Clinical Clarity vs. Terminological Order – The Readiness of SNOMED CT Concept Descriptors for Primary Care

Zhe He\textsuperscript{a}, Michael Halper\textsuperscript{a}, Yehoshua Perl\textsuperscript{a}, Gai Elhanan\textsuperscript{b}

\textsuperscript{a} Structural Analysis of Biomedical Ontologies Center
Department of Computer Science
New Jersey Institute of Technology, Newark, NJ

\textsuperscript{b} Halfpenny Technologies, Inc, Blue Bell, PA
Motivation

- Health Information Technology for Economics and Clinical Health (HITECH) Act
  - Signed into law in 2009 to promote the adoption and meaningful use of health information technology
  - As part of the regulation of HITECH Act, SNOMED will be the exclusive terminological system for encoding problem list by 2015

- In light of the increasing role of SNOMED CT in clinical care, are the concept descriptors ready to be used as an interface terminology?
Overview

- SNOMED CT and concept descriptors
- Study: a simulated clinical scenario of selecting terms
- Method: the four samples of the study
- Results: Types and examples of SNOMED issues detected
- Possible implication and solutions
- Final remarks
SNOMED CT

- SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms)
  - Reference terminology for health care
  - Managed by IHDSTO (International Health Terminology Standards Development Organisation)
  - Already used in more than 50 countries
  - More than 311,000 active concepts
  - Concepts distributed in 19 top-level hierarchies
  - Based on Description Logic
SNOMED Concept Descriptors

- Each of SNOMED’s concepts has
  - A fully specified name (FSN) including a semantic tag
    - e.g., *hematoma (morphologic abnormality)*
  - A preferred term (PT)
    - e.g., *hematoma*
  - May have one or more synonyms
- Synonyms and preferred terms are not necessarily unique
- Acronyms are also considered synonymous terms
  - e.g., COPD and COLD are two of 15 synonyms of *chronic obstructive lung disease (disorder)*

Concept ID: 13645005
- chronic obstructive lung disease (disorder)
- chronic obstructive lung disease
- chronic obstructive pulmonary disease
- COPD
- CAO - Chronic airflow obstruction
- CAL - Chronic airflow limitation
- chronic airway disease
- chronic obstructive airway disease
- chronic airflow obstruction
- COLD
- CCOPD - Chronic obstructive pulmonary disease
- CCAA - Chronic obstructive airways disease
- CAFL - Chronic airflow limitation
- chronic airway obstruction
- chronic airflow limitation
- chronic irreversible airway obstruction
- COLD - Chronic obstructive lung disease
Importance of synonyms

- Synonyms are important for effective use in interface terminologies. [Chute et al. JAMIA 98], [Rosenbloom et al. JAMIA 06]

- SNOMED has a relative paucity of synonyms. 36% of SNOMED’s concepts have assigned synonyms. (0.51 synonyms per concept)

- Missing synonym was reported the second most encountered deficiency in SNOMED by 17% of the respondents in a survey [Elhanan et al. JAMIA 11].
UMLS is used as “diagnostic” tool for SNOMED concepts

- In the integration of SNOMED into the UMLS, there were numerous cases (13.4%) where two or more SNOMED concepts are duplicately mapped to the UMLS [Fung et al. JAMIA 05]

Reasons
- Strict separation between hierarchies and use of Description Logic
- Exceptionally fine level of granularity in some SNOMED concept
- Merging involved SNOMED concepts containing the “NOS” (Not otherwise specified) qualifier.
- Two equivalent concepts existing as distinct concepts
The Study: A Simulated Clinical Scenario to Select a Term

- A simulated clinical scenario was used to assess SNOMED’s concepts descriptors’ readiness (especially synonyms).

- Used the search mechanism of SNOMED’s CliniClue browser to select a medical term.

