Semiautomatic mapping of a national drug terminology to standardised OMOP drug concepts using publicly available supplementary information

Research | Open access | Published: 26 September 2025

Volume 25, article number 213, (2025) Cite this article

Florian Katsch, Ágota Mészáros, Tibor Héja, Rada Hussein & Georg Duftschmid


Austria


NO
KANGAROOS
IN AUSTRIA

Why was this work done?

Possibilities of joining OMOP in Austria (with drug data):

map the drugs by hand
(ASP-Liste: 19 273 drug concepts)

accept that there is only an
ATC mapping

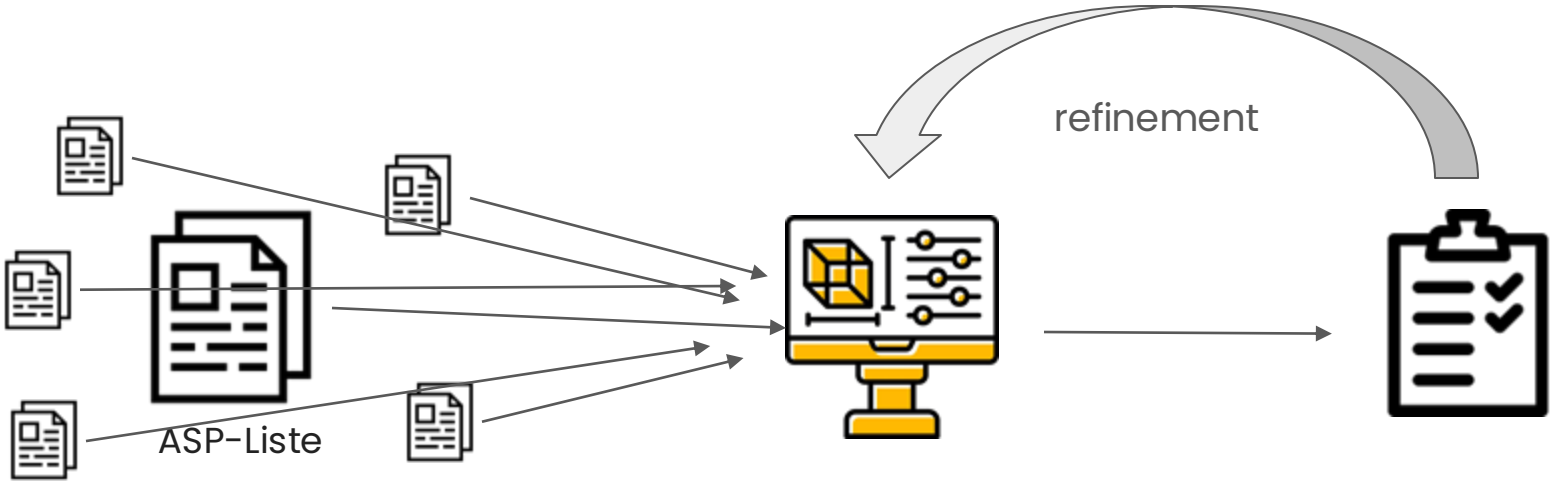
use the DrugMapping Tool by
OHDSI → only Clinical Drug class

