MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

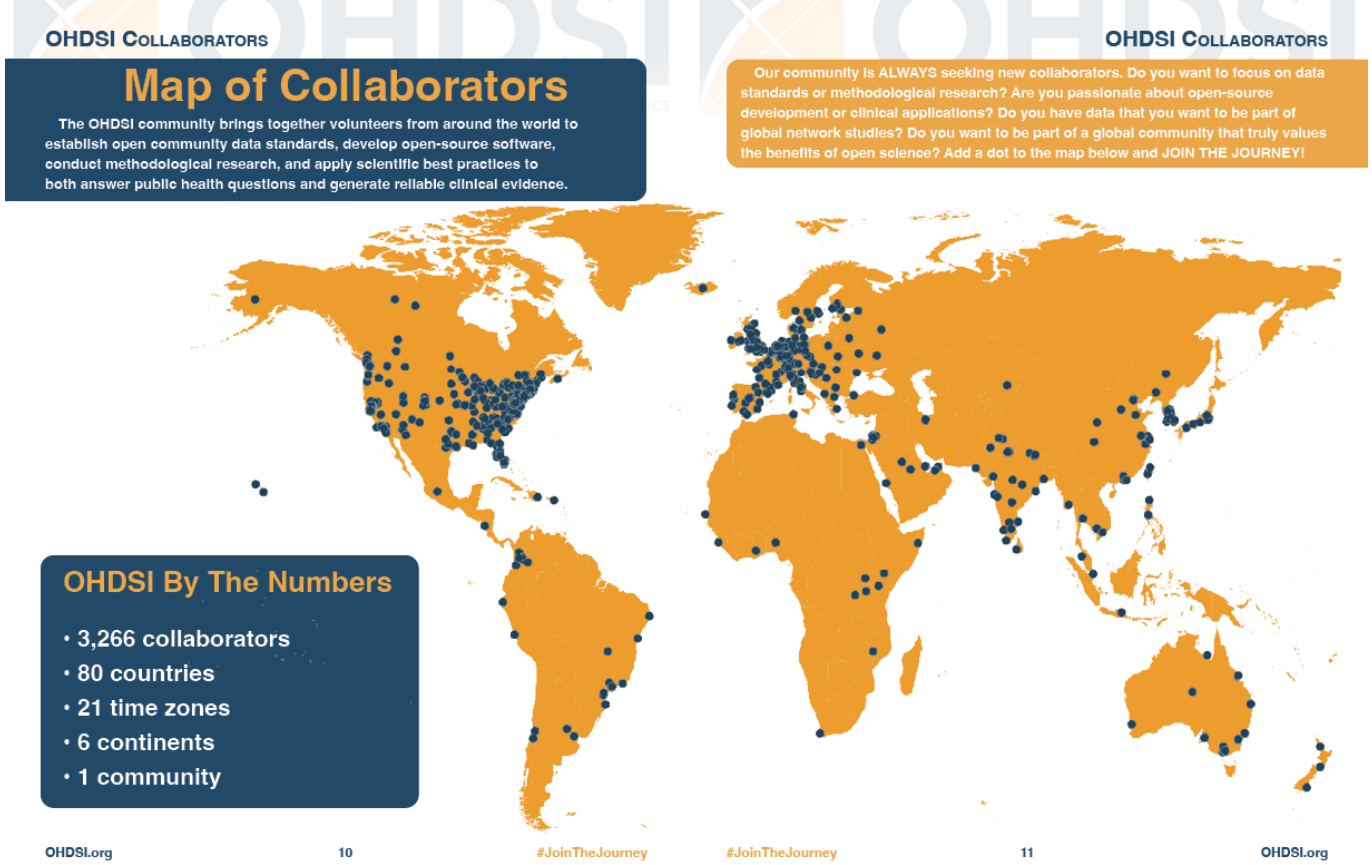
A high-fidelity combined ATC-RxNorm drug hierarchy for large-scale observational research

Anna Ostropolets

Columbia University, Odysseus Data Services

Co-authors: Polina Talapova, Marcel De Wilde, Hamed Abedtash, Peter Rijnbeek, Christian Reich

Background: OHDSI



The Observational Health Data Sciences and Informatics (OHDSI) – multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics.

