OMOP – PCORnet Alignment

Rimma Belenkaya
26-April-2016
Rimma.Belenkaya@gmail.com
Outline

• History of OMOP adoption for CDRN
• Development of crosswalk
  ✓ Approach
  ✓ Challenges
  ✓ Outcomes
  ✓ Remaining tasks
History of OMOP adoption for CDRN

• CDRN—Clinical Data Research Network
• PCORnet - the National Patient-Centered Clinical Research Network
  ✓ Initiated in 2014
  ✓ Funded by the Patient-Centered Outcomes Research Institute (PCORI)
  ✓ Integrates data from 13 networks for CER
  ✓ PCORnet CDM is based on the Mini-Sentinel CDM

• Out of initial 11 networks
  ✓ 7 adopted i2b2
  ✓ 4 adopted OMOP
  ✓ 0 adopted PCORnet
# PCORNet CDM V1

**PCORnet Common Data Model, Original v1.0**

**Fundamental basis**

<table>
<thead>
<tr>
<th>DEMOGRAPHIC</th>
<th>VITAL</th>
<th>ENROLLMENT</th>
<th>DIAGNOSIS</th>
<th>PROCEDURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>PATID</td>
<td>PATID</td>
<td>PATID</td>
<td>PATID</td>
<td>PATID</td>
</tr>
<tr>
<td>BIRTH_DATE</td>
<td>ENCRYPTED_DATE</td>
<td>ENR_START_DATE</td>
<td>ENCOUNTERID</td>
<td>ENCOUNTERID</td>
</tr>
<tr>
<td>BIRTH_TIME</td>
<td>MEASURE_DATE</td>
<td>ENR_END_DATE</td>
<td>ENC_TYPE</td>
<td>ENC_TYPE</td>
</tr>
<tr>
<td>SEX</td>
<td>MEASURE_TIME</td>
<td>CHART</td>
<td>Admit_DATE</td>
<td>Admit_DATE</td>
</tr>
<tr>
<td>HISPANIC</td>
<td>VITAL_SOURCE</td>
<td>ENR_BASIS</td>
<td>PROVIDERID</td>
<td>PROVIDERID</td>
</tr>
<tr>
<td>RACE</td>
<td>HT</td>
<td>PATID</td>
<td>FACILITYID</td>
<td>FACILITYID</td>
</tr>
<tr>
<td>BIOBANK_FLAG</td>
<td>WT</td>
<td>ENCOUNTERID</td>
<td>DISCHARGE_DISPOSITION</td>
<td>DISCHARGE_DISPOSITION</td>
</tr>
</tbody>
</table>

**Data captured from processes associated with healthcare delivery**

Data captured within multiple contexts: healthcare delivery, registry activity, or directly from patients

