Development of a National Distributed Research Network Data Infrastructure: Design of the PCORnet Common Data Model

Jeffrey S. Brown, PhD[1], Shelley A. Rusincovitch[2], Abel N. Kho, MD, MS[3], Keith Marsolo, PhD[4], and Lesley H. Curtis, PhD[5]
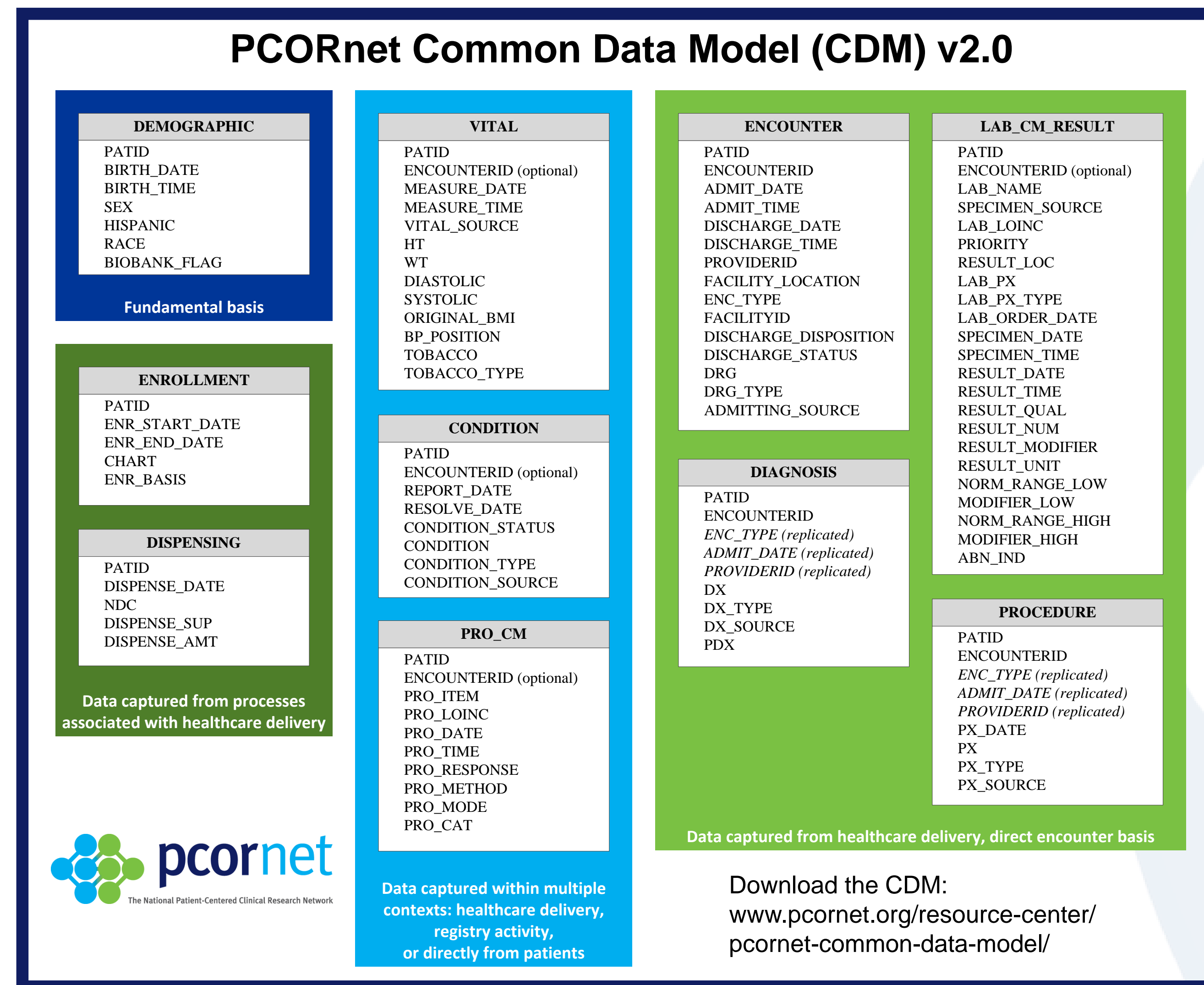
[1]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA; [2]Duke Translational Medicine Institute, Durham, NC; [3]Northwestern University Departments of Medicine and Preventive Medicine, Evanston, IL; [4]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; [5]Duke University School of Medicine, Department of Medicine, Durham, NC

