Comparison of a National and a Local Electronic Health Data Warehouse for Clinical Laboratory Quality Improvement Research

Jeremy Provance, M.S.[1]; Shivani Sivasankar[1]; An-Lin Cheng, Ph.D.[1]; Suman Sahil[1]; Kamani Lankachandra, M.D.[1,2]; Mark A. Hoffman, Ph.D.[1,3];

1. University of Missouri-Kansas City School of Medicine; 2. Truman Medical Center, Kansas City, MO; 3. Children's Mercy Hospital, Kansas City, MO

Introduction

- Increasing use of electronic health records (EHRs) over the past two decades has enabled large-scale health data research

- Since EHRs are used in daily clinical practice, data are often extracted and stored in electronic data warehouses (EDWs) to enable research and quality improvement initiatives

- The Cerner Health Facts EDW contains over 63 million unique patients from 863 participating health facilities

- The Cerner i2b2 EDW at University of Missouri-Kansas City (UMKC) and Truman Medical Center (TMC) contains over 700,000 unique patients from the three TMC facilities

- The Health Facts and i2b2 EDWs are used for research and quality improvement projects by physicians, researchers, and students at UMKC and TMC in Kansas City, Missouri

Objective

- We compared characteristics of Health Facts and i2b2 for use with quality improvement projects providing shared and unique benefits and limitations each presents
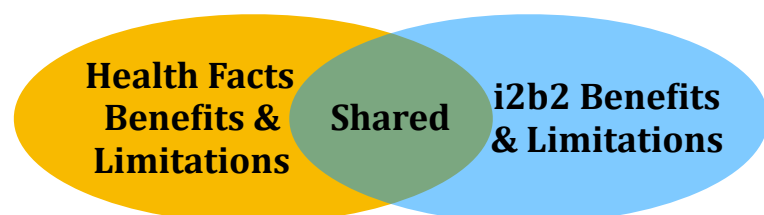
Results

	BENEFITS	LIMITATIONS
Shared Between Health Facts & i2b2	- Available to all trained physicians, researchers, and students - De-identification allows "Non-Human Subjects Research" designation for easier data access; Cerner has established HIPAA-compliant operating policies to establish de-identification for Health Facts	- Incorrect EHR data propagate through to EDWs - Limited to participating Cerner facilities (i.e. no EPIC sites) - Significant financial barriers for personnel and hardware to stand up and maintain databases
Unique to Health Facts	- Large patient volumes (over 63 million) - Access to all data tables through relational database - Uniquely positioned for longitudinal analysis (2000-2016) - Access multi-node computer through remote desktop	- Requires expert database analyst to manage - SAS user interface with proc SQL data access queries - Missing and null values are prevalent - Data are updated and refreshed annually

Results Continued

	BENEFITS	LIMITATIONS
Unique to i2b2	- User-friendly, drag-and-drop query building - Data curated by Cerner professionals - Data are current; new data loading every other week - Available anywhere via web browser	- Data limited to available, mapped concepts - Data views and analysis limited to the i2b2 program plug-ins - Constrained longitudinal analysis due to limited views - No back-end database access

Conclusions

- As more research centers around EDW health data analysis, it will be crucial to understand how each EDW is optimized for specific analysis.

- There are trade-offs that must be considered with EDW data access and end-user ease-of-use. EDW users and administrators must be in conversation with EDW managers and creators to create warehouses that are optimal for analysis.

Disclosures