**Data captured from healthcare delivery, direct encounter basis**
OMOP CDM v4
## Rationale for Choosing OMOP CDM

<table>
<thead>
<tr>
<th></th>
<th>C32 CCD</th>
<th>MiniSentinel</th>
<th>OMOP V5</th>
<th>PCORI CDM V1</th>
<th>SCILHS PCORI SHRINE</th>
<th>HARVARD SHRINE</th>
<th>NYC DMH QUERY HEALTH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient Demographics</strong></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>Provider</strong></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Encounters</strong></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Encounter cost</strong></td>
<td>O</td>
<td>DRG</td>
<td>O</td>
<td>DRG</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Locations</strong></td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Care sites</strong></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Organizations</strong></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>Reason for encounter</strong></td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Problem List(Condition)</strong></td>
<td>X (SNOMED CT, ICD-*; CM)</td>
<td>X (SNOMED CT; MedDRA)</td>
<td>X (Snomed CT)</td>
<td>X (via relevant_condition_concept_id)</td>
<td></td>
<td>X (ICD-9-CM)</td>
<td>X (ICD-9-CM)</td>
</tr>
<tr>
<td><strong>Encounter Diagnoses</strong></td>
<td>X (ICD-9-CM; SNOMED CT; ICD-10-CM)</td>
<td>X (ICD-9-CM; SNOMED CT; ICD-10-CM)</td>
<td>X (ICD-9-CM; SNOMED CT; ICD-10-CM)</td>
<td>X (ICD-9-CM; SNOMED CT; ICD-10-CM)</td>
<td></td>
<td>X (ICD-9-CM)</td>
<td>X (ICD-9-CM)</td>
</tr>
<tr>
<td><strong>Principal Dx</strong></td>
<td>X</td>
<td>X</td>
<td>O</td>
<td>O</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Cause of death</strong></td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>O</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Order diagnosis</strong></td>
<td>X (SNOMED CT)</td>
<td>X (SNOMED CT)</td>
<td>X</td>
<td>X</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>Family History</strong></td>
<td>X</td>
<td>O</td>
<td>X</td>
<td>O</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>Social History</strong></td>
<td>O</td>
<td>X (SNOMED CT)</td>
<td>X</td>
<td>O</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Review of systems</strong></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>Alerts/Allergies</strong></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Vital Signs</strong></td>
<td>X</td>
<td>X (SNOMED CT, RxNORM)</td>
<td>X (SNOMED CT)</td>
<td>X (LOINC)</td>
<td></td>
<td>X (LOINC)</td>
<td>X (LOINC)</td>
</tr>
<tr>
<td><strong>Physical findings</strong></td>
<td>X</td>
<td>X (SNOMED CT, LOINC)</td>
<td>X (LOINC, SNOMED CT)</td>
<td>X (some, in progress)</td>
<td></td>
<td>X (LOINC)</td>
<td>X (LOINC)</td>
</tr>
<tr>
<td><strong>Lab Results</strong></td>
<td>X (LOINC)</td>
<td>X (LOINC, Specimen)</td>
<td>X (LOINC,UCUM)</td>
<td>X (LOINC)</td>
<td></td>
<td>X (LOINC) X (LOINC)</td>
<td>X (LOINC, NQF)</td>
</tr>
<tr>
<td><strong>Radiology results</strong></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td></td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td><strong>Medication history</strong></td>
<td>X (RxNORM)</td>
<td>X (RxNORM, CPT, Specimen)</td>
<td>X (RxNORM; NDF-RT)</td>
<td>X (RxNORM, NDF-RT)</td>
<td></td>
<td>X(NDC,NF,RxNORM)</td>
<td>X (NDC, NQF, RXNORM)</td>
</tr>
<tr>
<td><strong>Medications prescribed(output)</strong></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>Administered</strong></td>
<td>X</td>
<td>O</td>
<td>O</td>
<td>0</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Rx dispensed</strong></td>
<td>O</td>
<td>X</td>
<td>O</td>
<td>0</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Medication costs</strong></td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>0</td>
<td></td>
<td>O</td>
<td>X(NDC)</td>
</tr>
</tbody>
</table>
Rationale for Choosing OMOP CDM

• **Data model** - more reflective of clinical domain, granular, better structured
• **Vocabulary** - uniformly structured and well curated
• **Information Model** - formalized connection between data model and conceptual model (vocabulary)
• **Supportive Community** of developers and researches
• **Tools** for data integration, characterization, and analysis
Development of CDM Crosswalk

Step 1. Addressing PCORnet Requirements in OMOP CDM v5

Step 2. Establishing interoperability standards
**PCORnet requirements in OMOP CDM v5**

✓ New attributes and concepts added to Person, Visit, and Observation Period domains

