Iterative Design and Modeling for the PCORnet Common Data Model v3.0

Shelley A. Rusincovitch[1], Jeffrey S. Brown, PhD[2], Keith Marsolo, PhD[3], Jenny Ibarra, RN, MSN[4], Abel N. Kho, MD, MS[5], Daniella Meeker, PhD[6,7], Lesley H. Curtis, PhD[4,8]

[1]Duke Translational Research Institute, Durham, NC; [2]Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA; [3]Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; [4]Duke Clinical Research Institute, Durham, NC; [5]Northwestern University Departments of Medicine and Preventive Medicine, Evanston, IL; [6]University of Southern California, Los Angeles, CA; [7]Information Sciences Institute, University of Southern California, Marina Del Rey, CA [8]Duke University School of Medicine, Durham, NC

Abstract

The PCORnet Common Data Model (CDM) is a standardized representation for data generated through healthcare delivery. Using an iterative design process, 3 versions of the PCORnet CDM were developed and released within Phase I of PCORnet. The CDM version 3.0 had 5 primary design drivers, and the assessment of medication data domains was a key activity. The CDM modeling and development of shared conventions and interpretations has been refined by incorporating additional channels for feedback.
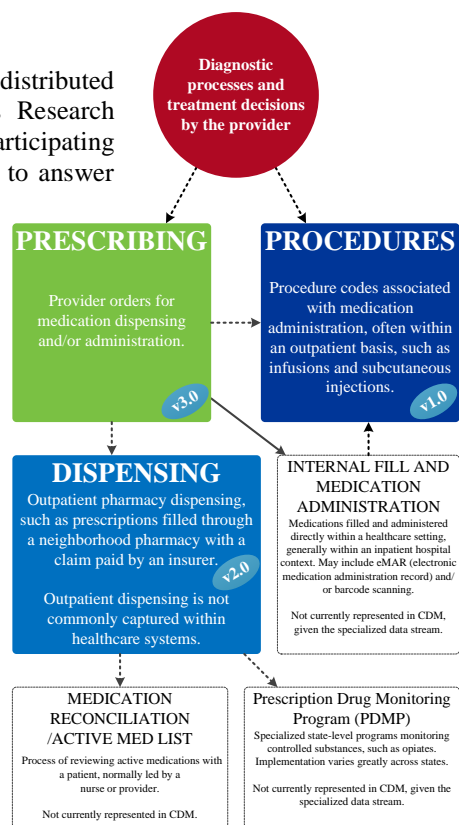
Introduction

The Patient-Centered Outcomes Network (PCORnet) is a national distributed research network (DRN) funded by the Patient-Centered Outcomes Research Institute (PCORI). This "network of networks" brings together participating institutions across the United States to form an infrastructure intended to answer important scientific questions in a high-throughput model, and allows data from multiple data partners to be queried securely[1]. An important component of PCORnet Phase I (early 2014 – late 2015) was to develop the PCORnet Common Data Model (CDM): a standardized representation of data elements and domains, selected and structured to optimize rapid implementation of distributed analytic functionality, to support PCORnet DRN objectives[2].

Using an iterative design process, 3 versions of the PCORnet CDM were developed and released during an approximately 16-month development period from February 2014 through May 2015, strongly based on existing standards and related initiatives. The PCORnet v3.0, released on June 1, 2015, is the final version within PCORnet Phase I.

Methods

The PCORnet CDM v3.0 had 5 primary design drivers: 1. Support of PCORnet pragmatic clinical trials; 2. Support of PCORnet observational studies; 3. Expanding the capabilities to represent the different workflows in medication data generation for either trial or observational purposes; 4. Support of rapid querying capabilities; and 5. Incremental improvements based on feedback in newly-introduced feedback channels from users and implementers.

The assessment and modeling of data-generating clinical activities into domains has been a key component of the PCORnet CDM development[3], given that the infrastructure is primarily based upon

Figure 1. A framework of medication data in the context of clinical data-generating activities, used as the basis for the PCORnet CDM medication domain modeling.

secondary data generated through healthcare delivery, including both clinical and administrative data in electronic health records (EHRs), health plans, and claims data sources. Incorporation of the prescribing data domain into the CDM was a significant addition for v3.0 (Figure 1) and identified related data streams for future assessment.

