Session III: Application and Authoring

Luke Rasmussen
Feinberg School of Medicine, Northwestern University

luke.rasmussen@northwestern.edu
@lrasmus

AMIA TBI/CRI 2015 Tutorial
Disclaimers

• I receive funding from:
  – NIH: NHGRI, NIGMS, NCATS
What have we learned?

• LOTS of potential in EHR data
  – Respect and understand (don’t fear) the data

• Growing set of tools and standards to help in this area
  – MAT, KNIME, i2b2, etc. – authoring UI
  – QDM / HQMF – represent logic
  – QRDA, FHIR, CIMI – represent data

• Collaboration is important, therefore transportability is important, therefore standards are important
Approach to Authoring

• Basic steps
  – Define the phenotype
  – Execute the phenotype
  – Validate the phenotype
  – Evaluate the results
  – Iterate

• Do this
  – Locally
  – One other place
  – Multiple other places

Need to think about transportability and standards before you progress, what data you want to ask for
Defining an Algorithm

Patients who have taken metformin after being diagnosed with T2DM

• Watch out for “blank canvas syndrome”
• Start with what you know
• Think about it in an iterative manner
• Don’t over-specify attributes
• Don’t under-specify attributes
eCQM Resources

• eCQI Resource Center

• QDM Specification
Use a Template

- **Category**
- **Code System**
  - Codes
  - Value Set Name
- **Data Type**
  - Data Type Attributes: (Depends on data type)
- **Data Flow Attributes:**
  - Health record field
  - Source
  - Recorder
Patient has T2DM

- Start with the Category
  - “What is this thing?”
- QDM Category:
  - Condition/Diagnosis/Problem

**Category**

- Code System
  - Codes
- Value Set Name

**Data Type**

- Data Type Attributes:
  - (Depends on data type)

**Data Flow Attributes:**

- Health record field
- Source
- Recorder
### Patient has T2DM

**Condition/Diagnosis/Problem**

<table>
<thead>
<tr>
<th>Code System</th>
<th>Codes</th>
<th>Value Set Name</th>
</tr>
</thead>
</table>

**Data Type**

| Data Type Attributes: | (Depends on data type) |

**Data Flow Attributes:**

- Health record field
- Source
- Recorder

- **Think about the data type**
  - “Do I want any mention?”
  - “Do I want the patient to still have it?”
- **QDM Data Type:**
  - Active
Patient has T2DM

- Think about how it’s identified
  - “What coding system(s)?”
  - “What’s a comprehensive list?”
  - “What’s historical & what’s the future?”

- Value Set:
  - Group contains ICD-9, ICD-10, SNOMED-CT
Value Set Authority Center

# Value Set Authority Center


<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Code System</th>
<th>Steward</th>
<th>OID</th>
</tr>
</thead>
<tbody>
<tr>
<td>diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>Grouping</td>
<td>ICD10CM, ICD9CM</td>
<td>NCQA</td>
<td>2.16.840.1.113883.3.464.1003.103.12.1001</td>
</tr>
<tr>
<td>Diabetes Visit</td>
<td>Grouping</td>
<td>CPT, SNOMEDCT</td>
<td>NCQA</td>
<td>2.16.840.1.113883.3.464.1003.103.12.1012</td>
</tr>
<tr>
<td>Gestational Diabetes</td>
<td>Grouping</td>
<td>ICD10CM, SNOMEDCT</td>
<td>NCQA</td>
<td>2.16.840.1.113883.3.464.1003.103.12.1010</td>
</tr>
<tr>
<td>Gestational Diabetes</td>
<td>Extension</td>
<td>ICD10CM</td>
<td>NCQA</td>
<td>2.16.840.1.113883.3.464.1003.103.11.1014</td>
</tr>
</tbody>
</table>

# Value Set Authority Center

**Metadata**

- **Name:** Diabetes
- **Type:** Extensional
- **Steward:** National Committee for Quality Assurance

**OID:**

2.16.840.1.113883.3.464.1003.103.11.1002

**Definition ID:**

20140501

**Program:**

CMS, MU2 EP Update 2014-07-01 using this value set

## Value Set Members

**Expansion ID:** 20121025

### Expanded Code List

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>E10.10</td>
<td>Type 1 diabetes mellitus with ketoacidosis without coma</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
<tr>
<td>E10.11</td>
<td>Type 1 diabetes mellitus with ketoacidosis with coma</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
<tr>
<td>E10.21</td>
<td>Type 1 diabetes mellitus with diabetic nephropathy</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
<tr>
<td>E10.22</td>
<td>Type 1 diabetes mellitus with diabetic chronic kidney disease</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
<tr>
<td>E10.29</td>
<td>Type 1 diabetes mellitus with other diabetic kidney complication</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
<tr>
<td>E10.311</td>
<td>Type 1 diabetes mellitus with unspecified diabetic retinopathy with macular edema</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
<tr>
<td>E10.319</td>
<td>Type 1 diabetes mellitus with unspecified diabetic retinopathy without macular edema</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
<tr>
<td>E10.321</td>
<td>Type 1 diabetes mellitus with mild nonproliferative diabetic retinopathy with macular</td>
<td>ICD10CM</td>
<td>2014</td>
<td>2.16.840.1.113883.6.90</td>
</tr>
</tbody>
</table>

**Value Set Center**

Patient has T2DM

- Think about other specific attributes you care about
  - “Want things that didn’t happen?”
  - “Was it the principle diagnosis?”