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Required</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation_period_id</td>
<td>integer</td>
<td>Yes</td>
<td>A system-generate unique identifier for each observation period</td>
</tr>
<tr>
<td>person_id</td>
<td>integer</td>
<td>Yes</td>
<td>A foreign key identifier to the person who is experiencing the condition. The demographic details of that person are stored in the person table.</td>
</tr>
<tr>
<td>observation_period_start_date</td>
<td>date</td>
<td>Yes</td>
<td>The start date of the observation period for which data are available from the data source</td>
</tr>
<tr>
<td>observation_period_end_date</td>
<td>date</td>
<td>Yes</td>
<td>The end date of the observation period for which data are available from the source.</td>
</tr>
<tr>
<td>period_type_concept_id</td>
<td>integer</td>
<td>Yes</td>
<td>A foreign key identifier to the predefined concept in the Standardized Vocabularies reflecting the source of the observation period information</td>
</tr>
</tbody>
</table>

---

44814722  Period while enrolled in insurance'
44814723  Geography based period'
44814725  Period inferred by algorithm'
44814724  Period covering healthcare encounters'
PCORnet requirements in OMOP CDM v5

✔ Observations and fact relationships used as a workaround for unsupported attributes

<table>
<thead>
<tr>
<th>VISIT_OCCURRENCE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Field</td>
<td>Value</td>
<td></td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>555555</td>
<td></td>
</tr>
<tr>
<td>person_id</td>
<td>804860</td>
<td></td>
</tr>
<tr>
<td>visit_start_date</td>
<td>10/10/2012</td>
<td></td>
</tr>
<tr>
<td>visit_start_time</td>
<td>23:00</td>
<td></td>
</tr>
<tr>
<td>visit_end_date</td>
<td>10/14/2012</td>
<td></td>
</tr>
<tr>
<td>visit_end_time</td>
<td>14:15</td>
<td></td>
</tr>
<tr>
<td>provider_id</td>
<td>666</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OBSERVATION</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Field</td>
<td>Value</td>
<td></td>
</tr>
<tr>
<td>observation_id</td>
<td>238923849</td>
<td></td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>555555</td>
<td></td>
</tr>
<tr>
<td>person_id</td>
<td>804860</td>
<td></td>
</tr>
<tr>
<td>observation_date</td>
<td>10/14/2012</td>
<td></td>
</tr>
<tr>
<td>observation_time</td>
<td>14:15</td>
<td></td>
</tr>
<tr>
<td>provider_id</td>
<td>666</td>
<td></td>
</tr>
<tr>
<td>observation_concept_id</td>
<td>4137274</td>
<td>Discharge to establishment</td>
</tr>
<tr>
<td>value_as_concept_id</td>
<td>44814680</td>
<td>Residential Facility</td>
</tr>
</tbody>
</table>

