SNOMED CT and Clinical Genomics: Use case for precision medicine

James R. Campbell MD
W. Scott Campbell PhD
Jay Pedersen MS

Departments of Internal Medicine and Pathology
University of Nebraska Medical Center
Acknowledgements

- Allison Cushman-Vokoun MD PhD
- Geoffrey Talmon MD
- Audrey Lazenby MD
- Raj Dash MD
- Mary Kennedy
- Observable project team
- iPaLM SIG
Outline

- Terminology challenges created by Precision Medicine
- Harmonized model for observables and observations
- Precision medicine terminology use cases
- Progress report: Molecular Pathology content development and deployment
“Doctors have always recognized that every patient is unique, and doctors have always tried to tailor their treatments as best they can to individuals. You can match a blood transfusion to a blood type — that was an important discovery. What if matching a cancer cure to our genetic code was just as easy, just as standard? What if figuring out the right dose of medicine was as simple as taking our temperature?”

- President Obama, January 30, 2015
Revolution in healthcare

- Sequencing of human genome has led to flood of new observational data appearing in scientific literature and now...clinical records
- Majority of this clinical data is unstructured and not useful for decision support in the EHR or clinical research

- Understanding of genetic and molecular basis of human disease now having impact on diagnosis and therapy
- Challenge posed by precision medicine - to have sufficiently detailed molecular genetics observations and diagnostic data in treatment of cancer, infectious disease and pharmacogenetics
Limitations of ONC terminologies for Human Genomics

Research community
- HUGO; Human Gene Nomenclature Committee
- UniProt; Ensemble; Cosmic
- NCBO: OMIM; Orphanet; Protein Ontology

Clinical community
- SNOMED CT:
  - No concept model for subcellular anatomy or molecular structure
  - No concept model for Observable entity or molecular basis of disease in Clinical findings
  - Little content, all primitives
- LOINC 2.54:
  - 1275 PCR; 1406 MOLGEN; 116 FISH; 1502 CELL MARKERS
  - Concept model inadequate to fully define what is being result
  - Provides only tag-level interoperability of molecular data

- **No meaningful bridge joining genetic scientific findings with clinical concept models**
LOINC – SNOMED CT
Harmonization of Observable entities

- 2008 agreement between Regenstrief (RI) and IHTSDO
  https://loinc.org/collaboration/ihtsdo/agreement.pdf
  - Extend and harmonize a shared concept model for 363787002|Observable entity|
  - Map and instantiate LOINC parts in SNOMED CT content
  - Jointly publish expression data set defining (lab) LOINC concepts within the harmonized concept model
  - Technology preview alpha January 2016

- What is needed to support clinical decision making and research in molecular genetics is a unified domain ontology for Observable entities spanning clinical and research content including genomics
UNMC: Project for structured encoding of Pathologist cancer reports

- Objective: Detailed structured reporting of all anatomic and molecular pathology observations for all CAP synoptic cancer worksheets (82 types of malignancies)

- Proposal: Analyze detailed semantics of CAP worksheets; apply harmonized concept model to develop terminology requirements; deploy as real-time structured reporting from COPATH system interfaced to tissue biobank and EPIC

- Tooling: Nebraska Lexicon© extension namespace; SNOWOWL authoring platform; SNOMED CT International + US Extension + Observables Technology preview; ELK 0.4.1 DL classifier

- Penciled into IHTSDO workplan for 2017
UNMC: Project for structured encoding of Pathologist cancer reports

Objective: Detailed structured reporting of all anatomic and molecular pathology observations for all CAP synoptic cancer worksheets (82 types of malignancies)

Proposal: Analyze detailed semantics of CAP worksheets; apply harmonized concept model to develop terminology requirements; deploy as real-time structured reporting from COPATH system interfaced to tissue biobank and EPIC

Tooling: Nebraska Lexicon© extension namespace; SNOWOWL authoring platform; SNOMED CT International + US Extension + Observables Technology preview; ELK 0.4.1 DL classifier

Penciled into IHTSDO workplan for 2017
SNOMED CT Content Extensions for Precision Medicine

