SNOMED CT Web Series

WELCOME!
We will begin shortly
Clinical Webinars
Implementation Webinars
Research Webinars

snomed.org/web-series
Upcoming Webinars

April Clinical Webinar

SNOMED CT for Clinicians

22 April 12-15:00 UTC

Presenters: TBC
April Research Webinar

Using SNOMED CT relationships for data exploration and discovery in rare diseases - An example in urea cycle disorders

Wednesday, 28 April
17:00 UTC

Dr. Rachel Richesson
Professor of Learning Health Sciences
University of Michigan
Join our SNOMED CT Research Reference Group!

Be notified of upcoming Research Webinars and other SNOMED CT research-related news.

Email Suzy (sro@snomed.org) to join
CALL FOR ABSTRACTS

#SCTEXPO21

Submit at snomedexpo.org before April 30, 2021.
Utilizing the SNOMED CT hierarchy to generate patient-friendly clarifications: challenges and opportunities

Mr. Hugo J.T. van Mens
Utilizing the SNOMED CT hierarchy to generate patient-friendly clarifications: challenges and opportunities

SNOMED CT Research Webinar

Hugo van Mens, MSc
Software developer, R&D Care Portal, ChipSoft
PhD candidate, Department of Medical Informatics, Amsterdam UMC

Wednesday, March 17, 2021, 14:00 UTC
Content

• Patient access to EHRs
• Patient-Friendly Extension
• Generalization method
• Validation study
• Further research
PET scan
CAT scan

Computer on wheels (COW)

https://www.pantagraph.com/lifestyles/health-med-fit/medical-jargon-can-be-confusing-misinterpreted/
Access to health records by patients

• Associated with efficiency, self-management and patient engagement

• Stimulated by subsidies and legal requirements

• Medical data are registered by clinicians for the clinical process itself
  o Specialist knowledge, acronyms, abbreviations, spelling errors, typos, and formal language; jargon
  o Latin and Greek words more formal, unfamiliar in Dutch (than languages derived from Latin for example)

https://www.opennotes.org/onc-federal-rule/
Access to health records by patients

• Most patients have no medical background
  o 29 to 53% of NL has a low “health literacy”

• Especially certain groups have difficulties
  o People with disabilities
  o Low education
  o Elderly and children
  o Non-native speakers
  o Recently diagnosed, not chronic

• How can we help patients understand their medical data?
Netherlands Patient-Friendly Extension

---

**Resultaten (50 van 483)**

- **slagader**
  - pneumonie (aandoening)
  - pneumonie
  - longontsteking
- **mamma-augmentatie (verrichting)**
  - mamma-augmentatie
  - vergroting van borsten
  - augmentatie van mammee
  - borstvergroting
  - longontsteking
  - augmentation mammoplasty
  - Breast augmentation
  - Pneumonia

---

**structuur van arterieel vaatstelsel**

- **arterie**
  - structuur van arterie
  - arteria
  - slagader
  - De bloedvaten waarin bloed van het hart af stroomt.
  - Arterial structure
  - Artery

---

**abces (aandoening)**

- abces
  - abcess
  - Dit is een niet-natuurlijke holte gevuld met pus bestaande uit celafval en (dode) bacteriën.

---

**terminologie.nictiz.nl**
malignancy stomach

cancer in stomach

malignant neoplasm of stomach
Single data-entry, multiple use

24,966 diagnoses

Diagnosethesaurus

18,579 mapped to

SNOMED CT

714 synonyms

361 definitions

3.8%

Patient-friendly extension

11788

Contusio cerebri

262689001

Contusion of cerebrum

Bruise of the brain

This is damage to brain tissue because the brain is shaken back and forth by a sudden, violent movement.
Pneumonia caused by SARS-CoV-2 is COVID-19 and lung inflammation.

Cardiomyopathy caused by SARS-CoV-2 is a viral infection and heart muscle disease.
Clarification software

- Rule-based system with text concatenation
  - Imports Diagnosethesaurus, SNOMED CT
  - Supports previous terminology versions (Full release)
  - Browse and search functionality
  - Export clarifications and statistics

- Software
  - ASP .NET Core Web API
  - Neo4j graph database
Generalization algorithm

• Clarification
  o Synonym + supertypes + supertype definitions

• Transitive closure
  o Through graph database traversal

• Redundancy
  o Remove redundant supertypes that are supertypes of supertypes
  o Remove patient-friendly synonyms and supertypes contained in medical term already

Sepsis = blood poisoning + infection

Angiosarcoma of liver = liver cancer + definition of angiosarcoma

<table>
<thead>
<tr>
<th>Subtype (descendant)</th>
<th>Supertype (ancestor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>233604007</td>
<td>Pneumonia</td>
</tr>
<tr>
<td>233604007</td>
<td>Pneumonia</td>
</tr>
<tr>
<td>233604007</td>
<td>Pneumonia</td>
</tr>
<tr>
<td>312342009</td>
<td>Infective pneumonia</td>
</tr>
</tbody>
</table>

e.g. aortic valve stenosis is a heart valve disorder, is a heart disease, is a disorder of cardiovascular system

e.g. Acne is acne

e.g. Aneurysm of heart is an aneurysm + definition of aneurysm
Problem definition