...
PCORnet requirements in OMOP CDM v5

✓ PCORnet vocabulary created and mapped to standard vocabulary

<table>
<thead>
<tr>
<th>CONCEPT_ID</th>
<th>CONCEPT_NAME</th>
<th>CONCEPT_LEVEL</th>
<th>CONCEPT_CLASS</th>
<th>VOCABULARY_ID</th>
<th>CONCEPT_CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>44814649</td>
<td>Other</td>
<td>0</td>
<td>Generic</td>
<td>PCORNet</td>
<td>44814649-NI</td>
</tr>
<tr>
<td>44814650</td>
<td>No information</td>
<td>0</td>
<td>Generic</td>
<td>PCORNet</td>
<td>44814650-UN</td>
</tr>
<tr>
<td>44814653</td>
<td>Unknown</td>
<td>0</td>
<td>Generic</td>
<td>PCORNet</td>
<td>44814653-OT</td>
</tr>
<tr>
<td>44814647</td>
<td>Available in biobank</td>
<td>0</td>
<td>Biobank Flag</td>
<td>PCORNet</td>
<td>44814647-Y</td>
</tr>
<tr>
<td>44814648</td>
<td>Unavailable in biobank</td>
<td>0</td>
<td>Biobank Flag</td>
<td>PCORNet</td>
<td>44814648-N</td>
</tr>
<tr>
<td>44814651</td>
<td>Hispanic</td>
<td>0</td>
<td>Hispanic</td>
<td>PCORNet</td>
<td>44814651-Y</td>
</tr>
<tr>
<td>44814652</td>
<td>Not Hispanic</td>
<td>0</td>
<td>Hispanic</td>
<td>PCORNet</td>
<td>44814652-N</td>
</tr>
<tr>
<td>44814654</td>
<td>American Indian or Alaska Native</td>
<td>0</td>
<td>Race</td>
<td>PCORNet</td>
<td>44814654-01</td>
</tr>
<tr>
<td>44814655</td>
<td>Asian</td>
<td>0</td>
<td>Race</td>
<td>PCORNet</td>
<td>44814655-02</td>
</tr>
<tr>
<td>44814656</td>
<td>Black or African American</td>
<td>0</td>
<td>Race</td>
<td>PCORNet</td>
<td>44814656-03</td>
</tr>
<tr>
<td>44814657</td>
<td>Native Hawaiian or Other Pacific Islander</td>
<td>0</td>
<td>Race</td>
<td>PCORNet</td>
<td>44814657-04</td>
</tr>
<tr>
<td>44814658</td>
<td>White</td>
<td>0</td>
<td>Race</td>
<td>PCORNet</td>
<td>44814658-05</td>
</tr>
<tr>
<td>44814659</td>
<td>Multiple race</td>
<td>0</td>
<td>Race</td>
<td>PCORNet</td>
<td>44814659-06</td>
</tr>
<tr>
<td>44814660</td>
<td>Refuse to answer</td>
<td>0</td>
<td>Race</td>
<td>PCORNet</td>
<td>44814660-07</td>
</tr>
<tr>
<td>44814664</td>
<td>Ambiguous</td>
<td>0</td>
<td>Sex</td>
<td>PCORNet</td>
<td>44814664-A</td>
</tr>
<tr>
<td>44814665</td>
<td>Female</td>
<td>0</td>
<td>Sex</td>
<td>PCORNet</td>
<td>44814665-F</td>
</tr>
<tr>
<td>44814666</td>
<td>Male</td>
<td>0</td>
<td>Sex</td>
<td>PCORNet</td>
<td>44814666-M</td>
</tr>
<tr>
<td>44814670</td>
<td>Adult foster home</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814670-AF</td>
</tr>
<tr>
<td>44814671</td>
<td>Assisted living facility</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814671-AL</td>
</tr>
<tr>
<td>44814672</td>
<td>Ambulatory visit</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814672-AV</td>
</tr>
<tr>
<td>44814673</td>
<td>Emergency department</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814673-ED</td>
</tr>
<tr>
<td>44814674</td>
<td>Home health</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814674-HH</td>
</tr>
<tr>
<td>44814675</td>
<td>Home / self care</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814675-HO</td>
</tr>
<tr>
<td>44814676</td>
<td>Hospice</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814676-HS</td>
</tr>
<tr>
<td>44814677</td>
<td>Other acute inpatient hospital</td>
<td>0</td>
<td>Admitting Source</td>
<td>PCORNet</td>
<td>44814677-IP</td>
</tr>
</tbody>
</table>
CDM Interoperability Collaborative

• Core members
  • Rimma Belenkaya
  • Rob Follett
  • Ritu Khare
  • Mark Khayter
  • Toan Ong
  • Lisa Schilling
  • Don Torok

• Weekly meetings

• Deliverables: standard specifications
Approach

• Satisfy PCORnet requirements
• Preserve granularity of source data
• Preserve integrity of OMOP CDM
• Enforce uniformity of workarounds

Crosswalk from source – to – OMOP CDM – to – PCORNet CDM

OMOP CDM Population Conventions

Patient -> Person

PCORnet CDM ETL Standards

Person -> Demographic

Source EHR data

OMOP Data Model

PCORnet Data Model

☑ Convert data in the OMOP CDM to PCORnet CDM via a set of mechanistic transformation rules
Examples

