Pragmatic Data Domain Selection for a National Distributed Research Network: The PCORnet Common Data Model Strategy

Shelley A. Rusincovitch[1], Abel N. Kho, MD, MS[2], Jon E. Puro, MPA:HA[3], Daniella Meeker, PhD[4], Pedro Rivera, MSCS[3], Aaron A. Sorensen, MA[5], Jeffrey S. Brown, PhD[6], and Lesley H. Curtis, PhD[7]

[1]Duke Translational Medicine Institute, Durham, NC; [2]Northwestern University Departments of Medicine and Preventive Medicine, Evanston, IL; [3]OCHIN, Inc., Portland, OR; [4]Department of Health, RAND Corporation, Santa Monica, CA; [5]Temple University School of Medicine, Philadelphia, PA; [6]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA; [7]Duke University School of Medicine, Department of Medicine, Durham, NC

Introduction

- The PCORnet Common Data Model (CDM) specifies the data foundation for PCORnet, and is developed with a **phase-based approach**
- Each phase incorporates **new concepts and data tables** to support distributed clinical research (observational and interventional)
- In order to establish priorities for subsequent CDM development, it was necessary to establish a **method of assessing new concepts** and **making decisions for inclusion** to serve the functional, pragmatic focus of the initiative

Methods

- The assessment was **organized by data domains**; i.e., the high-level concepts of data organization based upon existing data sources, workflows, and processes
- Assessment included best practices established by **existing data models** and **advice from external experts**
- Close attention to **PCORnet-specific requirements**

Contact Information

Shelley A. Rusincovitch
Project Leader in Applied Informatics
Duke Medicine
shelley.rusincovitch@duke.edu

Data Domain Evaluation Table
The use of "stoplight" colors conveys favorable (green), caution (yellow), or unfavorable (red)

Domain	Effort Needed to Acquire Data	Analytic Utility/Value	Ability to Standardize Across Sites	Availability Across Networks (Anecdotal)
Allergies and/or Contraindications	HIGH	MOD	LOW	MOD
Patient-reported Outcome (PRO) Common Measures *CDM v2.0*	MOD	MOD	HIGH	N/A (prospective)
Condition *CDM v2.0*	LOW	MOD	MOD	HIGH
Death and Death Cause	HIGH	HIGH	HIGH	LOW
Facility	MOD	MOD	LOW	HIGH
Family history	MOD	MOD	LOW	LOW
Inpatient Medication Administration	HIGH	LOW	LOW	LOW
Laboratory Result Common Measures *CDM v2.0*	MOD	HIGH	LOW	HIGH
Medication Reconciliation	MOD	MOD	MOD	HIGH
Outpatient Pharmacy Dispensing *CDM v2.0*	MOD	HIGH	HIGH	LOW
Primary Care Provider (PCP)	MOD	MOD	MOD	MOD
Provider Orders (including Medication Orders)	MOD	MOD	LOW	MOD
Social History & Lifestyle Choices	HIGH	LOW	LOW	MOD
State Vaccine	HIGH	LOW	LOW	MOD
Study Enrollment	LOW	HIGH	MOD	N/A (prospective)
Study Visits	MOD	MOD	MOD	N/A (prospective)

Data Model Landscape Scanning

- Mini-Sentinel Common Data Model, v4.0
- i2b2 Data Repository Cell, v1.7.00
- OMOP Common Data Model, v5.0
- HMORN VDM, v3.2
- ESPnet Data Form, 2013
- National Quality Forum, Quality Data Model, Version 4.1.1.

Results

- Assessment resulted in recommendation for **prioritization of data domains**
- Initiative has subsequently **formalized a process** for stakeholder review, facilitated discussion, and the approval process for adoption

Discussion

- Key lesson learned: **Importance of identifying and articulating foundational strategic decisions**, including **interoperability** within the analytic framework
- Limitation: Dimensions were **assessed by the CDM Working Group**, rather than by a formal survey of all participants; however, the working group represented individuals with the deep expertise necessary to make informed recommendations

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

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