• To what extent can we increase content coverage by generalizing diagnoses to supertype concepts with patient-friendly terms and definition in the extension, using the SNOMED CT hierarchy?
  o Compared to 3.8% using synonyms only
## Generalization results

<table>
<thead>
<tr>
<th>Supertypes</th>
<th>Diagnoses</th>
<th>Clarifications</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.686</td>
<td>531</td>
<td>Factor X deficiency is a type of <strong>blood clotting disorder</strong>.</td>
</tr>
<tr>
<td>2</td>
<td>5.007</td>
<td>1.404</td>
<td>Arachnoid cyst of pituitary is a type of <strong>cerebral cyst</strong> and <strong>hormonal disorder</strong>.</td>
</tr>
<tr>
<td>3</td>
<td>1.111</td>
<td>541</td>
<td>Rheumatoid carditis is a type of <strong>autoimmune disorder, joint inflammation</strong>, and <strong>heart muscle disorder</strong>.</td>
</tr>
<tr>
<td>4</td>
<td>242</td>
<td>156</td>
<td>Lemierre syndrome is a type of <strong>inflammation of veins, bacterial infection, blood poisoning</strong>, and <strong>thrombosis</strong>.</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>44</td>
<td>Bartonella-henselae-neuroretinitis is a type of <strong>retinal disease, bacterial infection, eye infection, eye inflammation</strong>, and <strong>optic neuritis</strong>.</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>11</td>
<td>Kinky hair disease is an <strong>inborn</strong> and <strong>hereditary disease</strong> of <strong>bone, skin, nervous system</strong>, and <strong>metabolism</strong>.</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>3</td>
<td>Ataxia telangiectasia is an <strong>inborn and hereditary disease</strong> of <strong>heart and veins, brains, skin, immune system</strong>, and <strong>spinal cord</strong>.</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>16.124</strong></td>
<td><strong>2.690</strong></td>
<td>Using only <strong>620</strong> supertypes</td>
</tr>
</tbody>
</table>

86.8% of diagnoses: 22 times more than with synonyms!
Validation – background

• Problem definition
  o Loss of detail, granularity
  o Multiple supertypes might be confusing
  o Dependency on terminology mappings, SNOMED CT hierarchy and modeling of defining relationships
  o Patient friendliness of the extension synonyms and definitions

• Objective: test medical validity of generalization method
  o Identify errors and unacceptable results
  o Assess requirements for implementation into patient portals and personal health records problem lists

620 supertypes...

...to clarify 16.124 diagnoses?!
Validation – method

- Random sample of 1200 clarifications
  - Covering each supertype’s patient-friendly term and definition
  - Different combinations of those 620 supertypes

- Two raters from Nictiz
  - Medical background and involved in SNOMED CT translation
  - Instructions with 6 cases and discussion after first 60 cases

- Ratings
  - Errors and feedback, to identify error types
  - Acceptability, Completeness, Relevance, Clarity on a scale
    - From e.g. 1 very unacceptable to 5 very acceptable (where 3 is not unacceptable or acceptable)
Validation – results

- 12.7% (10.8-14.7%) contained errors,
  - As identified by one or both raters
- 14.2% (12.2-16.2%) were unacceptable
  - According to one or both raters
    - 1. Very unacceptable
    - 2. Unacceptable
- Moderate interrater reliability
**Validation results: Mapping errors**

**Duchenne female carrier**

A kind of birth defect and limb-girdle muscular dystrophy. It is hereditary. Limb-girdle muscular dystrophy: This is a hereditary condition in which the muscles of the shoulders, upper arms, hips and thighs in particular do not function or do not function sufficiently.

**Reason**

Carrier (not further specified) mapped on manifesting form: 240050008 | Manifesting female carrier of X-linked muscular dystrophy |
Validation results:
Errors in defining relationships

40425004 | Postconcussion syndrome |
A type of dementia.

Reason
Defined as ‘is a’ 230282000 | Post-traumatic dementia |

https://browser.ihtsdotools.org/?perspective=full&conceptId1=40425004&edition=MAIN/2021-01-31&release=&languages=en
Validation results:
Definition not applicable to all subtypes

20751000146109 | Recurrent nasopharyngeal carcinoma |
A type of chronic illness and cancer in the throat.
Chronic illness: Conditions with no prospect of full recovery and with a relatively long duration of illness, often requiring long-term care.

Comments
Should not consider recurrent disease a chronic disease

Reason
Intermittent diseases are also defined as types of chronic diseases
Validation results:
Errors in synonyms in extension

32644009 | Braxton Hicks contractions |  
A type of childbirth

Comments  
This is actually not a childbirth, but false labor pains.

Reason  
Patient-friendly extension synonym ‘childbirth’ for ‘118216002 |Labor finding|’
Validation results:
Algorithm splitting changes semantics

94360002 | Secondary malignant neoplasm of kidney |
A type of cancer in kidney and metastases.

