Ontologies support semantic interoperability in healthcare

Two use cases in information exchange and analytics

Olivier Bodenreider
Lister Hill National Center for Biomedical Communications
Bethesda, Maryland - USA
Disclaimer

The views and opinions expressed do not necessarily state or reflect those of the U.S. Government, and they may not be used for advertising or product endorsement purposes.
Outline

◆ The ecosystem of ontologies for healthcare
  ● Clinical ontologies
  ● Ontology integration systems

◆ Two use cases
  ● Information exchange
    ◾ Value sets for Clinical Quality Measures
  ● Analytics
    ◾ Analysis of large distributed clinical data warehouses
The ecosystem of ontologies for healthcare
Biomedical ontologies

◆ Structural perspective
  ● What are they (vs. what are they for)?

◆ “High-impact” biomedical ontologies
  ● International Classification of Diseases (ICD)
  ● Logical Observation Identifiers, Names and Codes (LOINC)
  ● SNOMED Clinical Terms
  ● Foundational Model of Anatomy
  ● Gene Ontology
  ● RxNorm
  ● Medical Subject Headings (MeSH)
  ● NCI Thesaurus
  ● Unified Medical Language System (UMLS)

[J. Cimino, YBMI 2006]
Biomedical ontologies

◆ Functional perspective
  ● What are they for (vs. what are they)?

◆ “High-impact” biomedical ontologies

◆ 3 major categories of use
  ● Knowledge management (indexing and retrieval of data and information, access to information, mapping among ontologies)
  ● Data integration, exchange and semantic interoperability
  ● Decision support and reasoning (data selection and aggregation, decision support, natural language processing applications, knowledge discovery).

[Bodenreider, YBMI 2008]
LOINC

- Type: Controlled terminology*
- Domain: Laboratory and clinical observations
- Developer: Regenstrief Institute
- Availability
  - Publicly available: Yes
  - Repositories: UMLS, BioPortal
- Size: > 50k codes
- Uses: information exchange (e.g., HL7 messages)
- URL: [www.regenstrief.org/loinc/loinc.htm](http://www.regenstrief.org/loinc/loinc.htm)
**LOINC Example**

- **Sodium:**SCnc:Pt:Ser/Plas:Qn (2951-2)
  
  [the molar concentration of sodium is measured in the plasma (or serum), with quantitative result]

<table>
<thead>
<tr>
<th>Axis</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td>Sodium</td>
</tr>
<tr>
<td>Property</td>
<td>SCnc – Substance Concentration (per volume)</td>
</tr>
<tr>
<td>Timing</td>
<td>Pt – Point in time (Random)</td>
</tr>
<tr>
<td>System</td>
<td>Ser/Plas – Serum or Plasma</td>
</tr>
<tr>
<td>Scale</td>
<td>Qn – Quantitative</td>
</tr>
<tr>
<td>Method</td>
<td>--</td>
</tr>
</tbody>
</table>
SNOMED CT

- **Type:** Reference terminology / ontology
- **Domain:** Clinical medicine
- **Developer:** IHTSDO

**Availability**
- Publicly available: Yes* (in member countries)
- Repositories: UMLS, BioPortal

**Size:** > 300k concepts

**Uses:** clinical documentation, analytics

**URL:** [http://www.ihtsdo.org/](http://www.ihtsdo.org/)
RxNorm

- Type: Controlled terminology
- Domain: Drug names
- Developer: NLM

Availability
- Publicly available: Yes*
- Repositories: UMLS, BioPortal

Size: > 10k ingredients; 19k clinical drugs
Uses: e-prescription, information exchange

URL: http://www.nlm.nih.gov/research/umls/rxnorm/
Azithromycin

Azithromycin 250 MG
Azithromycin Oral Tablet

Zithromax

Azithromycin 250 MG
Azithromycin Oral Tablet

{6 (Azithromycin 250 MG Oral Tablet) } Pack

Z-PAK
Unified Medical Language System

- Type: Terminology integration system
- Domain: Biomedicine
- Developer: NLM
- Availability
  - Publicly available: Yes*
  - Repositories: also available through BioPortal
- Size: 140 source vocabularies; 3.1M concepts
- Uses: annotation, NLP, cross-walk, …
- URL: http://www.nlm.nih.gov/research/umls/
UMLs Example

Addison's disease (363732003)

Clinical repositories

SNOMED CT

Genetic knowledge bases

OMIM

Biomedical literature

Addison Disease (D000224)

Other subdomains

NCBI Taxonomy

Model organisms

Anatomy

FMA

GO

Genome annotations

MeSH

Clinical repositories

Biomedical literature

Addison Disease (D000224)
Use case #1
Information exchange

Value sets for Clinical Quality Measures
Information exchange

◆ “Meaningful Use” incentive program

- Use of certified electronic health record (EHR) systems
- Requires use of select biomedical terminologies
  - For information exchange (e-prescribing, lab results)
  - For quality purposes (clinical quality measures)
    - Rely on reference value sets
Reference value sets

**eMeasure Title**: Venous Thromboembolism Prophylaxis

**Description**: This measure assesses the number of patients who received VTE prophylaxis or have documentation why no VTE prophylaxis was given the day of or the day after hospital admission or surgery end date for surgeries that start the day of or the day after hospital admission.

- `$MedicationVTEProphylaxis` = Union of:
  - "Medication, Administered: Low Dose Unfractionated Heparin for VTE Prophylaxis"
  - "Medication, Administered: Low Molecular Weight Heparin for VTE Prophylaxis"
  - "Medication, Administered: Injectable Factor Xa Inhibitor for VTE Prophylaxis"
  - "Medication, Administered: Warfarin"

**Value Set Details**

- **Name**: Low Dose Unfractionated Heparin for VTE Prophylaxis
- **OID**: 2.16.840.1.113762.1.4.1045.39

**Metadata**

<table>
<thead>
<tr>
<th>Code</th>
<th>Descriptor</th>
<th>Code System</th>
</tr>
</thead>
<tbody>
<tr>
<td>1361568</td>
<td>heparin sodium, porcine 2000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361574</td>
<td>heparin sodium, porcine 20000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361577</td>
<td>heparin sodium, porcine 2500 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361615</td>
<td>heparin sodium, porcine 5000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1361853</td>
<td>0.5 ML heparin sodium, porcine 10000 UNT/ML Prefilled Syringe</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1362831</td>
<td>heparin sodium, porcine 10000 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
<tr>
<td>1362837</td>
<td>heparin sodium, porcine 12500 UNT/ML Injectable Solution</td>
<td>RXNORM</td>
</tr>
</tbody>
</table>

Use case #2
Analytics

Analysis of large distributed clinical data warehouses
Analytics

◆ Clinical data warehouses
  ● Distinct from EHR systems

◆ “ETL” (extract – transform – load) processes
  ● Data normalized to "standards"
  ● Local data mapped to ontologies
    ▪ Facilitated by ontology integrations systems (e.g., UMLS)

◆ Analysis leverages hierarchical and other relations
  ● Transitive closures

◆ Facilitates analysis of large-scale data repositories
  ● Including distributed repositories across institutions
“Multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics”

OMOP Common Data Model
- Standard vocabularies
  (e.g., LOINC, SNOMED CT, RxNorm)

Investigation of treatment pathways
- For 3 chronic diseases (3 year-follow up)
  - > 1M patients with hypertension
- Across multiple clinical institutions
- In several countries

http://www.ohdsi.org/
References  Review articles