Invent a method
and map them
semi-automatically

ATC mapping

Semi-automated mapping

Methods

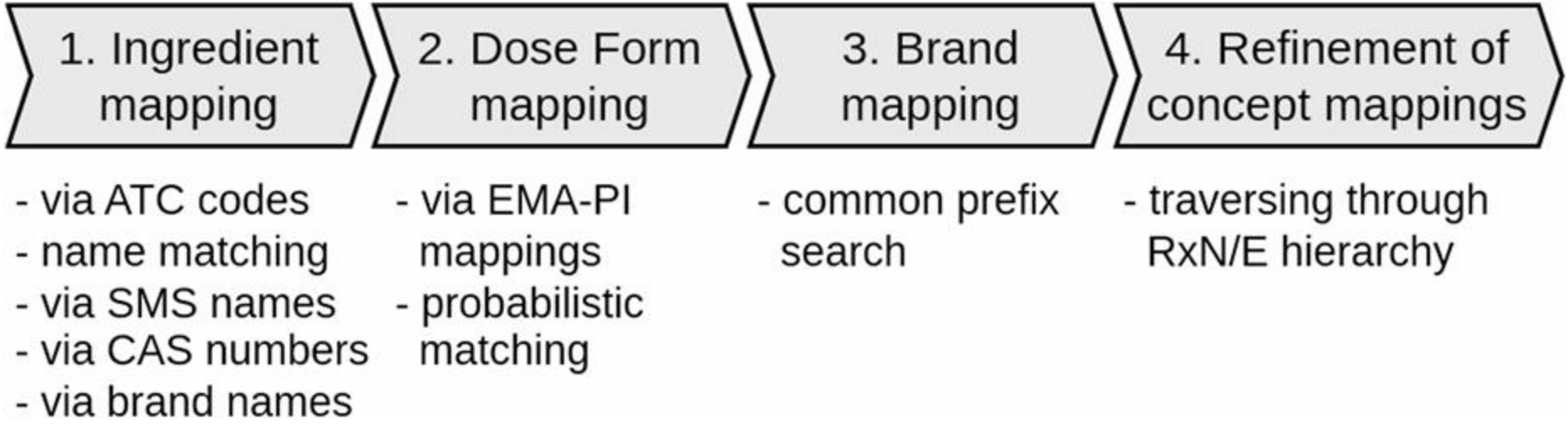


several (mostly) public databases

semi-automatic mapping, **Python software, shared on GitLab**

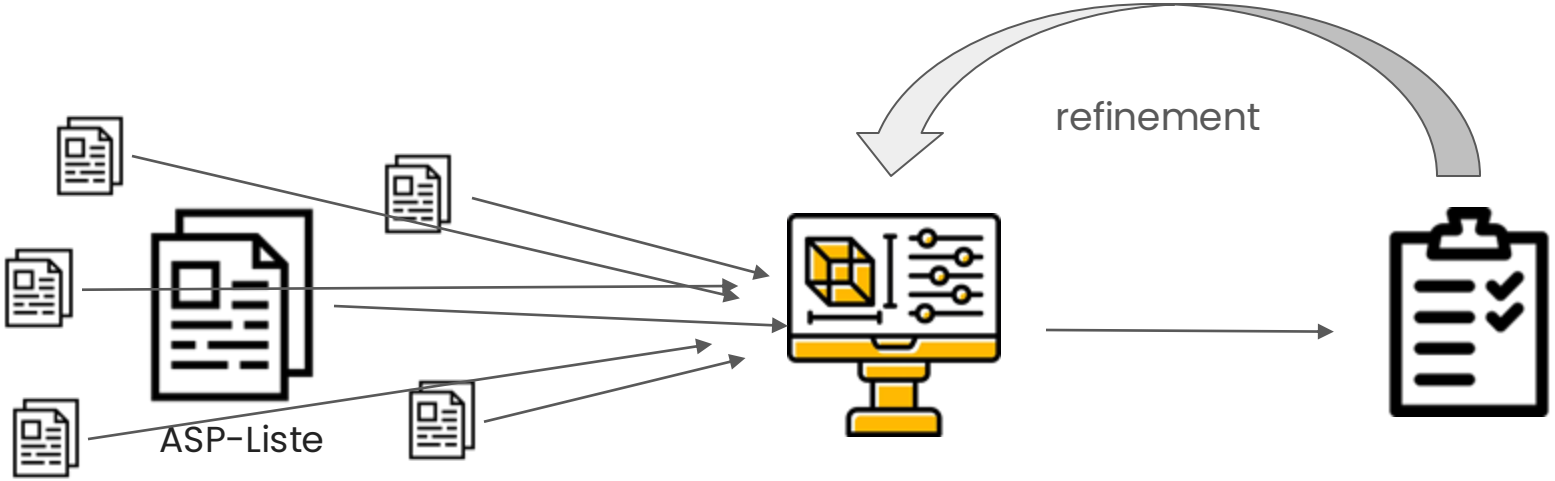
mapping results

Mapping sequence



SMS - substance translations by European Medicines agency
CAS - Chemical Abstract Service Registry Number