94184004 | Secondary malignant neoplasm of base of tongue |
A type of cancer in the throat, cancer in the mouth, cancer in the larynx and metastases.

Reasons
Loss of the relationship between the supertype concepts by splitting it into supertypes
Splitting into several body parts
Validation results:
Incomplete

234482009 | Amegakaryocytic thrombocytopenia |
A type of disease. Disease: This is a deviation from the healthy state of the body or mind, which manifests itself in symptoms and disorders.

125666000 | Burn |
A type of disease. Disease: This is a deviation from the healthy state of the body or mind, which manifests itself in symptoms and disorders.
Validation results: Irrelevant

225564006 | Pain of nose |
A kind of facial pain. Facial pain: This includes facial pain such as pain in the jaw, pain around the eyes, headache radiating to the face or nerve pain from the triplet nerve (trigeminal neuralgia).

262830005 | Pancreatic duct injury |
A pancreatic disease. Pancreatic disease: Disorders of the pancreas, such as pancreatitis and pancreatic cancer.
Validation results:
Unclear

19011000146107 | Lumbar spinal stenosis with cauda equina-compression |
A type of cauda equina syndrome.

232357009 | Post-surgical epistaxis |
Nosebleed after surgery

3 subtypes of cauda equina

217 bleedings
654 complications
Validation results: Majority acceptable clarifications

652005 | Gangrenous tonsillitis | A type of inflammation of tonsils.

10698009 | Herpes zoster iridocyclitis | A type of shingles, neuropathy, eye infection and eye inflammation.

59393003 | Apocrine acne | A type of bacterial infection, hidradenitis and skin infection.
Bacterial Infection: Infectious diseases caused by a bacteria.
Hidradenitis: These are chronic inflammations of the hair follicles, sebaceous glands and sweat glands, especially in the groin and / or armpits, which lead to cysts, fistulas and abscesses.
Validation – discussion

• Found errors and unacceptable clarifications
  o Issues with underlying terminologies
  o Problems with the algorithm
  o Majority without errors and acceptable

• Implications
  o Remove or modify unacceptable clarifications
    ▪ Validation of complete final set necessary
  o Recommendations for terminology auditing
    ▪ Mappings and modelling
    ▪ Patient-friendly synonyms and definitions
Validation – discussion

• Advantages to generalization approach
  o Reuse of patient-friendly terms and definitions
  o Costly to develop good clarifications
  o Simplification leaves out some details too
  o Supports newly introduced concepts

• Disadvantages to generalization approach
  o Dependency on underlying terminologies
  o Quality assurance and validation required
  o Loss of semantic relationships between supertypes
Further research: Implementation study

- Evaluation in real-time patient portal
  - Baseline for further development

- Outcome measures
  - Diagnoses viewed on problem list
  - Clicks on infobutton
  - Ratings and feedback

- Aggregated data extraction
  - Audit trail data
  - Diagnosis data
Further research:
Towards dynamic, post-coordinated clarifications

• Knowledge representation
  o Information not present in descriptions only

• Fully defined concepts could be fully clarified
  o Thousands of infections and cancers
  o Reuse of explanation of body parts
  o Some generalization
    ▪ e.g. meninges -> in the brain
    ▪ SARS-CoV -> a coronavirus
Defining relationships

372003004 | Primary malignant neoplasm of pancreas |

Malignant, cancerous tumor originating in the pancreas.

https://browser.ihtsdotools.org/?perspective=full&conceptId1=372003004&edition=MAIN/2021-01-31&release=&languages=en
Inflammation of the heart muscle. Caused by an infection with a bacteria called Coxiella burnetii.

https://browser.ihtsdotools.org/?perspective=full&conceptId1=233867000&edition=MAIN/2021-01-31&release=&languages=en
A form of diabetes that develops as a result of disorders in the secretion of insulin and / or the insufficient utilization of the insulin present by tissues occurs at all ages, but especially in the elderly and in overweight people.
A type of congenital condition and blood clotting disorder. It is hereditary.
Inflammation in the brain due to infection with the bacteria Streptococcus
Conclusions

• Generalization helps to increase content coverage
• Identified several problems with diagnoses generalization to patient-friendly terms and definitions
  o Most of which can be addressed
  o Majority was considered acceptable

• Unexpected input for terminology auditing and quality improvement

• Further research
  o Implementation for baseline study
  o Finding appropriate level of detail, granularity
  o Using defining relationships
Thanks to project members and contributors

**Validation:** Savine Martens, Malou Paiman  
**Case-report forms:** Sander Mertens  
**Promotors:** Ronald Cornet, Nicolette de Keizer  
**Manager:** Remko Nienhuis  

**Support in carrying out study:**  
Annemarieke Homan, Natasha Krul, Feikje Hielkema, Theo van Mens, Frank van Oosterhout, Pim Volkert
Thanks for your attention!


Thank you for attending!

THANK YOU

Hugo J.T. van Mens

Suzy Roy
sro@snomed.org
info@snomed.org