CT
Medications	RxNORM, NDC, ATC, NDF-RT, FDB, Medi-span, Multium, FDA, MeSH, SNOMED CT
Demographics	CDC race, HL7 administrative
Findings/conditions	SNOMED CT, MEDRA, ICD-10, ICD-9
Procedures/diagnsoses	HCPCS, ICD-10, ICD-9, CPT
Observations	LOINC, SNOMED CT



The beauty about standards is that there are many to choose from…

http://www.howtogeek.com/geekers/up/sshot50509a6b8cb11.jpg

Best available Standards

- Labs and vitals
  - LOINC
- Medications
  - RxNorm
- Observations
  - SNOMED CT

- https://www.healthit.gov/policy-researchers-implementers/section-i-best-available-vocabularycode-setterminology-standards-and

Labs – LOINC

- Logical Observation Identifiers Names and Codes (LOINC)
- Preferred for labs and vital signs
- Developed by Regenstrief
- More info at
  - loinc.org
  - Search.loinc.org

Labs – LOINC

- Component
  - Creatinine, INR, sodium, ejection fraction, etc.
- Property
  - Mass, pressure, length, temp, etc.
- Timing
  - Point in time, max, min, time range, etc.
- System
  - Urine, blood, serum, tricuspid valve, etc.
- Scale
  - Quantitative, ordinal, nominal, etc.
- Method (optional)
  - Siting, estimated, count, etc.

LOINC® from Regenstrief

creatinine

Search

LOINC	LongName	Component	Property	Timing	System	Scale	Method
20624-3	Creatinine [Mass/volume] in 24 hour Urine	Creatinine	MCnc	24H	Urine	Qn	
2159-2	Creatinine [Mass/volume] in Amniotic fluid	Creatinine	MCnc	Pt	Amnio fld	Qn	
38483-4	Creatinine [Mass/volume] in Blood	Creatinine	MCnc	Pt	Bld	Qn	
12190-5	Creatinine [Mass/volume] in Body fluid	Creatinine	MCnc	Pt	Body fld	Qn	
2160-0	Creatinine [Mass/volume] in Serum or Plasma	Creatinine	MCnc	Pt	Ser/Plas	Qn	
2161-8	Creatinine [Mass/volume] in Urine	Creatinine	MCnc	Pt	Urine	Qn	
35674-1	Creatinine [Mass/volume] in Urine collected for unspecified duration	Creatinine	MCnc	XXX	Urine	Qn	

Medications – RxNorm

- Unique identifiers for drugs and ingredients
- Crosswalk between other terminologies
- Produced by National Library of Medicine (NLM)
- More info at
  - **RxNorm**
    - http://www.nlm.nih.gov/research/umls/rxnorm/
  - RxNav
    - Visual navigation of RxNorm
    - http://mor.nlm.nih.gov/download/rxnav/RxNavDoc.html
  - RxMix
    - API for RxNorm and RxImageAccess
    - http://mor.nlm.nih.gov/RxMix/

Observations – SNOMED CT

- Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT)
- Contains
  - Concepts, descriptions, and relationships
- Maintained by the International Health Terminology Standards Development Organization (IHTSDO)
- More info at
  - https://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html
  - http://browser.ihtsdotools.org/

IHTSDO SNOMED CT Browser

Release: United States Edition 20150901 | Perspective: Full | Feedback | About

© IHTSDO 2015

(finding)

- Kidney disease (disorder)
  - Absent renal function (disorder)
  - Acquired immune deficiency syndrome-related nephropathy (disorder)
  - Acute nephropathy (disorder)
  - Adhesions of kidney (disorder)
  - Amyloid nephropathy (disorder)
  - Anemia of renal disease (disorder)
  - Caliectasis (disorder)
  - Calyceal fistula (disorder)
  - Complication of transplanted kidney (disorder)
  - Congenital anomaly of fetal kidney (disorder)
  - Congenital anomaly of the kidney (disorder)
  - Congenital pelviureteric junction obstruction (disorder)
  - Congenital renal atrophy (disorder)

Concept Details

Summary | Details | Diagram | Refsets | Members | References

Stated | Inferred

Parents

- Renal impairment (disorder)

Renal failure syndrome (disorder)

SCTID: 42399005

42399005 | Renal failure syndrome (disorder) |

Renal failure syndrome (disorder)
RF - Renal failure
Renal failure
Renal failure syndrome
Renal insufficiency
Renal insufficiency syndrome

Finding site → Kidney structure

Children (13)

- Acute renal failure syndrome (disorder)
- Anemia secondary to renal failure (disorder)
- Chronic renal failure syndrome (disorder)

Diagnosis codes – ICD-10

- International Statistical Classification of Diseases and Related Health Problems, 10th revision Maintained by the World Health Organization (WHO)
- More info at

http://www.icd10data.com/ICD10PCS/Codes

ICD10Data.com

ICD-9: 250 Diabetes mellitus

- 250.0 Diabetes mellitus without mention of complication
- 250.1 Diabetes with ketoacidosis
- 250.2 Diabetes with coma
- 250.3 Diabetes with renal manifestations
- 250.4 Diabetes with ophthalmic manifestations
- 250.5 Diabetes with neurological manifestations
- 250.6 Diabetes with peripheral circulatory disorders
- 250.7 Diabetes with other specified manifestations
- 250.9 Diabetes with unspecified complications

ICD-10: E10-E14 Diabetes Mellitus

- E10 Insulin-dependent diabetes mellitus
- E11 Non-insulin-dependent diabetes mellitus
- E12 Malnutrition- related diabetes mellitus
- E13 Other specified diabetes mellitus
- E14 Unspecified

.0 With coma
.1 With ketoacidosis
.2+ With renal complications
.3+ With ophthalmic complications
.4+ With neurological complications
.5 With peripheral circulatory complications
.6 With other specified complications
.7 With multiple complications
.8 With unspecified complications
.9 Without complications

Procedures – CPT

- Current Procedural Terminology (CPT)
- Part of Healthcare Common Procedural Coding System
  - HCPCS codes by CMS
- Contains
  - Billable procedures
- Maintained by the American Medical Association (AMA)
- More info at
  - http://www.ama-assn.org/ama/pub/physician-resources/solutions- managing-your-practice/coding-billing-insurance/cpt.shtml
  - http://www.cms.gov/Medicare/Coding/MedHCPCSGenInfo/HCPCS_Coding_Questions.html

CPT code examples

92526 Treatment of swallowing dysfunction and/or oral function for feeding

92507 Treatment of speech, language, voice, communication, and/or auditory processing disorder, individual

92521 Evaluation of speech fluency (eg, stuttering, cluttering)

92522 Evaluation of speech sound production (eg, articulation, phonological process, apraxia);

92524 Behavioral and qualitative analysis of voice and resonance

Questions

Issues in Drug Product Identification

Dan Malone, RPh, PhD
Professor
University of Arizona

Pharmaceutical Claims

- National Council for Prescription Drug Programs (NCPDP) Standard
- Telecommunication between pharmacies and payers

Mandatory
Transaction Header Segment
Segment Separator
Insurance Segment
Situational
Segment Separator
Patient Segment
Mandatory - first Claim/Encounter
Group Separator
Segment Separator
Claim Segment
Segment Separator
Pricing Segment
Situational
Segment Separator
Pharmacy Provider Segment
Segment Separator
Prescriber Segment
Segment Separator
Coordination of Benefits/Other Payments Segment
Segment Separator
Workers' Compensation Segment
Segment Separator
DUR/PPS Segment
Segment Separator
Coupon Segment
Segment Separator
Compound Segment
Segment Separator
Clinical Segment
Segment Separator

Pharmaceutical Claims

- Real-time pharmaceutical claims permit assessment of drug exposure
- Common elements in prescription claims:
  - Date of service (Mandatory)
  - Provider (pharmacy)
  - Cardholder ID (Mandatory)
  - Person code (Situational)
  - Product identifier
  - Product quantity
  - Prescriber identifier

Pharmaceutical claim REQUIRED data elements

- Product service ID (NDC)
- Prescription number
- Quantity dispensed
- Days supply
- Date prescription written
- Ingredient cost
- Gross amount due

Pharmaceutical claim OPTIONAL data elements

- Patient first name
- Number of refills authorized
- Unit of measure
- Prior authorization number
- Patient paid amount
- Prescriber ID; Prescriber Last Name

Drug Knowledge Databases

- Proprietary companies build and maintain drug knowledge databases
- Primary purpose: Drug pricing
- Secondary purposes (not exhaustive):
  - Pharmacy reimbursement
  - Clinician decision support
    - Drug-drug interactions
    - Auxiliary labels
    - Consumer drug information

Drug Product Identification

- How do I identify the products of interest for my study?
  - Drug name
    - Multiple spellings
  - National Drug Codes
    - Useful for limited number of products
    - Very specific
  - Therapeutic Classes
    - Useful for many “similar” products
    - Not specific

Methods to Identify Exposure

- Drug Name
- Manufacturer NDC codes
  - Single source – limited manufacturers
  - Multiple source (generic) – many manufacturers
- Therapeutic classification

Product Name from Medicaid Claims

Product Name	Frequency
COUMADIN	270
COUMADIN TAB	13
COUMADIN TABLET	215
JANTOVEN	130
JANTOVEN TAB	24
JANTOVEN TABLET	262
WARFARIN	1,093
WARFARIN TAB	763
WARFARIN SODIUM	8,717
WARFARIN SODIUM TA	95
WARFARIN SODIUM TAB	516
WARFARIN SODIUM TABL	587
WARFARIN SODIUM TABLET	8,555

Product Names for Warfarin in Medi-Span Drug Database

Product	Frequency	Percentage
Coumadin	117	28.2
Jantoven	42	10.1
Warfarin Sodium	253	61.0
Warfarin Sodium Amorphous	2	0.5
Warfarin Sodium Clathrateform	1	0.2
Total NDCs	415	100.0

Select "Warfarin" Observations from Medi-Span File

prodNam	totPkgQty	pkgDesc	mfgName	labelerTyp	fmtIdNbr	genericNam
WARFARIN SODIUM	1,000.00	BOTTLE	TEVA PHARMACEUTICALS USA	G	00555-0831-05	Warfarin Sodium
JANTOVEN	100.00	BOTTLE	UPSHER-SMITH	O	00832-1211-00	Warfarin Sodium
WARFARIN SODIUM	30.00	BOTTLE	DISPENSEXPRESS	G	68115-0359-30	Warfarin Sodium
WARFARIN SODIUM	15.00	BOTTLE	PHYSICIANS TOTAL CARE	O	54868-4349-02	Warfarin Sodium
COUMADIN	100.00	BOTTLE	B-M SQUIBB U.S. (PRIMARY CARE)	B	00056-0169-70	Warfarin Sodium
WARFARIN SODIUM	100.00	BOTTLE	TARO	G	51672-4027-01	Warfarin Sodium

Brand name product

NDC

Unique Warfarin Manufacturers Listed in Medi-Span

Manufacturers/Labelers	Manufacturers/Labelers
A-S MEDICATION SOLUTIONS	NUCARE PHARMACEUTICALS
AMERICAN HEALTH PACKAGING	PALMETTO STATE PHARMACEUTICALS
AQ PHARMACEUTICALS	PCCA
B-M SQUIBB U.S. (PRIMARY CARE)	PDRX PHARMACEUTICAL
CORE PHARMACEUTICAL	PHYSICIAN PARTNER
DHS INC. WORKMAN COMP	PHYSICIANS TOTAL CARE
DISPENSEXPRESS	PREPAK SYSTEMS
DISPENSING SOLUTIONS INC.	QUALITY CARE
DRX	SANDOZ
GENPHARM LP	SOUTHWOOD PHARMACEUTICALS
H.J. HARKINS COMPANY, INC.	SPECTRUM
MALLINCKRODT PHARM	TARO
MCKESSON PACKAGING SERVICES	TEVA PHARMACEUTICALS USA
MEDISCA	UPSHER-SMITH
MEDVANTX	VA CMOP DALLAS
	ZYDUS PHARMACEUTICALS (USA)

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

NDC Elements

XXXXX - XXXX - XX

Manufacturer

Product

Packaging

NDC Forms

9999-9999-99 (4-4-2)

99999-999-99 (5-3-2)

99999-9999-9 (5-4-1)

NDC Characteristics

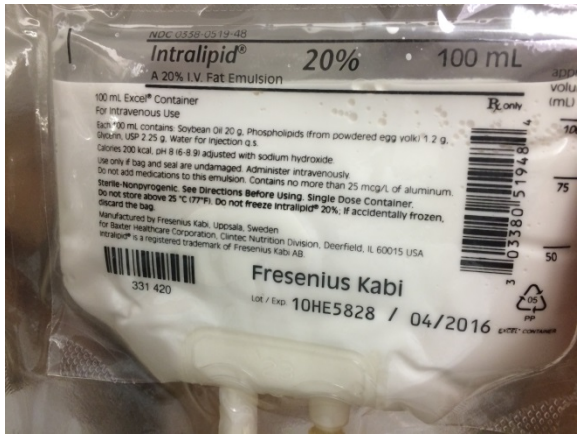
- 11 Digit code (leading zero for 4-4-2 format)
- Hyphens between segments are missing in claims transmission (Field 407 in NCPDP claim format)
- NDC codes set by the manufacturer/labeler
- Approximately 200 NDC's added/deleted per month (Source: First DataBank: AMIA 2002 annual meeting)
- Product codes 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

