Mapping U.S. FDA National Drug Codes to Anatomical-Therapeutic-Chemical Classes using RxNorm

Fabrício Kury, MD[1], Olivier Bodenreider, MD, PhD[1]

[1]U.S. National Library of Medicine, Bethesda, MD

Background

The U.S. Food and Drug Administration (FDA) National Drug Codes (NDCs) are the official identifiers for drug products in the U.S. To analyze datasets from a clinical perspective, it is often useful to identify NDCs by drug classes. The Anatomical-Therapeutic-Chemical (ATC) drug classification system[1] is among the most popular. Here we demonstrate how to use the U.S. National Library of Medicine RxNorm[2] application programming interface (API) to map large volumes of NDCs to ATC classes, and present statistics from mapping two datasets.

Methods and Results

We produced an R script, which is freely available at https://github.com/fabkury/ndc_map, to completely automate the mapping process. The script ingests a CSV file with the NDCs (and their dates), queries the RxNorm API to obtain their fourth-level ATC class (or classes) if any is available, then outputs the results into a new CSV file. Tables 1 and 2 provide statistics about the mapping of the two datasets under investigation.

Table 1. Results of the mapping process.

	Medicare Part D	All-payer (Partners Healthcare©)
Total NDCs	71,309	134,580
Years	2006 - 2013	2011 - 2012
Mapped NDCs	55,565 *(77.9%)*	97,961 *(72.8%)*
NDCs with no RxCUI	10,413 *(14.6%)*	29,160 *(21.7%)*
RxCUI but no ATC-4	5,331 *(7.5%)*	7,459 *(5.5%)*
Unique NDC—ATC-4	114,069	187,426
Unique ATC-4 classes	517 *(58.62%)*	540 *(61.2%)*
Mapped prescriptions	97.6%	*Not available.*
Running time	65 minutes	146 minutes

Table 2. Tallies of ATC-4 classes per NDC.

ATC-4 per NDC	Medicare	All-payer
0	15,744 *(22.1%)*	36,619 *(27.2%)*
1 to 2	45,266 *(63.5%)*	78,414 *(58.3%)*
3 to 8	8,887 *(12.5%)*	17,667 *(13.1%)*
9 to 21	1,412 *(2%)*	1,880 *(1.4%)*
≥ 22	0	0

Discussion

In the Medicare dataset, only 77.9% of the NDCs were mapped to at least one ATC class, but those represented 97.6% of all prescriptions. From an informal analysis, the majority of the non-mapped prescriptions were for over-the-counter medications or non-drug supplies (e.g. syringes, gauze). Individual NDCs can get mapped to multiple ATC-4 classes because the script performs the mapping via the drug ingredients alone. Some drugs (NDCs) have more than one active ingredient (e.g. acetaminophen + hydrocodone), and furthermore ATC can provide different classes for one same ingredient depending on its usage (e.g. miconazole can be an oral, gynecological, dermatological, or otological drug). It is possible, however, to solve most ambiguities in semi-automated manner using other RxNorm resources leveraging other details of each NDC beyond its ingredients. If not addressed, the NDC-to-ATC ambiguity can create issues in analyzing prescription datasets, such as duplication of data. Our approach leveraged RxNorm to provide a very straightforward, fully automated and scalable solution.

1. World Health Organization (WHO) Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2016. 2016. Accessed from http://www.whocc.no/atc_ddd_index/ on March 9, 2017.

2. U.S. National Library of Medicine. RxNorm APIs. Accessed from https://rxnav.nlm.nih.gov/APIsOverview.html on March 9, 2017.

Source code *(also in QR code)*:
https://github.com/fabkury/ndc_map

Contact:
Fabrício Kury – kuryf@nih.gov

U.S. National Library of Medicine