- Do concept descriptors provide sufficient differentiation to enable appropriate concept selection between similar terms?
Example of a Simulated Clinical Scenario

Concept ID: 3424008
- tachycardia (finding)
- rapid heart beat
- heart rate fast

Definition: Fully Defined as...
- is a
- finding
- cardiac conducting system structure
- has interpretation
- increased
- heart rate

Concept ID: 6285003
- tachyarrhythmia (disorder)
- heart rate fast
- tachycardia
- rapid heart beat
- increased heart rate

Definition: Fully Defined as...
- is a
- finding
- cardiac conducting system structure
- has definitional manifestation
- tachycardia
In accordance with [Fung et al. JAMIA 05], four random samples were defined:

**Sample A (SSP):** 65 Same String Pairs of concepts mapped to the same UMLS concept.

**Sample B (NoSSP):** 81 No Same String Pairs of concepts mapped to the same UMLS concept.

Excluded from Samples A and B due to common occurrences:

- **X (substance)** and **Y (product)**
  - e.g. aspirin (product) and aspirin (substance)
- **X (disorder)** and **Y (morphologic abnormality)**
Four Random Samples

- **Sample C (SynCtrl):** 50 SNOMED concepts with at least one synonym such that each does not share a UMLS concept with any other SNOMED concept.

- **Sample D (Ctrl):** 100 individual SNOMED concepts randomly selected without regard to their number of synonyms

General information about the samples
- 442 concepts in all four samples.
- Data from Jan 2010 release of SNOMED.
- All samples were chosen to be mutually disjoint.
Four-point Scale Grade of difficulty

- Quantified the degree of difficulty (four-point scale) that a user may face in making a decision
  - Grade 0 indicates non-issue
  - Grade 1 indicates a minimal issue
  - Grade 2 indicates a moderate issue
  - Grade 3 indicates a significant/critical issue
  - “no (issue)”
  - “yes”
## General sample characteristics

<table>
<thead>
<tr>
<th></th>
<th>General SNOMED</th>
<th>Sample A (SSP)</th>
<th>Sample B (NoSSP)</th>
<th>Sample C (SynCtrl)</th>
<th>Sample D (Ctrl)</th>
</tr>
</thead>
<tbody>
<tr>
<td># concepts</td>
<td>291,205</td>
<td>130</td>
<td>162</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Percentage of concepts with synonyms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>31%</td>
</tr>
<tr>
<td></td>
<td>35.7%</td>
<td>68.5%</td>
<td>50.6%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Average # synonyms</td>
<td>0.51</td>
<td>1.39</td>
<td>1.22</td>
<td>2.80</td>
<td>0.51</td>
</tr>
<tr>
<td>Average # synonyms for concepts with synonyms</td>
<td>1.42</td>
<td>2.05</td>
<td>2.40</td>
<td>2.80</td>
<td>1.65</td>
</tr>
<tr>
<td>Min / max # synonyms</td>
<td>0 / 27</td>
<td>0 / 7</td>
<td>0 / 8</td>
<td>2 / 8</td>
<td>0 / 5</td>
</tr>
</tbody>
</table>
Grade 3 findings across the four samples

<table>
<thead>
<tr>
<th></th>
<th>Sample A (SSP)</th>
<th>Sample B (NoSSP)</th>
<th>Sample C (SynCtrl)</th>
<th>Sample D (Ctrl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>65 (pairs)</td>
<td>81 (pairs)</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Grade 3 Issues</td>
<td>40</td>
<td>14</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>% Grade 3</td>
<td>&quot;62%&quot;</td>
<td>&quot;17%&quot;</td>
<td>&quot;2%&quot;</td>
<td>&quot;1%&quot;</td>
</tr>
<tr>
<td>Synonym Errors</td>
<td>7</td>
<td>–</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>Duplicate Concepts</td>
<td>8</td>
<td>7</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Container Classes</td>
<td>11</td>
<td>3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Other</td>
<td>14</td>
<td>4</td>
<td>–</td>
<td>1</td>
</tr>
</tbody>
</table>
Grade 3 (Significant/Critical) Issues

- Erroneous synonym (8 cases)
- Duplicate concepts (15 cases)
- Container classes (14 cases)
An Example of Erroneous Synonym

*balanoplasty* is a surgical reconstruction of the glans penis.