Issues with NDCs

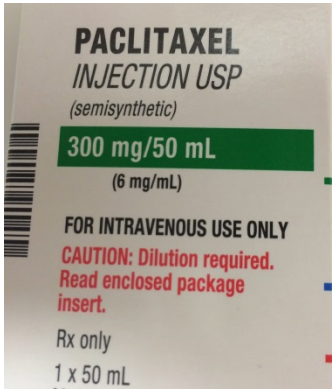
- NDC is specific to the manufacturer
- Corporate mergers will affect the NDC value (sometimes)
- Bulk purchasers and relabelers must use a new NDC code
- NDC codes can be re-used after 5 years
- Manufacturers may not follow coding "rules"
  - See Simonaitis and McDonald (AJHP 2009)

Issues with NDCs

- Reused NDC codes
  - 00074-4335-01(Liposyn)



  - 00074-4335-01 (Paclitaxel)



Lau and Shakib, Towards Data Interoperability: Practical Issues in Terminology Implementation and Mapping. AHiMA. http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_028677.hcsp?dDocName=bok1_028677

Other Considerations in using NDC

- "Old" NDCs deleted
  - Some vendors provide comprehensive database of all NDCs ever used
- Non-prescription items
  - Most payer databases won't contain OTC medications

Therapeutic Classes: An Alternative Methods for Product Identification

- Therapeutic classes are present in all proprietary databases
- Class and sub-class designation varies by vendor
- Medications might fall into multiple therapeutic categories

Some Proprietary Drug Knowledge Databases (DKB)

- First DataBank (National Drug Data File Plus)
- Wolter Kluwer (Medi-Span Master Drug Data Base)
- Cerner-Multum (Multum Lexicon)
- Thomson Corporation (Micromedex Red Book)
- Others…

Medi-Span's GPI Code

- 14 digit code to represent drug product
- Hierarchical
- Able to drill down to very specific drugs

Digits	Meaning	Example
58	Drug group	Antidepressants
58-20	Drug class	Tricyclic agents
58-20-00	Drug sub-class	
58-20-00-60	Drug name	Nortriptyline
58-20-00-60-10	Drug name extension	Hydrochloride
58-20-00-60-10-01	Dosage form	Capsule
58-20-00-60-10-01-05	Strength	10mg

First DataBank Generic Sequence Number

- Other names
  - Clinical formulation ID
  - Formerly GCN Sequence number
- 6 Digits
- No Hierarchy
- Multiple classification systems
  - Enhanced Therapeutic Classification System used for group at organ system, therapeutic class and subclass, and active ingredient (with/without salt)
  - Hierarchical Ingredient code (HIC)

First DataBank Examples

- GCN
  - 16033 - Morphine sulfate (10mg/ml Rapi-Ject)
  - 16034 - Morphine sulfate (5mg/ml Rapi-Ject)
  - 25793 - Warfarin Sodium (5mg Tablet)
- Hierarchical Ingredient Code

Characters	Information	Example
1	Organ system	Nervous system
1-2	Pharmacological class	Analgesics
1-3	Therapeutic class	Narcotic Analgesics
1-4	Ingredient (base)	Morphine
1-6	Ingredient (specific)	Morphine Sulfate

H3ADSU

AHFS Codes

- American Hospital Formulary Service therapeutic classification system
- Hierarchical classification system
  - 2 digit sequence up to 8 digits
  - Leading zero

Non-Proprietary Classification Systems

- ATC
- Department of Veterans Affairs

Anatomical Therapeutic Chemical Classification (ATC)

- Created and maintained by World Health Organization
- Based on organ system and pharmacological action
  - Single chemical can have more than one code
    - ASA
      - A01AD05 (local oral treatment)
      - B01AC06 (platelet inhibitor)
      - N02BA01 (analgesic / pyretic)

Anatomical Therapeutic Chemical Classification (ATC)

- Five different levels
  - Level One – 14 Anatomical groups

Code	Description
A	Alimentary tract and metabolism
B	Blood and blood forming organs
C	Cardiovascular
D	Dermatologicals
H	Hormonal prep (excluding sex hormones and insulin)
J	Antiinfectives for systemic use
L	Antineoplastic and immunomodulating agents

Code	Description
M	Musculoskeletal system
N	Nervous system
P	Anti-parasitic products
R	Respiratory system
S	Sensory organs
V	Various

Anatomical Therapeutic Chemical Classification (ATC)

- Five different levels
  - Level One – 14 Anatomical groups
  - Level Two - Therapeutic main group
    - 2 digits
    - Ex. A10 - drug used in diabetes

Anatomical Therapeutic Chemical Classification (ATC)

- Five different levels
  - Level One – 14 Anatomical groups
  - Level Two - Therapeutic main group
  - Level Three – Pharmacological/therapeutic subgroups
    - 1 digit
    - Ex. A10B – Blood glucose lowering (excluding insulin)

Anatomical Therapeutic Chemical Classification (ATC)

- Five different levels
  - Level One – 14 Anatomical groups
  - Level Two - Therapeutic main group
  - Level Three – Pharmacological/therapeutic subgroups
  - Level Four – Chemical subgroup
    - 1 digit
    - Ex. A10BA - Biguanides

Anatomical Therapeutic Chemical Classification (ATC)

- Five different levels
  - Level One – 14 Anatomical groups
  - Level Two - Therapeutic main group
  - Level Three – Pharmacological/therapeutic subgroups
  - Level Four – Chemical subgroup
  - Level Five – Chemical substance
    - 2 digits
    - Ex. A10BA02 - Metformin

Anatomical Therapeutic Chemical Classification (ATC)

- Characteristics of ATC
  - Immediate and slow release have same code
  - Different codes for different formulations and route of administration
  - ATC is not strictly a therapuetic classification system
  - New products not belonging to existing 4 level group will be assigned a code of X
  - Combination products typically have 5th level code as 20 or 30

VANDF Classifications

- Department of Veterans Affairs (VA) Drug Classification
  - 5 digit code
  - 2 alpha characters
    - Major classifications
      - AM – antimicrobials
      - HS – Hormones
  - 3 numeric characters
      - AM111 – Penicillins, amino acid derivatives
      - AM113 – Penicillins, extended spectrum penicillins

VA Class Sources

- Available from Pharmacy Benefits Management Services
  - http://www.pbm.va.gov/nationalformulary.asp
- Not all NDCs are assigned a VA class code
  - Some products not used in the Department of Veterans Affairs

Summary

- Defining drug exposure can be difficult
- Numerous approaches to selecting drugs of interest
- No approach is 100% accurate
- Don’t limit NDCs to originator’s codes

RxNorm in action

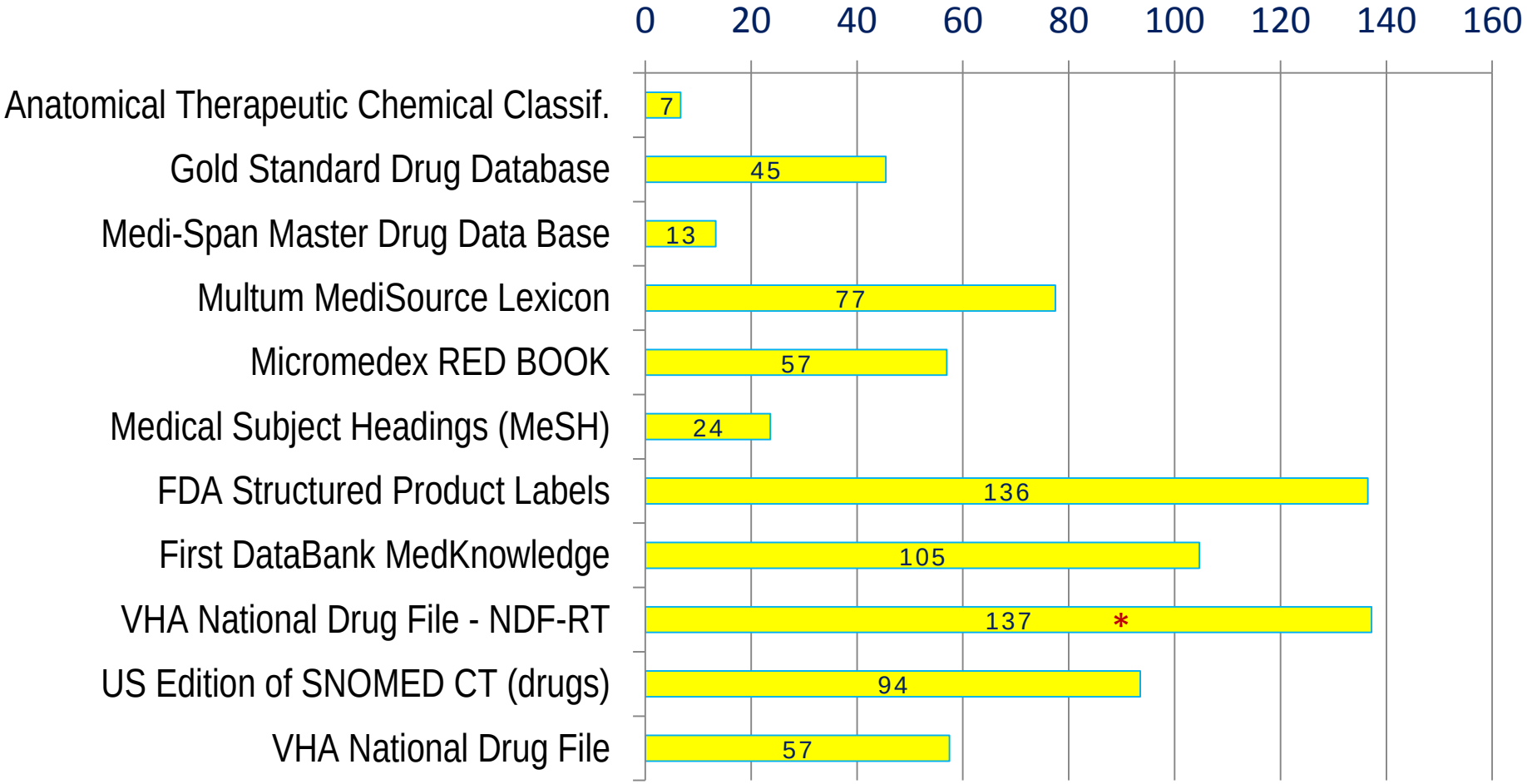
Olivier Bodenreider, MD, PhD

Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA

Interoperability among drug vocabularies

- Exchange of information requires standardized names
  - Ordering drugs
  - Checking interactions
  - Inventory management
- No standard naming conventions for drugs
- Integrating drug vocabularies
- Unique identifiers for drugs
- Specify relations among drug entities

Source vocabularies in RxNorm



(terms in thousands, as of March 2016)

Normalization Lexical level

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

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

Normalized form

Strength | **Ingredient** | **Dose form**

1 MG | *Warfarin Sodium* | *Oral Tablet*

Strength | **Ingredient**

Warfarin Sodium 1 MG (855287)

Clinical drug component

Clinical drug form

Ingredient | **Dose form**

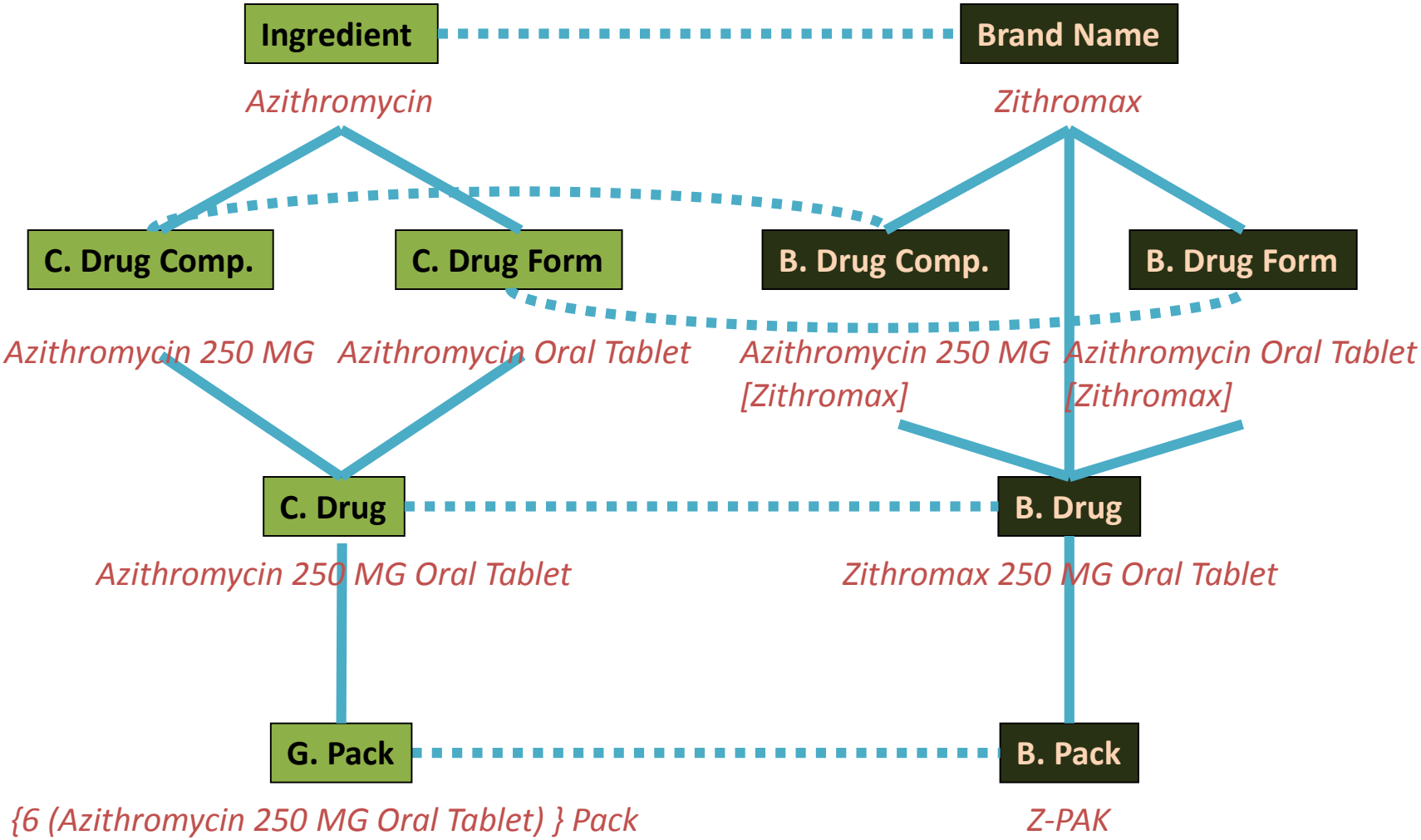
Warfarin Oral Tablet (374319)