Introduction

- PCORnet is a national Distributed Research Network (DRN) sponsored by the **Patient-Centered Outcomes Research Institute (PCORI)**
- Formed in late 2013 with 29 networks located across the United States
- Objective is to create a "network of networks" to form a **highly representative national infrastructure for research**
- Uses the **PopMedNet™** software infrastructure
- Architecture will allow queries and analytic programs to be **disseminated to individual data stores**
- The PCORnet Common Data Model (CDM) has been developed to provide **a standard organization and representation**

Methods

- Development began with the **Guiding Principles,** drafted as part of the charter of the PCORnet Data Standards, Security, and Network Infrastructure (DSSNI) Task Force
- The PCORnet CDM is a **stand-alone model** as a result of the **strategic direction** laid out in the Guiding Principles
- Heavily derived from the **Mini-Sentinel** Common Data Model and informed by other initiatives
- **Developed in incremental phases**, allowing agility in deployment, and new data domains to be incorporated
- Modeling is **optimized for analytic functionality** and to be **intuitive to investigators**

PCORnet Common Data Model (CDM) v2.0

DEMOGRAPHIC
- PATID
- BIRTH_DATE
- BIRTH_TIME
- SEX
- HISPANIC
- RACE
- BIOBANK_FLAG

Fundamental basis

ENROLLMENT
- PATID
- ENR_START_DATE
- ENR_END_DATE
- CHART
- ENR_BASIS

DISPENSING
- PATID
- DISPENSE_DATE
- NDC
- DISPENSE_SUP
- DISPENSE_AMT

Data captured from processes associated with healthcare delivery

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

PRO_CM
- PATID
- ENCOUNTERID (optional)
- PRO_ITEM
- PRO_LOINC
- PRO_DATE
- PRO_TIME
- PRO_RESPONSE
- PRO_METHOD
- PRO_MODE
- PRO_CAT

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

ENCOUNTER
- PATID
- ENCOUNTERID
- ADMIT_DATE
- ADMIT_TIME
- DISCHARGE_DATE
- DISCHARGE_TIME
- PROVIDERID
- FACILITY_LOCATION
- ENC_TYPE
- FACILITYID
- DISCHARGE_DISPOSITION
- DISCHARGE_STATUS
- DRG
- DRG_TYPE
- ADMITTING_SOURCE

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

LAB_CM_RESULT
- PATID
- ENCOUNTERID (optional)
- LAB_NAME
- SPECIMEN_SOURCE
- LAB_LOINC
- PRIORITY
- RESULT_LOC
- LAB_PX
- LAB_PX_TYPE
- LAB_ORDER_DATE
- SPECIMEN_DATE
- SPECIMEN_TIME
- RESULT_DATE
- RESULT_TIME
- RESULT_QUAL
- RESULT_NUM
- RESULT_MODIFIER
- RESULT_UNIT
- NORM_RANGE_LOW
- MODIFIER_LOW
- NORM_RANGE_HIGH
- MODIFIER_HIGH
- ABN_IND

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

Data captured from healthcare delivery, direct encounter basis

Download the CDM:
www.pcornet.org/resource-center/pcornet-common-data-model/

Results

PCORnet CDM v1.0
- Released on **May 30, 2014**
- **276 discrete comments** received and incorporated

PCORnet CDM v2.0
- Released on **February 27, 2015**
- **265 discrete comments** received and incorporated in 2 feedback cycles
- Two stakeholder sessions with **79 and 60 attendees**, respectively
- **Unanimous vote for adoption** by the Steering Committee

PCORnet CDM v3.0
- Currently in development

Discussion

- The CDM uses a strong basis of **secondary data** generated by **healthcare delivery processes**
- Development takes a **pragmatic approach,** supporting the **current state** of these data
- Area of future development is for data domains that **support multiple contexts**: healthcare delivery, registries and clinical studies, and data generated and received directly from patients
- **Participation** in review cycles, comment submission, and attendance demonstrates **high engagement within the PCORnet community**

Contact Information

Jeff Brown, PhD, Associate Professor
Harvard Medical School,
Harvard Pilgrim Health Care Institute
jeff.brown@harvardpilgrim.org

2015 Joint Summits on Translational Science, March 23-27, 2015