✓ Extension of OMOP vocabulary to preserve null-values

<table>
<thead>
<tr>
<th>NULL-values</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>A data field is not present in the source system</td>
</tr>
<tr>
<td>No Information</td>
<td>A data field is present in the source system, but the source value is null or blank</td>
</tr>
<tr>
<td>Unknown</td>
<td>A data field is present in the source system, but the source value explicitly denotes an unknown value</td>
</tr>
<tr>
<td>Other</td>
<td>A data field is present in the source system, but the source value cannot be mapped to the CDM</td>
</tr>
</tbody>
</table>
Examples

✓ Convention for enforcing PCORnet referential integrity and cardinality
✓ Preserving integrity of OMOP CDM
✓ Enforcing uniformity of workarounds

<table>
<thead>
<tr>
<th>VISIT_OCCURRENCE</th>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>visit_occurrence_id</td>
<td>555555</td>
<td></td>
</tr>
<tr>
<td>person_id</td>
<td>804860</td>
<td></td>
</tr>
<tr>
<td>visit_start_date</td>
<td>10/10/2012</td>
<td></td>
</tr>
<tr>
<td>visit_start_time</td>
<td>23:00</td>
<td></td>
</tr>
<tr>
<td>visit_end_date</td>
<td>10/14/2012</td>
<td></td>
</tr>
<tr>
<td>visit_end_time</td>
<td>14:15</td>
<td></td>
</tr>
<tr>
<td>provider_id</td>
<td>666</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OBSERVATION</th>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>observation_id</td>
<td>238923849</td>
<td></td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>555555</td>
<td></td>
</tr>
<tr>
<td>person_id</td>
<td>804860</td>
<td></td>
</tr>
<tr>
<td>observation_date</td>
<td>10/14/2012</td>
<td></td>
</tr>
<tr>
<td>observation_time</td>
<td>14:15</td>
<td></td>
</tr>
<tr>
<td>provider_id</td>
<td>666</td>
<td></td>
</tr>
<tr>
<td>observation_concept_id</td>
<td>4137274</td>
<td>Discharge to establishment</td>
</tr>
<tr>
<td>value_as_concept_id</td>
<td>44814680</td>
<td>Residential Facility</td>
</tr>
</tbody>
</table>

Only one record with the same observation_concept_id per visit occurrence
Examples

✔ Convention for enforcing data completeness

<table>
<thead>
<tr>
<th>VISIT_OCCURRENCE</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field</td>
<td>Value</td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>555555</td>
</tr>
<tr>
<td>person_id</td>
<td>804860</td>
</tr>
<tr>
<td>visit_start_date</td>
<td>10/10/2012</td>
</tr>
<tr>
<td>visit_start_time</td>
<td>23:00</td>
</tr>
<tr>
<td>visit_end_date</td>
<td>10/14/2012</td>
</tr>
<tr>
<td>visit_end_time</td>
<td>14:15</td>
</tr>
<tr>
<td>provider_id</td>
<td>666</td>
</tr>
<tr>
<td>Inpatient visit</td>
<td>visit_concept_id</td>
</tr>
<tr>
<td>….</td>
<td>OBSESSION</td>
</tr>
<tr>
<td>Field</td>
<td>Value</td>
</tr>
<tr>
<td>observation_id</td>
<td>238923849</td>
</tr>
<tr>
<td>visit_occurrence_id</td>
<td>555555</td>
</tr>
<tr>
<td>person_id</td>
<td>804860</td>
</tr>
<tr>
<td>observation_date</td>
<td>10/14/2012</td>
</tr>
<tr>
<td>observation_time</td>
<td>14:15</td>
</tr>
<tr>
<td>provider_id</td>
<td>666</td>
</tr>
<tr>
<td>observation_concept_id</td>
<td>4137274</td>
</tr>
<tr>
<td>value_as_concept_id</td>
<td>44814680</td>
</tr>
<tr>
<td>….</td>
<td></td>
</tr>
</tbody>
</table>