Strength | **Ingredient** | **Dose form**

Warfarin Sodium 1 MG Oral Tablet (855288)

Clinical drug

Relations among drug entities



Example Mapping NDCs to ATC drug classes

- NDCs are attached to a clinical drug (SCD) or a branded (drug)
- Branded drugs are mapped to clinical drugs
- Clinical drugs are linked to their ingredient
- Many drug classification systems link classes to ingredient-level drugs (e.g., ATC, NDF-RT, EPC, MeSH pharmacologic action)

A **ALIMENTARY TRACT AND METABOLISM**
A02 **DRUGS FOR ACID RELATED DISORDERS**
A02B **DRUGS FOR PEPTIC ULCER AND GASTRO-OESOPHAGEAL REFLUX DISEASE (GORD)**
A02BC **Proton pump inhibitors**

ATC code	Name	DDD	U	Adm.R	Note
A02BC05	esomeprazole	30	mg	O	
		30	mg	P	

Esomeprazole (283742) → *Esomeprazole (A02BC05)*

Esomeprazole 40 MG Delayed Release Oral Capsule (606730) ← *Esomeprazole 40 MG Delayed Release Oral Capsule [Nexium] (606731)*

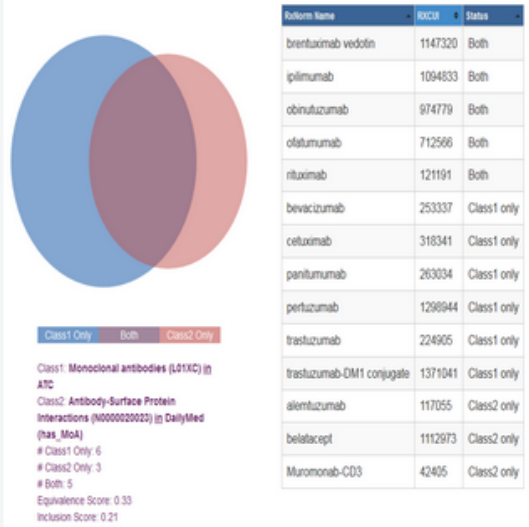
00186504031

0186-5040-31

Home APIs Applications Documentation Data Sources Terms Of Service Contact Us

RxClass

The RxClass Browser is a web application for exploring and navigating through the class hierarchies to find the RxNorm drug members associated with each class.



RxNorm Name	RXCUI	Status
brentuximab vedotin	1147320	Both
ipilimumab	1094833	Both
obinutuzumab	974779	Both
ofatumumab	712566	Both
rituximab	121191	Both
bevacizumab	253337	Class1 only
cetuximab	318341	Class1 only
panitumumab	263034	Class1 only
pertuzumab	1298944	Class1 only
trastuzumab	224905	Class1 only
trastuzumab-DM1 conjugate	1371041	Class1 only
alemtuzumab	117055	Class2 only
belatacept	1112973	Class2 only
Muromonab-CD3	42405	Class2 only

RxNav **APIs** **RxMix** **RxClass**

http://rxnav.nlm.nih.gov/

News Statistics

RxNav brand names

Recent brand names added

B-Donna	Berri-Freez Plus
BP 10-Wash	Cefotan
Clear Eyes ACR	Clinoril

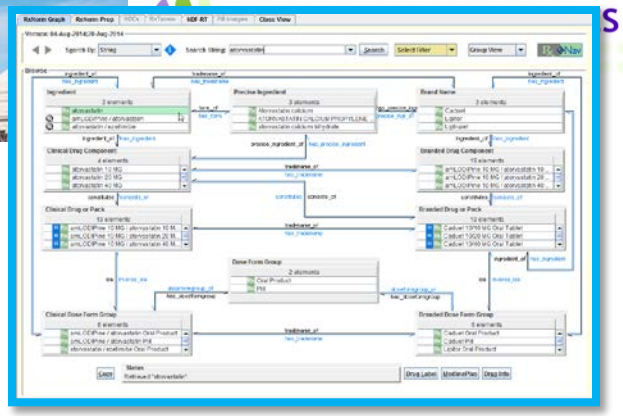
DailyMed API in RxMix

Functionality of the DailyMed API has been added to RxMix. Included are functions to retrieve drug classes, manufacturer drug names, National Drug Codes (NDCs), packaging information and drug class

MedEx added to RxMix

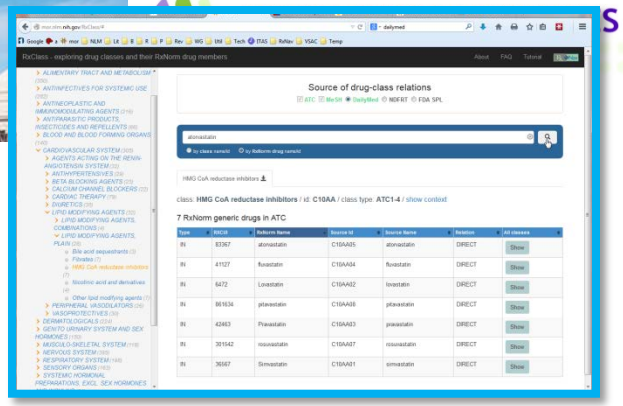
MedEx, a program to process text to recognize drug information such as medication names, drug dosage, frequency, route and duration, is available in RxMix. More details

RxNav



- Drug-centric browser
  - RxNorm
    - Prescribable subset
  - RxTerms
  - NDF-RT
  - Pill images
  - Drug-drug interactions
- Supports navigation to the rich RxNorm and NDF-RT graphs
- Links to other drug resources
  - DailyMed, MedlinePlus, NLM Drug Information Portal
- Drug-centric “class view”
- Leverages the drug APIs

RxClass

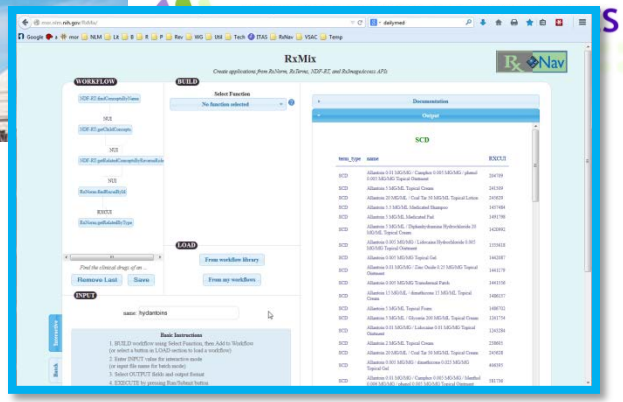


- Class-centric browser for RxNorm drugs
  - ATC
  - DailyMed (Mechanism of action, Physiologic effect, Chemical structure, FDA classes)
  - MeSH (Pharmacologic actions)
  - NDF-RT (VA classes, Diseases for indications/contra-indications)
- Supports search by drug or by class
- Features
  - Display and navigation
    - All the drugs for a class
    - All the classes for a drug
  - Compute similarity among drug classes (based on shared drug members)

APIs

- Expose the content of RxNorm, RxTerms and NDF-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, NDF-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

RxMix



- Graphical interface to the drug APIs
  - RxNorm, NDF-RT, RxTerms, RxImageAccess, 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

RxNorm use cases

- E-prescribing
  - *NCPDP SCRIPT standard for e-prescribing requires RxNorm*
- Information exchange
  - *DoD and VA rely on RxNorm to mediate drug information across their electronic medical record systems*
- Formulary development
  - *CMS uses RxNorm in their Formulary Reference File, as part of the guideline for Medicare drug benefits*
- Reference value sets
  - *The drug value sets used in clinical quality measures for Meaningful Use are defined in reference to RxNorm*
- Analytics
  - *OHDSI, the Observational Health Data Sciences and Informatics research group, uses RxNorm to analyze prescription data*

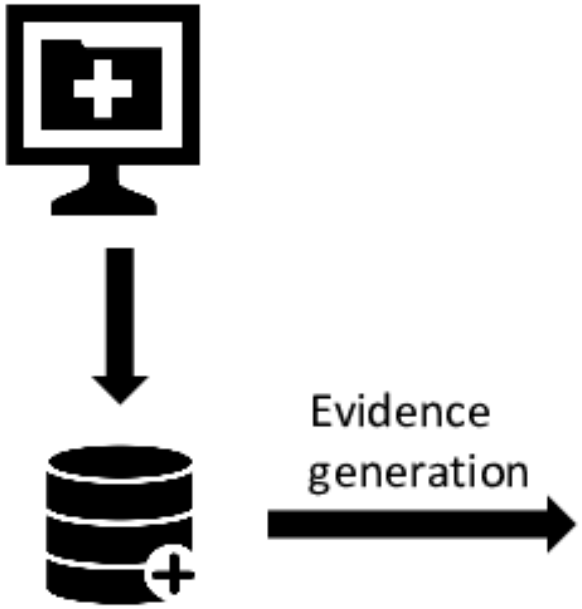
Example of the value of interoperable data for clinical research:

Richard D. Boyce[1]

[1]Department of Biomedical Informatics, University of Pittsburgh

May 21st, 2016

Databases for research



- Electronic Health Records
- Insurance claims
- Registries

- Diseases
- Health care provided
- Effects of treatments
- Differences between patients
- Personalized medicine

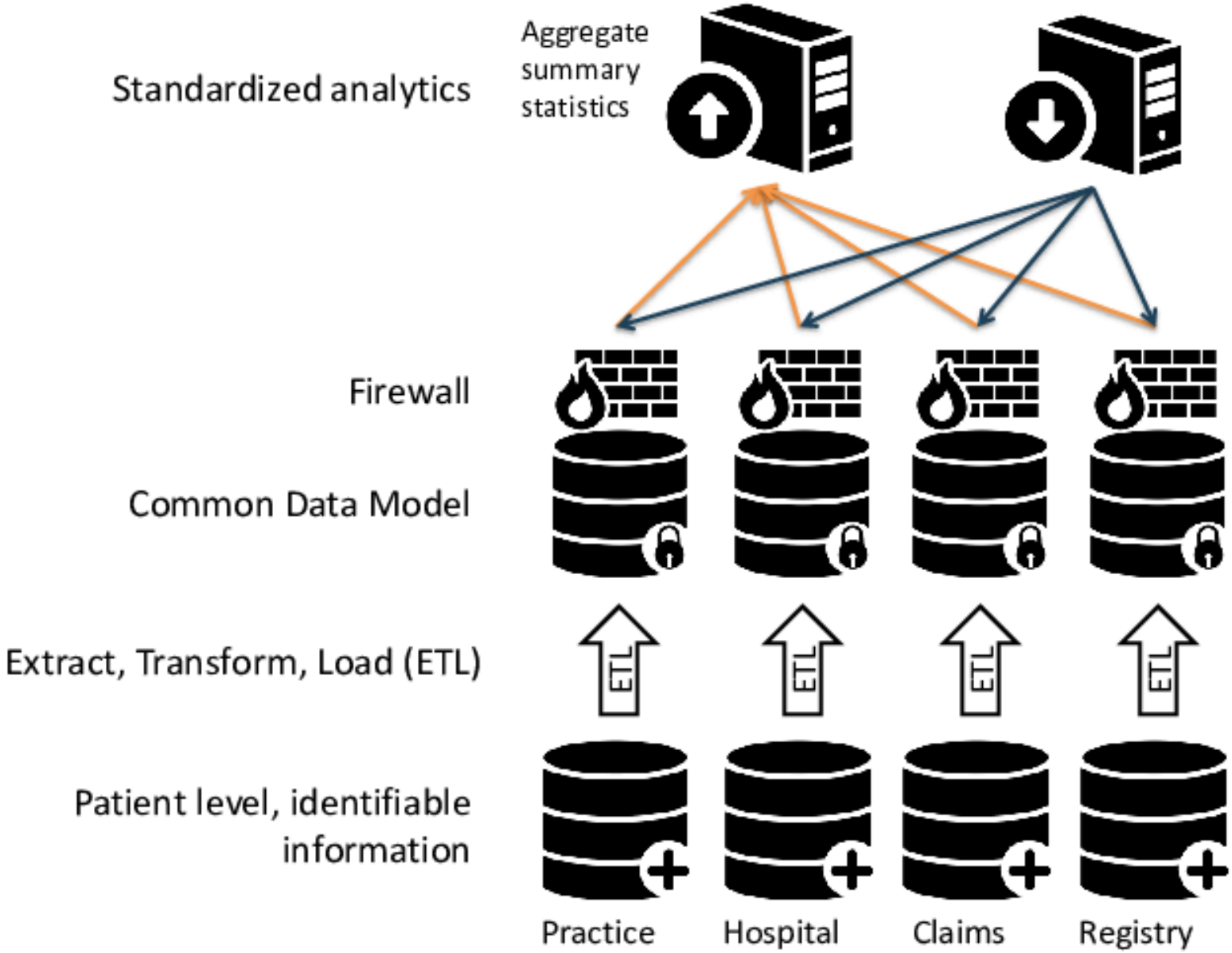
Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