An erroneous synonym for the concept *repair of penis (procedure)*
An Example of Duplicate Concepts

Concept ID: 336623009
- oxygen nasal cannula (physical object)
- Definition: Primitive
  - is a
  - oxygen equipment

Concept ID: 39849001
- nasal oxygen catheter, device (physical object)
- nasal oxygen catheter
- nasal oxygen catheter, device
- Definition: Primitive
  - is a
  - catheter
  - is a
  - respiratory equipment
Another Example of Duplicate Concepts

Concept ID: 69960004
- antineoplastic chemotherapy regimen (procedure)
- antineoplastic chemotherapy regimen
- chemotherapy regimen
- cancer chemotherapy regimen
- cancer chemotherapy
- chemotherapy

Concept ID: 367336001
- chemotherapy (procedure)
- chemotherapy
- CT - Chemotherapy

Concept ID: 363688001
- administration of antineoplastic agent (procedure)
- administration of antineoplastic agent
- chemotherapy
- administration of medication
- method
- administration - action
- direct substance
- drug or medicament
- antineoplastic agent
- chemotherapy
An Example of Container Class

This issue resulted from the fact that one or both of the involved concepts were container classes that serve to group together and subsume collections of more refined, sibling concepts.
Another Example of Container Class

- SNOMED uses container-class concepts clearly created to subsume a group of other concepts under the “same roof.”

- Can be algorithmically detected and avoided altogether by a disciplined editorial approach
- Unintended use of higher-level concepts can lead to reasoning mistakes by algorithmic decision-support systems
General Conclusions

- General population of SNOMED concepts carries a relatively low rate of major issues.

- SNOMED concepts duplicately mapped to the same UMLS concept may exhibit significant rate of synonym issues.

- May lead users to erroneously select a concept that does not necessarily apply to their patient.
Grade 3 findings across the four samples

<table>
<thead>
<tr>
<th></th>
<th>Sample A (SSP)</th>
<th>Sample B (NoSSP)</th>
<th>Sample C (SynCtrl)</th>
<th>Sample D (Ctrl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>65 (pairs)</td>
<td>81 (pairs)</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Grade 3 Issues</td>
<td>40</td>
<td>14</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>% Grade 3</td>
<td>62%</td>
<td>17%</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>Synonym Errors</td>
<td>7</td>
<td>–</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>Duplicate Concepts</td>
<td>8</td>
<td>7</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Container Classes</td>
<td>11</td>
<td>3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Other</td>
<td>14</td>
<td>4</td>
<td>–</td>
<td>1</td>
</tr>
</tbody>
</table>
Possible Implication

Due to successful leadership and adoption initiatives, SNOMED has already passed the tipping point of clinical adoption

- Average users may find it hard to select a concept
- Novice users of SNOMED cannot be expected to know the inherent structure and underlying description logic.
- Most likely average users do not desire to use terminological tools to discern the differences between SNOMED’s concepts.
Reference Terminology versus Interface Terminology

- IHTSDO does not expect SNOMED to be used as an interface terminology
- However, many EHR vendors attempt to utilize it that way
- The complexity of a reference terminology should be balanced against its clinical usefulness
Possible Solutions

- Editorial policies
  - Container class
  - Redundant concepts
  - Incorrect synonym

- Local extensions
  - The extension mechanism requires a resource intensive, coordinated effort
  - The complexity by design is not likely to be resolved by local extensions

- Well-curated dataset
  - The Department of Veterans Affairs (VA) and Kaiser Permanente (KP) spent years on it
  - Not everybody can afford it
Limitation

- This study is qualitative
- A single evaluator reviewed the samples
- In order to minimize the subjectivity
  - Only expose grade 3 findings (clear-cut issues)
- Use of CliniClue as a simulated clinical scenario
  - It’s unlikely that many of the more than 1000 current EHR vendors will offer a significantly better tool
Final Remarks

- In light of SNOMED’s increasing role in primary care, more attention should be focused on pragmatic usability aspects

- Closer attention should be paid to practical clinical use cases
  - Editorial policies to better address practical clinical needs and reduce structural complexity
Thank you!
Appendix:

CliniClue Screenshots
Airplane accidents
airplane accident

Concept Status: current

Descriptions
- Lang.: *:LANG
  - Concept not translated
- Lang.: *:LANG
  - Concept not translated
- Lang.: *:LANG
  - Concept not translated
- Lang.: *:LANG
  - Concept not translated
- Lang.: *:LANG
  - Concept not translated
- Lang.: *:LANG
  - Concept not translated
- Lang.: *:LANG
  - Concept not translated

