

## Historically Comprehensive Medications Metadata for i2b2 Data Warehouses

Pedersen, Jay G. MA<sup>1</sup>, Gardner, Bret J. <sup>1</sup>, Campbell, James R. MD<sup>1</sup>, Campbell, W. Scott PhD<sup>1</sup>, Geary, Carol PhD<sup>1</sup>; Peters, Lee<sup>2</sup>; Bodenreider, Olivier PhD<sup>2</sup>; McClay, James C. MD<sup>1</sup>

<sup>1</sup>University of Nebraska Medical Center, Omaha, NE

<sup>2</sup>National Library of Medicine, Bethesda, MD

### Introduction

In the U.S., the Office of the National Coordinator guidelines recommend using the RxNorm controlled medical vocabulary for clinical drug orders and National Drug Codes (NDC) for dispense and administration events within electronic health records (EHRs) (1, 2). Use of RxNorm to index and track medication events supports clinical research data needs for medications. The terminology is updated frequently to include new medications and remove discontinued drugs. To maintain current and historically accurate medication data, methods are necessary to manage RxNorm updates within research data warehouses.

The RxNorm vocabulary was created in 2005 and is maintained by the National Library of Medicine (NLM). RxNorm specifies the set of ingredients, clinical and branded drugs and drug packs that are orderable by clinicians. The NLM publishes these components, including the linkage of ingredients to the clinical drugs that employ them and the packs that manufacturers provide. Each ingredient and medication is assigned an RxNorm concept unique identifier (RxCUI).

The Informatics for Integrating Biology & the Bedside (i2b2) (3) clinical data warehouse is used extensively by research hospitals to facilitate clinical research. Numerous domains of clinical data are maintained in an i2b2 system, including procedures, diagnoses, demographics, and medications. Medication order data are commonly stored using RxCUI codes. Querying these facts requires metadata to refer to the same RxCUI codes. This paper describes a methodology for creating historically comprehensive RxNorm-based medication metadata.

As the NLM updates RxNorm, most of the information exposed by the NLM's REST API interface for RxNorm (4) and the native RxNorm files [<https://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html>] is restricted to current RxCUIs. This information does not include retired RxCUI codes and may not fully encompass the historic medication usage contained within a clinical data warehouse. If the medication metadata in an i2b2 data warehouse is derived solely from the current publication of the RxNorm vocabulary, historic drug facts may remain undiscovered in research queries. This leads to an incomplete picture of medication usage.

The approach used at UNMC for RxNorm-based medication metadata exposes both current and historic RxCUI drug codes. Within the metadata, RxCUI drug codes are children of their associated ingredient RxCUI codes. This allows queries by a drug ingredient to reliably find ordered drugs encoded with any current or historic RxCUI drug code associated with the ingredient.

### Methods

#### (1) Obtain a historically comprehensive set of RxCUI codes.

UNMC worked with NLM to obtain a historically comprehensive list of RxCUI codes as of the end of October 2017. The RxCUI History API (<https://rxnav.nlm.nih.gov/RxcuiHistoryAPIs.html#>) was systematically queried for every RxCUI ever published in RxNorm, whether currently active or not. The information provided for each RxCUI code included the status, descriptive name, term type (TTY), and start and end dates.

#### (2) Create medications metadata with a historically comprehensive set of RxCUI codes.

i2b2 metadata for RxNorm-coded medications was produced using historical RxCUI information from the previous step and current RxCUI information from the NLM's REST API for the RxNorm terminology (4) (<https://rxnav.nlm.nih.gov/RxNormAPIs.html>). The medications metadata is arranged as a hierarchy with drug ingredients at the top level and drug RxCUI codes as children of their respective ingredient codes.

#### (3) Update the metadata table periodically to maintain currency with RxNorm.

The UMLS publishes RxNorm releases (5) on a monthly basis with synchronous exposure to the NLM's REST API for RxNorm (4). To maintain currency, the medication metadata for i2b2 is updated regularly.

## Results

We found that 22% of the RxCUI-encoded facts in our i2b2 instance contained retired RxCUI codes.

*Table 1 RxCUI-encoded facts with retired codes in our i2b2 data warehouse*

Total RxCUI-encoded facts	Retired RxCUI-encoded facts	Retired percentage
17,864,298	3,959,064	22.2%

The following table shows the differences in RxCUI code counts between the comprehensive historical list of October 2017 and the list reported by the RxNorm API “all concepts” endpoint on Feb 24, 2018.

*Table 2 Historically Comprehensive RxCUI code counts*

Term Type	IN	MIN	SCD	GPCK	SBD	BPCK
Description	Ingredient	Multiple Ingredient	Semantic Clinical Drug	Generic Drug Pack	Semantic Branded Drug	Branded Drug Pack
Historically Comprehensive RxCUIs	12,612	4,211	52,044	768	39,158	941
Currently active RxCUIs	11,636	3,970	18,432	379	10,419	451
Difference	976	241	33,612	389	28,739	490

*Figure 1 Part of the Acetaminophen metadata, showing retired RxCUI codes*

Acetaminophen [4,220,515 facts; 220,123 patients]

- (Retired) Acetaminophen 250 MG / Hydrocodone Bitartrate 10 MG Oral Tablet [≤10 facts]
- (Retired) Acetaminophen 300 MG / Codeine 15 MG Oral Tablet (52006-62010) [768 facts; 107 patients]
- (Retired) Acetaminophen 300 MG / Codeine 30 MG Oral Tablet [16 facts]
- (Retired) Acetaminophen 300 MG / Codeine 30 MG Oral Tablet (52006-62010) [111,605 facts; 16,614 patients]

## Discussion

Maintaining accurate medication metadata for an i2b2 data warehouse is challenging. The set of RxCUI codes changes rapidly with additions of new drugs and removals of retired drugs. The RxCUI codes within an i2b2 data warehouse are likely to be a mixture of currently active and retired RxCUI codes. Accurate data analysis requires consideration of all these RxCUI codes, requiring the metadata to be historically comprehensive.

The NLM has agreed to create a REST API endpoint to publish the historically comprehensive set of RxCUI codes. Presently, it is possible to determine this set from published RxNorm release files (5).

## Acknowledgement

This work was supported by the Intramural Research Program of the NIH, National Library of Medicine.

## Conclusion

We successfully developed and deployed historically comprehensive RxNorm-based medications metadata for i2b2.

## References

- (1) Office of the National Coordinator for Health IT. 2016 Interoperability Standards Advisory. 2016.
- (2) Nelson, Stuart J., et al. "Normalized names for clinical drugs: RxNorm at 6 years." *Journal of the American Medical Informatics Association* 18.4 (2011): 441-448.
- (3) Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc* 2010 Mar-Apr;17(2):124-130.
- (4) Peters, Lee, and Olivier Bodenreider. "Using the RxNorm web services API for quality assurance purposes." *AMIA Annual Symposium Proceedings*. Vol. 2008. American Medical Informatics Association, 2008.
- (5) <https://www.nlm.nih.gov/research/umls/rxnorm/docs/rxnormfiles.html>