Research networks

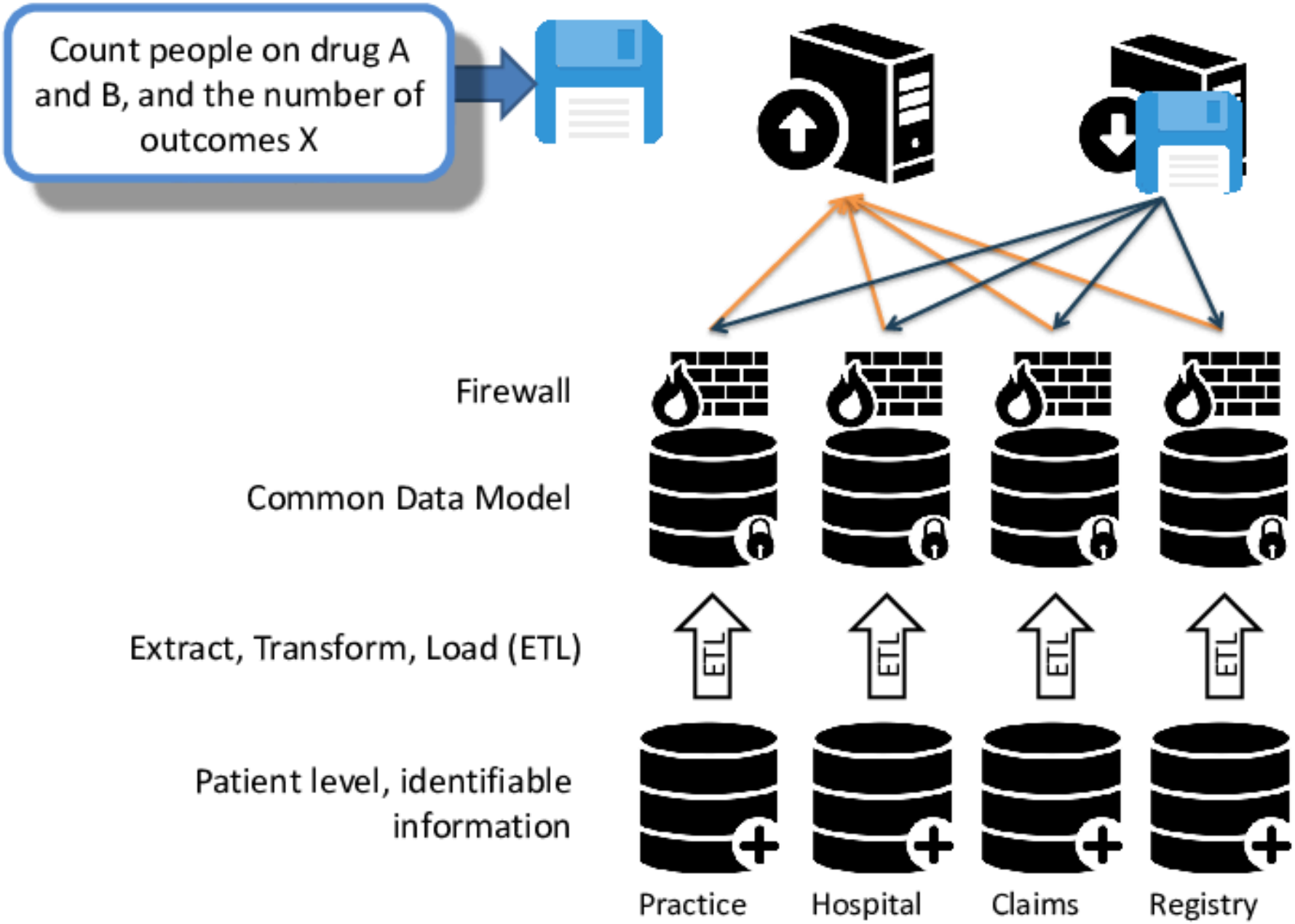
- Data may be at different sites
- Sites often cannot share data at the patient level
- Data can be in very different formats

Patient level, identifiable information

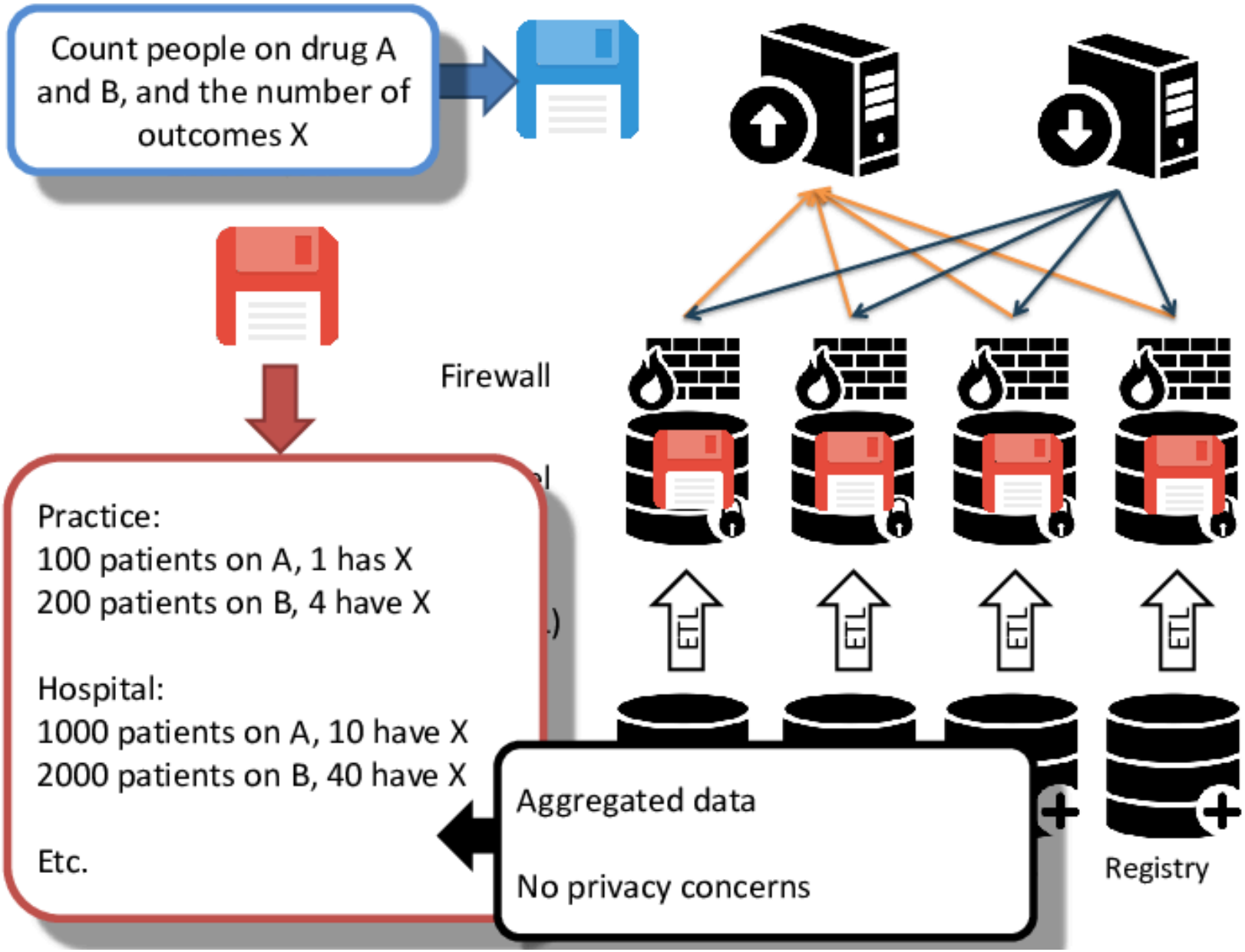
Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.



Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.



Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.



Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

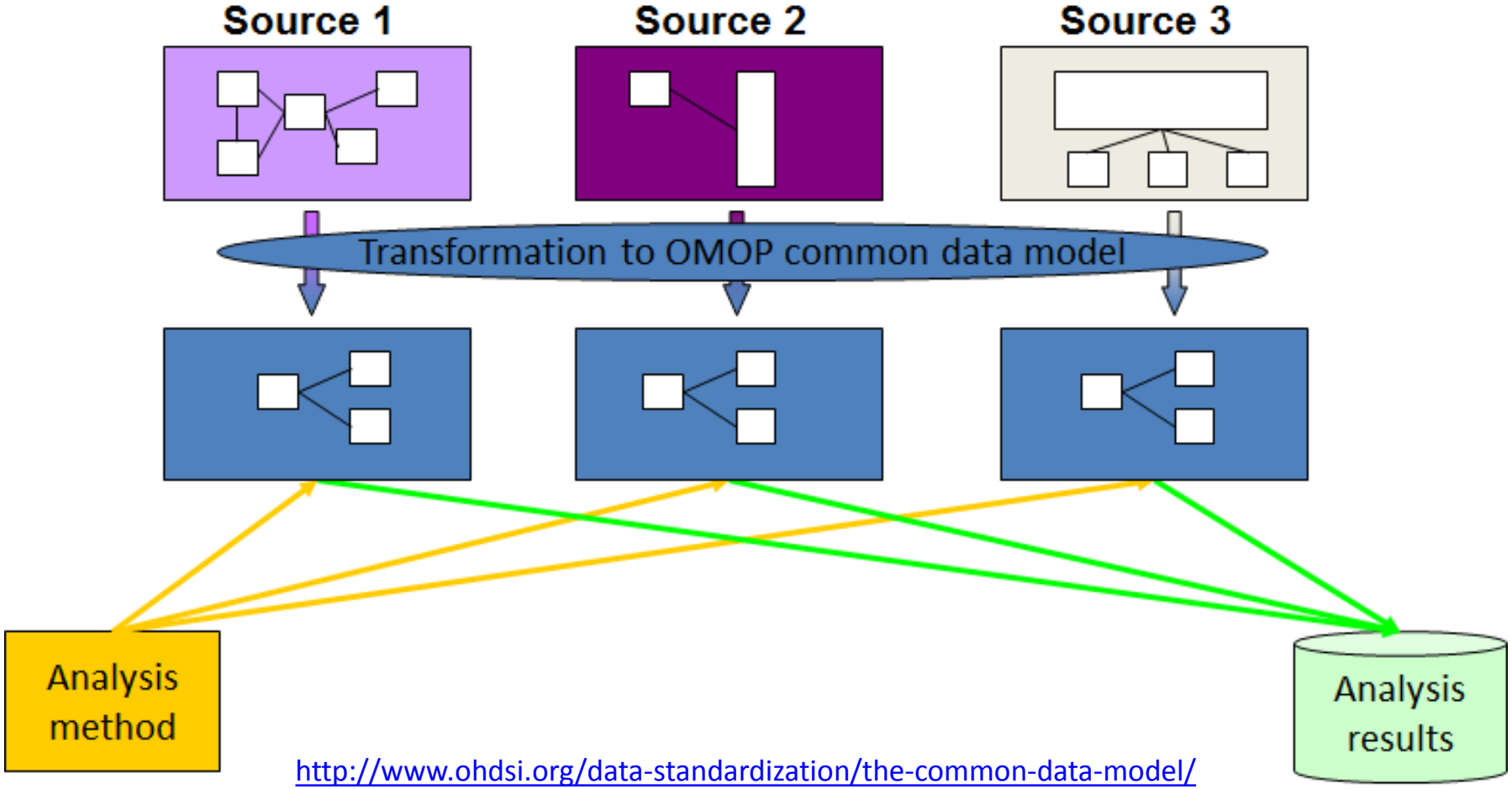
"Common Data Models" and Interoperability

- What is a "Common Data Model" (CDM)

  "A **common data model** allows for the systematic analysis of disparate observational databases.

