SNOMED CT Physician Specialty Subsets:
From development to refinement with an eye on implementation and maintenance

Thursday, October 27, 2016

Robyn Kuropatwa & Kathy Giannangelo
IHTSDO Showcase Presentations

- 2014 Amsterdam
  - Moving from ICD-9-CM Legacy to SNOMED CT Based Clinical Content – An Implementation in Progress...

- 2015 Montevideo
  - Adopting SNOMED CT with Physician Specialty Subsets

- 2016 Wellington
  - SNOMED CT Physician Specialty Subsets: From development to refinement with an eye on implementation and maintenance
OUTLINE

- Introduction
- Describing Content in SNOMED CT Subsets
- Enhancing & Refining SNOMED CT Subsets
- SNOMED CT Subset Outputs
- Challenges, Lessons Learned & Next Steps
- Questions
INTRODUCTION

Background
Purpose
Physician Specialties
The Problem
Nova Scotia Project Goals

- To modernize insured services by implementing a framework to accurately define clinical diagnoses and services that reflect physician practices
- To update policies and processes to support physician, organization and accountability needs with improved data capture and reuse
- To determine which SNOMED CT concepts and descriptions to include in the physician specialty subsets in order to accurately define clinical diagnoses that reflect physician practices
Purpose

- Use of SNOMED CT
  - Enables physicians to use standardised terms

- Development of Physician Specialty Subsets
  - Concept subsets
    - Ensure clinically relevant content is included
  - Description subsets
    - Ensure only relevant descriptions are displayed to physicians
PHYSICIAN SPECIALTIES

1. Anaesthesia*
2. Cardiology
3. Clinical Immunology & Allergy
4. Community Medicine*
5. Dermatology
6. Diagnostic Imaging*
7. Emergency Medicine
8. Endocrinology & Metabolism
9. Gastroenterology
10. General Practitioner
11. General Surgery
12. Genetics
13. Geriatric Medicine
14. Haematology
15. Infectious Diseases
16. Internal Medicine
17. Medical Microbiology*
18. Medical Oncology
19. Nephrology
20. Neurology
21. Neurosurgery
22. Obstetrics & Gynaecology
23. Ophthalmology
24. Orthopaedic Surgery
25. Otolaryngology
26. Paediatrics
27. Pathology
28. Physical Medicine & Rehabilitation
29. Plastic Surgery
30. Psychiatry
31. Radiation Oncology
32. Respiratory Medicine
33. Rheumatology
34. Surgery
35. Urology

* Denotes unique clinical characteristics
The Problem

- Challenge describing content in subsets
  - Example
    - Total concepts: 769
    - Top-level hierarchy
      - Clinical finding: 768
      - Event: 1
  - What does it mean?
  - What concepts are actually contained in the subset?
- Inadequate to just
  - State the total number of concepts
  - State the number of concepts by top-level hierarchy
Describing Content in SNOMED CT Subsets

Nova Scotia Subset Factsheet
Terminologists’/Analysts’ Point of View
Clinicians’ Point of View
Nova Scotia Subset Factsheet

- Introduction
- Subset Overview
  - Hierarchy (Tiers of Granularity)
    - Disorder Types
    - Disorders by Body System
    - Finding Types
    - Findings by Body Site
  - Descriptions
    - Descriptions Per Concept
    - Length of Descriptions
    - Levels of Granularity
  - Relationship Between Concepts
  - Leaf vs Non-leaf Concepts
Clinicians’ Point of View: Tiers of Granularity

1. Top-level Hierarchy
2. Sub-hierarchy
3. View (physiology/anatomy)
4. Concept Group
5. Cluster
6. Individual Concept

- Clinical Diagnoses Subset
  - Clinical Finding
  - Event
  - Situation with Explicit Context

- Disorders
  - Disorder Types
  - Disorders by Body System

- Findings
  - Finding Types
  - Findings by Body Site
TIERS OF GRANULARITY: EXAMPLE

#1: Top-level Hierarchy

#2 Sub-hierarchy

#3: View (physiology/anatomy)

#4: Concept Groups (examples)

#5: Cluster (examples)

#6: Individual Concept (examples)

