mCODE and Clinical Genomics – A Foundation for Oncology Precision Medicine

May Terry, MITRE Corporation
Topics

• Brief overview of mCODE
• mCODE Clinical Genomics Elements
• mCODE Genomics Examples
• Comparison with other models
  • HL7 Clinical Genomics Reporting FHIR IG
  • OMOP Oncology
mC\textsuperscript{ODE}, or Minimal Common Oncology Data Elements, is a data standard that can be widely adopted. It holds promise to greatly increase high-quality data for all cancer types.

A standard health record for oncology

The minimal set of data elements applicable to all cancers, and collected for:

- Standardized information exchange
- Use-case driven and targeted use

Oncology data element domains: patient, disease, treatment, outcomes, genomics, lab/vital

mC\textsuperscript{ODE}™
What is a FHIR Accelerator?
The HL7 FHIR Accelerator Program is designed to assist implementers across the health care spectrum in the creation of FHIR Implementation Guides (IGs) and other documents.

Other FHIR Accelerators:

CodeX is following the successful Da Vinci approach to legal, organizational, funding, and governance models.

http://www.hl7.org/about/davinci/members.cfm

https://confluence.hl7.org/display/COD/CodeX+home
mCODE 1.0 Conceptual Model

- mCODE FHIR IG Release 1 STU1
  - balloted in Sept 2019
  - published in Mar 2020
Precision medicine — a working definition:

• A form of medicine that uses information about a person’s own genes or proteins to prevent, diagnose, or treat disease. In cancer, precision medicine uses specific information about a person’s tumor to help make a diagnosis, plan treatment, find out how well treatment is working, or make a prognosis.

Source: NCI Dictionary
mCODE and Oncology Precision Medicine

mCODE 1.0 – Using FHIR for aligning patient-centric clinical oncology data with genomics research data
mCODE 1.0 model – Biomarker-related elements

Genomics Report: Sequencing, NGS, etc

Tumor Marker: IHC, FISH, CISH, etc. + serum tumor markers
Example 1: BRCA2 Positive

- EHR receives a genomics report unstructured document (text, PDF); user transcribes the result into an EHR form.

- **Cancer Patient**
  - identifier: ####
  - birthDate: 1980-01-01
  - gender: F

- **Genetic Specimen**
  - identifier: 44-55-66
  - type: Blood
  - collectedDateTime: 2016-06-11
  - receivedTime: 2016-12-06

- **US Core Organization**
  - identifier: 249
  - name: Ambry Genetics

- **Cancer Genomics Report**
  - subject: Reference (Cancer Patient)
  - specimen: Reference (Genetic Specimen)
  - performer: Reference (Organization)
  - result: Reference (Cancer Genetic Variant)

- **Cancer Genetic Variant**
  - gene-studied: BRCA2
  - valueCodeableConcept: Present
  - interpretation: Positive
Example representation of CancerGeneticVariant

id: mCODECancerGeneticVariantExample01

status: final

category: Laboratory (Details: (http://terminology.hl7.org/CodeSystem/observation-category code 'laboratory' = 'Laboratory'))

code: Genetic variant assessment (Details: (LOINC code '69548-6' = 'Genetic variant assessment'))

subject: Generated Summary; id: mCODEPatientExample01; Medical Record Number = m123 (USUAL); John B. Anyperson; gender: male; birthDate: 1951-01-20

effective: Apr 1, 2019, 12:00:00 AM

value: Present (Details: (LOINC code 'LA9633-4' = 'Present', given as 'Present'))

method: Sequencing (Details: (LOINC code 'LA26308-0' = 'Sequencing', given as 'Sequencing'))

code: Gene studied [ID] (Details: (LOINC code '48018-6' = 'Gene studied [ID]'))

value: STK11 (Details: (http://www.genenames.org/genetid code 'HGNC:11389' = 'HGNC:11389', given as 'STK11'))

component:

code: Gene studied [ID] (Details: (LOINC code '81252-9' = 'Gene studied [ID]'))

value: NC_000019.8:g.1171707G>A (Details: (http://www.ncbi.nlm.nih.gov/clinvar code '619728', given as 'NC_000019.8:g.1171707G>A'))

code: Discrete genetic variant (Details: (LOINC code '81290-9' = 'Discrete genetic variant'))

value: NC_000019.8:g.1171707G>A (Details: (http://www.ncbi.nlm.nih.gov/clinvar code '619728', given as 'NC_000019.8:g.1171707G>A'))

component:

code: Genomic DNA change (pHGVS) (Details: (LOINC code '69545-b' = 'Genomic DNA change (pHGVS)'))

value: NC_000019.8:g.1171707G>A (Details: (http://www.genenames.hgvs.org code 'NC_000019.8:g.1171707G>A'))

content:

http://hl7.org/fhir/us/mcode/Observation-mCODECancerGeneticVariantExample01.html
mCODE Genomics-related Preferred Terminologies