  The concept behind this approach is to **transform** data contained within disparate databases **into a common format** (data model), and then perform systematic analyses using a library of standard analytic routines that have been written based on the common format." - http://www.ohdsi.org/data-standardization/the-common-data-model/

Going from multiple databases to one



http://www.ohdsi.org/data-standardization/the-common-data-model/

Common Data Model

- A common **structure**

Person
- person_id
- year_of_birth
- month_of_birth
- day_of_birth
- gender_concept_id

- A common **vocabulary**

How do we store gender?
- M, F, U
- 0, 1, 2
- 8507 (male), 8532 (female), 8851 (unknown gender), 8570 (ambiguous gender)

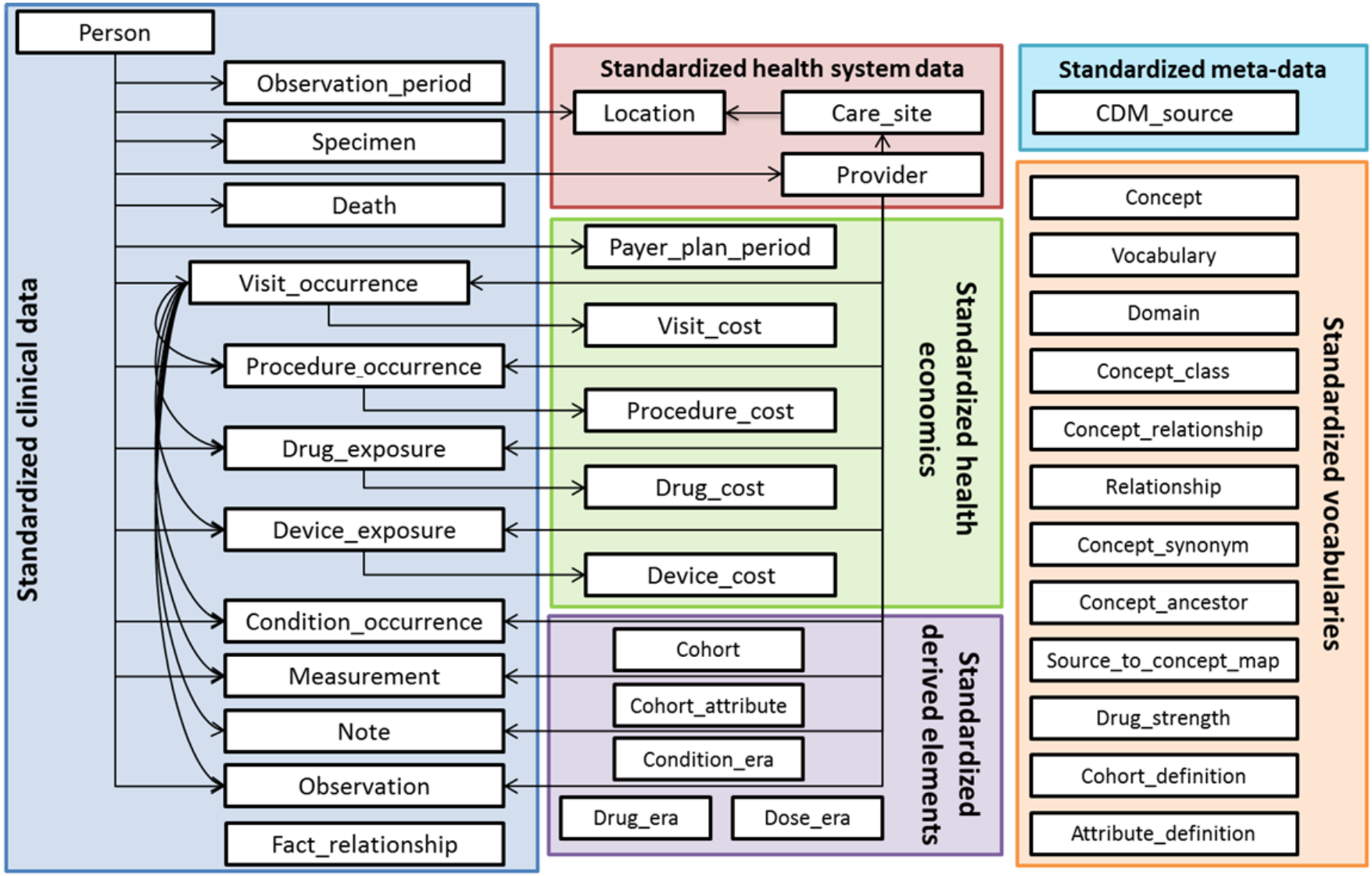
Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

OMOP Common Data Model

- Designed for **various types of data** (EHR, insurance claims) in **various countries**
- Developed by the OMOP / OHDSI community
- Currently on 5th version
- 'Easy to get data in, easy to get data out'

Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

OMOP CDM Version 5



OMOP Collaborative. OMOP Common Data Model Specifications Version 5.0. 10/14/2014. https://goo.gl/aeWsuh . Last accessed 5/5/2016.

CDM principles

- Person centric
- Data is split into domains (e.g. conditions, drugs, procedures)
- Preserve source values and map to standard values
- Store the source of the data
- Separate verbatim data from inferred data

Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

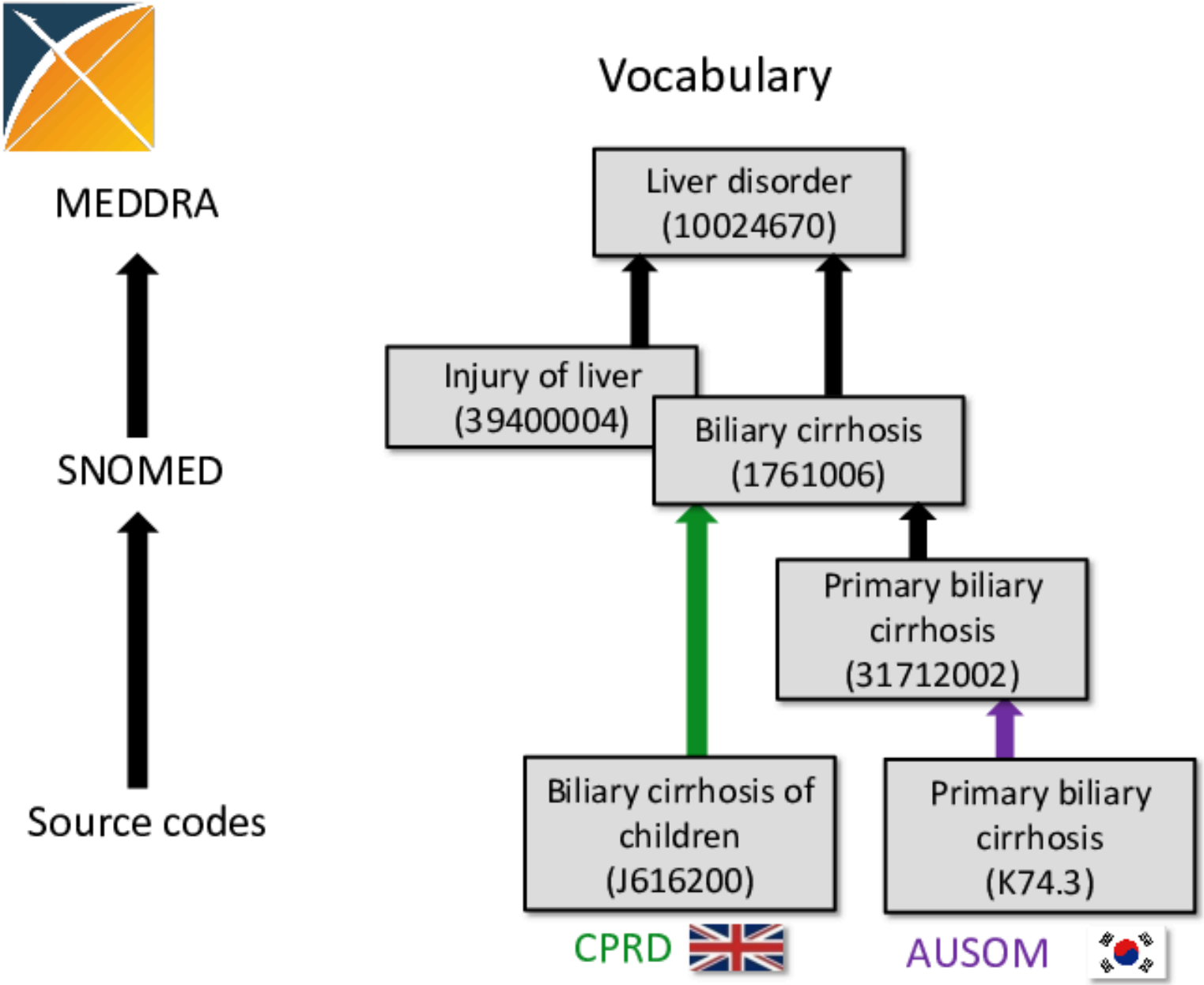
OMOP Vocabulary

All codes are mapped to **standard coding systems**

- Drugs: RxNorm
- Conditions: SNOMED
- Measurements: LOINC
- Procedures: ICD9Proc, HCPCS, CPT-4

Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

Vocabulary



Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

An example with Drug data

Source Data – drug dispensing

`ptId|gpi|ndc|start|end|drug name|form|strength|sig|prn`

Drug Exposure

DRUG_EXPOSURE_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_EXPOSURE_START_DATE	DATE	No
DRUG_EXPOSURE_END_DATE	DATE	Yes
DRUG_TYPE_CONCEPT_ID	NUMBER(38,0)	No
STOP_REASON	VARCHAR2(20 BYTE)	Yes
REFILLS	NUMBER(38,0)	Yes
QUANTITY	FLOAT	Yes
DAYS_SUPPLY	NUMBER(38,0)	Yes
SIG	CLOB	Yes
ROUTE_CONCEPT_ID	NUMBER(38,0)	Yes
EFFECTIVE_DRUG_DOSE	FLOAT	Yes
DOSE_UNIT_CONCEPT_ID	NUMBER(38,0)	Yes
LOT_NUMBER	VARCHAR2(50 BYTE)	Yes
PROVIDER_ID	NUMBER(38,0)	Yes
VISIT_OCCURRENCE_ID	NUMBER(38,0)	Yes
DRUG_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DRUG_SOURCE_CONCEPT_ID	NUMBER(38,0)	Yes
ROUTE_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DOSE_UNIT_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes

Drug Era

COLUMN_NAME	DATA_TYPE	NULLABLE
DRUG_ERA_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_ERA_START_DATE	DATE	No
DRUG_ERA_END_DATE	DATE	No
DRUG_EXPOSURE_COUNT	NUMBER(38,0)	Yes
GAP_DAYS	NUMBER(38,0)	Yes

An example with Drug data cont…

ptId …

Drug Exposure

DRUG_EXPOSURE_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_EXPOSURE_START_DATE	DATE	No
DRUG_EXPOSURE_END_DATE	DATE	Yes
DRUG_TYPE_CONCEPT_ID	NUMBER(38,0)	No
STOP_REASON	VARCHAR2(20 BYTE)	Yes
REFILLS	NUMBER(38,0)	Yes
QUANTITY	FLOAT	Yes
DAYS_SUPPLY	NUMBER(38,0)	Yes
SIG	CLOB	Yes
ROUTE_CONCEPT_ID	NUMBER(38,0)	Yes
EFFECTIVE_DRUG_DOSE	FLOAT	Yes
DOSE_UNIT_CONCEPT_ID	NUMBER(38,0)	Yes
LOT_NUMBER	VARCHAR2(50 BYTE)	Yes
PROVIDER_ID	NUMBER(38,0)	Yes
VISIT_OCCURRENCE_ID	NUMBER(38,0)	Yes
DRUG_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DRUG_SOURCE_CONCEPT_ID	NUMBER(38,0)	Yes
ROUTE_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DOSE_UNIT_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes

Drug Era

COLUMN_NAME	DATA_TYPE	NULLABLE
DRUG_ERA_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_ERA_START_DATE	DATE	No
DRUG_ERA_END_DATE	DATE	No
DRUG_EXPOSURE_COUNT	NUMBER(38,0)	Yes
GAP_DAYS	NUMBER(38,0)	Yes

An example with Drug data cont…

… gpi|ndc …

RxNorm clinical drug

RxNorm ingredient

Drug Exposure

DRUG_EXPOSURE_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_EXPOSURE_START_DATE	DATE	No
DRUG_EXPOSURE_END_DATE	DATE	Yes
DRUG_TYPE_CONCEPT_ID	NUMBER(38,0)	No
STOP_REASON	VARCHAR2(20 BYTE)	Yes
REFILLS	NUMBER(38,0)	Yes
QUANTITY	FLOAT	Yes
DAYS_SUPPLY	NUMBER(38,0)	Yes
SIG	CLOB	Yes
ROUTE_CONCEPT_ID	NUMBER(38,0)	Yes
EFFECTIVE_DRUG_DOSE	FLOAT	Yes
DOSE_UNIT_CONCEPT_ID	NUMBER(38,0)	Yes
LOT_NUMBER	VARCHAR2(50 BYTE)	Yes
PROVIDER_ID	NUMBER(38,0)	Yes
VISIT_OCCURRENCE_ID	NUMBER(38,0)	Yes
DRUG_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DRUG_SOURCE_CONCEPT_ID	NUMBER(38,0)	Yes
ROUTE_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DOSE_UNIT_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes

Drug Era

COLUMN_NAME	DATA_TYPE	NULLABLE
DRUG_ERA_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_ERA_START_DATE	DATE	No
DRUG_ERA_END_DATE	DATE	No
DRUG_EXPOSURE_COUNT	NUMBER(38,0)	Yes
GAP_DAYS	NUMBER(38,0)	Yes

An example with Drug data cont…

How to create code mappings:

