Introduction to Big Data Topics

CRISP Seminar Series

September 19, 2016

Beth McManus, PT, MPH, ScD
Overview & Acknowledgments

Acknowledgments:
ACCORDS and Educational Core
Chase Cameron
(chase.cameron@ucdenver.edu)

Overview for Today
1. What is Big Data?
2. Big Data Methods Workshop
3. CRISP seminar series
Big Data

• Different definitions that cross many sectors
  – “Large data sets, structured and unstructured” (Ernest Davis)
  – Volume of data: How big is “big”?
  – Not new concept, but potential for analytics is
Big Data

• Large Observational Data Sets
  – Electronic health records
  – Claims data
  – Genomics

• Accessing multiple databases to track patients, outcomes, and improve clinical decision-making and inform good health policy
Continued Questions and Concerns

- Large observational health databases have vast potential, but issues remain:
  - Ethical and legal consideration
  - Data management and storage challenges
  - Appropriate analytics to make casual inferences, applying predictive models to dynamic health service use and policy
CRISP Seminar Series

• Highlight the extensive expertise on this campus
• Foster collaborations
• Provide resources
• Build a community of researchers and increase national reputation
Advanced Methods Workshop

• “Big Data: The Power of Large Observational Data Sets”
• Thursday October 13, 2016 8AM-5PM
• Lilly Marks Boardroom (UPI Building)
• An intermediate / advanced applied methods workshop
• Link to register: On the CRISP website! https://redcap.ucdenver.edu/surveys/?s=PK4HKCWP3W
Advanced Methods Workshop

• Geared toward researchers who are using large obs datasets (LODS) and want more information and resources
• Format will be 3 expert panels with applied examples of the opportunities and challenges of using LODS
• In-depth and applied discussion of dissemination and implementation
• Ethical considerations and navigating COMIRB
Workshop Schedule Overview

• Keynote (Dr. Gosh): *Overview of the Data “Life-Cycle” approach* (8AM-9:30)

• **Panel #1: Data Management Considerations**
  – Local vs Cloud / distributed
  – Record linkages
  – National data networks
  – Models for data extraction, standardization, and data quality
  – Dr.’s Bennett, Colborn, (other panelists to be confirmed)
  – 10:00-12:00
Workshop Schedule Overview

• Working Lunch (provided): Dr. Devine, COMIRB (12P-1P)

• **Panel #2: Data Analytics (1p-3p)**
  – Causal inference from LODS
  – Machine learning
  – Imputation
  – Appropriately addressing patient heterogeneity
  – Drs. Lindrooth, Feinstein, Carlson, Dickinson
Workshop Overview

• **Panel #3: Dissemination and Implementation**
  - Dr. Morrato and colleagues
  - How does D & I apply to LODS?
  - When and how do we think about D & I?
  - What are the resources available at AMC?
  - 3:30p-5:00p
Questions?

• Beth McManus, ACCORDS Education Program Lead
  – Beth.mcmanus@ucdenver.edu

• Chase Cameron, ACCORDS Education Program Coordinator
  chase.cameron@ucdenver.edu
CRISP Seminar Topics

• Occurs monthly, 3rd Monday
• Will continue theme of LODS
  – Machine Learning
  – D & I
  – Causal inference (RDD)
  – D4 (Large Transformational grant)
  – Using LOS for individualized care, clinical decision-making