Use of SNOMED in HL7 Messaging

PHIN Conference – August 2008

James T. Case MS, DVM, PhD
Professor, Clinical Diagnostic Informatics
California Animal Health and Food Safety Laboratory
University of California, Davis

Setting the stage

'When I use a word,' Humpty Dumpty said, in a rather scornful tone, ' it means just what I choose it to mean, neither more nor less.'

'The question is,' said Alice, 'whether you can make words mean so many different things.'

'The question is,' said Humpty Dumpty, 'which is to be master - that's all.'

Goals of Health Information Standards

- **Interoperability – the ability to exchange information between organizations**
- **Comparability – the ability to ascertain the equivalence of data from different sources**
- **Data quality – the measurement of completeness, accuracy and precision**

Missed Opportunities: Public Health, without (information) standards

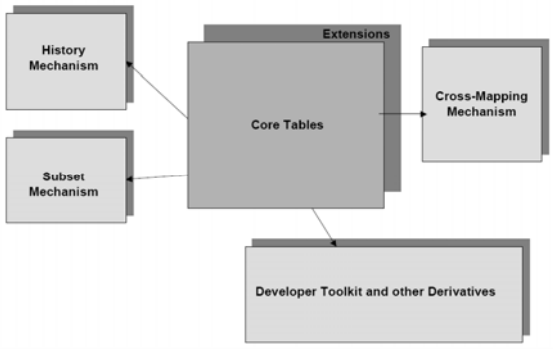
- We cannot transmit information **reliably**.
- We cannot **reliably** aggregate, collate and analyze data.
- We **cannot integrate** external sources of information into patient care (e.g. Structured Product Label)
- We **throw away information** that might lead to better patient care, increased profitability, improved biosecurity, etc…
- We cannot **reliably** contribute to pools of biomedical data (i.e., Adverse reaction reporting; clinical trials; evidence-based medicine).

What is SNOMED?

- SNOMED CT is a comprehensive clinical terminology that provides clinical content and expressivity for clinical documentation and reporting. It is a concept-based terminology, which means that each medical concept is uniquely identified and can have multiple descriptions.

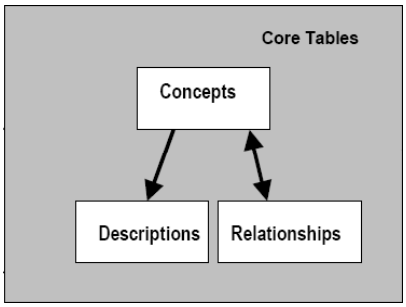
SNOMED CT® Technical Reference Guide – July 2007 International Release

SNOMED CT Data Structure Overview



SNOMED CT® Technical Reference Guide – July 2007 International Release

Core SNOMED Tables

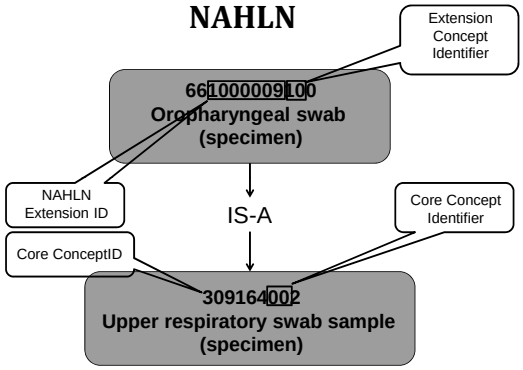


SNOMED CT® Technical Reference Guide – July 2007 International Release

SNOMED Extensions

- Rule #1 for terminologies is "Content, content, content
- "Extensions consist of terminology developed for a particular geographic or administrative requirement that is not part of the SNOMED CT Core."
- Extensions are authorized by SNOMED, but not necessarily managed by them.
- Allow for rapid creation of new terms that comply with SNOMED modeling rules
- Many extension terms are submitted for inclusion in the core

Example Extension Term NAHLN



What is HL7?

- HL7 is a standard for exchanging information between medical applications and is an abbreviation of "Health Level Seven"...a protocol for data exchange. It defines the format and the content of the messages that applications must use when exchanging data with each another in various circumstances.

http://www.interfaceware.com/manual/what_is_hl7.html

SNOMED: Why, When, What and How

- ***Why*** **choose to use SNOMED?**
  - One of the few actively maintained general reference terminologies for health
  - Provides a broad scope of terminological coverage
  - Based on a solid information model that ensures longevity
  - Conforms to good vocabulary practices
  - Provides tools for mapping and subsetting
  - Allows extensions for domain specific terms

SNOMED: Why, When, What and How

- ***When*** **to consider SNOMED for HL7 Messaging**
  - When the field is populated with a coded value (CWE data type)
  - Examples:
    - Marital Status
    - Specimen Type
    - Specimen Source
    - Species/Taxonomy
    - Observation Value