Results

The v3.0 development began with expert guidance from the CDM Working Group, which played an essential role in determining direction and feasibility, and initial presentations of design drivers for PCORnet task force and principal investigator forums. The draft CDM was disseminated in a formal feedback cycle where 236 discreet comments were received from investigators, technical teams, and other key individuals. Each comment was assessed and tagged into 23 thematic categories, then distilled into 12 themes presented during 2 stakeholder sessions with 57 and 60 attendees, respectively. The final product, integrating feedback from all stakeholders, was presented to and approved by the PCORnet Steering Committee. The PCORnet CDM v3.0 includes five new tables and 10 new fields (Figure 2).
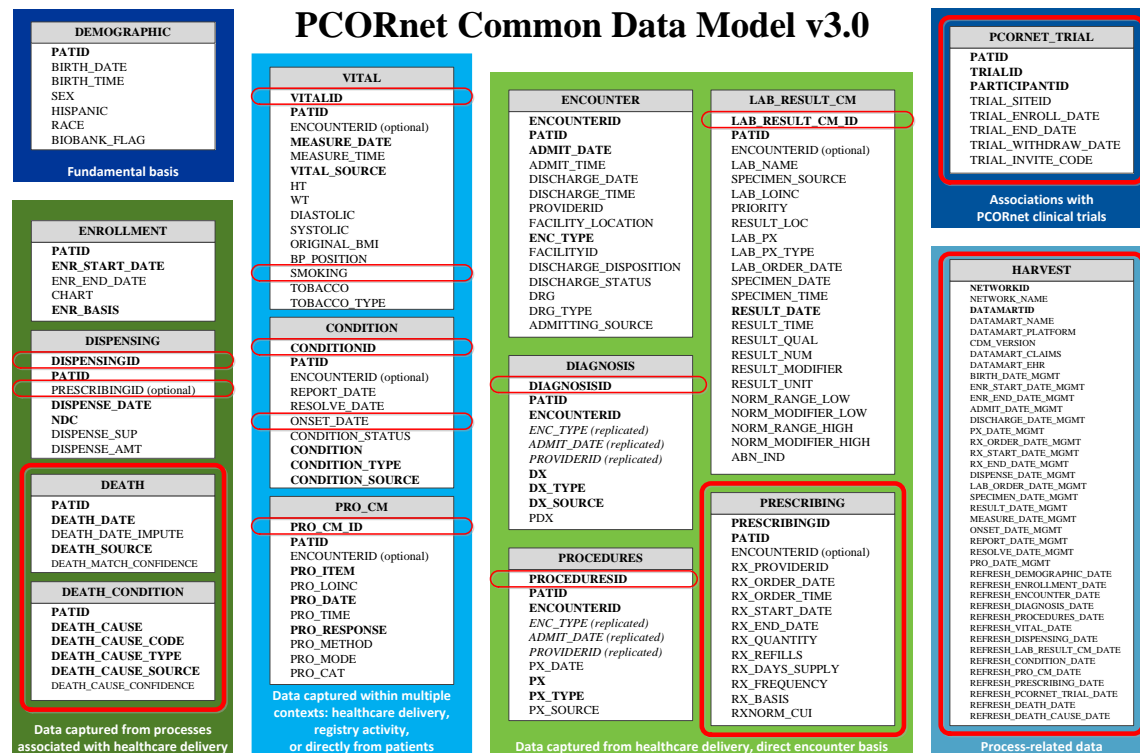
Figure 2. The PCORnet Common Data Model (CDM) v3.0 release. Bold font indicates fields that cannot be null due to primary key definitions or record-level constraints. Areas circled in red indicate new tables and fields.

Discussion

By assessing design drivers and adding additional channels for feedback from prospective users and current implementers on technical teams, the CDM Working Group has been able to refine modeling and development of shared conventions and interpretations across the PCORnet consortium. PCORnet Phase II, beginning in fall 2015, will be an important progression as the CDM and other key infrastructure developed in Phase I supports activity, including pragmatic clinical trials and observational studies, to "explore the questions that matter most to patients and their families."[4]

References

1.   Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. Journal of the American Medical Informatics Association : JAMIA. 2014;21(4):578-82.

2.   Brown JS RS, Kho AN, Marsolo K, Curtis LH. Development of a National Distributed Research Network Data Infrastructure: Design of the PCORnet Common Data Model. AMIA Jt Summits Transl Sci Proc. 2015;2015:302.

3.   Rusincovitch SA KA, Puro JE, Meeker D, Rivera P, Sorensen AA, Brown JS, Curtis LH. Pragmatic Data Domain Selection for a National Distributed Research Network: The PCORnet Common Data Model Strategy. AMIA Jt Summits Transl Sci Proc 2015. 2015:384.

4.   PCORnet. Why PCORnet exists [cited 2015 September 23]. Available from: http://www.pcornet.org/why-pcornet-exists/.