INTRODUCTION

The secondary use of routinely collected drug and health data is becoming more widespread, especially since there is a growing interest from pharma companies and medicine agencies as well - for example, the Data Analysis and Real World Interrogation Network (DARWIN EU) project. Our University previously manually mapped a part of the Hungarian medical procedures vocabulary. National drug code vocabularies seem to be more analogous to each other than procedure vocabularies; therefore, an automated or semi-automated mapping method is more feasible.

AIMS

In this work, we aimed to explore the most efficient way to map Semmelweis University’s drug data to RxNorm and RxNorm Extension.

METHODS

Our database contains the patient-level inpatient medication and prescription drug data from Semmelweis University’s Hospital Information System, e-MedSolution covering the last 14 years. We performed data cleaning to determine which drugs are to be mapped and which are not: our target is to map those drugs used at our university that were authorized in Hungary or the European Union (EU) (See Figure 1.)

Figure 1. Source data

Relevant stakeholders were contacted in the process to gather information about drug data and develop the mapping methods and standards of operations. Besides e-MedSolution, other databases and code systems were also explored.

RESULTS

INITIAL DATASET

Initially, data on 22540 drugs and drug-related items were downloaded from e-MedSolution. After deduplication, 13505 distinct drugs and drug-related items were found, which were prescribed or administered at Semmelweis University from 1 January 2010 to 1 December 2023.

DATA CLEANING

We cleaned out the rows not containing TTT code (Social Insurance Supported Product Identifier), the Formulæ Normales and magistral drugs, the Pharmacopoeia Hungarica drugs, the feeding formulas or medicinal food products, and the medical aids or devices, thus 10678 distinct drugs remained. The flowchart showing the data cleaning can be found in Figure 2.

Figure 2. Flowchart of data cleaning

CONNECTED WITH ATC5 CODES

Based on the ATC5 code, these 10678 drugs match 1361 distinct chemical ingredients.

REFERENCES


163 rows with no TTT code

95 Pharmacopoeia Hungarica drugs

313 Formulæ Normales and magistral drugs

522 feeding formulas and medicinal food products

1734 medical aids and devices

CONCLUSION

For the mapping of medicines authorized only at the national level, the data from multiple databases can help to map them efficiently. However, for those medicines that have an EU authorization number, the use of a common EU medicines database mapping would be a possible solution to make the mapping automated and more standardized.

After we mapped our drug codes, we plan to find a way to make this mapping available for everyone, thus with 1676 codes starting the freely usable mapping between the EU authorization number and RxNorm or RxNorm Extension.