1. Prepare the data
   - codes
   - descriptions
   - Frequencies
2. Use existing mappings if available
   - Free mappings
     - RxNorm maps to NDC, NDF-RT, SPL, ATC
     - Others exist in the UMLS (https://www.nlm.nih.gov/research/umls/)
   - Proprietary mappings (e.g., RxNorm to Medi-Span ©)
3. If necessary, create a new mapping
   - Run USAGI https://github.com/OHDSI/Usagi

Usagi in action

File Edit View Help

Status	Source code	Source term	Frequency	Match score	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code
Unchecked	180677	ABACAVIR 600 ...	0	0.97	45775747	abacavir 600 ...	Drug	Clinical Drug	RxNorm	1546888
Unchecked	14369	ABCIXIMAB 10 ...	0	0.85	43536091	abciximab 10...	Drug	Clinical Drug	Multilex	1019961
Unchecked	70305	ABILIFY 2 MG T...	0	0.97	757759	aripiprazole 2 ...	Drug	Branded Drug	RxNorm	615172
Approved	36437	ABILIFY 5 MG T...	0	0.97	19102836	aripiprazole 5 ...	Drug	Branded Drug	RxNorm	404602
Unchecked	8939	ACEBUTOLOL 2...	0	0.93	1320038	Acebutolol 20...	Drug	Clinical Drug C...	RxNorm	328656
Approved	113	ACETAZOLAMID...	0	0.97	19018781	Acetazolamide...	Drug	Clinical Drug	RxNorm	197304
Unchecked	8962	ACETAZOLAMID...	0	0.83	19072239	Acetazolamide...	Drug	Clinical Drug	RxNorm	307701
Unchecked	128434	ACTEMRA IV	0	0.75	40171291	tocilizumab 20 ...	Drug	Branded Drug	RxNorm	895764
Unchecked	171177	ACYCLOVIR 200...	0	0.93	19072293	Acyclovir 40 M...	Drug	Clinical Drug	RxNorm	307730
Unchecked	8969	ACYCLOVIR 200...	0	0.97	19018786	Acyclovir 200 ...	Drug	Clinical Drug	RxNorm	197310
Unchecked	8971	ACYCLOVIR 400...	0	0.97	1703691	Acyclovir 400 ...	Drug	Clinical Drug	RxNorm	197311
Approved	8972	ACYCLOVIR 800...	0	0.97	1703743	Acyclovir 800 ...	Drug	Clinical Drug	RxNorm	197313

Source code

Source code	Source term	Frequency
180677	ABACAVIR 600 MG-DOLUTEGRAVIR 50 MG-LAMIVUDINE 300 MG ...	0

Target concepts

Synonym	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Valid start date	Valid end date	Invalid reason
abacavir 600 MG...	45775747	abacavir 600 MG...	Drug	Clinical Drug	RxNorm	1546888	20141201	20991231	

Remove concept

Search

Query

- (●) Use source term as query
- ( ) Query:

Filters

- [ ] Filter by automatically select concepts
- [ ] Filter by concept class: Admin Concept
- [ ] Filter by domain: Condition
- [ ] Filter invalid concepts
- [ ] Filter by vocabulary: APC

Results

Score	Synonym	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Valid start date	Valid end date	Invalid reason
0.97	abacavir 600 ...	45775747	abacavir 600 ...	Drug	Clinical Drug	RxNorm	1546888	20141201	20991231	
0.85	abacavir / dolu...	45775746	abacavir / dolu...	Drug	Clinical Drug F...	RxNorm	1546886	20141201	20991231	
0.84	abacavir 600 ...	45775751	abacavir 600 ...	Drug	Branded Drug	RxNorm	1546894	20141201	20991231	
0.83	abacavir 600 ...	45775749	abacavir 600 ...	Drug	Branded Drug ...	RxNorm	1546890	20141201	20991231	
0.80	abacavir 600 ...	19122306	abacavir 600 ...	Drug	Clinical Drug	RxNorm	602393	20060226	20991231	
0.70	abacavir / dolu...	45775750	abacavir / dolu...	Drug	Branded Drug ...	RxNorm	1546891	20141201	20991231	

Replace concept | Add concept

Approve

Approved / total: 387 / 1575 ⍰% of total frequency

An example with Drug data

… start|end …

The date of a single exposure

Dates of continuous exposure

Drug Exposure

DRUG_EXPOSURE_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_EXPOSURE_START_DATE	DATE	No
DRUG_EXPOSURE_END_DATE	DATE	Yes
DRUG_TYPE_CONCEPT_ID	NUMBER(38,0)	No
STOP_REASON	VARCHAR2(20 BYTE)	Yes
REFILLS	NUMBER(38,0)	Yes
QUANTITY	FLOAT	Yes
DAYS_SUPPLY	NUMBER(38,0)	Yes
SIG	CLOB	Yes
ROUTE_CONCEPT_ID	NUMBER(38,0)	Yes
EFFECTIVE_DRUG_DOSE	FLOAT	Yes
DOSE_UNIT_CONCEPT_ID	NUMBER(38,0)	Yes
LOT_NUMBER	VARCHAR2(50 BYTE)	Yes
PROVIDER_ID	NUMBER(38,0)	Yes
VISIT_OCCURRENCE_ID	NUMBER(38,0)	Yes
DRUG_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DRUG_SOURCE_CONCEPT_ID	NUMBER(38,0)	Yes
ROUTE_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DOSE_UNIT_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes

Drug Era

COLUMN_NAME	DATA_TYPE	NULLABLE
DRUG_ERA_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_ERA_START_DATE	DATE	No
DRUG_ERA_END_DATE	DATE	No
DRUG_EXPOSURE_COUNT	NUMBER(38,0)	Yes
GAP_DAYS	NUMBER(38,0)	Yes

An example with Drug data

… drug name|form|strength|sig|prn …

- Drug name, form, strength handled by the vocabulary
- SIG entered directly
- ‘PRN’ – by concatenating with the SIG string

Drug Exposure

DRUG_EXPOSURE_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_EXPOSURE_START_DATE	DATE	No
DRUG_EXPOSURE_END_DATE	DATE	Yes
DRUG_TYPE_CONCEPT_ID	NUMBER(38,0)	No
STOP_REASON	VARCHAR2(20 BYTE)	Yes
REFILLS	NUMBER(38,0)	Yes
QUANTITY	FLOAT	Yes
DAYS_SUPPLY	NUMBER(38,0)	Yes
SIG	CLOB	Yes
ROUTE_CONCEPT_ID	NUMBER(38,0)	Yes
EFFECTIVE_DRUG_DOSE	FLOAT	Yes
DOSE_UNIT_CONCEPT_ID	NUMBER(38,0)	Yes
LOT_NUMBER	VARCHAR2(50 BYTE)	Yes
PROVIDER_ID	NUMBER(38,0)	Yes
VISIT_OCCURRENCE_ID	NUMBER(38,0)	Yes
DRUG_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DRUG_SOURCE_CONCEPT_ID	NUMBER(38,0)	Yes
ROUTE_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes
DOSE_UNIT_SOURCE_VALUE	VARCHAR2(50 BYTE)	Yes

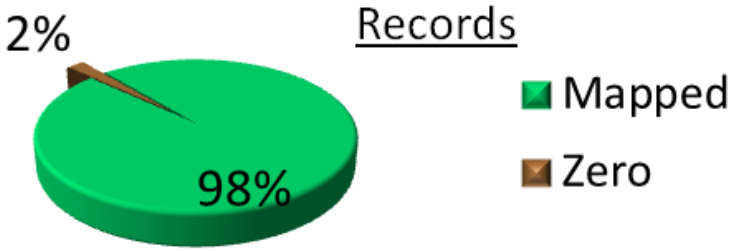
Drug Era

COLUMN_NAME	DATA_TYPE	NULLABLE
DRUG_ERA_ID	NUMBER(38,0)	No
PERSON_ID	NUMBER(38,0)	No
DRUG_CONCEPT_ID	NUMBER(38,0)	No
DRUG_ERA_START_DATE	DATE	No
DRUG_ERA_END_DATE	DATE	No
DRUG_EXPOSURE_COUNT	NUMBER(38,0)	Yes
GAP_DAYS	NUMBER(38,0)	Yes

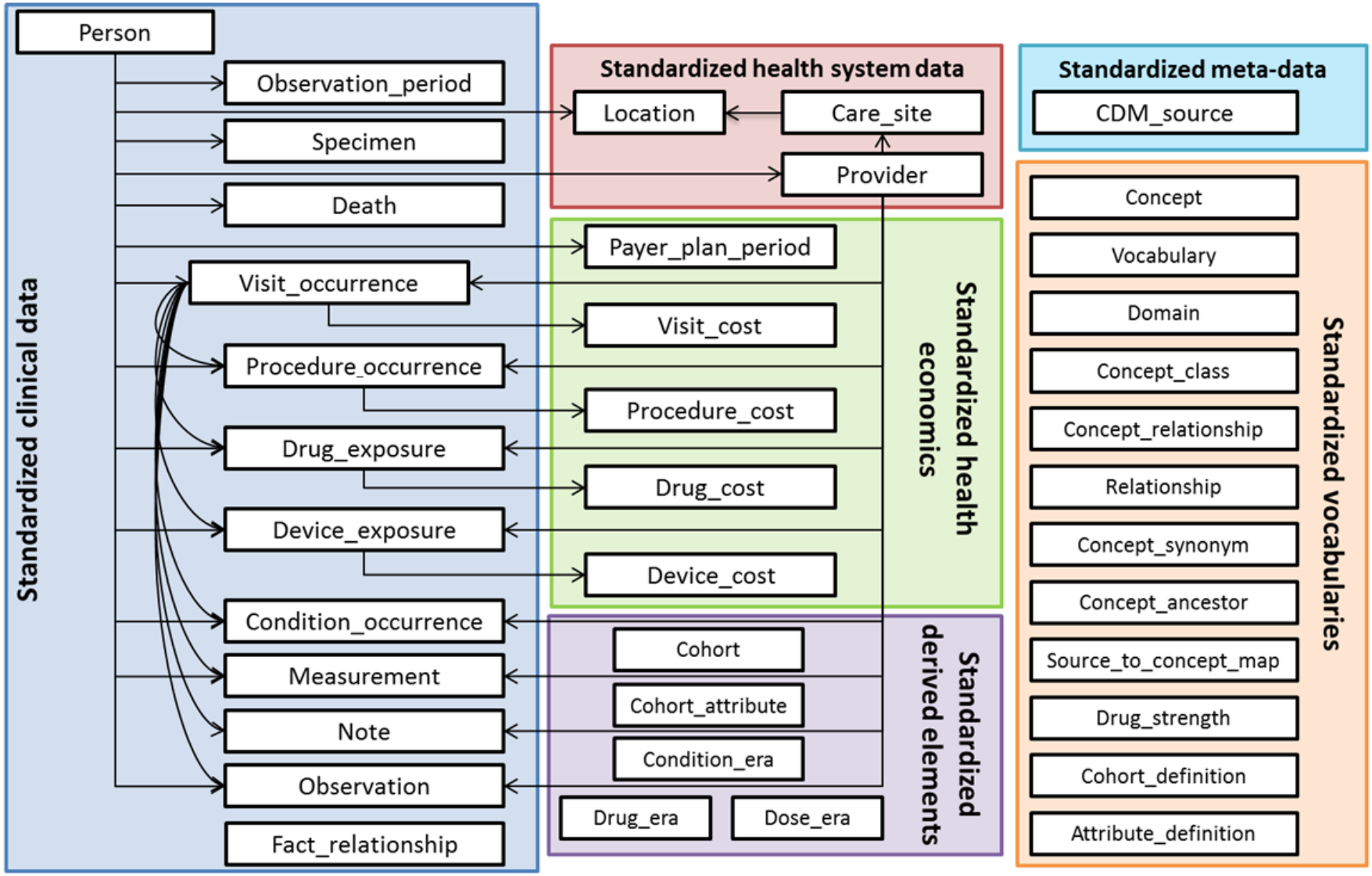
An example with drug data cont…

- Real world cases:
  - Nursing home dataset (5K patients)
    - 430,910 dispensing records
      - 40,950 (9.5%) records dropped for not providing a GPI that could be mapped to RxNorm
        - » the great majority were for gastrointestinal aids
  - Truven CCAE (2003 - 2011, https://goo.gl/pR7IUl)

