Drug Identifier Standardization for FDA-Approved Drug Labeling Using RxNorm

Isaac Harris, Taylor Ingle, Zhichao Liu, Leihong Wu, Junshuang Yang, Stephen Harris, Joshua Xu, Weida Tong, and Hong Fang

National Center for Toxicological Research, U.S. Food & Drug Administration, Jefferson, AR 72079, USA

Abstract

The FDA regulates and approves all prescription drugs marketed in the US. The drug labeling, which must be available at the time of approval, contains rich drug safety and efficacy information widely used in healthcare practices. Drug identifiers can be presented in various forms in FDA systems, databases, and publications due to the inherent complexity of a drug. These conceptual inconsistencies can present challenges when comparing data within the FDA and with outside drug sources. This study investigates the different naming systems of a drug in FDA approved drug labeling and mapping of these identifiers to RxNorm (National Library of Medicine), a normalized drug naming system widely used in clinical medicine and research. We mapped different identifiers of a drug from the FDA labeling (e.g., UNII, Application ID, NDC, SPL Set ID) to an external identifier, RxNorm's SCD (Semantic Clinical Drug), and created a translator across the various drug identifiers to accurately represent standardized drug content. Mapping results between FDA drug labeling and RxNorm will allow easier identification and searching of drug information under the standard names, codes, vocabularies, and drug product representations, to advance drug related clinical decisions, research, and ultimately, drug safety for public health. Ultimately, RxNorm has enormous potential for the future development of FDALabel; as such it will benefit regulatory, clinical, and research applications of FDA Drug Labeling documents.

Introduction

Drug identifiers can be difficult to understand due to the inherent complexity of a drug and its approval processes. FDA Drug Labeling uses four primary drug identifiers: Unique Ingredient Identifier (UNII) for Active Ingredient, Application ID (Appl. ID), Structured Product Labeling (SPL) Set ID, and National Drug Code (NDC). **(1) UNII** codes are designed to identify any substance and broad enough to include anything from an atom to an organism. **(2) Appl. IDs** are designed for submission of drug products for FDA approval. **(3) SPLs** are written in XML and designed as a mechanism to exchange product and facility information. As such, SPLs are prepared by the drug product manufacturer and submitted to FDA for approval. Each SPL contains a drug labeling that can contain multiple drug products. **(4) NDCs** are designed to identify a singular drug product. RxNorm provides a multitude of drug concepts/identifiers (***Figure 1***, ***Workflow***). Here, we used two identifiers: Semantic Clinical Drug (**SCD**) and Semantic Branded Drug (**SBD**). **SCDs** are defined by RxNorm to be Ingredients + Strength + Dose Form and represent generic drugs; **SBDs** are defined by RxNorm to be Ingredients + Strength + Dose Form + Brand Name. Through RxNorm, the multiple drug identifiers can be standardized to be translated and interpreted in an accurate fashion to advance the understanding of drug labeling information and its applications in research, medicine, and public health.

Methodologies

RxNorm's normalized drug naming system incorporates a multitude of information sources such as MTHSPL (FDA Drug Labeling), SNOWMED CT, VANDF (Veterans Administration), to provide a carefully curated concept of a drug. We used this concept to standardize the many drug identifiers used within FDALabel (FDA Drug Labeling Tool). First, we constructed a relational table model in a PostgreSQL server to assist us in extracting information from RxNorm (***Figure 2***). Next, we mapped FDALabel drug identifiers to RxNorm's Semantic Clinical Drug (SCD)/Generic Drug (***Figure 3***) to present a single drug identifier that represents a multitude of drug products. The slight difference between RxNorm's concept of Ingredient and MTHSPL's Substances on a conceptual level will be explored for future standardization on the substance level.

Disclaimer: The views presented in this poster do not necessarily reflect the current or future opinion or policy of the U.S. Food and Drug Administration. Any mention of commercial products is for clarification and not intended as endorsement.

Results

Figure 1: Workflow for Mapping FDALabel Identifiers to RxNorm Standards

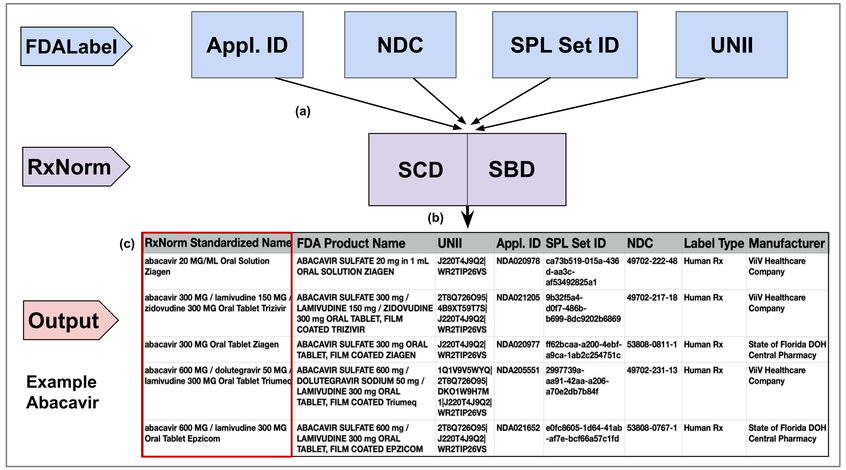

Figure 1: (a) FDALabel's four drug identifiers are mapped to RxNorm's Semantic Clinical Drug (SCD) or Semantic Branded Drug (SBD). There are more SCDs than SBDs and every SBD is related to an SCD, but not vice versa. (b) This mapping is then used to standardize the drug identifier. (c) By standardizing the drug identifier, we have a unified drug name to represent multiple drug products and information pertaining to said drug products. Note, for simplification, the example information does not show the additional drug products mapped to the RxNorm standardized name.