SNOMED: Why, When, What and How

- ***What*** **part of SNOMED should be used?**
  - Only the parts that you need
  - Terms from the appropriate hierarchy
    - Cannot match on words alone
      - Swab (SCTID 257261003) – Specimen
      - Swab (SCTID 408098004) – Physical Object
      - Swab (SCTID 420401004) – Qualifier value
    - Consider the top level IS-A parent IS-A
  - Terms of the proper specificity
    - Body structure – differentiate entire organs from parts of an organ
    - Liver structure (10200004) vs. entire liver (181268008) vs. liver part (119216005)

SNOMED: Why, When, What and How

- ***How*** **do I actually use SNOMED in HL7?**
  - Understand the semantics of the HL7 field to be populated
  - Map local terminologies to SNOMED Concepts
    - Match local concepts with concepts in SNOMED
    - Identify missing concepts and request them from IHTSDO
    - Add missing concepts to local extension
    - Will require post-coordination in some instances
    - Populate messages from mapping tables
  - Use SNOMED terms as interface terminology (subsetting)
    - Use no more of SNOMED than you need.
      - Large portions of SNOMED are irrelevant to a particular purpose.
      - Entire hierarchies may be isolated.
    - Subsets need to be small
    - Even with provided tools, will require manual labor to refine

Post-coordination of results in HL7 messages

- Scenario – Only a subset of possible types of nutritional deficiencies are represented as pre-coordinated terms within SNOMED
  - Copper deficiency – 19577007
  - Calcium deficiency - 238117007
- Some deficiencies are not represented as pre-coordinated terms
  - Taurine deficiency – no pre-coordinated code

Two Mechanisms for Post-Coordination

- SNOMED-CT compositional grammar
  - Taurine Deficiency
    - 70241007: 47429007 = 10944007

70241007|Nutritional deficiency|: 47429007|Associated with|= 10944007|Taurine|

    - Can be transmitted as a unit in OBX.5
    - Requires specific site-to-site negotiation to interpret
- Observation SubID

OBX|1|CE|29308-4|1|70241007^Nutritional deficiency^SCT|...

OBX|2|CE|29308-4|1.1|47429007 ^Associated with^SCT|...

OBX|3|CE|29308-4|1.1.1| 10944007^Taurine^SCT|...

Also would require site-to-site negotiation to interpret

Specific Issues with Specimen

- Specimen types in Public Health are often complex
  - Nearly anything that physically exists
- HL7 specimen tables are incomplete and contain ambiguous, mixed concepts
- SNOMED does not contain sufficient pre-coordinated specimen content
- SPM segment (HL7 version 2.5 or greater) partially addresses lack of pre-coordinated terms

Specimen representation with HL7 and SNOMED

- Abscess Swab
  - SPM.4 code = 258497007 Abscess swab (specimen)
- Corneal swab
  - SPM.4 type = 257261003 Swab (specimen)
  - SPM.8 source = 28726007 Cornea (Body structure)
- Nasal Swab in Transport Media
  - SPM.4 type = 257261003 Swab (specimen)
  - SPM.8 source = 279549004 Nasal Cavity (Body Structure)
  - SPM.6 Additive = 430028007 Michel transport Medium (substance)

Changes to the Organism Hierarchy in SNOMED (ongoing)

- 2006 Taxonomy "summit" held at UC Davis
- Clean up hierarchy to strict Linnaean
  - Retire ambiguous or concepts
    - Pathogenic organism
    - Infectious agent
    - Domestic mammal
- Use Latin names as FSN
  - Provide "common name" descriptions
    - 388747002/1462925016 – Puma concolor
    - 388747002/1493413011 – Cougar
    - 388747002/1493412018 – Mountain Lion
- Develop new "animal in context" hierarchy
- Restrict bacterial and viral names to intrinsic characteristics
  - Develop a mechanism to address functional characteristics

Legacy Codes vs. ConceptID

- Many implementations of SNOMED use legacy Alphanumeric codes
- SNOMED-CT "replaced" these with Concept Identifiers and Description Identifiers
  - Concepts still assigned legacy codes
- Advantages of SCT CIDs and DIDs
  - Can be used with extensions
    - Assures uniqueness with extension IDs
  - Provides for unique IDs of descriptions

Example – Anthrax SCTID - 409498004

Term	Legacy ID	Description ID
Anthrax	DE-1009A	2471280015
Charbon	DE-1009A	2477449018
Infection due to Bacillus anthracis	DE-1009A	2477447016
Milzbrand	DE-1009A	2477446013
Splenic Fever	DE-1009A	2477448014

Caution! – Descriptions are NOT Synonyms

Summary

- Identify the CWE field(s) in the HL7 message and fully understand its value domain
- Identify the proper SNOMED hierarchy from which to draw concepts
- Subset the hierarchy for the particular needs (prune severely)
- Post-coordinate as needed
- Evaluate the need for a local extension
- Submit missing concepts to SNOMED