<table>
<thead>
<tr>
<th>mCODE Preferred</th>
<th>Examples</th>
<th>Other Alternatives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cancer type</strong></td>
<td>SNOMED, ICD-10-CM</td>
<td>ICD-O-3, OncoTree</td>
</tr>
<tr>
<td><strong>Gene</strong></td>
<td>HGNC ID</td>
<td>NCBI ID</td>
</tr>
<tr>
<td><strong>Variation Code</strong></td>
<td>ClinVar</td>
<td>dbVar, COSMIC</td>
</tr>
<tr>
<td>• DNA Change</td>
<td>HGVS Sequence Variant Nomenclature</td>
<td>• NC_000017.11:g.43037405_43048457del</td>
</tr>
<tr>
<td>• Amino Acid Change</td>
<td></td>
<td>• p.Leu858Arg</td>
</tr>
</tbody>
</table>
Example Persona
Eve Anyone

- Eve Anyone is a 67 year old female diagnosed with Stage IIIA Breast Cancer - ER positive, PR negative, HER2 positive).
- Genomics testing further revealed she was BRCA2 positive.
- She was started on trastuzumab on TBD date. After 12 weeks of treatment, results showed her condition was improving.
mCODE Representation
(Only relevant parts shown for illustration purposes)

CancerPatient

Patient
id = PatientId01
name.given = Eve
name.family = Anyone
birthDate = 1953-11-07
gender = female

PrimaryCancerCondition

Condition
id = CancerCondition01
valueCodeableConcept.code.coding.code = 372137005
valueCodeableConcept.code.coding.display = malignant neoplasm of breast
subject = Patient/PatientId01
bodySite.coding.code = 76752008
bodySite.coding.display = Breast Structure
stage.assessment = Observation/myStageGroup

CancerGenomicsReport

DiagnosticReport
id = CancerGenomicsReport01
valueCodeableConcept.coding.code = LA9633-4
valueCodeableConcept.coding.display = malignant neoplasm of breast
subject = Patient/PatientId01
specimen = Observation/GeneticSpecimen01
result:CancerGeneticVariant = Observation/CancerGeneticVariant01
result:CancerGeneticVariant = Observation/CancerGeneticVariant02

CancerGeneticVariant

Observation
id = CancerGeneticVariant01
subject = Patient/PatientId01
valueCodeableConcept.coding.code = LA9633-4
valueCodeableConcept.coding.display = malignant neoplasm of breast
component:genomicSourceClass.valueCodeableConcept.coding.code = LA6684-0
component:genomicSourceClass.valueCodeableConcept.coding.display = Somatic
collection.bodySite.coding.code = LA9633-4
component:genomicSourceClass.valueCodeableConcept.coding.display = Present
component:geneStudied.valueCodeableConcept.coding.code = HGNC:3430
component:geneStudied.valueCodeableConcept.coding.display = ERBB2
interpretation.coding.code = POS

CancerMedStatement

MedicationStatement
id = CancerMedStatement01
valueCodeableConcept.coding.code = LA9633-4
valueCodeableConcept.coding.display = malignant neoplasm of breast
subject = Patient/PatientId01
mcode-treatment-intent.code.coding.code = 1922516
mcode-treatment-intent.code.coding.display = trastuzumab 21 MG/ML Injectable Solution

TNMClinicalStageGroup

Observation
valueCodeableConcept.coding.code = c3A
collection.bodySite.coding.code = Stage IIIA
collection.bodySite.coding.display = TISS
collection.bodySite.coding.display = Tissue
collection.bodySite.coding.display = Breast Structure

GeneticSpecimen

Observation
id = GeneticSpecimen01
subject = Patient/PatientId01
type.coding.code = TISS
type.coding.display = Tissue
collection.bodySite.coding.code = 76752008
collection.bodySite.coding.display = Breast Structure

CancerRelatedMedicationStatement

MedicationStatement
id = CancerMedStatement02
valueCodeableConcept.coding.code = LA9633-4
valueCodeableConcept.coding.display = malignant neoplasm of breast
subject = Patient/PatientId01
mcode-treatment-intent.code.coding.code = 1922516
mcode-treatment-intent.code.coding.display = trastuzumab 21 MG/ML Injectable Solution

CancerDiseaseStatus

Observation
id = CancerDiseaseStatus01
subject = Patient/PatientId01
valueCodeableConcept.coding.code = 268910001
valueCodeableConcept.coding.display = Patient condition improved (finding)

Approved for Public Release; Distribution Unlimited. Case Number 16-1988
mCODE and the Patient Journey

**minimal depends on when mCODE data is accessed.**
mCODE and other Models
Models are *fit-for-purpose*...

**EHR Data**
- Patient / encounter point of care specific data
- Example Actors: hospital, ambulatory care, specialty services
- focus on direct patient care and workflow/continuity of care
- EHR-to-EHR data ideally minimized lossyness.

**Research Datasets**
- De-identified patient data & cohorts
- Example Actors: registries, pharma, research organizations
- focus on inbound data processing for aggregation and research data analytics
- inbound data detailed; outbound data is likely lossy
mCODE is a small subset of the HL7 CG Reporting IG
OMOP G-CDM

- Extending OMOP's Common Data Model with Genomics elements...