*Inpatient visits must have ‘Discharge to establishment’*
Examples

Granularity of source data preserved

- More granular Race representation
- Multiple encounters to the same provider on the same date
- Storing diagnosis date/time different from encounter date/time
Challenges

• **Missing attributes** - workarounds, overhead in ETL
• **Missing concepts (‘NULL’ values)** - necessity to extend vocabulary
• **Lab mappings** – PCORnet targets not defined
• **PCORnet requirement to provide source values** – domain mismatch
• **Different versions** of OMOP and PCORnet CDM, necessity to reconcile
• **Time consuming** process of creating conventions
Outcomes

- **Interoperability standards** - from source (i.e. electronic health records) to OMOP CDM v4 (Specific to PCORI CDM) and from source to OMOP CDM v5 (specific to PCORI CDM)
- **A conventions document** - for populating the OMOP CDM, and an extract-transform-load specifications document for transforming to the PCORnet CDM v1, for both OMOP CDM v4 and OMOP CDM v5. The documents will be publicly available for the OHDSI community
- **OMOP CDM extended** – new important use cases
  - Added visit time
  - ADT – admission/discharge handling
  - Tobacco conventions – underway
- **OMOP CDM supports CDRN phenotypes** – ADAPTABLE study
Evolution of PCORnet CDMs

**PCORnet Common Data Model, Draft v2.0 Modifications**

**Fundamental basis**
- DEMOGRAPHIC
  - PATID
  - BIRTH_DATE
  - BIRTH_TIME
  - SEX
  - HISPANIC
  - RACE
  - BIOBANK_FLAG

**Data captured from processes associated with healthcare delivery**
- ENROLLMENT
  - PATID
  - ENR_START_DATE
  - ENR_END_DATE
  - CHART
  - ENR_BASIS

- DISPENSING
  - PATID
  - RX_DATE
  - NDC
  - RX_SUP
  - RX_AMT

**Data captured within multiple contexts: healthcare delivery, registry activity, or directly from patients**
- VITAL
  - PATID
  - ENCOUNTERID (optional)
  - MEASURE_DATE
  - MEASURE_TIME
  - VITAL_SOURCE
  - HT
  - WT
  - DIASTOLIC
  - SYSTOLIC
  - ORIGINAL_BMI
  - BP_POSITION
  - TOBACCO
  - TOBACCO_TYPE

- CONDITION
  - PATID
  - ENCOUNTERID (optional)
  - REPORT_DATE
  - RESOLVE_DATE
  - CONDITION_STATUS
  - CONDITION
  - CONDITION_TYPE
  - CONDITION_SOURCE

- DIAGNOSIS
  - PATID
  - ENCOUNTERID
  - ENC_TYPE (replicated)
  - ADMIT_DATE (replicated)
  - PROVIDERID (replicated)
  - DX
  - DX_TYPE
  - DX_SOURCE
  - PDX

- PRO_CM
  - PATID
  - ENCOUNTERID (optional)
  - CM_ITEM
  - CM_LOINC
  - CM_DATE
  - CM_TIME
  - CM_RESPONSE
  - CM_METHOD
  - CM_MODE
  - CM_CAT

- ENCRYPTED
  - PATID
  - SITID
  - ADMIT_DATE
  - ADMIT_TIME
  - DISCHARGE_DATE
  - DISCHARGE_TIME
  - PROVIDERID
  - FACILITYID
  - FACILITY_LOCATION
  - DISCHARGE_DISPOSITION
  - DISCHARGE_STATUS
  - DRG
  - DRG_TYPE
  - ADMITTING_SOURCE