Definition: Primitive
- Is a
  - D: air and space transport accident

Codes
- Original SNOMED Id: A-A23BC
- Read Code (Ctv3ld): XUWk6

Hierarchies
- Aircraft accident
  - Air and space transport accident
- Aircraft incident (event)
- Aircraft accident
- Aircraft accident
- Aircraft accidents
- Aeroplane accident
- Airplane accident
- Aircraft crash on takeoff
- Aircraft explosion on takeoff
- Aircraft fire on takeoff
- Aircraft accident
- Aircraft transport accident without damage to aircraft
  - Aircraft collision while landing
  - Aircraft crash on takeoff
  - Aircraft explosion on takeoff
  - Aircraft fire on takeoff
  - Airplane accident
  - Collision of aircraft with another aircraft whilst in transit
  - Collision of aircraft with bird, whilst in transit
  - Explosion on aircraft whilst in transit
  - Fall in, on, or from aircraft
  - Fire on aircraft while in transit
  - Fire on aircraft while landing
  - Hit by aircraft
  - Sucked into aircraft jet
Concept Id: 386680008
DescriptionId: 1477102017

Find:
- airplane accident
- airplane accidents

Concept Status: current

Descriptions

- Lang: *:LANG
- Concept not translated
- Lang: *:LANG
- Concept not translated
- Lang: *:LANG
- Concept not translated
- Lang: *:LANG
- Concept not translated
- Lang: *:LANG
- Concept not translated
- Lang: *:LANG
- Concept not translated
- Lang: *:LANG
- Concept not translated

Definition: Primitive
- is a

Codes
- Original SNOMED Id: A-A23BB
- Read Code (Ctv3Id): XUWk5

Hierarchy
- aircraft accident
  - airplane accident
- balanoplasty
• Nasal oxygen cannula
**Concept Id:** 338623009
**Description Id:** 439877015

**Words - any order**
- nasal oxygen cannula
- oxygen nasal cannula
- CPAP oxygen nasal cannula
- biPAP oxygen nasal cannula
- high flow oxygen nasal cannula

**Hierarchy**
- **oxygen equipment**
  - **oxygen nasal cannula**
    - **biPAP oxygen nasal cannula**
    - **CPAP oxygen nasal cannula**
    - **high flow oxygen nasal cannula**

**Concept Status:** current

**Descriptions**
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated
- **Lang:** *::LANG
  - Concept not translated

**Definition:** Primitive
- **is a**
  - **oxygen equipment**

**Codes**
- **Original SNOMED Id:** A-008C2
- **Read Code (Cti3Id):** pc5.
nasal oxygen cannula

Types:
- Physical object

Hierarchy:
- catheter
- respiratory equipment
- nasal oxygen cannula

Definitions:
- Primitive
- is a
  - catheter
  - respiratory equipment

Languages:
- Lang: *.LANG
  - Concept not translated
nasal oxygen cannula

Descriptions
- Lang: *:LANG
  - Concept not translated
- Lang: *:LANG
  - Concept not translated
- Lang: *:LANG
  - Concept not translated
- Lang: *:LANG
  - Concept not translated
- Lang: *:LANG
  - Concept not translated
- Lang: *:LANG
  - Concept not translated
- Lang: *:LANG
  - Concept not translated
- Lang: en
  - nasal oxygen catheter, device (physical object)
  - nasal oxygen catheter
  - nasal oxygen cannula
  - nasal oxygen catheter, device

Definition: Primitive
- is a
  - Dcatheter
- is a
  - Drespiratory equipment

Codes
- Original SNOMED Id : A-26833
- Read Code (Ch:3:Id) : XUDM
● chemotherapy
- megapode
• milk RAST
cow's milk RAST test

Concept Status: current

Descriptions
- Lang: *.LANG
  - Concept not translated
- Lang: *.LANG
  - Concept not translated
- Lang: *.LANG
  - Concept not translated
- Lang: en
  - cow's milk specific immunoglobulin E antibody measurement (procedure)
  - cow's milk RAST test
  - cow's milk specific immunoglobulin E antibody measurement

Definition: Primitive
- is a
  - milk RAST
- has specimen
  - D serum specimen
- component
  - D cow's milk specific immunoglobulin E
- method
  - D measurement - action

Qualifiers
- priority
  - D priorities

Codes
- Original SNOMED Id: F3-85378
- Read Code (Ctv3Id): Xalc9