Clinical Finding

Disorders

Disorder Types

Acute disease

Cardio-vascular system

Pericarditis

Disorders by Body System

Acute pericarditis

Acute tuberculous pericarditis

Acute myocardial pericarditis

Acute hepatitis

Acute hepatitis C

Acute viral hepatitis

Acute pericarditis

Chronic pericarditis

Clinical Finding

Disorders

Disorder Types

Acute disease

Cardio-vascular system

Pericarditis

Disorders by Body System

Acute pericarditis

Acute tuberculous pericarditis

Acute myocardial pericarditis

Acute hepatitis

Acute hepatitis C

Acute viral hepatitis

Acute pericarditis

Chronic pericarditis
Tier 2: Sub-hierarchy

- Clinical finding hierarchy is divided into:
  - **Disorders** – Clinical findings that are essentially abnormal
    - Includes diseases, functional disorders; anything that could be listed as a “diagnosis”
  - **Findings** – Clinical findings that refer to any observations that might be made in relation to the health of a person
    - Includes history, symptoms, signs, laboratory results, imaging findings, etc.
Tier 3: View

- Physiology
  - Disorder types – Represents groups of similar disorders or diseases
  - Finding types – Represents groups of similar findings

- Anatomy
  - Disorders by body systems – Represents groups of body organs or structures that together perform one or more vital functions
  - Findings by body sites – Represents groups of findings by their location
## Tier 4: Concept Group

- Represents related concepts (supertypes and subtypes) grouped together
- Facilitates reviewing a subset’s contents by like topics at a summary level

<table>
<thead>
<tr>
<th>Concept Group</th>
<th>Disorder Types</th>
<th>Disorders by Body Systems</th>
<th>Finding Types</th>
<th>Findings by Body Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children of</td>
<td>64572001</td>
<td>Disease</td>
<td>362965005</td>
<td>Disorder of body system</td>
</tr>
<tr>
<td>Exclude</td>
<td>123946008</td>
<td>Disorder by body site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>84</td>
<td>37</td>
<td>28</td>
<td>80</td>
</tr>
<tr>
<td>Examples</td>
<td>Acute disease</td>
<td>Chronic disease</td>
<td>Complication</td>
<td>Infectious disease</td>
</tr>
</tbody>
</table>
Still contains a large number of concepts

What does it mean when there are 54 concepts in the “ulcer” concept group?

How can we further describe the content within a concept group?
**Tier 5: Cluster (within a Concept Group)**

- **Cluster**: Collection of related concepts (supertypes / subtypes) within a concept group

The different colours indicate different clusters; Black are single concept clusters.
Tier 5: Cluster (within a Concept Group)

- If we remove the concepts that are not in the subset, we can more clearly see the clusters.

Ulcer concept group (54 concepts) in the gastroenterology subset consists of the clusters:

- Aphthous ulcer of mouth (n=1)
- Gastric ulcer (n=13) → Orange
- Peptic ulcer (n=15) → Red
- Pressure ulcer (n=1)
- Pyoderma gangrenosum (n=1)
- Ulcer of anorectal structure (n=3) → Purple
- Ulcer of buccal mucosa (n=1)
- Ulcer of esophagus (n=1)
- Ulceration of intestine (n=23) → Green
Tier 5: Cluster (within a Concept Group)

- Closer look at peptic ulcer cluster
Terminologists’/Analysts’ Point of View

- Levels of granularity
- Hierarchical relationships between concepts
- Leaf and non-leaf concepts
- Number of descriptions per concept
- Length of descriptions
Levels of Granularity

- Purpose: to summarise the granularity of concepts in a subset
- Example

<table>
<thead>
<tr>
<th>Level</th>
<th>Concepts</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>0.3%</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>2.0%</td>
</tr>
<tr>
<td>5</td>
<td>69</td>
<td>9.0%</td>
</tr>
<tr>
<td>6</td>
<td>152</td>
<td>19.8%</td>
</tr>
<tr>
<td>7</td>
<td>204</td>
<td>26.5%</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

The smallest level is used when a concept occurs at multiple levels.
Hierarchical Relationship Between Concepts

- **Purpose:** to indicate how the concepts in the subset are related to each other

- **Example**

<table>
<thead>
<tr>
<th>Relationship to Other Concepts within Subset</th>
<th>Concepts</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has supertype and/or subtype</td>
<td>660</td>
<td>85.8%</td>
</tr>
<tr>
<td>Has supertype and subtype</td>
<td>187</td>
<td>24.3%</td>
</tr>
<tr>
<td>Has supertype only</td>
<td>416</td>
<td>54.1%</td>
</tr>
<tr>
<td>Has subtype only</td>
<td>57</td>
<td>7.4%</td>
</tr>
<tr>
<td>Has no supertype and has no subtype</td>
<td>109</td>
<td>14.17%</td>
</tr>
</tbody>
</table>

