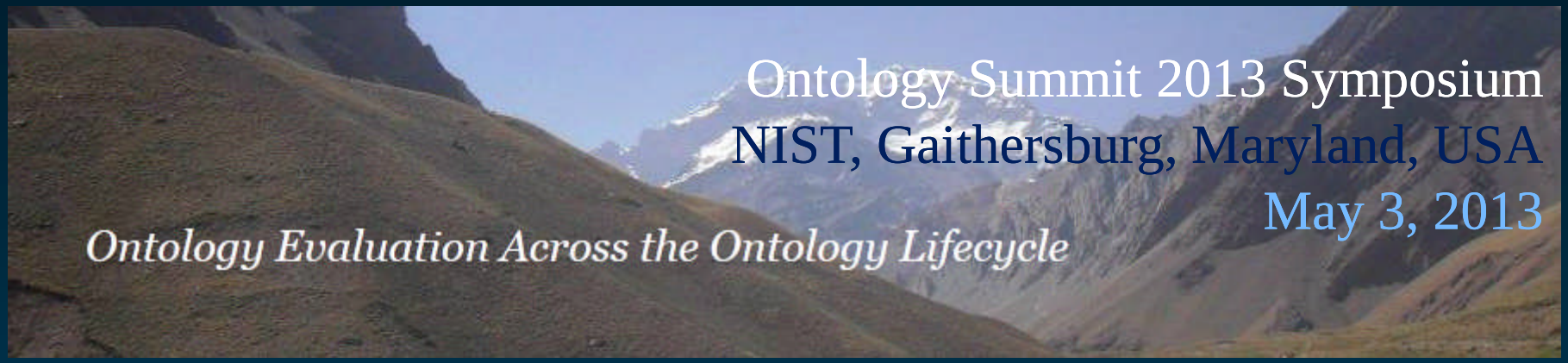


Quality assurance of biomedical ontologies (and derived artifacts) in the era of Meaningful Use

Olivier Bodenreider

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

Outline

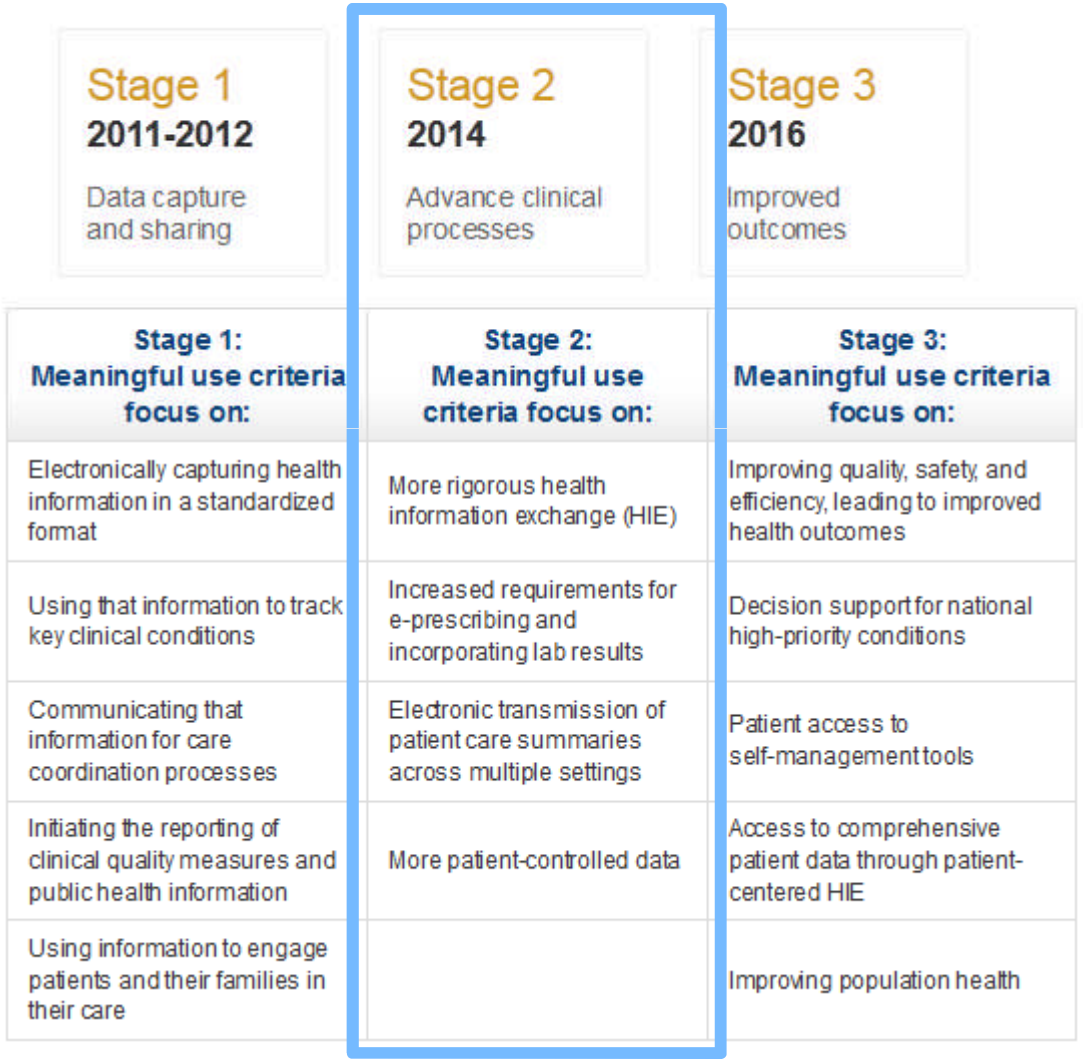
- Meaningful Use
- Standard vocabularies in Meaningful Use
- Value sets for clinical quality measures
- Quality assurance of biomedical terminologies
- Quality assurance of value sets

MEANINGFUL USE

“Meaningful Use”

- Health Information Technology for Economic and Clinical Health (HITECH) Act
  - Eligible health care professionals and hospitals can qualify for Medicare and Medicaid incentive payments when they adopt certified EHR technology and use it to achieve specified objectives
- Two sets of regulations
  - *Incentive Program for Electronic Health Records*
    Medicare and Medicaid Services (CMS)
  - *Standards and Certification Criteria for Electronic Health Records*
    Office of the National Coordinator (ONC)

Meaningful Use stages



Stage 1 2011-2012 Data capture and sharing	Stage 2 2014 Advance clinical processes	Stage 3 2016 Improved outcomes
Stage 1: Meaningful use criteria focus on:	**Stage 2: Meaningful use criteria focus on:**	**Stage 3: Meaningful use criteria focus on:**
Electronically capturing health information in a standardized format	More rigorous health information exchange (HIE)	Improving quality, safety, and efficiency, leading to improved health outcomes
Using that information to track key clinical conditions	Increased requirements for e-prescribing and incorporating lab results	Decision support for national high-priority conditions
Communicating that information for care coordination processes	Electronic transmission of patient care summaries across multiple settings	Patient access to self-management tools
Initiating the reporting of clinical quality measures and public health information	More patient-controlled data	Access to comprehensive patient data through patient-centered HIE
Using information to engage patients and their families in their care		Improving population health

Fact Sheets

Details for: CMS MEDICARE AND MEDICAID EHR INCENTIVE PROGRAMS: STAGE 2 FINAL RULE

Return to List

For Immediate Release: Thursday, August 23, 2012

Contact: CMS Office of Public Affairs
202-690-6145

CMS MEDICARE AND MEDICAID EHR INCENTIVE PROGRAMS: STAGE 2 FINAL RULE

On August 23, 2012, the Centers for Medicare & Medicaid Services (CMS) announced a final rule to govern Stage 2 of the Medicare and Medicaid Electronic Health Record (EHR) Incentive Programs. The rule specifies the Stage 2 criteria that eligible professionals (EPs), eligible hospitals, and critical access hospitals (CAHs) must meet in order to continue to participate in the EHR Incentive Programs.

Rule Provisions

Through the Stage 2 requirements of the Medicare and Medicaid EHR Incentive Programs, CMS seeks to expand the meaningful use of certified EHR technology. Certified EHR technology used in a meaningful way is one piece of a broader health information technology infrastructure needed to reform the health care system and improve health care quality, efficiency, and patient safety. Highlights of the rule's provisions follow.

Stage 2 Timing

In the Stage 1 meaningful use regulations, CMS established an original timeline that would have required Medicare providers who first demonstrated meaningful use in 2011 to meet the Stage 2 criteria in 2013. The Stage 2 rule gives providers more time to meet Stage 2 criteria. A provider that attested to Stage 1 of meaningful use in 2011 would attest to Stage 2 in 2014, instead of in 2013. Therefore, providers are not required to meet Stage 2 meaningful use before 2014. The table below illustrates the progression of meaningful use stages from the first year a Medicare provider begins participation in the program.

Clinical Quality Measures (CQMs)

Measure Sets and Reporting

The rule finalized that:

- EPs must report on 9 out of 64 total clinical quality measures (CQMs)
- Eligible hospitals and CAHs must report on 16 out of 29 total CQMs

In addition, all providers must select CQMs from at least 3 of the 6 key health care policy domains from the Department of Health and Human Services' National Quality Strategy:

- Patient and Family Engagement
- Patient Safety
- Care Coordination
- Population and Public Health
- Efficient Use of Healthcare Resources
- Clinical Processes/Effectiveness

The Office of the National Coordinator for Health Information Technology

ONC Fact Sheet:
2014 Edition Standards & Certification Criteria (S&CC) Final Rule

Summary

The 2014 Edition S&CC final rule completes the Office of the National Coordinator for Health IT's (ONC) second full rulemaking cycle to adopt standards, implementation specifications, and certification criteria for EHR technology. This final rule complements the newly released Centers for Medicare & Medicaid Services (CMS) final rule which establishes Stage 2 of the Medicare and Medicaid Electronic Health Record (EHR) Incentive Programs, updates Stage 1, and includes other program modifications.

The 2014 Edition S&CC final rule reflects ONC's commitment to reduce regulatory burden; promote patient safety and patient engagement; enhance EHR technology's interoperability, electronic health information exchange capacity, public health reporting, and security; enable clinical quality measure data capture, calculation, and electronic submission to CMS or States; and introduce greater transparency and efficiency to the certification process.

