NDFRT

Contents

Contents

Data Set Summary

- **Topic/focus:** National Drug File - Reference Terminology Database
- **Year(s):** 2015
- **Source:** NIH National Library of Medicine
- **Study and sample characteristics:** Drug information aggregated from RxNorm and VA NDF
- **Universe:** n/a
- **Variables:** n/a
- **Access:** Publicly Available
- **Cost:** Free
- **Key web links:** 2014AB UMLS National Drug File - Reference Terminology Source Information NDF-RT Documentation.pdf
- **Summary:** NDF-RT is an extension of the VHA National Drug File (NDF). It organizes the drug list into a formal representation. NDF-RT is used for modeling drug characteristics including ingredients, chemical structure, dose form, physiologic effect, mechanism of action, pharmacokinetics, and related diseases. NDF-RT combines the NDF hierarchical drug classification with a multi-category reference model. The categories are:
  - Cellular or Molecular Interactions [MoA]
  - Chemical Ingredients [Chemical/Ingredient]
  - Clinical Kinetics [PK]
  - Diseases, Manifestations or Physiologic States

[Disease/Finding]
- Dose Forms [Dose Form]
- Pharmaceutical Preparations
- Physiological Effects [PE]
- Therapeutic Categories [TC]
- VA Drug Interactions [VA Drug Interaction]

Data Set Details

Data processing

The NDF-RT is delivered as a hierarchical XML data structure that is a collection of code translation lookup tables (Kinds, Associations, Qualifiers, Properties, and Roles), and one giant Concepts table which is self-referential. This data structure is more suited for use in a NOSQL/MongoDB style database with various indexes setup to quickly locate concepts and then place the relationship burden on an application to perform searches, queries, and lookups.

The original data is split apart into multiple tables and a set of code lookup tables are provided when the user needs to get low level with the data. However, a set of views have been created to partially denormalize the data so that the retrieved information is easier to see when presented in a tabular format; that is, most of the codes are translated into their human readable equivalents (which are from the lookup tables themselves). This eliminates most of the ugly SQL joins necessary to make the data usable.

Data Tables

Lookup Tables

Five lookup tables convert Code identifiers (usually of the form C000000) into their textual representations based on where those codes are used. Initial analysis of these codes show that there is no actual overlap and while these could be combined into a single table, they have been left as separate tables to match their original delivered categorization.

Kinds

- All concepts are categorized into one of 7 kinds. They are the same as listed in the above summary. Each concept is associated with exactly one kind.

Associations

- A subset (16844 concepts) of concepts have association links that relate a concept to one or more other concepts. These concepts can be chained to generate a relationship tree that describes how that concept is related to other concepts. While these are similar to Defining Roles, this association tree that can

be generated is more related to the composition of various prescribable drugs and their optionally associated quantities.

Qualifiers

- Certain Properties and Associations have qualifiers that qualify their relationship to the referenced concept. In this data set:
    - Properties
        - VAND_Record properties have 3 qualifiers which simply break apart the VAND_Record string into their three constituent components:
        - VA.File Identifier (xx.xxx)
        - VA.IEN Identifier (Internal Entry Number) (xxxxx)
        - VA.Status indicator (Active/Master or Inactive/Master)

    - Associations
        - All associations have one or more qualifiers that consist of up to 3 qualifier attributes:
        - VA.IEN Identifier (xxxxx)
        - Strength - the strength or amount that this associated concept (drug) contributes to the primary concept (drug). e.g., "5" in 5%.
        - Unit - the unit of the strength value that this associated concept (drug) contributes to the primary concept (drug). e.g., "%" in 5%.

Properties

- All concepts are defined by one or more properties of which some properties will appear multiple times with the same property code (e.g., C136: Synonym)

Roles

- From the documentation: *Role relationships help to describe and define concepts according to their relationships with other concepts. Each role has a domain — the kind of concept whose definition may use the role, as well as a range — the kind of concept that the role can refer to. As definitional relationships, roles support tool-based automated classification and are inherited into descendant concepts in the resulting inferred taxonomy.*
    - Some roles are grouped together and define a roleGroup. All roles in this dataset are aggregated together. When a concept's role has a roleGroup of 0, the is no grouping of that role association with any other role association. When a concept's role has a roleGroup of 1 or higher, then all role associations with the same roleGroup for the specified concept must be consided together. e.g., A concept with a counterindication for a disease/finding may have a directly associated drug role that would indicate that the associated drug must not be used in conjunction with the primary concept when considered with that

disease/finding role.