*Purple: No supertype and subtype
*Green: Subtype only
*Blue: Supertype and subtype
*Red: Supertype only*
**Leaf and Non-leaf Concepts**

- **Purpose**: to determine the granularity of the concepts, not just in the subset, but all of SNOMED CT
- **Example**

<table>
<thead>
<tr>
<th>Leaf / non-leaf</th>
<th>Concepts</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf</td>
<td>194</td>
<td>25.2%</td>
</tr>
<tr>
<td>Non-leaf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With subtypes in subset</td>
<td>438</td>
<td>57.0%</td>
</tr>
<tr>
<td>Without subtypes in subset</td>
<td>137</td>
<td>17.8%</td>
</tr>
<tr>
<td>Total</td>
<td>575</td>
<td>74.8%</td>
</tr>
</tbody>
</table>

- **Leaf**: concept in SNOMED CT with no subtype
- **Non-leaf**: concept in SNOMED CT with a subtype
Number of Descriptions Per Concept

- **Purpose:** to demonstrate breath of synonyms and to identify concepts with “too many” synonyms
- **Example**

<table>
<thead>
<tr>
<th># of Desc.</th>
<th>Concepts</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>385</td>
<td>50.1%</td>
</tr>
<tr>
<td>2</td>
<td>180</td>
<td>23.4%</td>
</tr>
<tr>
<td>3</td>
<td>74</td>
<td>9.6%</td>
</tr>
<tr>
<td>4</td>
<td>46</td>
<td>6.0%</td>
</tr>
<tr>
<td>5</td>
<td>31</td>
<td>4.0%</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

- 1233754019 | GI - Gastrointestinal bleed |
- 1233753013 | Gastrointestinal bleed |
- 123689014  | GI bleeding |

- 1218401013 | GI - Gastrointestinal hemorrhage |
- 123688018  | Gastrointestinal hemorrhage |
- 123690017  | GI hemorrhage |

- 1218402018 | GIH - Gastrointestinal hemorrhage |
- 1218403011 | GIT - Gastrointestinal tract hemorrhage |
- 123695010  | Gastrointestinal bleeding |
Purpose: to summarise the length of descriptors and identify short and long descriptions

Example

<table>
<thead>
<tr>
<th>Length of Desc.</th>
<th>Concepts</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0.1%</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.1%</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.2%</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0.6%</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>0.8%</td>
</tr>
<tr>
<td>8</td>
<td>22</td>
<td>1.3%</td>
</tr>
<tr>
<td>9</td>
<td>23</td>
<td>1.3%</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>80</td>
<td>1</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

**2 characters**

293531018|CF| Should be excluded as it is a standalone initialism

**78 characters**

30686011|Duodenal ulcer without hemorrhage AND without perforation but with obstruction| Probably should be excluded as physicians rarely would indicate the “without” conditions
Enhancing & Refining
SNOMED CT Subsets

Enhancing SNOMED CT Subsets
Refining SNOMED CT Subsets
## Enhancing SNOMED CT Subsets Using External Subsets (Examples)

<table>
<thead>
<tr>
<th>Specialty</th>
<th>Organisation</th>
<th>NS Subset</th>
<th>External Subset</th>
<th>Overlap</th>
<th>Not in NS</th>
<th>Included in NS</th>
<th>Percent Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiology</td>
<td>Kaiser Permanente</td>
<td>886</td>
<td>653</td>
<td>117</td>
<td>536</td>
<td>531</td>
<td>99%</td>
</tr>
<tr>
<td>Dermatology</td>
<td>Kaiser Permanente</td>
<td>638</td>
<td>691</td>
<td>110</td>
<td>574</td>
<td>547</td>
<td>95%</td>
</tr>
<tr>
<td><strong>Gastroenterology</strong></td>
<td><strong>British Gastroenterology Society</strong></td>
<td><strong>582</strong></td>
<td><strong>286</strong></td>
<td><strong>80</strong></td>
<td><strong>206</strong></td>
<td><strong>204</strong></td>
<td><strong>99%</strong></td>
</tr>
<tr>
<td>Genetics</td>
<td>Apelon</td>
<td>486</td>
<td>2,283</td>
<td>90</td>
<td>369</td>
<td>353</td>
<td>96%</td>
</tr>
<tr>
<td>Haematology</td>
<td>Kaiser Permanente</td>
<td>714</td>
<td>330</td>
<td>42</td>
<td>288</td>
<td>288</td>
<td>100%</td>
</tr>
<tr>
<td>Infectious Disease</td>
<td>Kaiser Permanente</td>
<td>2,202</td>
<td>1,101</td>
<td>85</td>
<td>1,016</td>
<td>692</td>
<td>68%</td>
</tr>
<tr>
<td>Ophthalmology</td>
<td>Kaiser Permanente</td>
<td>1,327</td>
<td>413</td>
<td>139</td>
<td>274</td>
<td>274</td>
<td>100%</td>
</tr>
<tr>
<td>Orthopaedic Surgery</td>
<td>Kaiser Permanente</td>
<td>1,306</td>
<td>167</td>
<td>16</td>
<td>151</td>
<td>91</td>
<td>60%</td>
</tr>
<tr>
<td>Otolaryngology</td>
<td>Kaiser Permanente</td>
<td>1,344</td>
<td>641</td>
<td>189</td>
<td>445</td>
<td>421</td>
<td>95%</td>
</tr>
<tr>
<td>Paediatrics</td>
<td>Kaiser Permanente</td>
<td>3,699</td>
<td>2,181</td>
<td>750</td>
<td>1,431</td>
<td>1,397</td>
<td>98%</td>
</tr>
<tr>
<td>Rheumatology</td>
<td>National Health Services</td>
<td>592</td>
<td>71</td>
<td>22</td>
<td>49</td>
<td>48</td>
<td>98%</td>
</tr>
<tr>
<td>Respiratory Medicine</td>
<td>Kaiser Permanente</td>
<td>114</td>
<td>511</td>
<td>39</td>
<td>454</td>
<td>375</td>
<td>83%</td>
</tr>
</tbody>
</table>