NATIONAL LIBRARY OF MEDICINE

NLM'S LONG RANGE PLAN 2006-16

Goal 3. Integrated Biomedical, Clinical, and Public Health Information Systems that Promote Scientific Discovery and Speed the Translation of Research into Practice

Continue/enhance standards work in response to U.S. government priorities and feedback from "real" use in electronic health records

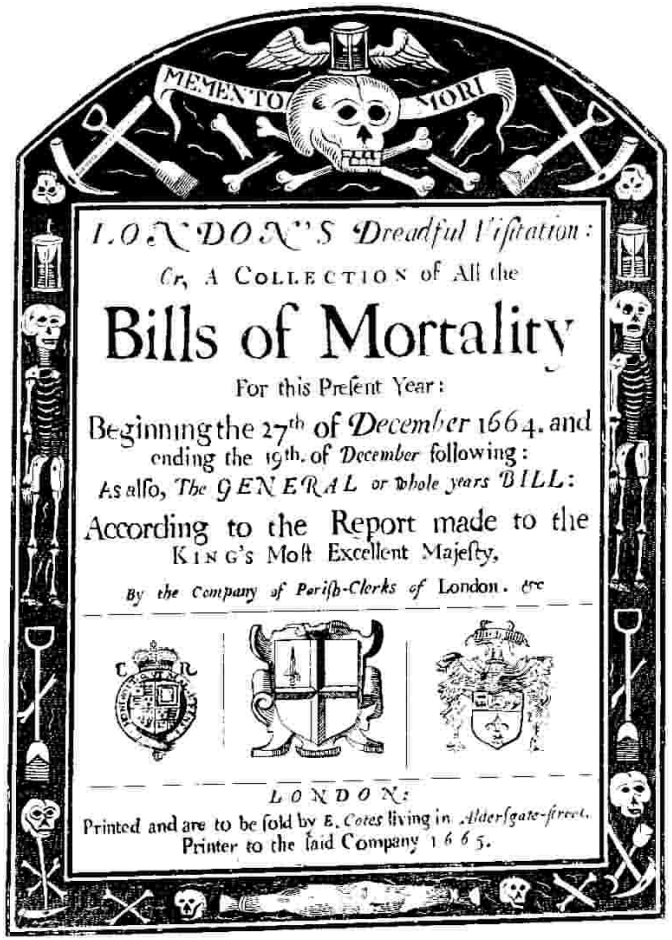
- e.g., Unified Medical Language System (UMLS), key clinical terminologies (SNOMED CT, LOINC, RxNorm)

NATIONAL LIBRARY OF MEDICINE

STANDARD VOCABULARIES

(Biomedical terminologies and ontologies)

London Bills of Mortality



Many biomedical terminologies

- Diagnoses / Diseases / Conditions
  - International classification of diseases (ICD)
  - SNOMED CT
- Procedures
  - Current Procedural terminology (CPT)
  - ICD10-PCS
  - SNOMED CT
- Drugs
  - RxNorm
- Laboratory tests
  - LOINC

Standard vocabularies for Meaningful Use

- Diagnoses / Diseases / Conditions
  - International classification of diseases (ICD)
  - SNOMED CT
- Procedures
  - Current Procedural terminology (CPT)
  - ICD10-PCS
  - SNOMED CT
- Drugs
  - RxNorm
- Laboratory tests
  - LOINC

SNOMED Clinical Terms

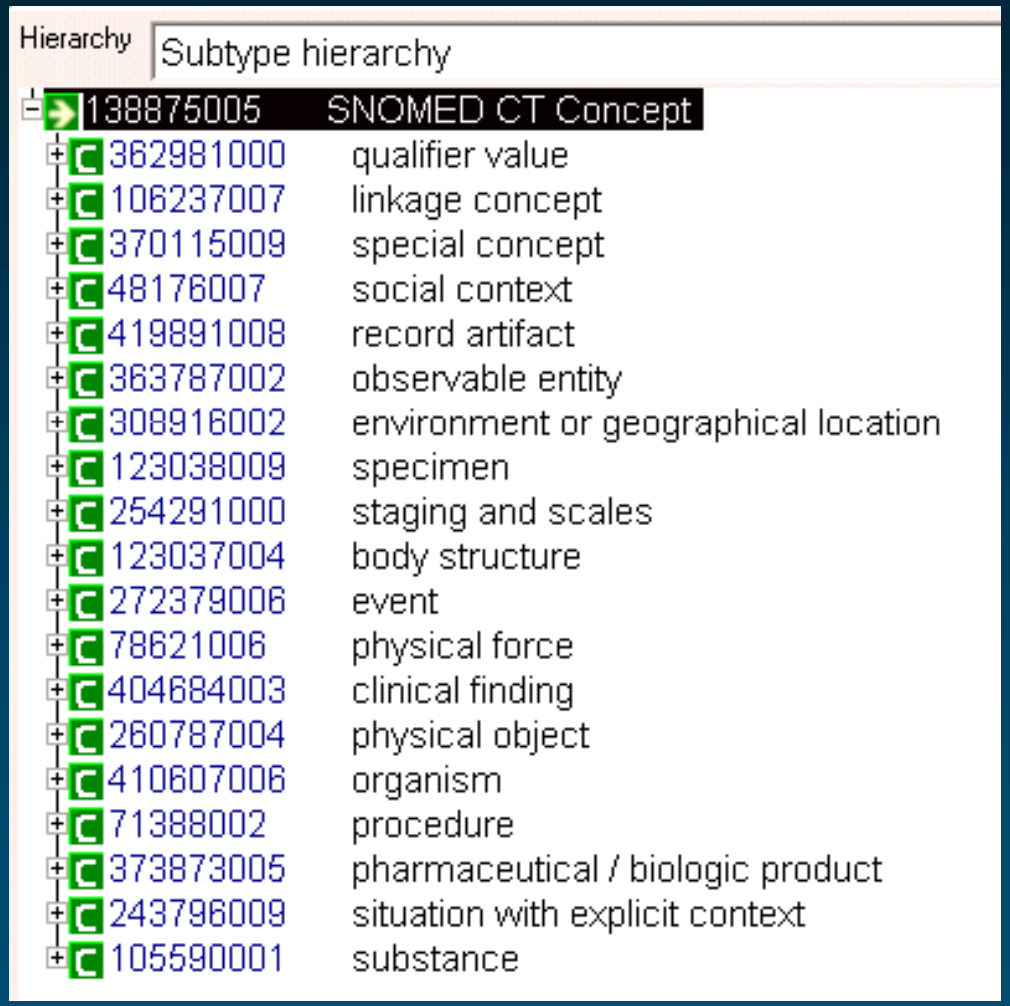
SNOMED CT Characteristics (1)

- Current version: January 31, 2013 (2 annual releases)
- Type: Reference terminology / ontology
- Domain: Clinical medicine
- Developer: IHTSDO
- Funding: IHTSDO
- Availability
  - Publicly available: Yes* (in member countries)
  - Repositories: UMLS
- URL: http://www.ihtsdo.org/

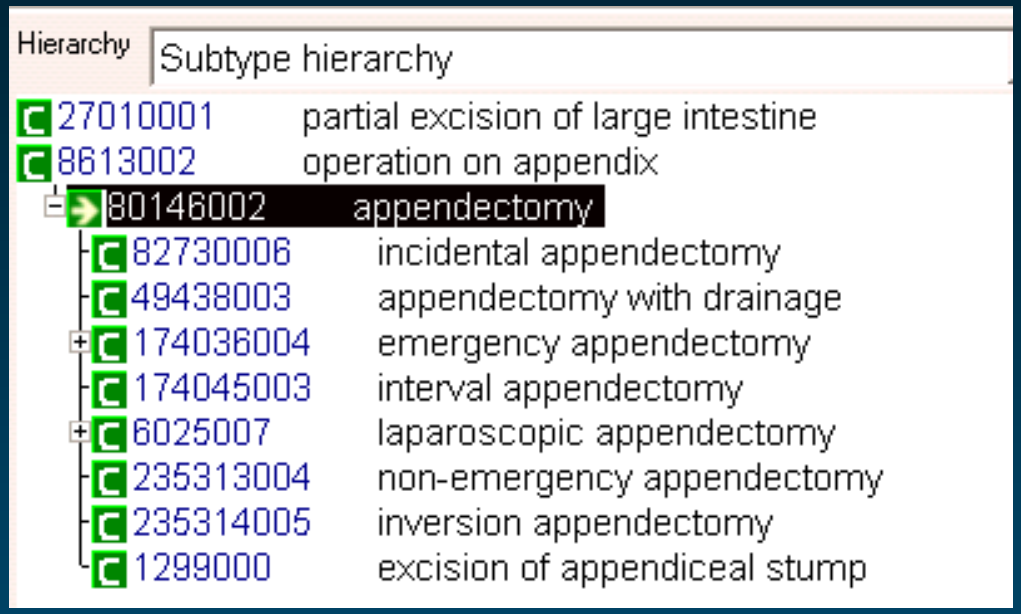
SNOMED CT Characteristics (2)

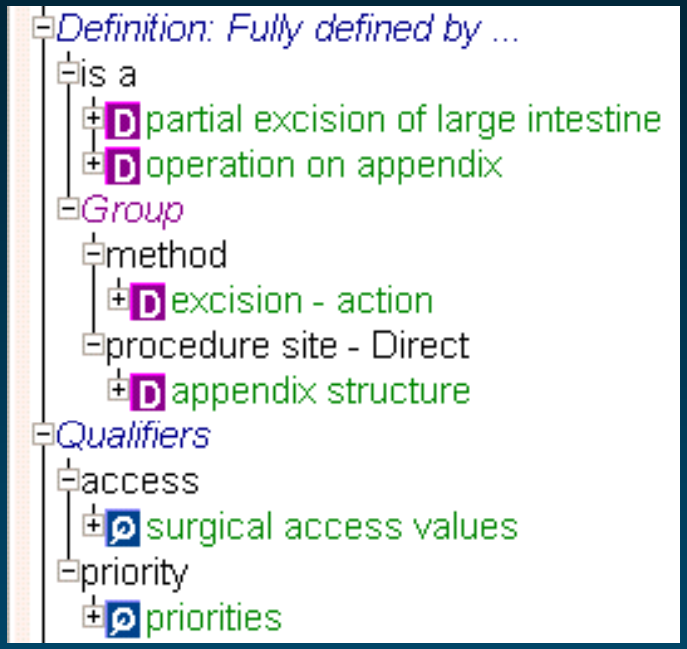
- Number of
  - Concepts: ~300,000 active concepts (Jan. 31, 2013)
  - Terms: ~1.1M active “descriptions”
- Major organizing principles:
  - Utility for clinical medicine (e.g., assertional + definitional knowledge)
  - Model of meaning (incomplete)
  - Rich set of associative relationships
  - Small proportion of defined concepts (many primitives)
- Formalism: Description logics (EL++)