EMA-PI: European Medicines Agency Product Information



several public databases

semi-automatic mapping: **Python software, shared on GitLab**

Results

95% drug concepts mapped, Usagi-styled CSV file

→ does not work well: herbal medicines, blood substitutes

Great, I would love to use it!

Happy to hear!

But where can I find publicly available data, which is structured enough?

EURIPID database
for mutual sharing of prices of medicines

470 000 medicinal products, 31 countries

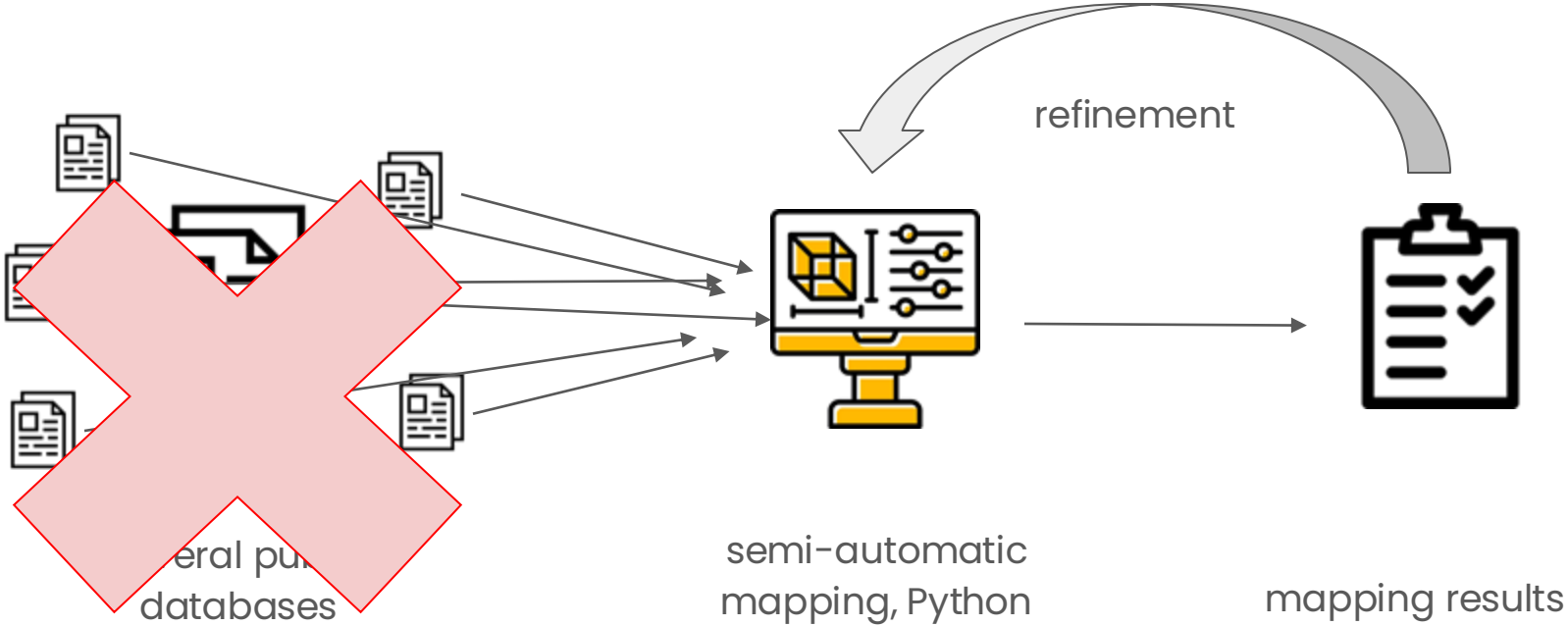
EURIPID - database structure

No	Product code	Country	Brand	Product name latest version	Package latest version	Dosage form
1	210184379	Hungary	intaxel	intaxel 30 mg/5 ml injekció	1x5 ml	concentrate for solution for inf
2	210184361	Hungary	intaxel	intaxel 100 mg/17 ml injekció	1x17 ml	concentrate for solution for inf
3	210141630	Hungary	synercid	synercid por oldatos infúzióhoz	1x	powder for solution for infusion