Figure 2: Drug Identifier Relationships for FDALabel using RxNorm Standard

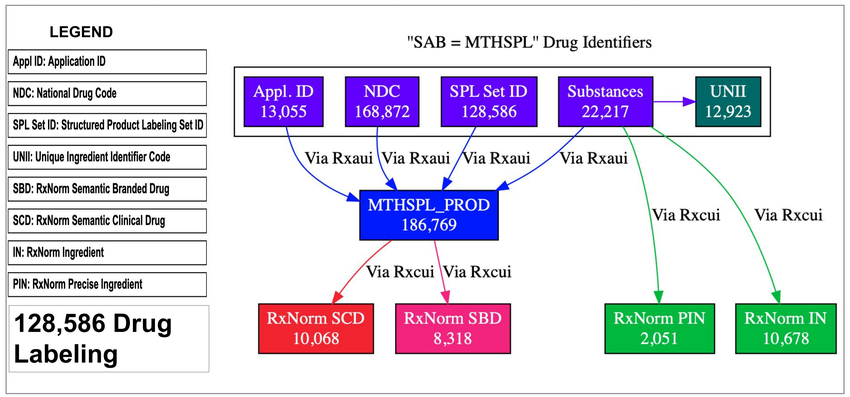

Figure 2: MTHSPL (FDALabel) is related to RxNorm via two different avenues. (1) Linking the rxcui (RxNorm concept unique identifier) between drug product and drugs forms a relationship between RxNorm and SPLs. From this relationship, the information and various identifiers contained within an SPL can be applied to an RxNorm drug concept. (2) Linking rxcui between RxNorm's ingredient concepts and FDA's substances. This allows us to associate the UNII codes from substances to RxNorm's ingredients. (3) To access the data within an SPL in MTHSPL, the relationship using RxNorm's rxaui (atom unique identifier) must be established. This relationship brings about a plethora of rich information. See ***Figure 1*** for example data.

Figure 3: Human Rx Drug Identifiers Mapped to More than One RxNorm Semantic Clinical Drug/Generic Drug

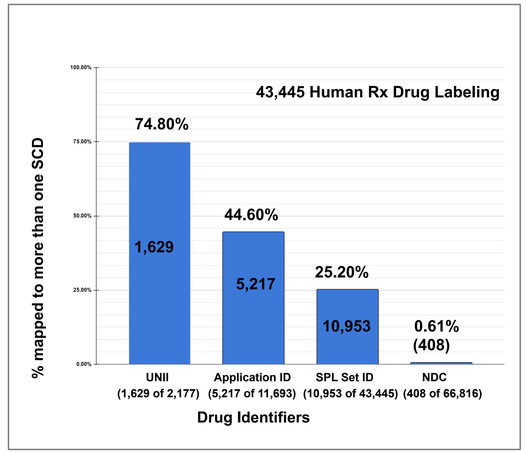

Figure 3: Chart illustrating the four primary drug identifiers used within FDALabel (total of 43445 Human Rx drug labeling) and the percentage that associated with more than one generic drug within RxNorm SCD: (1) Approximately 75% of all human Rx UNII codes (1,629 of 2,177); (2) Approximately 45% of all Human Rx application numbers (5,217 of 11,693); (3) Approximately 25% of all human Rx SPL Set IDs (10,953 of 43,445). (4) Approximately 0.61% of all Human Rx two-part sections NDC codes (408 of 66,816). As shown, NDCs are the best suited starting point for reaching a singular drug concept in RxNorm through the relationships expounded upon in ***Figure 2***.

Summary of Results

We investigated the four drug identifiers used by FDALabel (UNII, Appl. ID, SPL Set ID, NDC) and the various drug concepts/identifiers within RxNorm. (1) Due to their high percentage of association, UNII codes, Application ID, and SPL Set ID were found to be too broad of a concept to represent a singular drug. However, NDCs were found to be too narrow to represent a singular drug (***Figure 3***). No single identifier in FDALabel is as accurate as RxNorm SCD or SBD to describe a drug. We ascertained that the current drug identifiers are challenging for information integration within the FDA and with the public. (2) We explored RxNorm for its use in standardization for drug identifiers for FDALabel and found that it was possible to group multiple drug products and their relevant information under a single RxNorm SCD. As such, this will reduce redundancy in FDALabel and present a unified drug view for easier information access. (3) This standardization will help improve consistency resulting in higher quality drug labeling data. (4) RxNorm has enormous potential for future FDALabel application or as a standalone tool/application..

Conclusion

Standardization of the drug identifier for FDA-approved drug labeling using RxNorm will unify drug names and bolster interpretations between the regulatory interpretation of drugs and the clinical application of those drugs. This will allow pharmacists, clinicians, healthcare professionals, and researchers to efficiently utilize the rich resources (e.g., Drug Indications and Usage, Drug Interactions, Warnings and Precautions) included in the FDA drug labeling. In addition, standardization will allow products to be grouped under a unified drug identifier. This will improve the quality of information used in FDALabel and present a cleaner drug view for users. Ultimately, the mapping results between FDA drug labeling and RxNorm will allow easier identification and searching of drug information using standard names, codes, vocabularies, and drug product representations, and serve to advance drug related clinical decisions, research, and ultimately, drug safety for public health.

Acknowledgements

Acknowledgements: Isaac Harris is grateful for the support of the Summer Student Research Program at the National Center for Toxicological Research, U.S. FDA, administered by the Oak Ridge Institute for Science and Education.