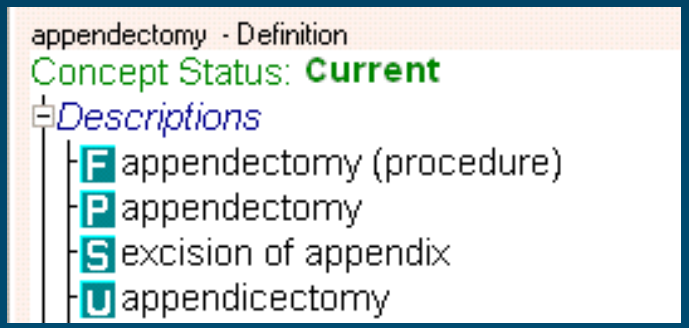
SNOMED CT Top level



SNOMED CT Example







RxNorm

RxNorm Characteristics (1)

- Current version: April 1, 2012 (monthly releases)
- Type: Controlled terminology
- Domain: Drug names
- Developer: NLM
- Funding: NLM
- Availability
  - Publicly available: Yes*
  - Repositories: UMLS
- URL: http://www.nlm.nih.gov/research/umls/rxnorm/

RxNorm Characteristics (2)

- Number of
  - Concepts: 213,500 drug entities (April 2013)
  - Terms: ~1.3 term per concept
- Major organizing principles:
  - Generic vs. brand
  - Combinations of Ingredient / Form / Dose
  - No hierarchical structure
  - Links to all major US drug information sources
  - No clinical information
- Formalism: UMLS RRF format

RxNorm Normalized form

Strength **Ingredient** **Dose form**

4mg/ml Fluoxetine Oral Solution

Strength **Ingredient**

Semantic clinical drug component

Ingredient **Dose form**

Semantic clinical drug form

Strength **Ingredient** **Dose form**

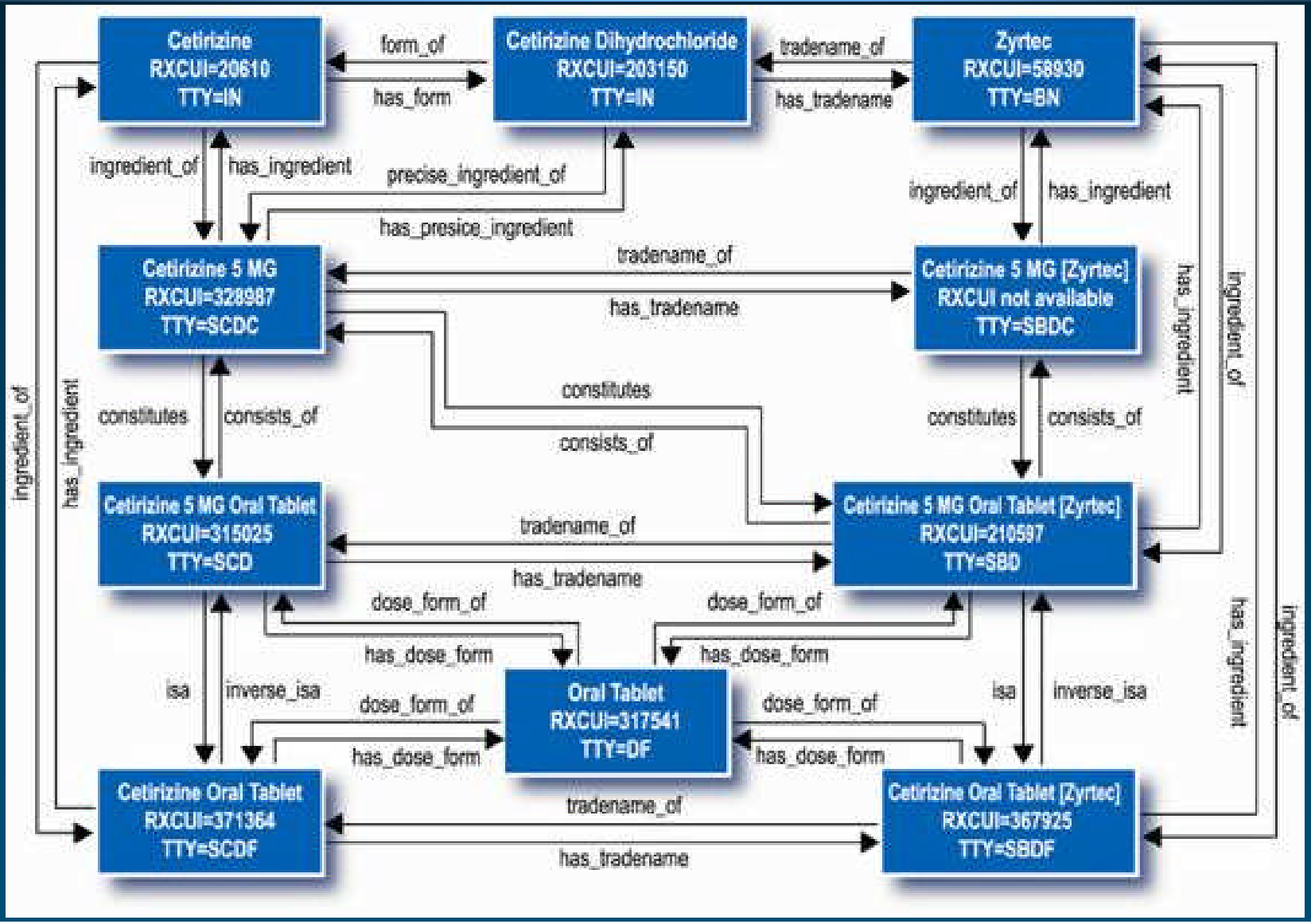
Semantic clinical drug

Rx Norm Generic *vs.* Brand

- Generic
  - Ingredient (IN)
  - Clinical drug form (SCDF)
  - Clinical drug component (SCDC)
  - Clinical drug (SCD)
- Brand
  - Brand name (BN)
  - Branded drug form (SBDF)
  - Branded drug component (SBDC)
  - Branded drug (SBD)

tradename_of

RxNorm Relations among drug entities



Logical Observation Identifiers, Names and Codes (LOINC)

LOINC Characteristics (1)

- Current version: 2.42 (Dec. 2012)
- Type: Controlled terminology*
- Domain: Laboratory and clinical observations
- Developer: Regenstrief Institute
- Funding: NLM
- Availability
  - Publicly available: Yes
  - Repositories: UMLS
- URL: www.regenstrief.org/loinc/loinc.htm

LOINC Characteristics (2)

- Number of
  - Concepts: ~70k active codes (2.42)
    (2 annual releases)
  - Terms: n/a*
- Major organizing principles:
  - No hierarchical structure among the main codes
  - 6 axes
    - Component (analyte [+ challenge] [+ adjustments])
    - Property
    - Timing
    - System
    - Scale
    - [Method]
- Formalism: “DL-like”

LOINC Example

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

Axis	Value
Component	Sodium
Property	SCnc – Substance Concen-tration (per volume)
Timing	Pt – Point in time (Random)
System	Ser/Plas – Serum or Plasma
Scale	Qn – Quantitative
Method	--

QUALITY ASSURANCE OF STANDARD VOCABULARIES

Analytical framework for QA research

- Special issue of JBI on “Auditing terminologies”
- Zhu et al. JBI 2009 review article
- Analytical framework
  - What is analyzed
  - Which source of knowledge
  - Which method

Zhu, X., J.W. Fan, D.M. Baorto, C. Weng, and J.J. Cimino, *A review of auditing methods applied to the content of controlled biomedical terminologies. J Biomed Inform, 2009.* ***42(3): p. 413-25.***

What is analyzed

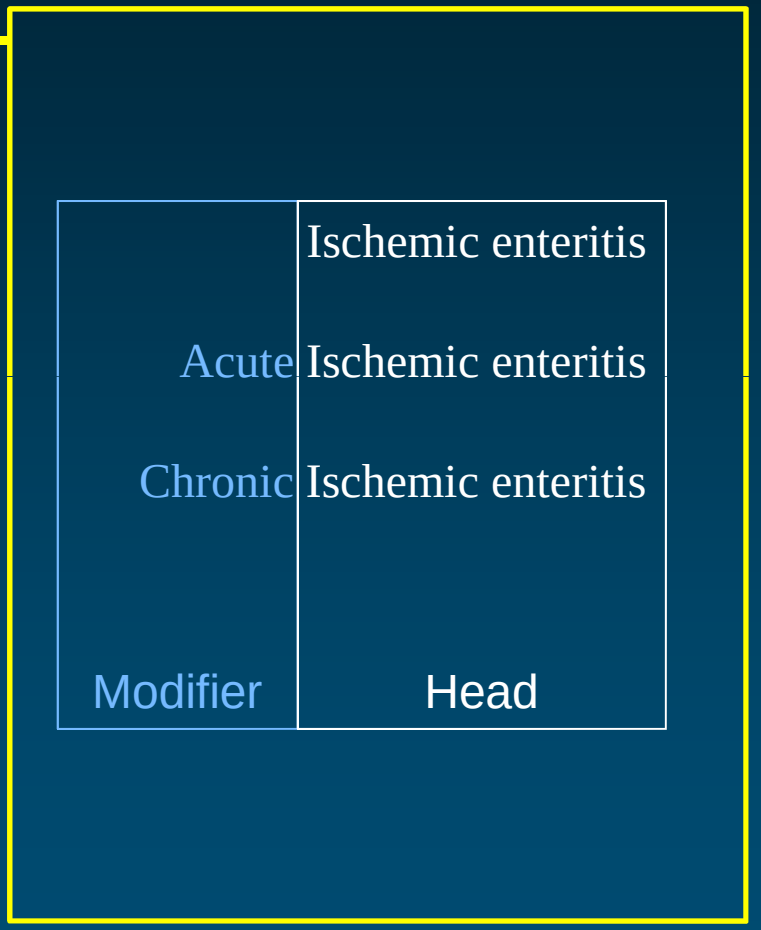
- Term/concept
  - Coverage (missing terms/concepts)
  - Wrong synonymy relation
  - Redundant concepts
- Relation
  - Missing relations
  - Inaccurate relations
- Categorization
  - Wrong categorization