Dosage form Group	ATC	RoA	INN & strength	Nou	Company latest version
parenteral, liquid form	L01CD01	parenteral	paclitaxel 30 mg	1	fresenius kabi hungary vegy-, g
parenteral, liquid form	L01CD01	parenteral	paclitaxel 100 mg	1	fresenius kabi hungary vegy-, g
parenteral, liquid form	J01FG02	parenteral	dalfopristin 350 mg; quinupristin 150 m	1	monarch pharmaceuticals irela
eye drops	S01GX01	ocular	cromoglicic acid 20 mg/ml	10	santen oy

→ some changes in the semi-automatic mapping software would be needed

Methods



EURIPID structured data

EURIPID database - participating countries

If you want to do (/redo) the mapping of drug codes from one of these countries, and you would like to use EURIPID data, please reach out!

nemeth.ger@neak.gov.hu, meszaros.agota@semmelweis.hu

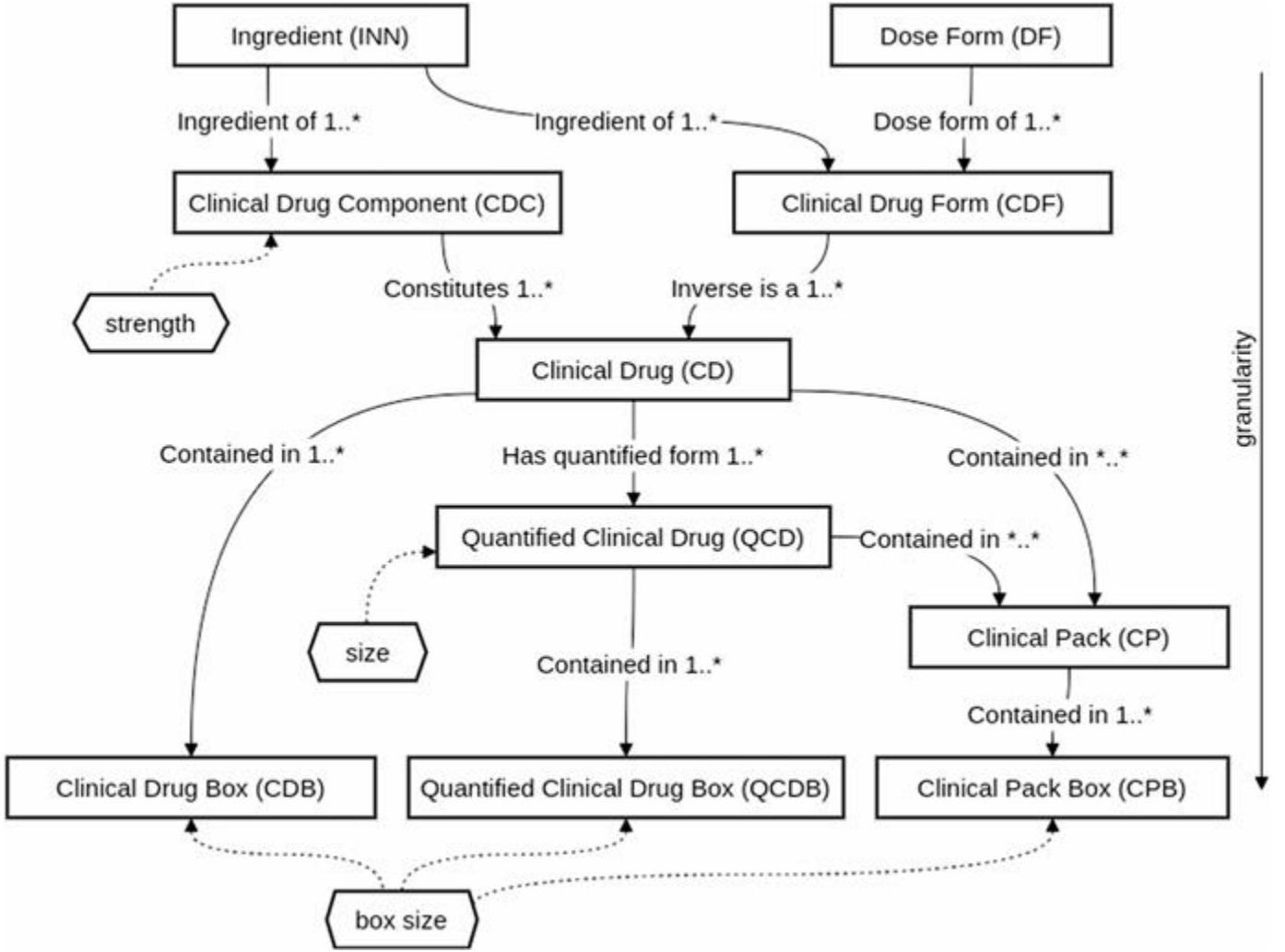
Thank you for your attention!