Some subsets had multiple specialties, which we manually split. Extension concepts from external subsets were not used in Nova Scotia. Concepts that were inactivated on July 31, 2015 were also excluded.
Refining SNOMED CT Subsets Using Tiers of Granularity

- 422014003 | Disorder associated with type 2 diabetes mellitus (disorder) |
- 420868002 | Disorder associated with type I diabetes mellitus (disorder) |
- 420270002 | Ketoacidosis in type I diabetes mellitus (disorder) |
- 421750000 | Ketoacidosis in type II diabetes mellitus (disorder) |
- 62479008 | Acquired immune deficiency syndrome (AIDS) (disorder) |
- 267543009 | Functional disorders of polymorphonuclear neutrophils (disorder) |
- 234532001 | Immunodeficiency disorder (disorder) |
- 190979003 | Selective immunoglobulin A deficiency (disorder) |

Early draft of Nova Scotia Gastroenterology Clinical Diagnoses Subset
Subset Outputs

Nova Scotia Department of Health & Wellness Terminology Distribution Format

SNOMED CT RF2
Human-readable subset snapshots
Machine-processable subset full
Transitive closure table
Historical relationships
HUMAN-READABLE SUBSET SNAPSHOTs
(Active Content)
Filter $\text{StartEffectiveTime} \leq \text{CurrentDate}$, $\text{EndEffectiveTime} > \text{CurrentDate}$ and $\text{Active} = 1$
<table>
<thead>
<tr>
<th>Extension Concepts</th>
<th>Extension Descriptions</th>
<th>Extension Relationships</th>
<th>Simple RefSet (Concepts)</th>
<th>Simple RefSet (Descriptions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Excel Sheet 1" /></td>
<td><img src="image2.png" alt="Excel Sheet 2" /></td>
<td><img src="image3.png" alt="Excel Sheet 3" /></td>
<td><img src="image4.png" alt="Excel Sheet 4" /></td>
<td><img src="image5.png" alt="Excel Sheet 5" /></td>
</tr>
</tbody>
</table>
Lessons Learned

1. Physician involvement
2. Multiple cycles of refinement
3. Use of SNOMED CT structure
4. Use of external subsets
5. User-friendly, capable tools
6. Variations in physician feedback
Challenges

- Physician ownership of their clinical content
- Subset depth – can physician use of clinical content be improved?
- Direct use of SNOMED CT descriptions for data capture versus use of interface terminology
- Integrating subsets across the health system
- Vendor requirements to support quality clinical data capture
- Ongoing subset refinement over time – informed by analyses of content use
**Next Steps**

- Conduct final subsets quality assurance checks
- Develop implementation specifications and guidance for vendors
- Integrating subsets across the health system
  - Health system policy, physician clinical workflow, education and change management, information management requirements and rules, data capture and storage, update current data uses and data analyses and address intersections outside of health
Questions?

Robyn Kuropatwa
<robynkuropatwa@shaw.ca>
Kathy Giannangelo
<kathy.giannangelo@gmail.com>