Which source of knowledge

- Intrinsic – the terminology itself
  - Terms/Concepts
  - Relations
  - Categorization
- Extrinsic – external resources
  - Corpus
    - Text corpus – identify terms in text
    - Annotation corpus – identify relations from co-occurring terms
  - Mapping

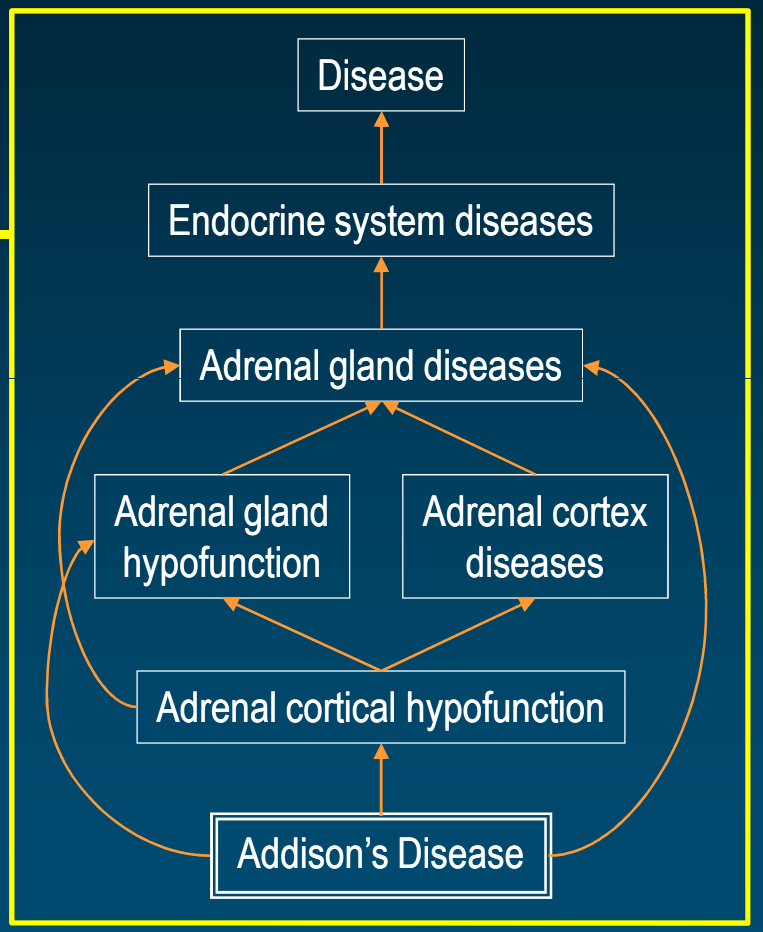
Which method Main categories

- Lexical
  - Properties of the term
- Structural
  - Properties of the organizational structure (relations)
- Semantic
  - Semantic properties of the concept (semantic type)
- Statistical
  - Associations among entities



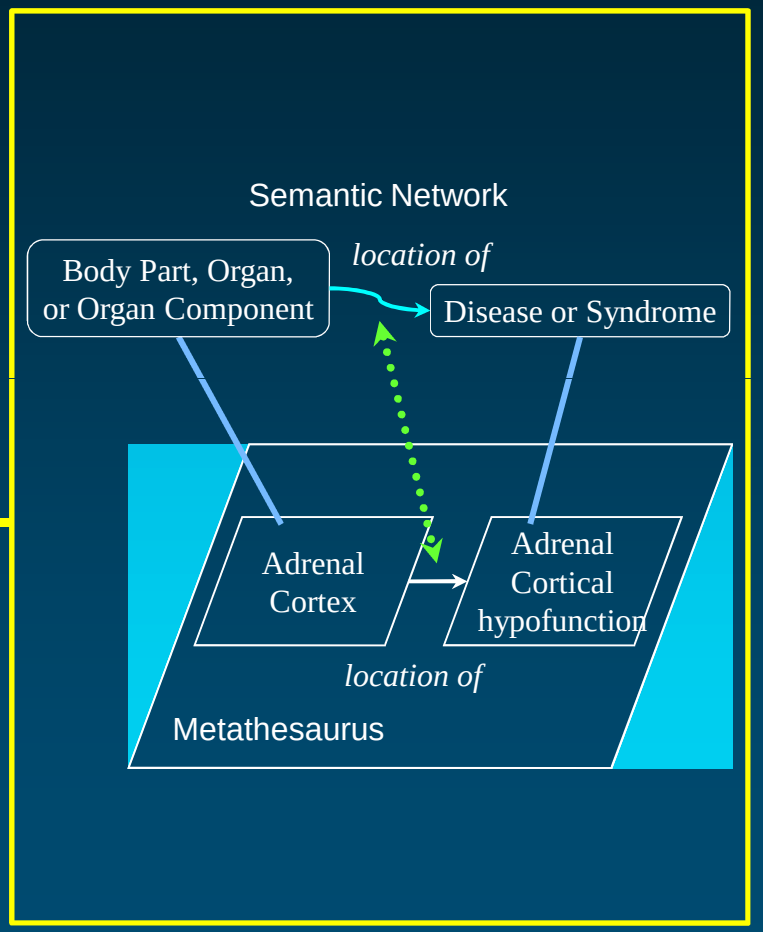
Which method Main categories

- Lexical
  - Properties of the term
- Structural
  - Properties of the organizational structure (relations)
- Semantic
  - Semantic properties of the concept (semantic type)
- Statistical
  - Associations among entities



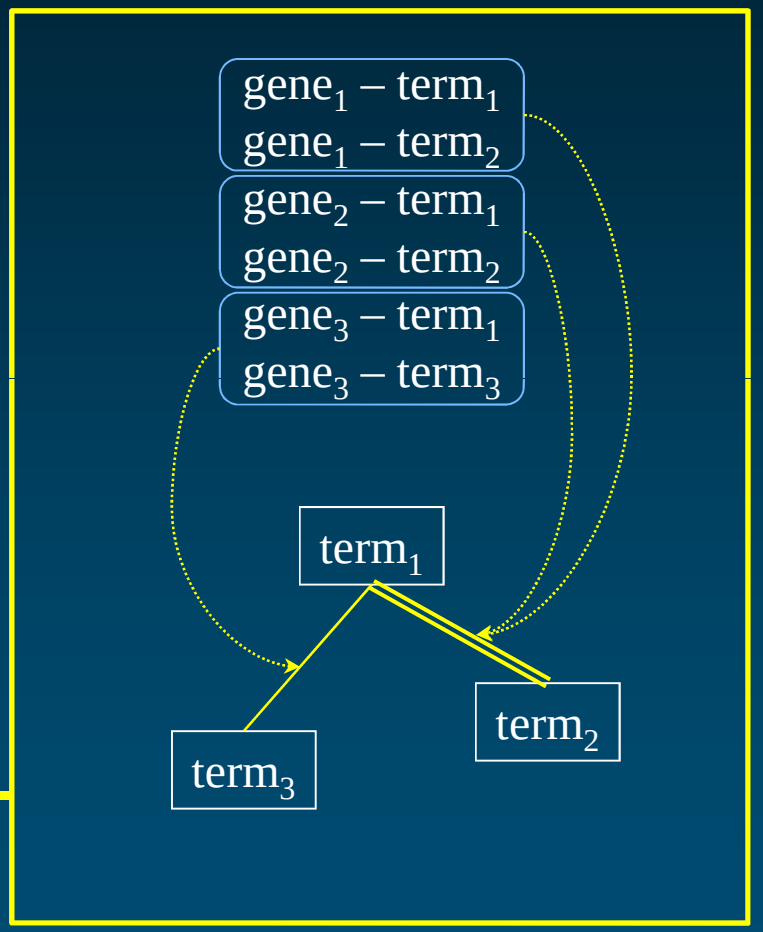
Which method Main categories

- Lexical
  - Properties of the term
- Structural
  - Properties of the organizational structure (relations)
- Semantic
  - Semantic properties of the concept (semantic type)
- Statistical
  - Associations among entities



Which method Main categories

- Lexical
  - Properties of the term
- Structural
  - Properties of the organizational structure (relations)
- Semantic
  - Semantic properties of the concept (semantic type)
- Statistical
  - Associations among entities