- Body structures >> Cell structures >> Nucleotide sequences and Named Genes
- Substances >> Proteins
- Qualifiers >> Techniques >> AP and MP methods
- Qualifiers >> Measurement Properties >> AP and MP properties
- Observable entity >> AP and MP observables
- Clinical findings >> Anatomic and molecular genetic observation results and disorders
SNOMED CT Content Extensions for PM

- Body structures >> Cell structures >> Nucleotide sequences and Genes

Define genes and related subcellular structures by mapped reference to scientific ontologies classified in conjunction with SNOMED CT & LOINC Observables
SNOMED CT Content Extensions for PM

- Body structures >> Cell structures >> Nucleotide sequences and Genes
- Substances >> Proteins

Define proteins and related molecular structures by mapped reference to scientific ontologies classified in conjunction with SNOMED CT & LOINC Observables.
SNOMED CT Content Extensions for PM

- Body structures >> Cell structures >> Nucleotide sequences and Genes
- Substances >> Proteins
- Qualifiers >> Techniques >> AP and MP methods
- Qualifiers >> Measurement Properties >> AP and MP properties
- Observable entity >> AP and MP observables
- Clinical findings >> Anatomic and molecular genetic observation results and disorders
SNOMED CT Content Extensions for PM

- Body structures >> Cell structures >> Nucleotide sequences and Genes
- Substances >> Proteins
- Qualifiers >> Techniques >> AP and MP methods
- Qualifiers >> Measurement Properties >> AP and MP properties
- Observable entity >> AP and MP observables
- Clinical findings >> Anatomic and molecular genetic observation results and disorders
889779351000004103|BRAF protein expression by immunoperoxidase staining of excised malignant neoplasm (observable entity)
Development for MP: Datatypes requirements

- 455350031000004100 | BRAF nucleotide sequence detected in excised malignant neoplasm (observable entity)
- 363787002 | Observable entity
- 923178251000004106 | Sequence variant property
- 367651003 | Malignant neoplasm (morphology)
- 100670521000004106 | BRAF gene locus (cell structure)
- 123029007 | Single point in time
- 59816316100000410 | Variant call format
- 59816316100000410 | Nucleotide sequencing technique
- 441652008 | Formalin fixed paraffin embedded tissue specimen
- 246501002 | Technique
- 704319004 | Inheres in
- 9216841000004100 | Inherent location
- 704318007 | Property type
- 370134009 | Time aspect
- 370132008 | Scale type
- 704327008 | Direct site
SNOMED CT Content Extensions for PM

- Body structures >> Cell structures >> Nucleotide sequences and Genes
- Substances >> Proteins
- Qualifiers >> Techniques >> AP and MP methods
- Qualifiers >> Measurement Properties >> AP and MP properties
- Observable entity >> AP and MP observables
- Clinical findings >> Anatomic and molecular genetic observation results and disorders
Molecular Pathology Finding
Fully defining observations of sequence variants

843509901000004102 | BRAF V600E mutation identified in excised malignant neoplasm (finding) |

441742003 | Evaluation finding |

363714003 | Interprets (attribute) |

455350031000004100 | BRAF nucleotide sequence detected in excised malignant neoplasm (observable entity) |

363698007 | Finding site (attribute) |

MCG.BRAF c1799T>A (V600E) |

116676008 | Associated morphology (attribute) |

100670521000004106 | BRAF gene locus (cell structure) |

687830031000004102 | Nucleotide sequence variant (morphology) |
Patient Condition

BRAF V600E mutation detected in malignant melanoma of skin (disorder)

- 843509901000004102 | BRAF V600E mutation identified in excised malignant neoplasm (finding)
- 93655004 | Malignant melanoma of skin (disorder)
  - 843509901000004102 | BRAF V600E mutation identified in excised malignant neoplasm (finding)
  - Due to
  - Finding site
    - Skin structure (Body structure)
  - Associated morphology
    - Malignant melanoma – category (morphology)
Precision Medicine Terminology Use Cases

- Health maintenance alert for BRCA1 or BRCA2 positive patients
- Retrieve all melanoma tissue specimens which had BRAF gene testing
- Alert cardiologist for CYP2C19 low metabolizer patients prior to angioplasty to adjust clopidogrel dosing
- Which patient with cystic fibrosis should be treated with Ivacaftor at $300,000 per year?
Patient condition: BRCA1 positive

- Extremely high lifetime risk of breast, ovarian and other cancers
- Thousands of mutations in BRCA gene
- Founder mutations in BRCA:
  - BRCA1*185delAG
  - BRCA1*5382insC
  - BRCA2*6174delT

Inventory of clinically relevant sequence variants leading to BRCA1/2 phenotype will be prepared using PM concepts for EHR documentation and to alert clinicians.
Melanoma is a potentially fatal skin cancer that is poorly treatable if not removed in time.