- LAB CM RESULT
  - PATID
  - ENCOUNTERID (optional)
  - LAB_NAME
  - SPECIMEN_SOURCE
  - SPECIMEN_DATE
  - SPECIMEN_TIME
  - RESULT_DATE
  - RESULT_TIME
  - RESULT_NUM
  - RESULT_MODIFIER
  - RESULT_UNITS
  - MODIFIER_LOW
  - NORM_RANGE_LOW
  - NORM_RANGE_HIGH
  - ABN_IND

- PROCEDURE
  - PATID
  - ENCOUNTERID
  - ENC_TYPE (replicated)
  - ADMIT_DATE (replicated)
  - PROVIDERID (replicated)
  - PX_DATE
  - PX
  - PX_TYPE

New to v2.0

**New to v2.0**
- PATID
- BIRTH_DATE
- BIRTH_TIME
- SEX
- HISPANIC
- RACE
- BIOBANK_FLAG
- PATID
- ENR_START_DATE
- ENR_END_DATE
- CHART
- ENR_BASIS
- PATID
- RX_DATE
- NDC
- RX_SUP
- RX_AMT
- PATID
- SITID
- PATID
- ENCOUNTERID
- MEASURE_DATE
- MEASURE_TIME
- VITAL_SOURCE
- HT
- WT
- DIASTOLIC
- SYSTOLIC
- ORIGINAL_BMI
- BP_POSITION
- TOBACCO
- TOBACCO_TYPE
- PATID
- ENCOUNTERID (optional)
- REPORT_DATE
- RESOLVE_DATE
- CONDITION_STATUS
- CONDITION
- CONDITION_TYPE
- CONDITION_SOURCE
- PATID
- ENCOUNTERID
- ENC_TYPE (replicated)
- ADMIT_DATE (replicated)
- PROVIDERID (replicated)
- DX
- DX_TYPE
- DX_SOURCE
- PDX
- PATID
- SITID
- PATID
- ENCOUNTERID
- MEASURE_DATE
- MEASURE_TIME
- VITAL_SOURCE
- HT
- WT
- DIASTOLIC
- SYSTOLIC
- ORIGINAL_BMI
- BP_POSITION
- TOBACCO
- TOBACCO_TYPE
- PATID
- ENCOUNTERID (optional)
- REPORT_DATE
- RESOLVE_DATE
- CONDITION_STATUS
- CONDITION
- CONDITION_TYPE
- CONDITION_SOURCE
- PATID
- ENCOUNTERID
- ENC_TYPE (replicated)
- ADMIT_DATE (replicated)
- PROVIDERID (replicated)
- DX
- DX_TYPE
- DX_SOURCE
- PDX
- PATID
- SITID
- PATID
- ENCOUNTERID
- MEASURE_DATE
- MEASURE_TIME
- VITAL_SOURCE
- HT
- WT
- DIASTOLIC
- SYSTOLIC
- ORIGINAL_BMI
- BP_POSITION
- TOBACCO
- TOBACCO_TYPE
- PATID
- ENCOUNTERID (optional)
- REPORT_DATE
- RESOLVE_DATE
- CONDITION_STATUS
- CONDITION
- CONDITION_TYPE
- CONDITION_SOURCE
- PATID
- ENCOUNTERID
- ENC_TYPE (replicated)
- ADMIT_DATE (replicated)
- PROVIDERID (replicated)
- PX_DATE
- PX
- PX_TYPE
Evolution of PCORnet CDMs