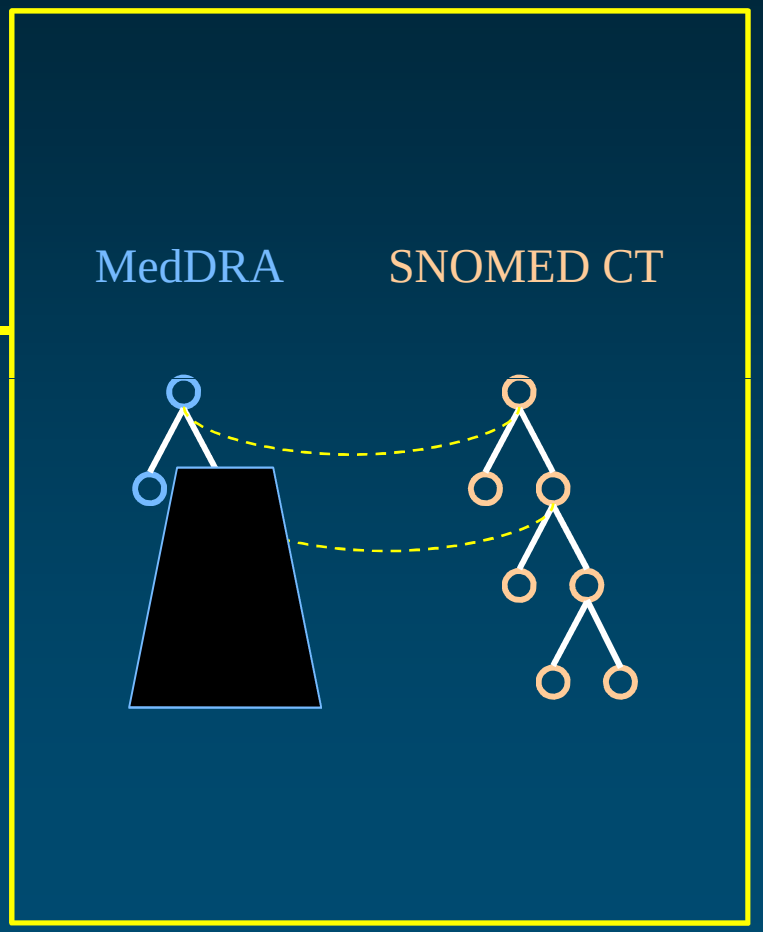
Which method Additional methods

- Compliance with ontological principles
  - Operational definitions
- Comparative
  - Comparisons between ontologies (mapping)
- Transformative
  - Representation formalism
- Use in an application

"Each concept, except for the root, must have (at least) one parent concept"

Which method Additional methods

- Compliance with ontological principles
  - Operational definitions
- Comparative
  - Comparisons between ontologies (mapping)
- Transformative
  - Representation formalism
- Use in an application



Which method Additional methods

- Co[illegible]gical

Heart in OWL DL

```
<owl:Class rdf:ID="Heart">
 <owl:equivalentClass>
   <owl:Class>
     <owl:intersectionOf rdf:parseType="Collection">
     <owl:Class rdf:about=
      "#Organ_with_cavitated_organ_parts"/>
     <owl:Restriction>
      <owl:onProperty
       rdf:resource="#constitutional_part" />
      <owl:someValuesFrom
        rdf:resource="#Wall_of_heart" />
     </owl:Restriction>
     …
 </owl:equivalentClass>
 <rdfs:subClassOf>
   <owl:Restriction>
    <owl:onProperty rdf:resource="#bounded_by"/>
      <owl:someValuesFrom
       rdf:resource="#Surface_of_heart"/>
   </owl:Restriction>
 </rdfs:subClassOf>
 <rdfs:subClassOf>
    <owl:Restriction>
     <owl:onProperty rdf:resource="#arterial_supply"/>
      <owl:someValuesFrom
       rdf:resource="#Right_coronary_artery" />
    </owl:Restriction>
  …
</owl:Class>
```

Heart in Protégé frames

Metaclass

```
(defclass Heart
  (is-a Organ_with_cavitated_organ_parts)
                    …
)
```

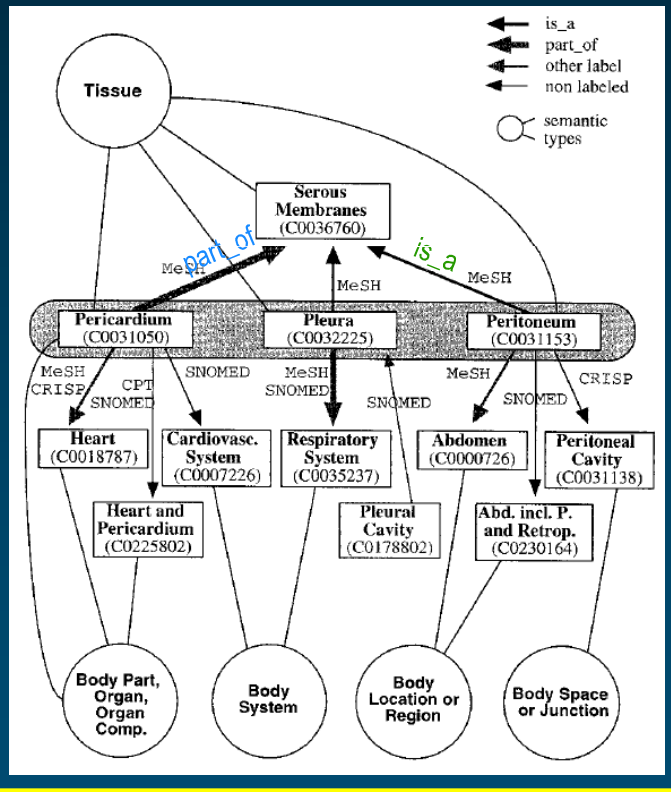
Class (Instance of Metaclass)

```
([Heart]
  of Organ_with_cavitated_organ_parts
  (constitutional_part
          Wall_of_heart
          Cavity_of_left_atrium
          Cavity_of_right_ventricle
          Cavity_of_left_ventricle
          Right_coronary_artery
          Left_coronary_artery
            …
  (bounded_by
          Surface_of_heart)
  (arterial_supply
          Right_coronary_artery
          Left_coronary_artery)
            …
)
```

Which method Additional methods

- Compliance with ontological principles
  - Operational definitions
- Comparative
  - Comparisons between ontologies (mapping)
- Transformative
  - Representation formalism
- Use in an application

MAOUSSC: Using UMLS for the description of medical procedures



Identifying errors in SNOMED CT

Rector AL, Brandt S, Schneider T, *Getting the foot out of the pelvis: modeling problems affecting use of SNOMED CT hierarchies in practical applications. J Am Med Inform Assoc, 2011. 18(4): p. 432-440.*

Alan Rector's motivation

- Knowledge source for clinical applications (industrial collaboration)
  - Retrieval of clinical records for a given diagnosis
  - Decision support
  - Eligibility for clinical trials
- Development of ICD 11 (WHO Health Information Modelling TAG)
  - Ontological component for ICD 11
  - ICD hierarchies derived from SNOMED CT

Investigation methods

- Technology-assisted review by domain experts
  - With focus on the CORE problem list (clinical relevance)
- Methods
  - Leveraging description logics (stated version converted to OWL, module extraction, OWL classifiers, scripting against OWL API)
  - Lexically suggest-logically refine (association between words from labels and roles)
  - Navigation (up and down the hierarchies for missing/extraneous ancestors/descendants)

Findings – Qualitative (1)

- Errors of omission
  - Primitive concepts
    - Ischemic heart disease: *due to* ischemia, but not defined
  - Missing axioms
    - Morphology
      - acute endocarditis (disorder): not *clinical course* = acute; *morphology*: **acute** inflammation
    - Other
      - Myocardial infarction: not *due to* ischemia

Findings – Qualitative (2)

- Errors of commission
  - Wrong
    - Diabetes mellitus *isa* Disease of the exocrine pancreas
      (true of type I, not type II)
    - Subdural hemorrhage: *finding site* Subdural space structure
      (intracranial is implied by usage, but not formally represented)
  - Wrong (but only indirectly)
    - Hypertension *isa* disorder of soft issue
      - Hypertension: *finding site* Systemic arterial structure (→ soft tissue)
        (not clinically thought of as a disorder of soft tissues)
    - Neoplasm of cranial nerve: *isa* Neoplasm of cranial nerve: *isa* Neuropathy
      (neuropathy denotes functional disorder of nerve)
  - Anatomy (branches inherit from root)
    - Structure of right popliteal artery *isa* […] *isa* Structure of pelvic region
      ("getting the knee out of the pelvis")

Consequences

Incorrect inferences

- By omission
  - Search on Ischemic heart disease fails to retrieve Myocardial infarction
- By commission
  - Search on Disorder of pancreas retrieves all cases of diabetes mellitus, including type II
- Missing equivalences from post-coordination
  - acute endocarditis (disorder):
    - Pre-coord: *morphology*: acute inflammation
    - Post-coord: *clinical course* = acute; *morphology*: inflammation