- Most concepts are also defined by a set of defining roles. This allows a graph relationship to be built to understand how this particular concept is classified and/or comprised of (depending on the type of concept). For example, a compound drug will typically be defined at least by the constituent drugs, but may also include other defining roles. Acetaminophen/Codeine will have at least two defining roles where one is Acetaminphen, and the other is Codeine.

Each lookup table is presented in its original form. The table layout is as follows:

ndfrt.kinds

Column	Type	Description
id	Integer	Numeric kinds type id (not used)
code	String	Kind code used by the concepts table to indicate the concept kind
name	String	Readable name of the concept kind

ndfrt.associations

Column	Type	Description
id	Integer	Numeric associations type id (not used)
code	String	Association Type code used by concept associations (concept_assoc table) table (only one used is C28035649040869)
name	String	Readable name of the concept association type (only one used is Product_Component)
inversename	String	reverse name code when used with a dataset that includes the A100 code (not used)

ndfrt.qualifiers

Column	Type	Description
id	Integer	Numeric qualifier type id (not used)
code	String	Property Qualifier code or Association Qualifier code used by the concept properties qualifiers (concept_prop_qual table) or concept associations qualifiers (concept_assoc table) to indicate the qualifier attribute type. The only ones

		used are C28476375667298, C28489595891377, Q147, C28345466776084, C28533803573638, and C28542169458967.
name	String	Readable name of the qualifier attribute type. The only ones used are VA_File, VA_IEN, VA_Status, VA.IEN, Strength, and Unit.

ndfrt.roles

Column	Type	Description
id	Integer	Numeric role type id (not used)
code	String	Role Type code used by concept roles (concept_roles and concept_defining tables) table (all roles are used except C46)
name	String	Readable name of the concept role type (all role types are used except metabolized_by_{NDFRT})
domain	String	The domain kind for this type of role (all of them are C8 - DRUG_KIND)
range	String	The range kind for this role (all 8 are used)

ndfrt.properties

Column	Type	Description
id	Integer	Numeric property type id (not used)
code	String	Property Type code used by concept properties (concept_prop and concept_prop_qual tables) table (all roles are used except P900)
name	String	Readable name of the concept property type (all property types are used except Value_Set)
range	String	The value type which is not actually a functional value (ignore this field because they're all "string", "long_string", and "real_long_string")

Table Definitions

ndfrt.concepts

This is the primary concepts table that house the primary attributes that uniquely describe a given concept. All the other

tables that start with ndfrt.concept_* reference this concepts table by the concept id.

Use the *ndfrt.v_concepts* view for a nicer data set where the kind column is in a readable format.

column	type	Foreign Key Reference	description
id	Integer		Numeric Concept Id (primary key)
code	String		Concept code id that is referenced by all other concept_* tables and external databases that reference concepts in the NDF-RT.
name	String		Readable description of what this concept is for; including: drug name, disease, classifier, etc.
kind	String	*ndfrt.kinds.code*	NDF-RT Kind code that is a foreign-key for the ndfrt.kinds table. use the views to avoid having to join this up to get a readable data set)

ndfrt.concept_assoc

This is the associations table for all concepts. This is effectively a bridge table that links concepts to concepts based on the association type relationship (only Product_Component is used). Use the **ndfrt.v_concept_assoc** for a more functional data table that expands both sides of the association relationship. Be aware that this is a *MANY:MANY* relationship; so a single concept that appears in this table (on the left side) may appear more than once.