PCORnet Common Data Model v3.0

- Demographic: PATID, BIRTH_DATE, BIRTH_TIME, SEX, RACE, HISPANIC, LANGUAGE
- Vital: VITALID, PATID, ENCOUNTERID (optional), MEASURE_DATE, MEASURE_TIME, VITAL_SOURCE, HT, WT, DX, SISTOC, SYSTOC, OXGENAL_BMI, BP, POSITION, SMOKE, TROACCO, TROACCO_TYPE
- Enrollment: PATID, ENR_START_DATE, ENR_END_DATE, CHART, ENR_BASIS
- Dispensing: PATID, PRESCRIPTIONID (optional), DISPENSE_DATE, DRTC, DISPENSE_SUP, DISPENSE_ANT
- Condition: PATID, ENCOUNTERID (optional), REPORT_DATE, RESOLVE_DATE, ONSET_DATE, CONDITION_STATUS, CONDITION_TYPE, CONDITION_SOURCE
- Pro-CM: PRO-CM_ID, PATID, ENCOUNTERID (optional), PRO_ITEM, PRO_HIGH, PRO_DATE, PRO_RESPONSE, PRO_METHOD, PRO_MODE, PRO_CAT
- Procedure: PROCESSES_ID, PATID, ENCOUNTERID (optional), RX_PROVIDER, RX_ORDER_DATE, RX_ORDER_TIME, RX_START_DATE, RX_END_DATE, RX_QUANTITY, RX_REFILLS, RX_DAYS_SUPPLY, RX_FREQUENC, RX_BETAS, RXNOMC_CUI
- Diagnosis: PATID, ENCOUNTERID, ADMDT_DATE (replicated), PROVIDERID (replicated), DX
- Lab Result CM: LAB_RESULT_CM_ID, PATID, ENCOUNTERID (optional), Lab_NAME, LAB_SOURCE, LAB_NAME, SPECIMEN_SOURCE, LAB_LOINC, RESULT_UNITS, RESULT_RANGE_LOW, RESULT_MODIFIER, RESULT_HIGH, RESULT_MODIFIER_LOW, RESULT_MODIFIER_HIGH, ABN_IND
- Prescribing: PRESCRIPTIONID, PATID, ENCOUNTERID (optional), RX_PROVIDER, RX_ORDER_DATE, RX_ORDER_TIME, RX_START_DATE, RX_END_DATE, RX_QUANTITY, RX_REFILLS, RX_DAYS_SUPPLY, RX_FREQUENC, RX_BETAS, RXNOMC_CUI
- Clinical Trials: PATID, TRIALID, PARTICIPANTID, TRAIL_SITEID, TRAIL_ENROLL_DATE, TRAIL_END_DATE, TRAIL_WITHDRAW_DATE, TRAIL_INVITE_CODE

Data captured from processes associated with healthcare delivery, registry activity, or directly from patients.

Data captured from healthcare delivery, direct encounter basis.

Process-related data

Associations with PCORnet clinical trials.
Remaining Tasks

• Complete OMOP v5 to PCORnet v3 CDM standards:
  – Death
  – Reconciliation between PCORnet v1, v2, and v3

• Make ETL code available in GitHub (pSCANNER)
Conclusions

- **Greater granularity of data representation** in OMOP CDM supports
  - Preservation of source data granularity
  - Straightforward transformation to less granular PCORnet

- **OMOP CDM is extendable** without breaking the CDM structure
  - Attributes that do not exist in OMOP can be represented without altering the standard table schema but require a set of documented conventions
  - OMOP CDM can support phenotype definitions not supported in PCORnet CDM
  - Adding non-standard/local concepts cover for missing vocabulary

- **OMOP CDM is a good intermediary data representation** when converting various healthcare datasets to PCORnet CDM
Great collaborative experience

NYC-CDRN
Parsa Mirhaji
Tom Campion
Rajan Chandras
Claudia Pulgarin
Ezra Fass
Karthik Natarajan
Uday Evani
Adler Perotte
Joan Leavey
Jim Singer

Patrick Ryan
Chris Reich

Mark Khayter
Don Torok
Lisa Schilling
Rob Follett
Daniella Meeker

Ritu Khare
Toan Ong
Michael Kahn

Shelley Rusincovitch

OHDSI
PCORN
PSANCER

The National Patient-Centered Clinical Research Network