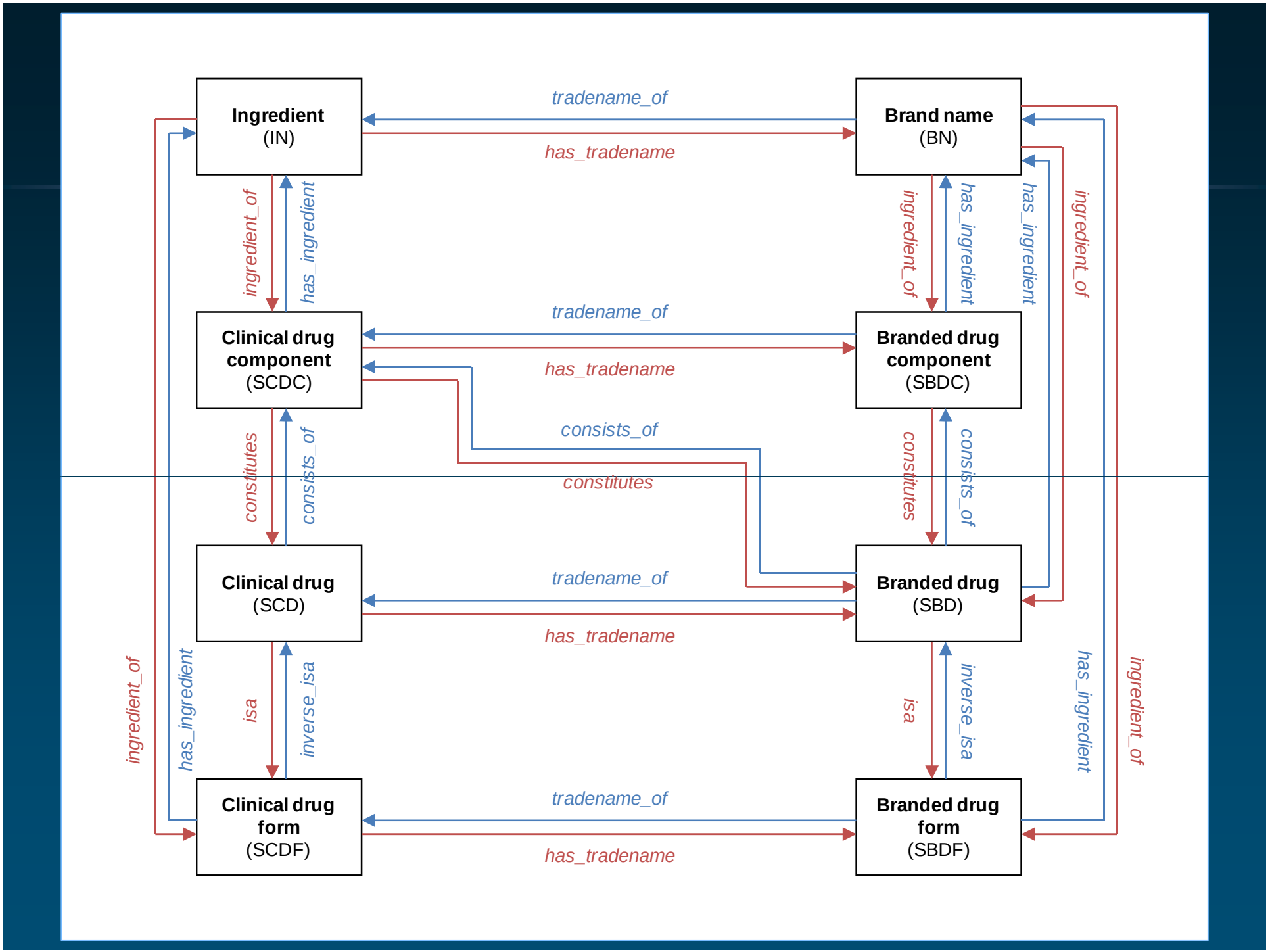
Identifying errors in RxNorm

Bodenreider, O. and L.B. Peters, *A graph-based approach to auditing RxNorm. Journal of Biomedical Informatics, 2009. 42(3): p. 558-570.*

Motivation

- Large terminology
- Relies heavily on human editors
- High quality

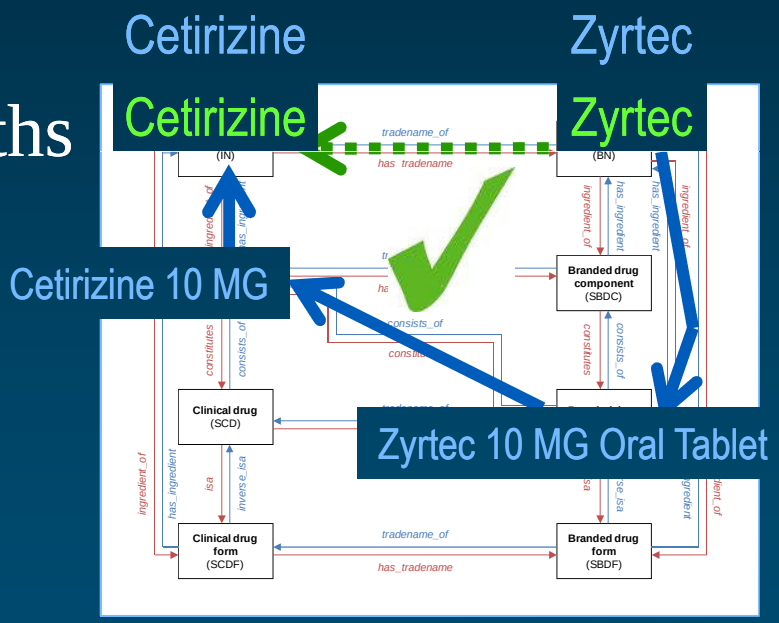
- Systematic evaluation
- Exploiting the graph structure


Ingredient
(IN)
Brand name
(BN)
Clinical drug component
(SCDC)
Branded drug component
(SBDC)
Clinical drug
(SCD)
Branded drug
(SBD)
Clinical drug form
(SCDF)
Branded drug form
(SBDF)
tradename_of
has_tradename
ingredient_of
has_ingredient
constitutes
consists_of
isa
inverse_isa

Methods

- Normalize multi-ingredient drugs
- Define “meaningful” paths between 2 nodes

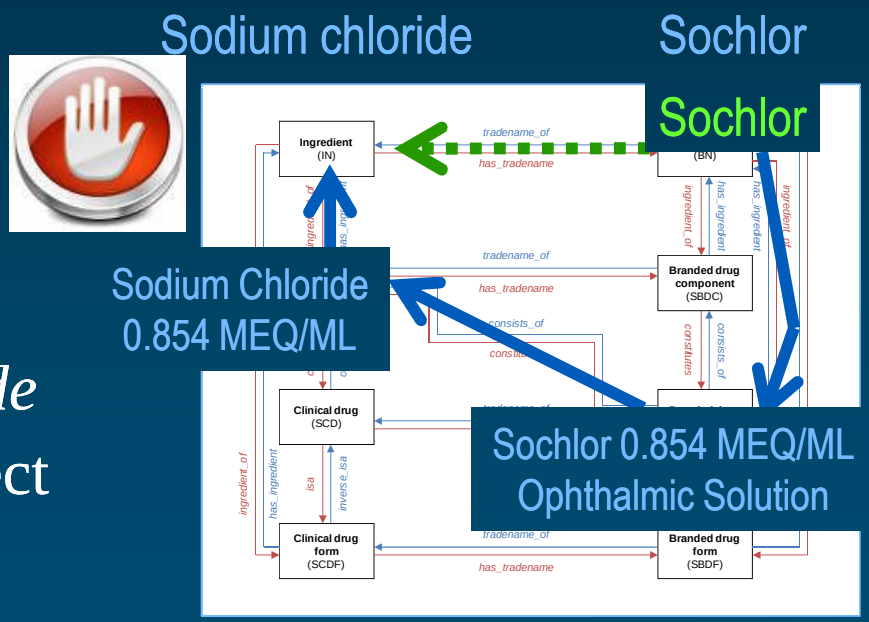
- Instantiate all meaningful paths
- Compare alternate paths
  - Alternate (meaningful) paths are expected to be functionally equivalent



Results

- 348 inconsistencies identified (April 2008)
- Reported to the RxNorm team
- 215 (62%) fixed (January 2009)

- Example (fixed)
  - missing link
    Sochlor → Sodium chloride
  - Brand name without a direct relation with an ingredient



VALUE SETS
IN
CLINICAL QUALITY MEASURES

CLINICAL QUALITY MEASURES (CQMs)

Tools that help measure and track the quality of healthcare services provided by eligible professionals, eligible hospitals and critical access hospitals within our health care system

CQMs measure many aspects of patient care including: health outcomes, clinical processes, patient safety, efficient use of healthcare resources, care coordination, patient engagements, population and public health, and clinical guidelines

[cms.gov]

93 CLINICAL QUALITY MEASURES
in 2014 Meaningful Use criteria

64

for

ELIGIBLE PROVIDERS

(NEED TO REPORT ON 9)

29

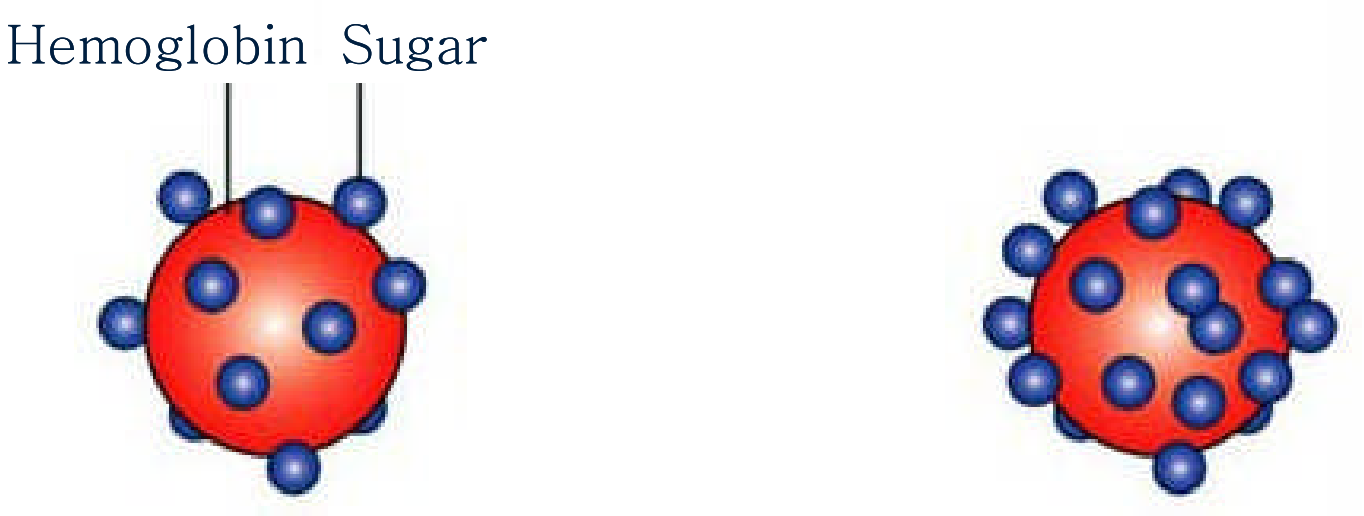
for

ELIGIBLE HOSPITALS

(NEED TO REPORT ON 16)

CLINICAL QUALITY MEASURE (example)

Hemoglobin A1c Test for Pediatric Patients



Normal glucose levels in blood
Low HbA1c concentration

High glucose levels in blood
High HbA1c concentration

CLINICAL RECOMMENDATIONS

1. American Association of Clinical Endocrinologists (2002): Recommends that a glycosylated hemoglobin be performed during an initial assessment and during follow-up assessments, which should occur at no longer than three-month intervals.

2. American Diabetes Association (2006): Recommends obtaining a glycosylated hemoglobin during an initial assessment and then routinely as part of continuing care. In the absence of well-controlled studies that suggest a definite testing protocol, expert opinion recommends glycosylated hemoglobin be obtained at least twice a year in patients who are meeting treatment goals and who have stable glycemic control and more frequently (quarterly assessment) in patients whose therapy was changed or who are not meeting glycemic goals.