**CARE_SITE**
- Column: care_site_id
- Sample: 8200001
- care_site_name: Ajou University Hospital
- place_of_service: Department of Pathology

**PROCEDURE_OCCURRENCE**
- Column: procedure_occurrence_id
- Sample: 57966701
- procedure_date: 2018-10-01

**GENOMIC_TEST**
- Column: genomic_test_id
- Sample: 101
- genomic_test_name: GeneScreen
- reference_genome: GRCh37 (hg19)

**TARGET_GENE**
- Column: target_gene_id
- Sample: 1
- chromosome: 7

**VARIANT_OCCURRENCE**
- Column: variant_occurrence_id
- Sample: 16
- reference_sequence: NM_020975.4
- rs_id: rs1028345

**VARIANT_ANNOTATION**
- Column: variant_annotation_id
- Sample: 1
- annotation_database: GnomAD

**G-CMD**

**GENOMIC_TEST**
- Technical Specification of Genomic Test performed in the Site

**TARGET_GENE**
- Gene list targeted in the Test

**HGNC & HGVS**

**VARIANT_OCCURRENCE**
- Description of a Variant occurred in the gene

**G-CDM**

**VARIANT_ANNOTATION**
- Clinical Interpretation of the variant

Source: [https://raw.githubusercontent.com/OHDSI/Genomic-CDM/master/ERD.png](https://raw.githubusercontent.com/OHDSI/Genomic-CDM/master/ERD.png)
mCODE and OMOP G-CDM

- G-CDM are additional tables to the base OMOP Common Data Model (CDM)

<table>
<thead>
<tr>
<th>mCODEProfile</th>
<th>G-CDM Table</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CancerPatient</td>
<td>PERSON</td>
<td></td>
</tr>
<tr>
<td>PrimaryCancerCondition, ComorbidCondition</td>
<td>CONDITION_OCCURRENCE</td>
<td>Alignment with SNOMED</td>
</tr>
<tr>
<td>CancerRelatedSurgicalProcedure</td>
<td>PROCEDURE_OCCURRENCE</td>
<td></td>
</tr>
<tr>
<td>CancerGenomicsReport</td>
<td>GENOMIC_TEST</td>
<td></td>
</tr>
<tr>
<td>CancerGeneticVariant</td>
<td>VARIANT_OCCURRENCE</td>
<td>Variants align with HGVS</td>
</tr>
<tr>
<td>CancerGeneticVariant</td>
<td>TARGET_GENE</td>
<td>HGNC</td>
</tr>
<tr>
<td>GeneticSpecimen</td>
<td>SPECIMEN</td>
<td></td>
</tr>
<tr>
<td>CancerRelatedMedicationStatement</td>
<td>VARIANT_ANNOTATION</td>
<td>Alignment with RxNorm</td>
</tr>
</tbody>
</table>

![OMOP-CDM Diagram](image-url)
mCODE and OncoKB

• OncoKB API – access to curated actionable genes and possible drug treatments
• concept maps and translation logic needed

mCODE to OncoKB Mappings

<table>
<thead>
<tr>
<th>mCODE</th>
<th>OncoKB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>PrimaryCancerCondition (SNOMED / ICD-10-CM)</td>
<td>TumorType (OncoTree)</td>
<td>possible alignment via UMLS with relationship maps to SNOMED and ICD-10-CM</td>
</tr>
<tr>
<td>CancerGeneticVariant (HGNC)</td>
<td>Gene (HGNC)</td>
<td>HGNC code to HGNC symbol</td>
</tr>
<tr>
<td>CancerMedicationStatement (RxNorm)</td>
<td>Drugs (NCI Thesaurus)</td>
<td>possible alignment via UMLS with RxNorm TTY=IN</td>
</tr>
</tbody>
</table>
mCODE and the Patient Journey (Revisited)

Prescriptive
mCODE and the Patient Journey (Revisited)

Retrospective

Learning Healthcare System was it effective?
What's Next?
Additional Use Cases and Areas of Consideration

• Clinical significance

• Tumor mutation burden (TMB)
  • The total number of mutations (changes) found in the DNA of cancer cells.[1]

• Tumor-normal
  • tests performed on multiple specimens, one normal, one tumor; variants compared between both
  • use case scenario included in the HL7 May 2020 FHIR Virtual Connectathon - Clinical Genomics Track

• Fusion Genes
  • A gene made by joining parts of two different gene (e.g.: BCR-ABL1)
  • represented by multiple gene-studied components in CancerGeneticVariant
Areas of Consideration (cont'd)

• Re-alignment with current directions in HL7 Genomics Reporting IG.
  • Diagnostic implications
  • Code bindings to SequenceOntology for representing variant concepts

• Possible adjustments for research-oriented use cases?
  • better handling of de-identified patient data
  • cohort identification
Other Activities

• mCODE Clinical Genomics Supplemental Guide

• Build the mCODE Genomics Community
  • Feedback
  • Engagement
  • Contributions