BRAF gene testing to refine and guide cancer treatment

- Melanoma is a potentially fatal skin cancer that is poorly treatable if not removed in time.
- BRAF is a gene that controls cell growth and plays a role in the development of melanoma and other cancers.
- Vemurafenib and dabrafenib have been approved by FDA for treatment of late-stage melanoma in BRAF positive tumors.
- BRAF sequence data is assessed as part of synoptic reporting. It can guide treatment and serve as a research database for future cancer trials.
CYP2C19
cytochrome P450 family 2 subfamily C member 19

- Clopidogrel is inactive pro-drug requiring metabolic activation to be anticoagulant
- Mutated alleles may have normal, increased, reduced or absent metabolic activity

Detailed genetic findings relevant to patient metabolic function can be stored directly in the EHR so that the cardiologist can be alerted when to adjust clopidogrel dosing after cardiac stent placement
Clopidogrel is inactive pro-drug requiring metabolic activation to be anticoagulant.

Mutated alleles may have normal, increased, reduced or absent metabolic activity.

Specific nucleotide variants can be fully defined in SNOMED CT-LOINC harmonized observables model.

Detailed genetic findings relevant to patient metabolic function can be stored directly in the EHR so that the cardiologist can be alerted when to adjust clopidogrel dosing after cardiac stent placement.
“Designer drug”: Ivacaftor

- Benefit documented for patient with specified Cystic Fibrosis genetic variants in CFTR gene:

- Nucleotide sequence findings for the patient may be documented and reported in full detail

- EHR genetic data will be integrated into decision support to assure that those with treatment indication get the therapy
Terminology development summary: CAP Colorectal and Breast Cancer checksheets

<table>
<thead>
<tr>
<th>SNOMED CT hierarchy</th>
<th>Anatomic Pathology Concepts/Primitives</th>
<th>Molecular Genetic Concepts/Primitives</th>
<th>Exemplar molecular extension concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observable entities</td>
<td>61/1</td>
<td>32/3</td>
<td>“BRAF nucleotide sequence detected in excised malignancy”</td>
</tr>
<tr>
<td>Body Structures</td>
<td>10/9</td>
<td>29/3</td>
<td>“BRAF gene locus”</td>
</tr>
<tr>
<td>Clinical findings</td>
<td>6/2</td>
<td>7/3</td>
<td>“BRAF V600E variant identified in excised malignancy”</td>
</tr>
<tr>
<td>Procedures</td>
<td>2/1</td>
<td>0</td>
<td>“Pyrosequencing”</td>
</tr>
<tr>
<td>Techniques</td>
<td>4/4</td>
<td>7/7</td>
<td>“Sequence property”</td>
</tr>
<tr>
<td>Property types</td>
<td>8/8</td>
<td>2/2</td>
<td>“Variant call format”</td>
</tr>
<tr>
<td>Scale types</td>
<td>0</td>
<td>9/9</td>
<td></td>
</tr>
<tr>
<td>Situations</td>
<td>1/0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Substances</td>
<td>0/0</td>
<td>11/11</td>
<td>“BRAF human cellular protein”</td>
</tr>
<tr>
<td>Attributes</td>
<td>2/2</td>
<td>3/3</td>
<td></td>
</tr>
<tr>
<td>Qualifiers</td>
<td>2/2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TOTALS</td>
<td>88/29</td>
<td>100/41</td>
<td></td>
</tr>
</tbody>
</table>

Genes modeled for SNOMED CT/LOINC extension:
APC, BRAF, BRCA, ERBB2, ESR1, ESR2, KIT, KRAS, MKI67, MLH1, MSH2, MSH6, NRAS, PGR, PIK3CA, PMS2, PTEN, SLC7A8 + codons + microsatellites
Questions?