CLINICAL QUALITY MEASURE (simplified)

Hemoglobin A1c Test for Pediatric Patients

$$= \frac{\text{\# diabetic patients [age 5-17] } \textit{tested for HbA1c}}{\text{\# diabetic patients [age 5-17]}}$$

CLINICAL QUALITY MEASURE (details)

Hemoglobin A1c Test for Pediatric Patients

Tests for HbA1c

$$= \frac{\text{\# diabetic patients [age 5-17] } \textit{tested for HbA1c}}{\text{\# diabetic patients [age 5-17]}}$$

- Type 1 or Type 2 diabetes
- Excludes gestational diabetes

- Requires date of birth

CLINICAL QUALITY MEASURE (implementation)

Hemoglobin A1c Test for Pediatric Patients

List of LOINC codes

Tests for HbA1c

diabetic patients [age 5-17] *tested for HbA1c*

=

diabetic patients [age 5-17]

Data element

- Type 1 or Type 2 diabetes
- Excludes gestational diabetes

- Requires date of birth

List of SNOMED CT or ICD 10 codes

Anatomy of a Clinical Quality Measure

Population criteria

- **Initial Patient Population** =
  - AND: "Patient Characteristic Birthdate: birth date" >= 5 year(s) starts before start of "Measurement Period"
  - AND: "Patient Characteristic Birthdate: birth date" <= 17 year(s) starts before start of "Measurement Period"
  - AND: "Diagnosis, Active: Diabetes" starts before or during (MOST RECENT :"Occurrence A of Encounter, Performed: Diabetes Visit" during "Measurement Period")
  - AND: "Encounter, Performed: Diabetes Visit" >= 12 month(s) starts before start of "Occurrence A of Encounter, Performed: Diabetes Visit"
- **Denominator** =
  - AND: "Initial Patient Population"
- **Denominator Exclusions** =
  - AND NOT: "Occurrence A of Diagnosis, Active: Gestational Diabetes" ends before start of "Measurement Period"
  - AND: "Occurrence A of Diagnosis, Active: Gestational Diabetes" starts before or during "Measurement Period"
- **Numerator** =
  - AND: "Laboratory Test, Result: HbA1c Laboratory Test (result)" during "Measurement Period"
- **Denominator Exceptions** =
  - None

Data criteria (QDM Data Elements)

- "Diagnosis, Active: Diabetes" using "Diabetes Grouping Value Set (2.16.840.1.113883.3.464.1003.103.12.1001)"
- "Diagnosis, Active: Gestational Diabetes" using "Gestational Diabetes Grouping Value Set (2.16.840.1.113883.3.464.1003.103.12.1010)"
- "Encounter, Performed: Diabetes Visit" using "Diabetes Visit Grouping Value Set (2.16.840.1.113883.3.464.1003.103.12.1012)"
- [redacted]
- "Patient Characteristic Birthdate: birth date" using "birth date LOINC Value Set (2.16.840.1.113883.3.56[illegible]

Value set = List of LOINC codes for HbA1c tests

NATIONAL LIBRARY OF MEDICINE

ASSOCIATED VALUE SET

Metadata | Measure | Grouping

Name:
HbA1c Laboratory Test

Type:
Grouping

Note:

OID:
2.16.840.1.113883.3.464.1003.198.12.1013

Developer:
National Committee for Quality Assurance

Value Set Members Expansion 20121025

Expanded Code List

View | Toggle | Clear | Page 1 of 1 | 20 | View 1 - 3 of 3

Code	Descriptor	Code System	Version
17855-8	Hemoglobin A1c/Hemoglobin.total in Blood by calculation	LOINC	2.40
17856-6	Hemoglobin A1c/Hemoglobin.total in Blood by HPLC	LOINC	2.40
4548-4	Hemoglobin A1c/Hemoglobin.total in Blood	LOINC	2.40

NATIONAL LIBRARY OF MEDICINE

Meaningful Use Criteria - 2014

93
CQMs

Developed by some 20 measure developers

3,011
value sets

1,520 unique

199,521
codes

83,723 unique

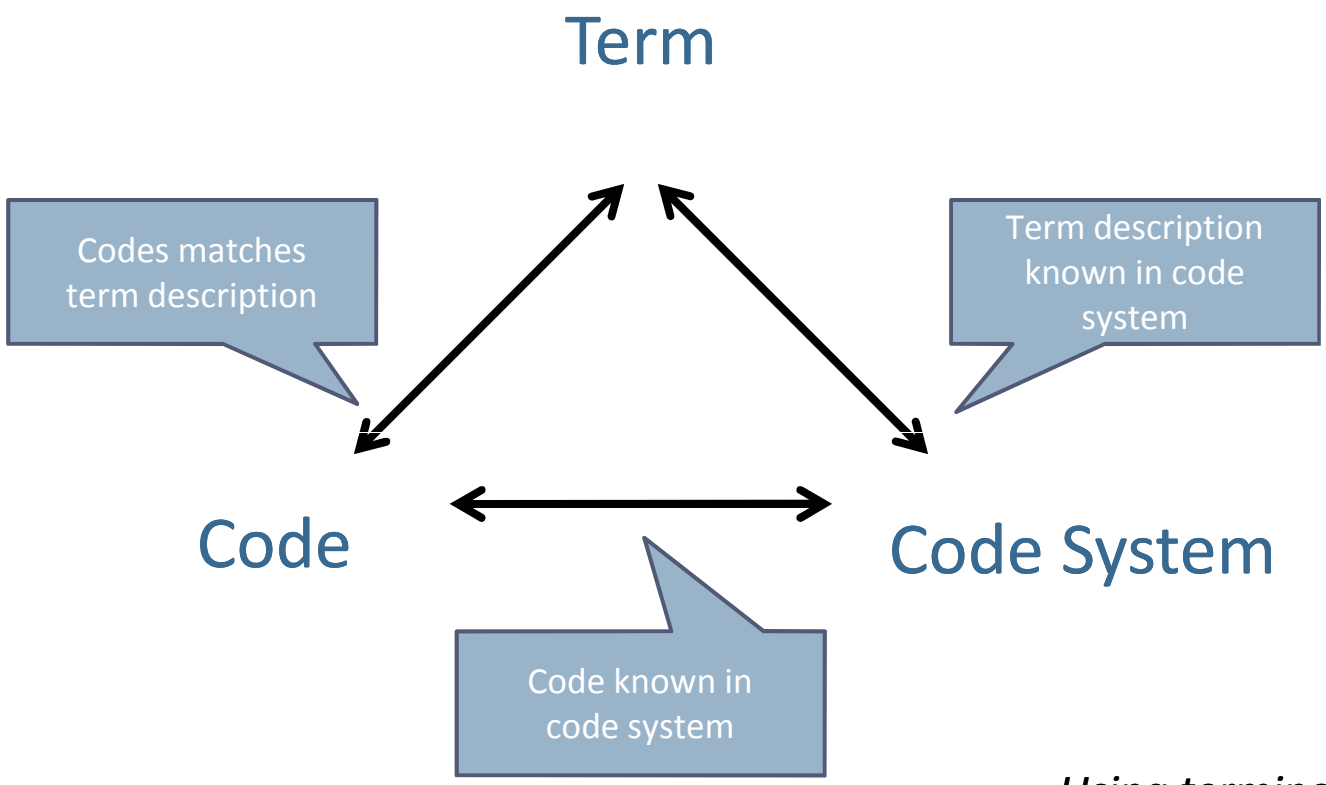
CURATING VALUE SETS

The NLM Value Set Authority Center

OBJECTIVES OF CURATION

- Ensure referential integrity
  - All codes in a VS are valid codes in the corresponding code system
  - Update VSs when the code systems are updated (no "stale" codes)
- Avoid duplication
  - Find value sets having similar members
- Ensure correctness and completeness
  - Compare intensional and extensional definitions

Code Validation Using Triangulation



Using terminology services to compare codes to reference terminologies

TYPES OF ERROR FOUND IN THE CODES

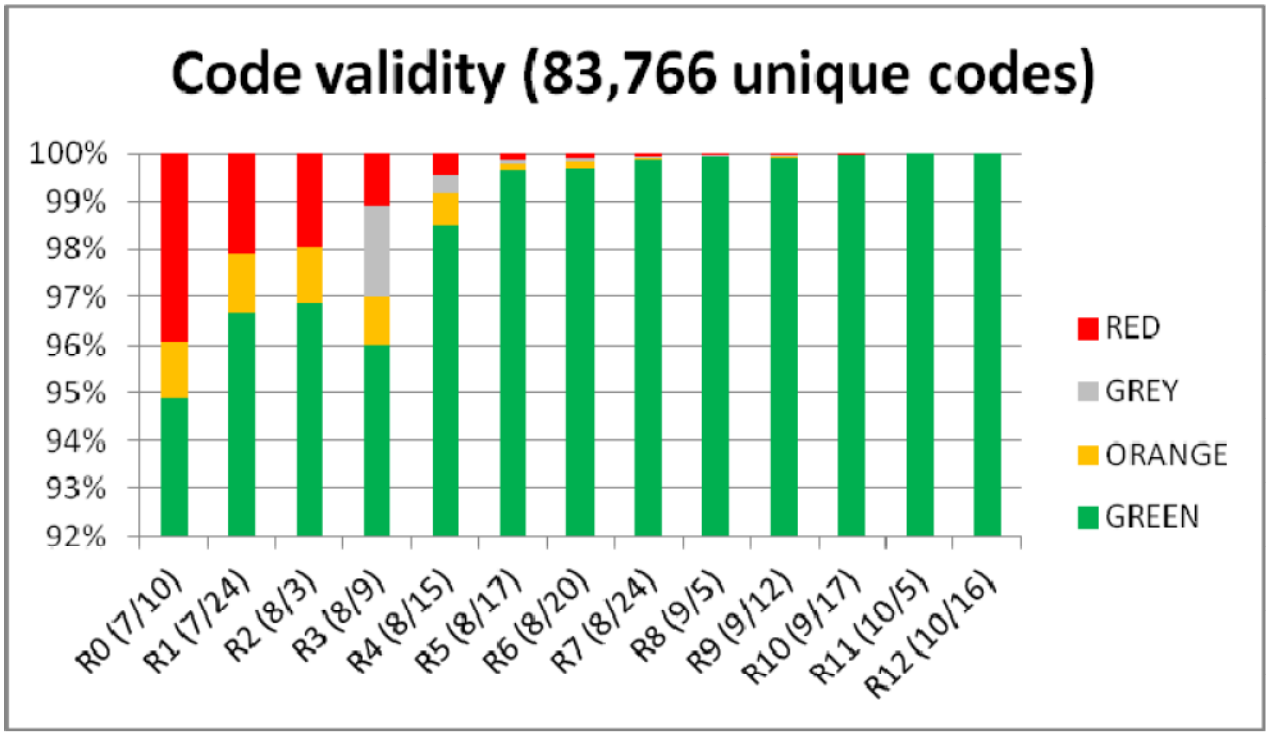
- Obsolete codes
  - Remap to the current code
- Typo / formatting issue in the code
  - Reformat
- Wrong code system listed
  - Fix code system
- Code/description mismatch
  - Small mismatch: Assign preferred term
  - Large mismatch: Send back to developers