Use the *ndfrt.v_concept_assoc* view for a nicer more usable data set (that internally joins this table together with everything).

column		type	Foreign Key Reference	
id		Integer	*ndfrt.concepts.id*	
assoc_type		String	*ndfrt.associations.code*	

LastPass
Enterprise

>

Software Tools

>

Training
Materials

assoc_concept		String		*ndfrt.concepts.code*	
va_ien		Integer			
strength		String			
unit		String			

ndfrt.concept_defining

This is the defining concepts table for all concepts that have defining concepts. This is bridge table that links concepts to other concepts with a defining relationship. That is, the primary concept (left side) is 'is a' or 'is a member of' the right side concept. You can "walk" a defining concept chain/tree such that you end up at a concept that has no more defining concepts. For example, walking the concept relationship for concept C12846 "CIDOFOVIR":

- C1286 - CIDOFOVIR
    - C288629 - C [Preparations]
        - C288626 - Drug Products by Generic Ingredient Combinations
            - C176 - Pharmaceutical Preparations
    - C634841 - Cytomegalovirus Nucleoside Analog DNA Polymerase Inhibitor [EPC]
        - C634794 - Drug Products by FDA Established Pharmacologic Class
            - C634798 - Prostacycline Vasodilator [EPC]
                - C634794 - Drug Products by FDA Established Pharmacologic Class
                    - C176 -

Pharmaceutical
Preparations

In this case CIDOFOVIR is defined by two concepts: C [Preparations], and Cytomegalovirus Nucleoside Analog DNA Polymerase Inhibitor [EPC]. Each of those two in turn are defined by additional defining roles (in this case, each had only one defining role) until they all terminated at C176 or Pharmaceutical Preparations.

Use the *ndfrt.v_concept_defining* view for a nicer more usable data set (that internally joins this table together with everything) and lets you extract both concept codes, concept names and concept kinds.

column		type		Foreign Key Reference		Has NULL
id		Integer				No
concept		String		*ndfrt.concepts.code*		No

ndfrt.concept_prop

This is the properties table for all concepts. This is an attributes table that provides 0 or more property attributes that *describe* the identified concept.

Use the *ndfrt.v_concept_prop* for a more functional data table that dereferences all the codes to provide a more readable properties list. Be aware that this is a *MANY:MANY* relationship; so a single concept that appears in this table (on the left side) may appear more than once.

Be aware that the property VANDF_Record has qualifiers that simply break down the VANDF_Record string in their constituent properties (which is the same as splitting the string on the "^").

column		type		Foreign Key Reference		Has NULL
id		Integer		*ndfrt.concepts.id*		No

name		String		*ndfrt.properties.code*		No
value		String		*... varies ...*		No

ndfrt.concept_prop_qual

This is the concept property qualifiers table for all properties that have qualifiers. In this data set, the only property type that has qualifiers is VANDF_Record. All VANDF_Record's have three qualifier attributes. They have been de-normalized and split out into their own column in this table.

Use the *ndfrt.v_concept_prop_qual* view for a more functional data table that lists the concept information and the three associated VANDF_Record qualifier attributes.

column		type		Foreign Key Reference		H N
concept_id		Integer		*ndfrt.concepts.id*		N
property_value		String				N
property_name		String		*ndfrt.properties.code*		N
va_file		String				N

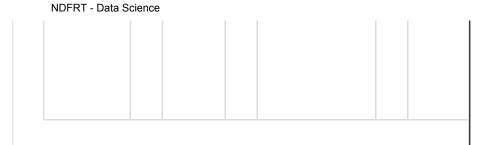|					
va_ien		Integer			N
va_status		String			N

ndfrt.concept_roles

This is the concept roles table that lists the roles based association between the primary concept and other concepts.

Use the *ndfrt.v_concept_roles* view for a more functional data table that lists the concept roles information.

column		type		Foreign Key Reference		Has NULL
concept_id		Integer		*ndfrt.concepts.id*		No
group_num		Integer				No
name		String		*ndfrt.roles.code*		No
value		String		*ndfrt.concepts.id*		No

Sources

All information relating to data elements descriptions and database summaries have been pulled from the database providers website located here: http://www.nlm.nih.gov/research/umls/sourcereleasedocs/cur