International federated network of researchers and observational health databases.

Background: drug studies in federated network

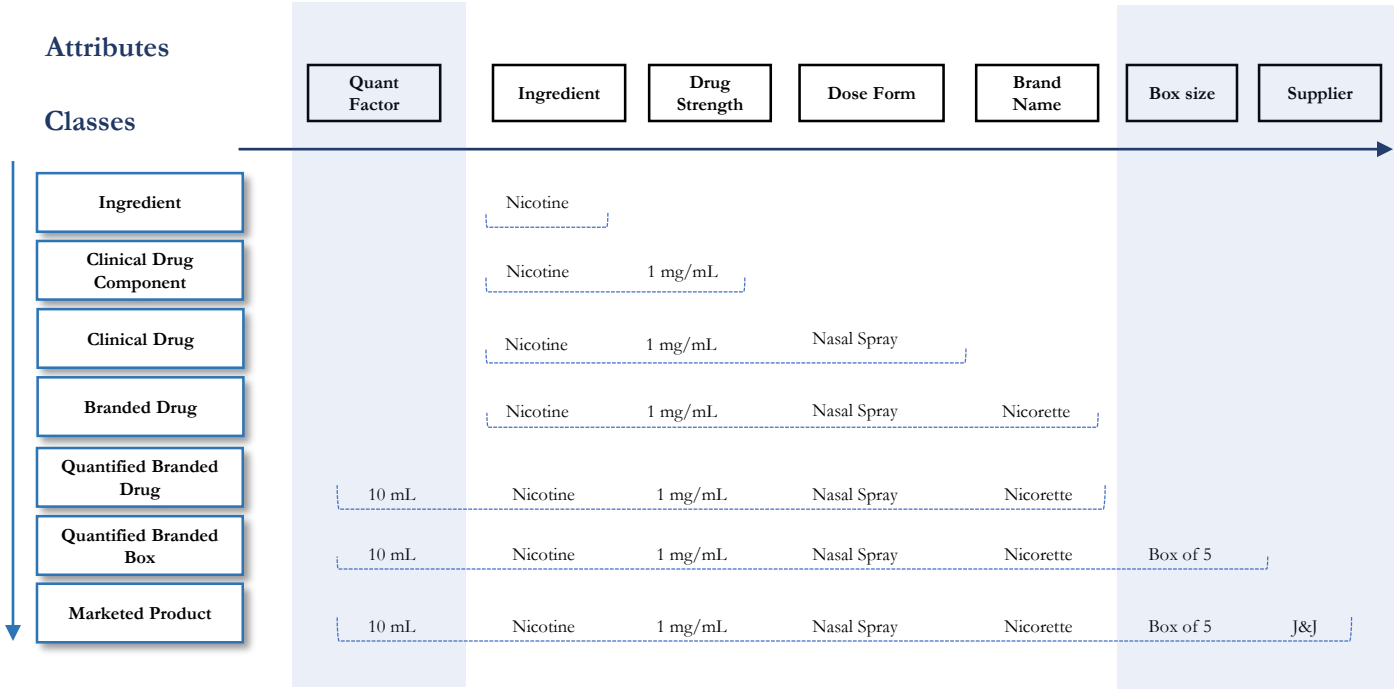
- Comparative effectiveness studies, safety studies, other studies that use drugs as features

- Source drug data: free text, local drug terminologies (e.g., dm+d in UK, NDC in US) mapped to standard OHDSI terminologies (RxNorm and RxNorm Extension)

- A huge need for classificational system to navigate extensive and complex drug space