Impact on Clinical Quality Measures

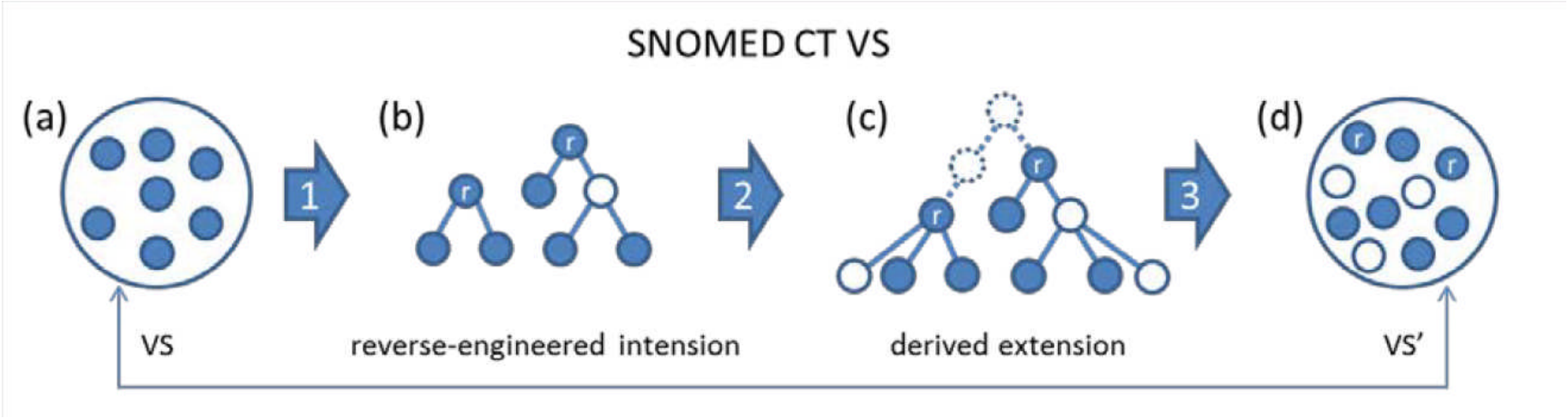
Iterative Analysis

- 13 rounds over 4 months
- Reports provided to measure developers
- Orange codes fixed automatically by NLM
- Red codes fixed by measure developers (and rechecked)

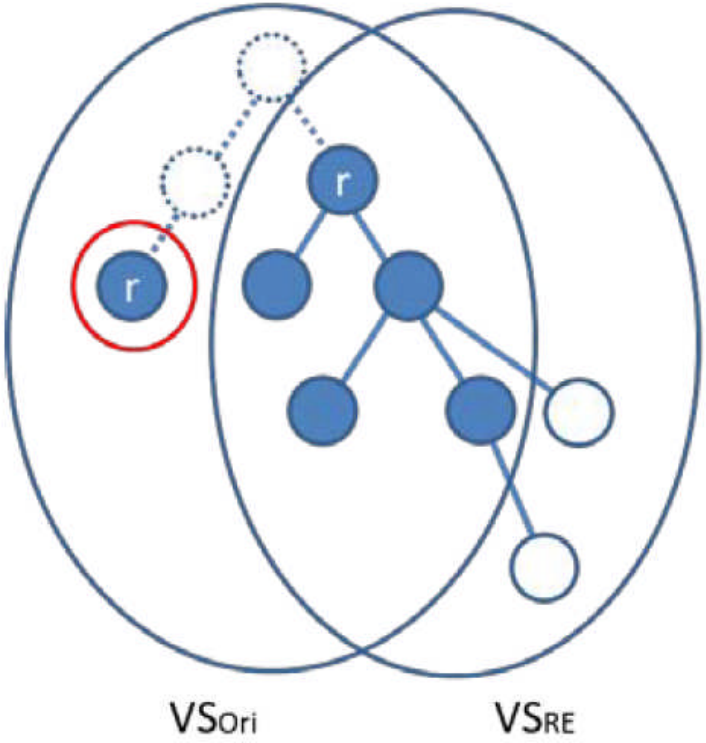
- ~4000 errors
- Affecting 70% of the value sets
- And 100% of the measures
- All fixed by October 2012



REVERSE-ENGINEERING OF THE INTENSION



QUALITY METRICS FOR VALUE SETS

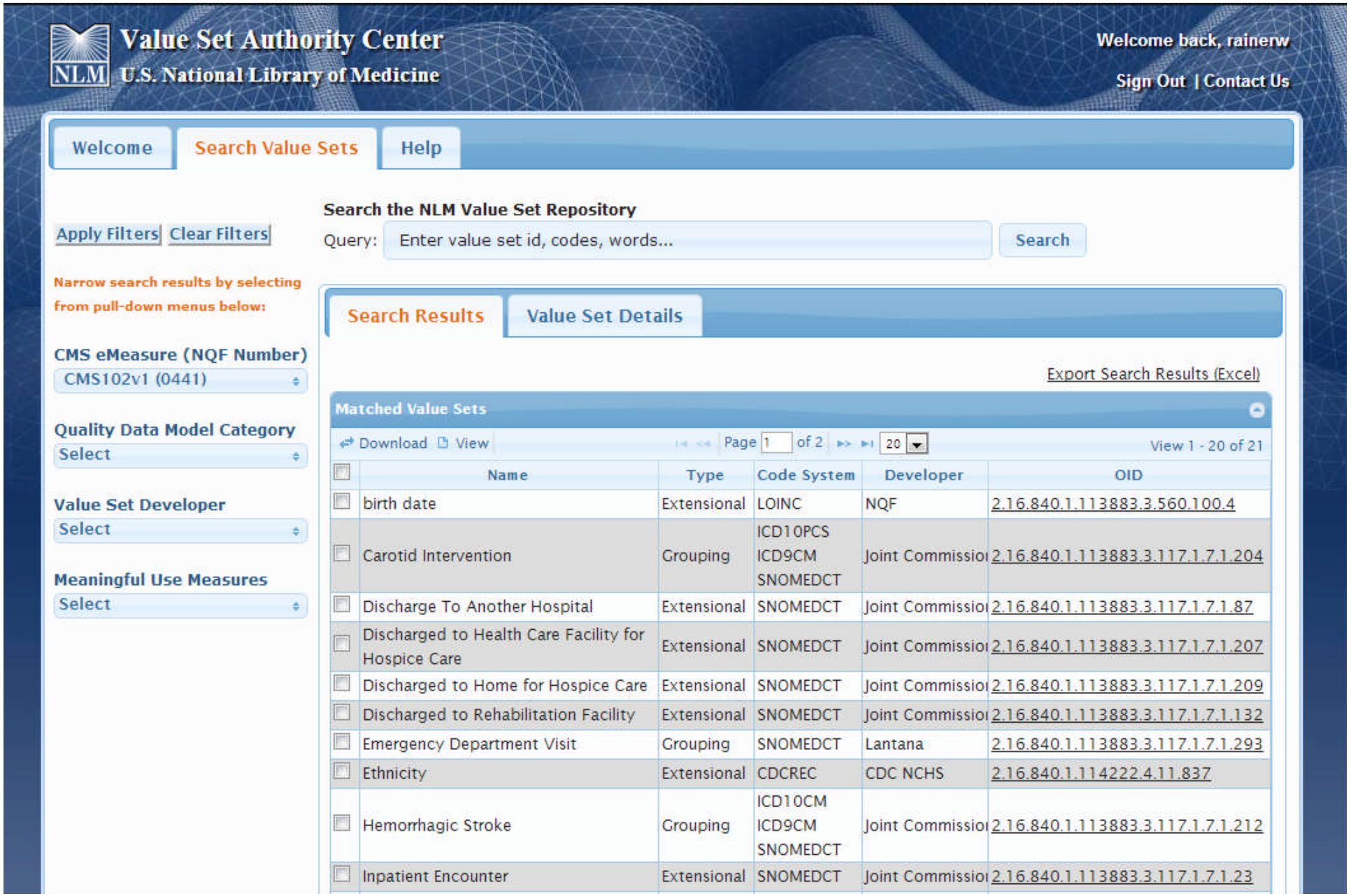


Completeness

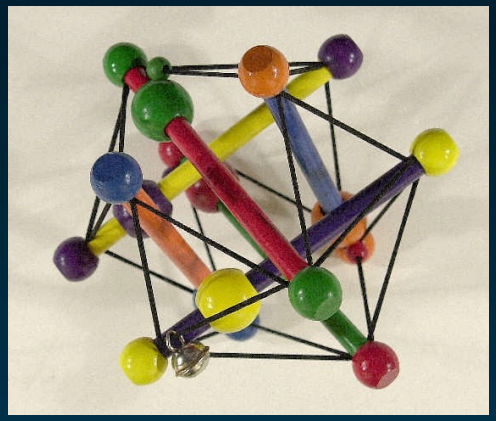
Compl (VS_{Ori}) = $|VS_{Ori} \cap VS_{RE}| \,/\, |VS_{RE}|$.

Correctness

Correct (VS_{Ori}) = $|VS_{Ori} \cap VS_{RE}^*| \,/\, |VS_{Ori}|$,
where $VS_{RE}^* \subseteq VS_{RE} \setminus \{\text{singleton nodes}\}$.



https://vsac.nlm.nih.gov/



Medical Ontology Research

Contact: olivier@nlm.nih.gov
Web: http://mor.nlm.nih.gov

NATIONAL LIBRARY OF MEDICINE

Olivier Bodenreider

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