Unique concepts	112,164
Mapped	76,633
Not Mapped	35,531



A similar process for other domains

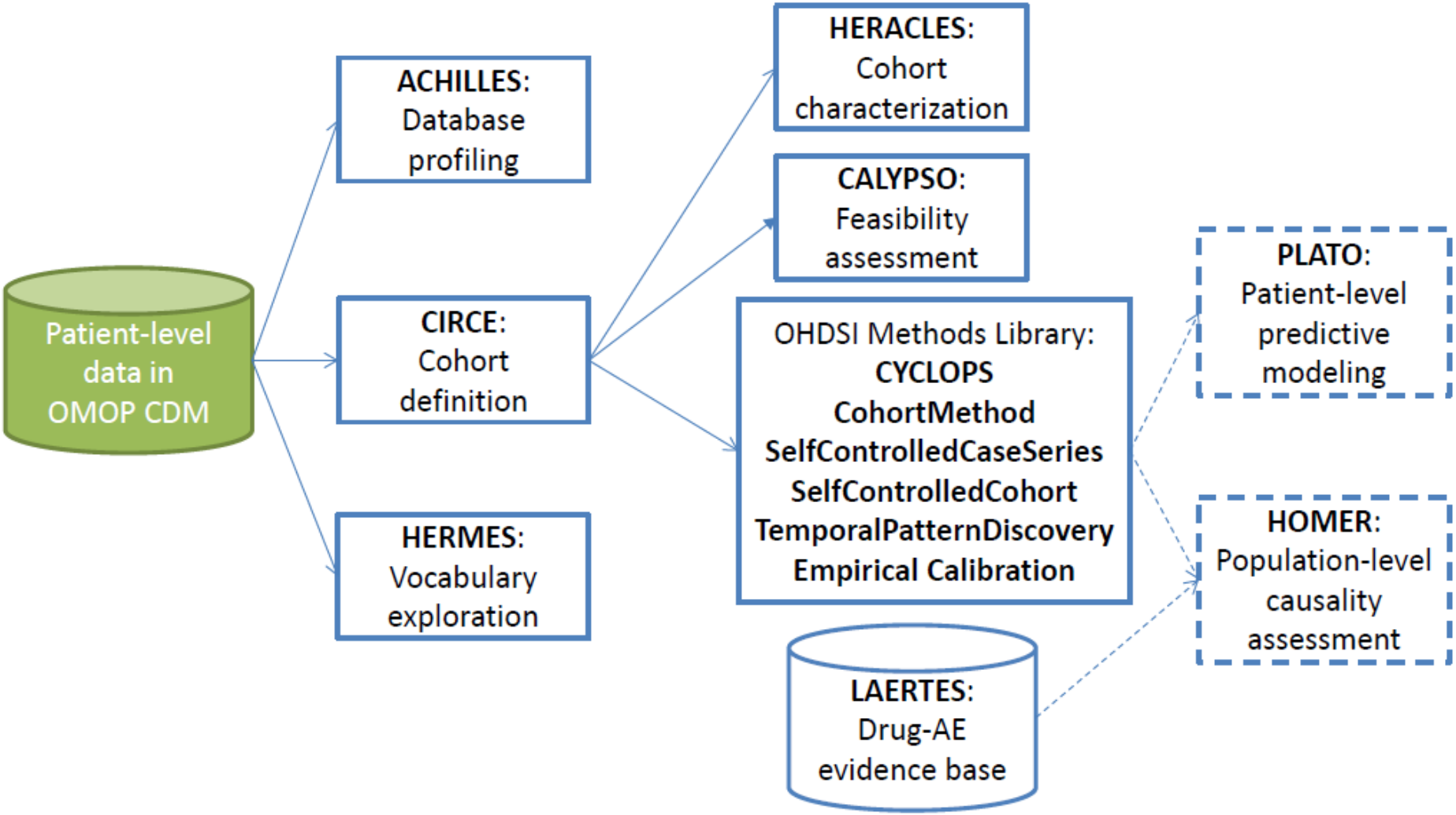


OMOP Collaborative. OMOP Common Data Model Specifications Version 5.0. 10/14/2014. https://goo.gl/aeWsuh . Last accessed 5/5/2016.

Pros and Cons of extraction, translation, and load (ETL) into the OHDSI CDM

- Pros
  - Our project now becomes a part of the OHDSI research network
    - International consortium of very skilled researchers to provide input
  - Sets us up nicely to do any number of other observational studies
    - Speeds up the generation of new evidence
- Cons
  - More work initially
    - potentially much less in the long run
  - Requires additional validation
    - E.g., records might be dropped because of no mapping to the CDM or vocabulary

OHDSI Tools Potentially Enabled by using the CDM a for a clinical dataset



Ryan, P., Hripsak, G., Duke, J. Open-Source Big Data Analytics in Healthcare, Presentation at Medinfo 2015. São Paulo, Brazil. August 2015. http://www.ohdsi.org/wp-content/uploads/2014/07/OHDSI-Tutorial-PreFinal-mod.pdf

Let's look at two example OHDSI tools

- Achilles
  - Automated Characterization of Health Information at Large-scale Longitudinal Evidence Systems
  - ETL quality assurance
- Atlas
  - a unified web interface that attempts to integrate features from various OHDSI applications into a single cohesive experience.

ACHILLES

Automated Characterization of Health Information at Large-scale Longitudinal Evidence Systems

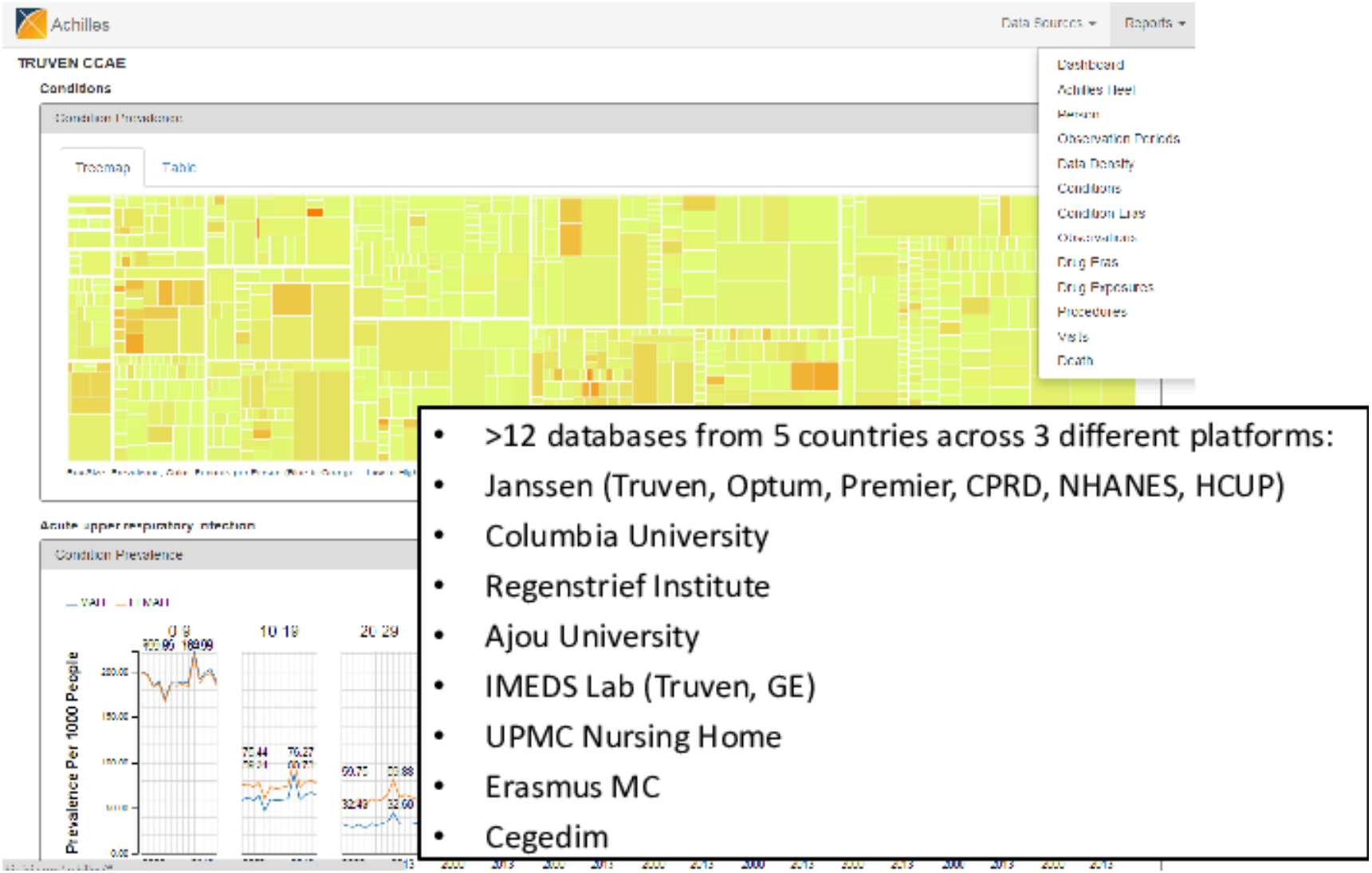
Data in CDM

Aggregate statistics:
- Persons
- Conditions
- Drugs
- Lab results
- ...

Explore statistics in a web browser

Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

ACHILLES



- >12 databases from 5 countries across 3 different platforms:
- Janssen (Truven, Optum, Premier, CPRD, NHANES, HCUP)
- Columbia University
- Regenstrief Institute
- Ajou University
- IMEDS Lab (Truven, GE)
- UPMC Nursing Home
- Erasmus MC
- Cegedim

Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

Achilles Heel

Achilles

Data Sources

Reports

HCUP

Achilles Heel Report

Data Quality Messages

Search:

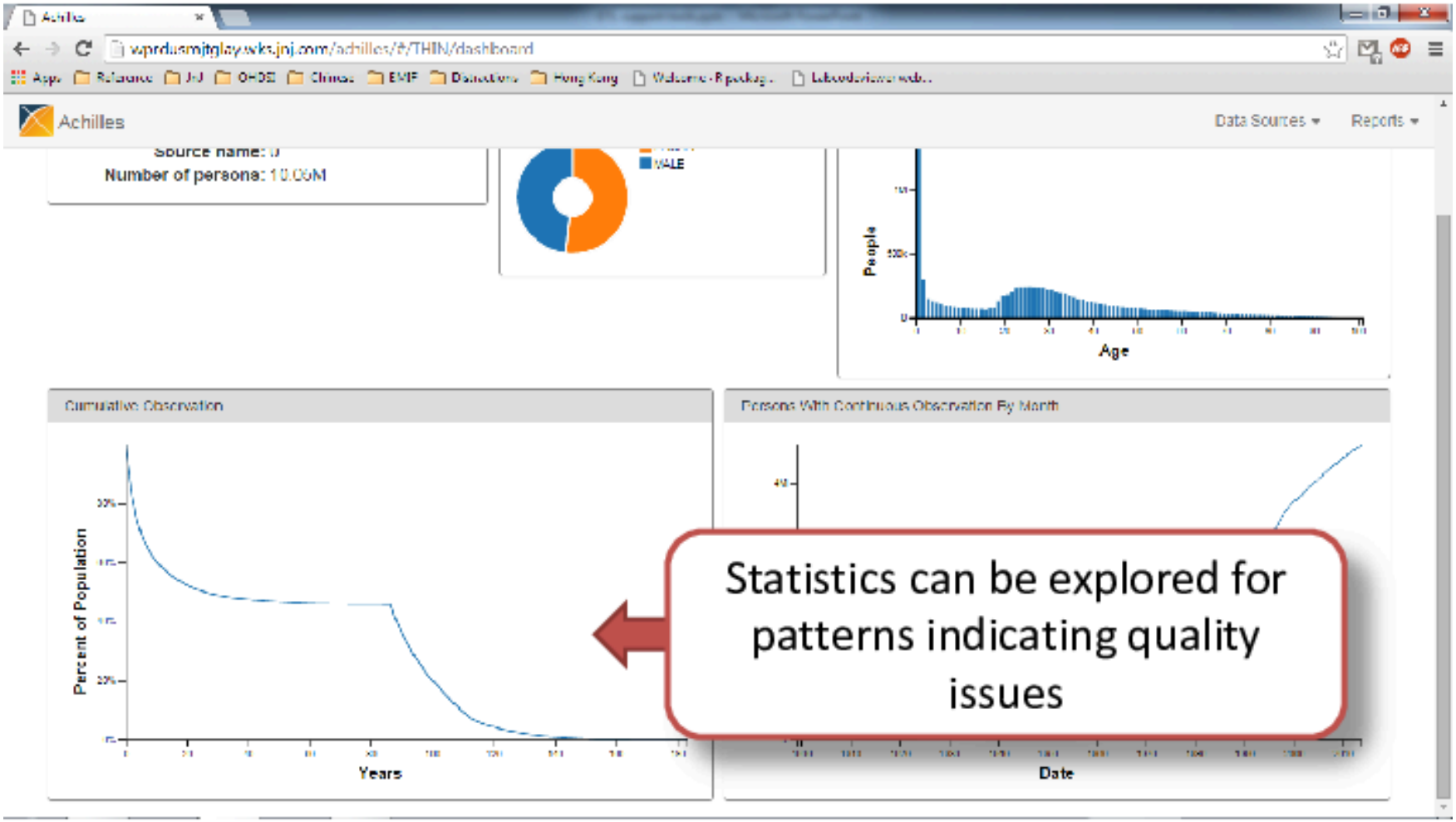
Message Type	Message
ERROR	3-Number of persons by year of birth; should not have year of birth < 1900 (n=684)
ERROR	101-Number of persons by age, with age at first observation period; should not have age > 100, (n=38866)
ERROR	400-Number of persons with at least one condition occurrence, by condition_concept_id; 121 concepts in data are not in correct vocabulary (SNOMED)
ERROR	1000-Number of persons with at least one condition era, by condition_concept_id; 121 concepts in data are not in correct vocabulary (SNOMED)
WA[illegible]S	5-Number of persons by ethnicity; data with unmapped concepts
WA[illegible]S	400-Number of persons with at least one condition occurrence, by condition_concept_id; data with unmapped concepts
	[illegible]t month, by condition_concept_id; 1,766 concepts have a 100% change in monthly count of events
	[illegible]ccurrence, by procedure_concept_id; data with unmapped concepts
	[illegible]t month, by procedure_concept_id; 777 concepts have a 100% change in monthly count of events
	[illegible], by condition_concept_id; 1023 concepts have a 100% change in monthly count of events
	[illegible]d, data with unmapped concepts

Previous 1 Next

Lists data quality issues

Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

Achilles



Schuemie , M., How to extract transform and load observational data? Janssen Research & Development. Beijing, China. May 2015. http://goo.gl/ArRX5u. Last accessed 5/5/2016.