- Attribute:
  - Ordinality (may be too restrictive)
Patient has T2DM

- Think about where the data came from
  - “Does it have to be physician-reported?”
  - “Do I care where it’s recorded?”

- Attribute:
  - Recorder
Patient has T2DM as Dx by Provider

Condition/Diagnosis/Problem

Diabetes
ICD9CM, ICD10CM, SNOMEDCT
2.16.840.1.113883.3.464.1003.103.1
2.1001

Diagnosis, Active

Data Type Attributes:

Data Flow Attributes:
Health record field
Source
Recorder: provider
Patient is actively taking metformin

Searching the VSAC returns no results. Don’t panic! Important part is defining the list of codes that will be used.

- **Medication**
  - RXNORM
    - 6809, 235743...
      - (Need to create one...)

- **Medication, Active**

- **Data Type Attributes:**

- **Data Flow Attributes:**
  - Health record field
  - Source
  - Recorder
Patients taking metformin after T2DM diagnosis

**Condition/Diagnosis/Problem**
- Code System
- Codes
- Value Set Name
- Diagnosis, Active
- Data Type Attributes:
- Data Flow Attributes:
  - Health record field
  - Source
  - Recorder: provider

**Medication**
- RXNORM
- 6809, 235743...
- (Need to create one...)
- Medication, Active
- Data Type Attributes:
- Data Flow Attributes:
  - Health record field
  - Source
  - Recorder
Ambiguity in Temporality

• Something happens “after”
• How is the data recorded?
  – Start time
  – End time
• Importance of event having ended
• Is end time reliably recorded?
• Consider “concurrency” too
Patients taking metformin after T2DM diagnosis

**Condition/Diagnosis/Problem**
- Diabetes
  - ICD9CM, ICD10CM, SNOMEDCT
  - 2.16.840.1.113883.3.464.1003.103.1
  - 2.1001
- Diagnosis, Active
- Data Type Attributes:
- Data Flow Attributes:
  - Health record field
  - Source
  - Recorder: provider

**Medication**
- RXNORM
  - 6809, 235743...
  - (Need to create one...)
- Medication, Active
- Data Type Attributes:
- Data Flow Attributes:
  - Health record field
  - Source
  - Recorder

Starts After or Concurrent With Start Of
Demo – Diabetes Example
Sharing the Results

• “Standards were meant to be re-invented”
• Data Dictionaries and Data Validation
  – PheKB
  – Local instance
Possible Impressions So Far

• “Seems like a lot of busy work”
• “Those concepts don’t match the way I think about the logic”
• “I could make a better standard”
• “Nobody would want to use my algorithm”
• “I can write it faster in [SQL, SAS, R]”
For Your Consideration

• “Seems like a lot of busy work”
  – It is admittedly more work
  – There’s a fixed cost to doing this – question is if you pay the cost in standardizing definition, there is downstream gain

• “Those concepts don’t match the way I think about the logic”
  – Like with new programming languages and paradigms, there is a learning curve
  – Need examples, tutorials and a support community
For Your Consideration

• “I could make a better standard”
  – Please don’t...

• “Nobody would want to use my algorithm”
  – You’d be surprised!
  – Post it to PheKB now and see

• “I can write it faster in [SQL, SAS, R]”
  – Encourage developing phenotypes in your favorite tool set
  – Standard definitions helpful at the end
Questions?
Hands On
Scenario

• [http://www.projectphema.org:8081](http://www.projectphema.org:8081)
• All patients need to have taken or be taking a statin.
  – Since not all patients may have a record confirming that they received the drug, you will look for patients who have had an order or prescription for a statin, or are known to have received it.
• For the patients that have had one or more of these statin medications, you also want to make sure they have had their HDL measured both before and after the first record of the medication.
  – You want to make sure they have had at least one HDL result any time before the first statin.
  – Next, you look for patients that have had at least one HDL result within the first year after the first statin.
• Finally, you want to see patients that did NOT have a MACE within 5 years after their first statin. You will be looking for patients that have had a cardiac procedure performed within 5 years after the first statin, or those that have a diagnosis recorded of a myocardial infarction (heart attack) within 5 years after the first statin. Then you will want to exclude those patients.