First author: Florian Katsch, florian.katsch@meduniwien.ac.at
Presenter: Agota Meszaros,
meszaros.agota@semmelweis.hu

References:
https://freeicons.io/
https://link.springer.com/article/10.1186/s12874-025-02669-0
http://euripid.eu/

Backup slides

RxNorm and Extension



Name	Content	Notes	Availability
OMOP CTD	Concepts from the Comparative Toxicogenomics Database (CTD) incorporated into OMOP's vocabularies. Concepts use the CAS numbers (Chemical Abstract Service Registry Number) as code	Used to assign CAS numbers via OMOP's "non standard to standard map" relationship to RxNorm and RxNorm Extension concepts of the ingredient concept class.	public – access via Athena [25]
OMOP ATC	Concepts from the World Health Organizations (WHO) ATC drug classification system. These concepts are augmented by OMOP with a variety of relationships to RxNorm and RxNorm Extension.	The "ATC – RxNorm pr lat" relationship is used to assign ATC codes to RxNorm and RxNorm Extension concepts of concept class ingredient and precise ingredient.	public – access via Athena [26]
Karapetian et al. terminology alignment of RxNorm and EDQM (EDQM-RX)	[27] contains a mapping from RxNorm Dose Forms to EDQM Dose Forms used for drugs described by the European Medicines Agency (EMA)	Used to assign RxN/E dose forms to EU drugs.	public – access via supplementary data in [27]
EMA – Product information URLs for member states (EMA-PI)	URLs to product information sheets for drug products recognized by EMA	Product information sheets are available as PDFs; EU drug registration numbers and EDQM dose forms are extracted from them. In combination with information from EDQM-RX, EU drugs can be assigned to RxN/E dose forms.	public – access via EMA website [28]
EMA – Substances Management Services (SMS)	Substance names in various European languages	Synonyms and translations of substance names share a common identifier; a preferred name in English language is designated. Used to derive synonyms and translations for ingredients.	public – access via EMA's SPOR website [29]
Raw data of substance names (BEZVO)	Substance names required to be used in regulatory affairs in Germany	Contains, among other information, substance names in German language and assigned CAS registry numbers. Mainly used to translate German ingredients and to gather synonyms.	public – access via German Federal Institute for Drugs and Medical Devices [30]