Background: RxNorm and RxNorm Extension

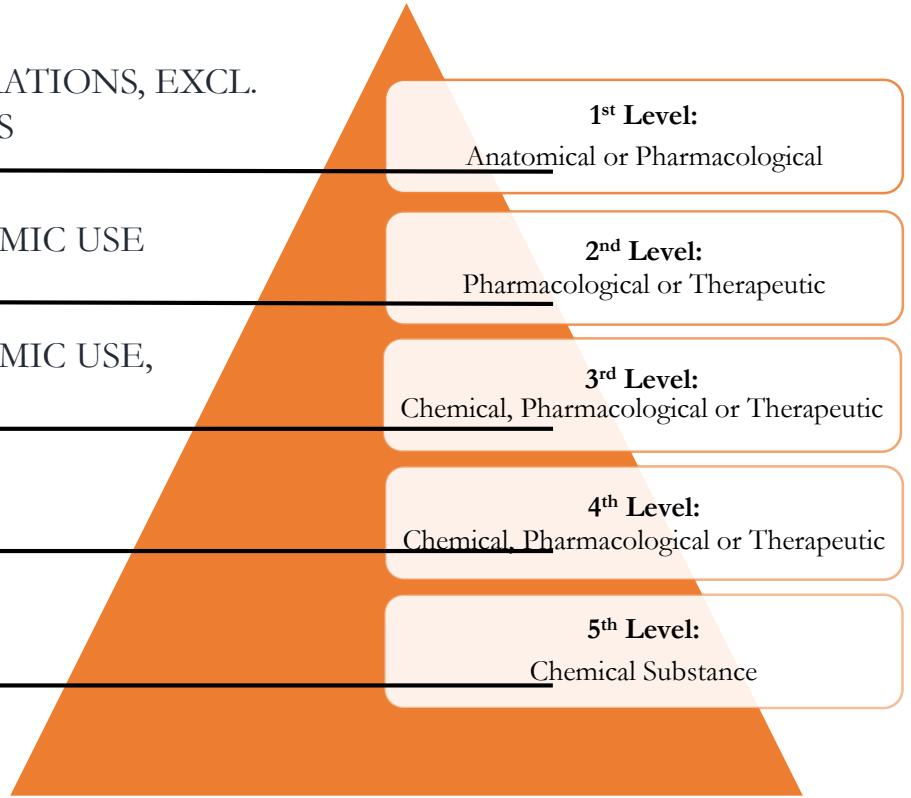
"10 ML Nicotine 1 MG/ML Nasal Spray [Nicorette] Box of 5 by Johnson & Johnson"



All drugs from national markets are mapped into RxNorm and RxNorm Extension in OHDSI

Background: Anatomic Therapeutic Classification (ATC)

H	SYSTEMIC HORMONAL PREPARATIONS, EXCL. SEX HORMONES AND INSULINS
H02	CORTICOSTEROIDS FOR SYSTEMIC USE
H02A	CORTICOSTEROIDS FOR SYSTEMIC USE, PLAIN
H02AB	Glucocorticoids
H02AB06	prednisolone oral/parenteral 10 mg



Task: build a systematic process to classify all drugs under ATC classes

Need to incorporate information about ingredient, form and indication systematically

Prednisolone can belong to 7 different ATC classes based on formulation and indication

Need to explicitly define classes in ATC may not be explicitly defined

- Sometimes Form has to be inferred from ATC 4th-1st
- Four different scenarios with combinations
- Some information is in free-text description

Methods: overview

1. Split ATC 5th into attributes and match them to RxNorm attributes
2. Match ATC 5th concepts to RxNorm drugs based on the combinations of attributes
3. Eliminate pairs if more precise match exists
4. Build the hierarchy

Methods: identify attributes and map them

ATC

S03BA02 prednisolone	A07EA01 prednisolone	H02AB06 prednisolone	C05AA04 prednisolone	D07XA02 prednisolone
Form: ophthalmic Dose: none Combo: no	Form: rectal (inferred) Dose: none Combo: no	Form: systemic Dose: 10 mg Combo: no	Form: topical, rectal (inferred) Dose: no Combo: no	Form: topical Dose: no Combo: yes

RxNorm Ingredient

prednisolone	coal tar	neomycin

RxNorm Dose Form

Ophthalmic Solution Ophthalmic Suspension	Rectal Solution Enema	Oral Tablet/Capsule Injection/Pen Injectable Solution	Topical Solution Topical Cream Rectal Cream	Topical Solution Topical Cream

Methods: eliminate less precise pairs for accurate hierarchy

ATC → Parent of → RxNorm

ATC	RxNorm
C05AA04 prednisolone; topical	Prednisolone mono
D07BA01 prednisolone and antiseptics; topical	Prednisolone + antiseptics but not antibiotics
D07CA03 prednisolone and antibiotics; topical	Prednisolone + antibiotics
D07XA02 prednisolone, topical (+ tar and neomycin)	Prednisolone + neomycin

Where does the rest of combos go?