- Demonstration of Achilles on a nursing home dataset

- Try it yourself on a simulated population!
  - http://www.ohdsi.org/web/achilles/

Atlas brings together several OHDSI tools

ATLAS

localhost/Atlas/#/home

ATLAS

- Home
- Data Sources
- Vocabulary
- Concept Sets
- Cohorts
- Jobs
- Configuration
- Feedback

Home

Welcome to ATLAS

ATLAS is an open source application developed as a part of OHDSI intended to provide a unified interface to patient level data and analytics.

The ATLAS user guide can be found here.

Getting Started

Define a New Cohort — Begin performing research by defining the group of people you intend to study

Search the Vocabulary — Search the different ontologies used to describe patient level data around the world

Release Notes

Version 1.0 Current Release Notes

- Integrated Data Source Characterization (project achilles)
- Integrated Cohort Definitions (project circe)
- Integrated Cohort Reporting (project heracles)
- Bug Fixes listed in Additional Details below

Additional Details:

This latest release contains **30** feature enhancements and issue resolutions.

- Error when opening cohort definition that hasn't been generated
- Concept Set JSON import problems
- Cohort report drilldown errors
- Atlas WebAPI Configuration
- Hide Patient Profile
- Concept set expression: include 'source codes' if non-standard concepts selected in conceptset expression
- Concept Set - Enforce Unique Names
- Cohort export to JSON
- Concept Set Saving Issues
- Remove Cohort Reporting left hand menu item
- Importing source codes, they do not get added to cart

Atlas- vocabulary exploration

ATLAS

localhost/Atlas/#/search/simvastatin

ATLAS

- Home
- Data Sources
- Vocabulary
- Concept Sets
- Cohorts
- Jobs
- Configuration
- Feedback

Vocabulary

Search | Advanced | Results | Import

Search Results for simvastatin

Vocabulary
NDC (2086)
SNOMED (320)
SPL (224)
RxNorm (113)
Multilex (86)
VA Product (73)
NDFRT (50)
Gemscript (43)

Class
11-digit NDC (1635)
9-digit NDC (451)
Pharma/Biol Product (314)
Prescription Drug (123)
Clinical Drug (111)
SPL (101)
VA Product (73)
Drug Interaction (45)

Domain
Drug (3057)
Observation (7)
Measurement (6)
Condition (5)

Standard Concept
Non-Standard (2841)
Classification (125)
Standard (109)

Invalid Reason
Valid (2941)
Invalid (134)

Has Records
false (3075)

Has Descendant Records
false (3075)

Show 15 entries

Filter:

Showing 1 to 15 of 3,075 entries

Previous 1 2 3 4 5 ... 205 Next

Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
45408439	96785839662900	*Simvastatin (Bulk) Powder***	GPI	timeout	timeout	Drug	GPI
19129329	761907	24 HR Niacin 1000 MG / Simvastatin 20 MG Extended Release Oral Tablet	Quant Clinical Drug	timeout	timeout	Drug	RxNorm
1518133	763228	24 HR Niacin 1000 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	Quant Branded Drug	timeout	timeout	Drug	RxNorm
45190421	54868590401	24 HR Niacin 1000 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
45071318	54868590400	24 HR Niacin 1000 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
44838579	00074331690	24 HR Niacin 1000 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
45281999	00074345590	24 HR Niacin 1000 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
40224128	999935	24 HR Niacin 1000 MG / Simvastatin 40 MG Extended Release Oral Tablet	Quant Clinical Drug	timeout	timeout	Drug	RxNorm
40224129	999939	24 HR Niacin 1000 MG / Simvastatin 40 MG Extended Release Oral Tablet [Simcor]	Quant Branded Drug	timeout	timeout	Drug	RxNorm
45128053	00074345790	24 HR Niacin 1000 MG / Simvastatin 40 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
45156278	54868616900	24 HR Niacin 1000 MG / Simvastatin 40 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
19129330	761909	24 HR Niacin 500 MG / Simvastatin 20 MG Extended Release Oral Tablet	Quant Clinical Drug	timeout	timeout	Drug	RxNorm
45162170	00074331207	24 HR Niacin 500 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
44952017	54868588600	24 HR Niacin 500 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC
44866903	54868588601	24 HR Niacin 500 MG / Simvastatin 20 MG Extended Release Oral Tablet [Simcor]	11-digit NDC	timeout	timeout	Drug	NDC

Atlas – Concept set generation

ATLAS

localhost/Atlas/#/conceptsets

ATLAS

- Beta-Blockers 37
- Home
- Data Sources
- Vocabulary
- Concept Sets
- Cohorts
- Jobs
- Configuration
- Feedback

Concept Sets

Concept Set Repository

Local

Add Concept Set

Show 10 entries

Filter Repository Concept Sets:

Id	Title
2523	Aldosterone Antagonists
1433	Amiodarones
240	Anticoagulants
2546	Aspirins
1572	Azoles
2562	Azoles w/o Fluconazoles
1894	Beta-Blockers
1861	Calciums
2587	Calciums (IV)
1129	CCA-Cognitive-Scale

Showing 1 to 10 of 63 entries

Previous 1 2 3 4 5 6 7 Next

ATLAS

localhost/Atlas/#/conceptset/2523/details

ATLAS

- Aldosterone Antagonists 7
- Home
- Data Sources
- Vocabulary
- Concept Sets
- Cohorts
- Jobs
- Configuration
- Feedback

Concept Set

Aldosterone Antagonists

Close Save

Concept Set Expression | Included Concepts 7 | Included Source Codes | Export

Show 25 entries

Search:

Showing 1 to 7 of 7 entries

Previous 1 Next

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
970283	198223	Spironolactone 50 MG Oral Tablet	Drug	Standard			
19079658	313096	Spironolactone 25 MG Oral Tablet	Drug	Standard			
970282	198222	Spironolactone 100 MG Oral Tablet	Drug	Standard			
974226	198224	Hydrochlorothiazide 25 MG / Spironolactone 25 MG Oral Tablet	Drug	Standard			
1309801	351257	eplerenone 50 MG Oral Tablet	Drug	Standard			
1309800	351256	eplerenone 25 MG Oral Tablet	Drug	Standard			
1309832	351258	eplerenone 100 MG Oral Tablet	Drug	Standard			

Classification Non-Standard Standard

Atlas – Study cohort generation

ATLAS

localhost/Atlas/#/cohortdefinitions

ATLAS

- Beta-Blockers 37
- Home
- Data Sources
- Vocabulary
- Concept Sets
- Cohorts
- Jobs
- Configuration
- Feedback

Cohorts

New Cohort

Cohort Definition Repository

Local

Last Modified

2+ Weeks Ago (20)

Author

system (20)

Show 15 entries

Filter:

Showing 1 to 15 of 20 entries

Previous 1 2 Next

Id	Name	Created	Updated	Author
138	Fall Risk	2/2/2016	2/25/2016	system
180	Tramadol - antidepressant	2/17/2016	2/17/2016	system
167	Fall Risk - High 4 levels - obs	2/11/2016	2/11/2016	system
166	Fall Risk - Med 4 levels (Q3)	2/10/2016	2/11/2016	system
165	Fall Risk - Low 4 level	2/9/2016	2/11/2016	system
168	Fall Risk - Med - 4 levels - obs (Q3)	2/11/2016	2/11/2016	system
164	Fall Risk - High 4 levels	2/9/2016	2/9/2016	system
130	metformin test 1	2/1/2016	2/1/2016	system
120	Test antibiotic 1	2/1/2016	2/1/2016	system
60	NH Data Test 7	1/19/2016	1/19/2016	system
1	NH Data test 1	1/9/2016	1/19/2016	system
28	NH data set test 6	1/11/2016	1/11/2016	system
8	NH Data Test 5	1/10/2016	1/10/2016	system
4	NH Data Test 4	1/9/2016	1/9/2016	system
2	NH Data test 2	1/9/2016	1/9/2016	system

Atlas – Study cohort generation cont…

ATLAS

localhost/Atlas/#/cohortdefinition/180

ATLAS

- Tramadol - antidepressant
- Home
- Data Sources
- Vocabulary
- Concept Sets
- Cohorts
- Jobs
- Configuration
- Feedback

Cohort

Tramadol - antidepressant

Save | Show SQL | Close | Copy | Delete Cohort

Definition | Concept Sets | Generation | Reporting | Print Friendly | JSON

Description:

Co-exposure to tramadol and an antidepressant

People having any of the following: Add Primary Event Filters...

a drug exposure of antidepressants

Add Filter... | Delete Filter

with observation at least 0 days prior and 0 days after index

Limit primary events to: all events per person.

For people matching the Primary Events, include:

People having all of the following criteria: Add New Criteria...

with at least 1 using all occurrences of:

Add Filter...

a drug exposure of tramadol

occurring between All days Before and All days After index

Delete Filter

Remove Additional Filters

Limit cohort expression results to: all events per person.

Atlas – Study cohort characterization

ATLAS

localhost/Atlas/#/cohortdefinition/180

ATLAS

Tramadol - antidepressant

antidepressants 1

Home

Data Sources

Vocabulary

Concept Sets

Cohorts

Jobs

Configuration

Feedback

Cohort

Tramadol - antidepressant

Save | Show SQL | Close | Copy | Delete Cohort

Definition | Concept Sets | Generation | Reporting | Print Friendly | JSON

	Source Name	Generation Status	Distinct People	Generated	Generation Duration
▶ Generate	dbmi ohdsi oracle	COMPLETE	253	2/17/2016 3:44:17 PM	3.025s

ATLAS

localhost/Atlas/#/cohortdefinition/180

ATLAS

Tramadol - antidepressant

Home

Data Sources

Vocabulary

Concept Sets

Cohorts

Jobs

Configuration

Feedback

Cohort

Tramadol - antidepressant

Save | Show SQL | Close | Copy | Delete Cohort

Definition | Concept Sets | Generation | Reporting | Print Friendly | JSON

Report	dbmi ohdsi oracle
Cohort Specific	
Condition	view
Condition Eras	view
Conditions By Index	
Death	
Drug Eras	view
Drug Exposure	
Drugs by Index	
Observation Periods	
Person	view
Procedure	
Procedures by Index	

Generate | Export

Condition Report (DBMI_OHDSI)

Condition Prevalence

Treemap | Table

Box Size: Prevalence, Color: Records per Person (Light to Dark = Low to High), Use Ctrl-Click to Zoom, Alt-Click to Reset Zoom

Atlas – Study cohort characterization cont…

ATLAS

localhost/Atlas/#/cohortdefinition/180

ATLAS

Tramadol - antidepressant

antidepressants 1

Home

Data Sources

Vocabulary

Concept Sets

Cohorts

Jobs

Configuration

Feedback

Cohort

Tramadol - antidepressant

Save | Show SQL | Close | Copy | Delete Cohort

Definition | Concept Sets | Generation | Reporting | Print Friendly | JSON

Report	dbmi ohdsi oracle
Cohort Specific	
Condition	view
Condition Eras	view
Conditions By Index	
Death	
Drug Eras	view
Drug Exposure	
Drugs by Index	
Observation Periods	
Person	view
Procedure	
Procedures by Index	

Blood and lymphatic system disorders
Red blood cell disorders
Red blood cell abnormal findings NEC
Red blood cell abnormality

Anemia

Prevalence: 49.41%

Number of People: 125

Records per Person: 5.84

Condition Report (DBMI_

Condition Prevalence

Treemap | Table

Box Size: Prevalence, Color: Records per Person (Light to Dark = Low to High), Use Ctrl-Click to Zoom, Alt-Click to Reset Zoom

- Demonstration of Atlas on a nursing home dataset

- Try it yourself on a simulated population!
  - http://www.ohdsi.org/web/atlas/#/home

Conclusion

- Structural and semantic interoperability are important for clinical research
- Placing your clinical data into a common data model and standard vocabulary are essential first steps
  - Perhaps more work initially
  - However, several benefits
    - Network research
    - Common tools and methods
      - Quality assurance
      - Study design
- The OHDSI CDM and tools can be very helpful

Learn more about interoperability for clinical research

- Observational Health Data Sciences and Informatics (OHDSI) collaborative: http://www.ohdsi.org/

  Voss EA, Makadia R, Matcho A, Ma Q, Knoll C, Schuemie M, DeFalco FJ, Londhe A, Zhu V, Ryan PB. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. J Am Med Inform Assoc. 2015 Feb PubMed PMID: 25670757.

- Informatics for Integrating Biology and the Bedside (i2b2) https://www.i2b2.org/

  Weber GM, Murphy SN, McMurry AJ, Macfadden D, Nigrin DJ, Churchill S, Kohane IS. The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories. J Am Med Inform Assoc. 2009 Sep-Oct;16(5):624-30. PMC2744712.

Acknowledgements

- Reagan-Udall Foundation for the FDA, project RUF-IMEDS-SA_0017
- The US National Institute on Aging (K01AG044433)
- National Library of Medicine (R01LM011838)

Discussion