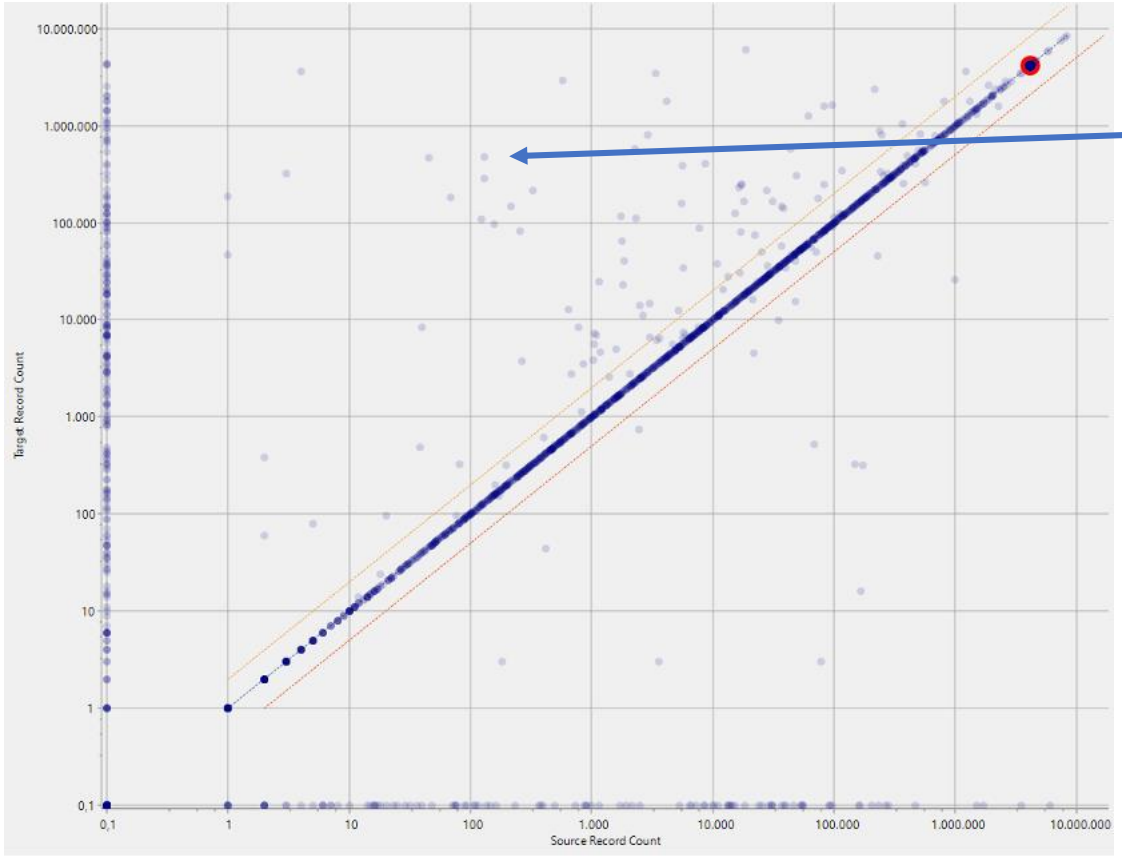
Results

- 89.2% (4,656) ATC 5th-level concepts linked to RxNorm ingredients,
- 79.6% (27,629) RxNorm and 65.3% (50,753) RxNorm Extension Clinical Drugs were linked to ATC 5th-level,
- 97.9% of ingredients in the data and 99.8% of patient records covered through ATC-RxNorm (E) joint hierarchy.

Download: athena.ohdsi.org

Results: comparison of code assignment

s



Source has one ATC per drug, we have multiple

Future work

- Adopt the data-driven approach of attribute selection based on the data sources that have ATC codes (Z index, JMDC, others),

- Identify discrepancies and similarities between code assignment in different data sources to establish more consistent and accurate mappings from ATC to RxNorm (Ext).

Questions?