

Integrated Approach to Secondary Data Analysis at Ohio State: Access, Research Support, and Education

Guy Brock, PhD¹; Amy Lehman, MAS¹; Tanya Mathew, BDS, MS²; Lang Li, PhD¹; Timothy Huerta, PhD, MS^{1,3}; Henry Xiang, MD, MPH, PhD⁴; Rebecca Jackson, MD⁵

¹ Department of Biomedical Informatics, The Ohio State University, Columbus, OH; ² Center for Clinical and Translational Science, The Ohio State University, Columbus, OH; ³ Department of Family Medicine, The Ohio State University, Columbus, OH; ⁴ Center for Pediatric Trauma Research & Center for Injury Research and Policy, Nationwide Children's Hospital, Columbus, OH; ⁵ Division of Endocrinology, Diabetes, and Metabolism, The Ohio State University, Columbus, OH

The Ohio State University CCTS is developing a comprehensive support system for researchers to work with existing datasets through *data curation and access, financial support* for innovative research ideas, and *education* about practical data issues and challenges.



DATA CURATION & ACCESS

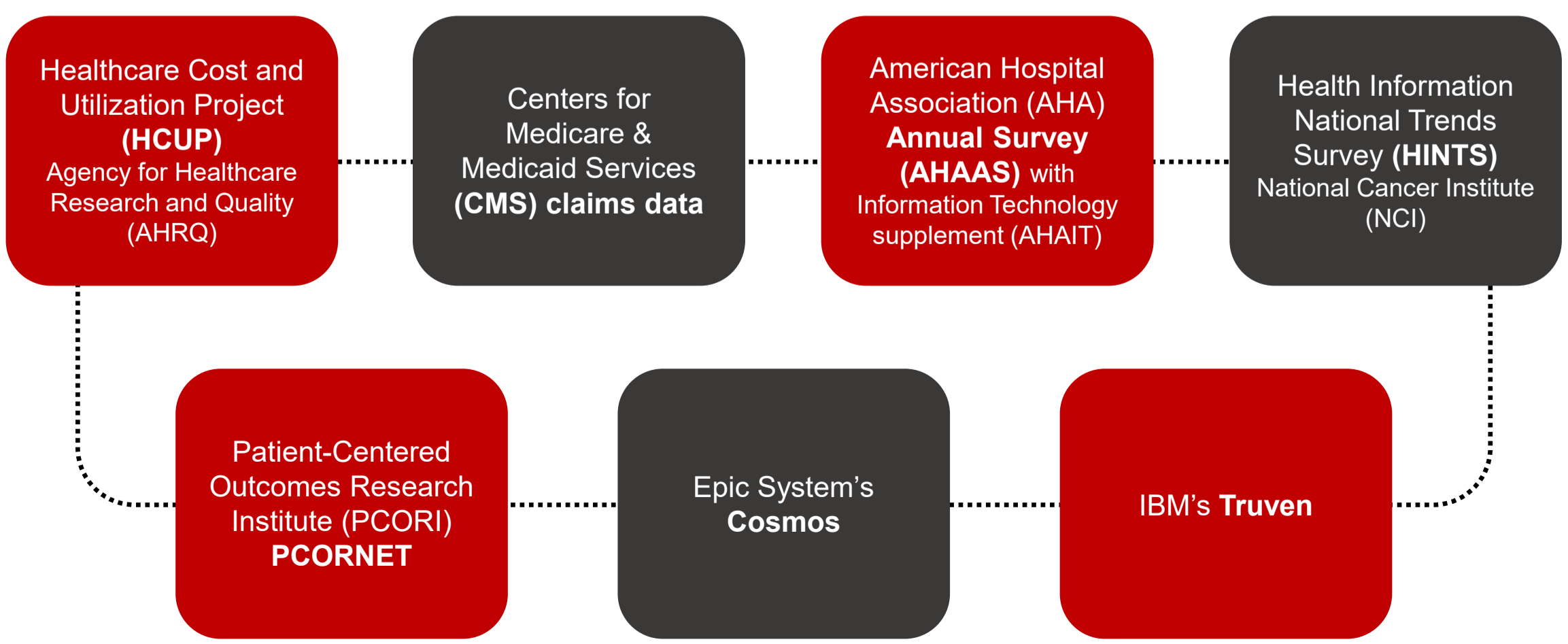


The Center for the Advancement of Team Science, Analytics, and Systems Thinking in Health Services and Implementation Science Research (CATALYST) DataCore

CATALYST was formed by The Ohio State University College of Medicine to advance T3 research at Ohio State. As part of that mission, the **CATALYST DataCore (CDC)** is being developed for researchers across the College of Medicine to streamline and expand access to data that accelerates discovery and leads to funded research.

The CDC will improve investigator access by using automation to select the data they want and then automatically pull it into the statistical analysis software of their preference, including clear instructions on how to use the data with a dataset that is already cleaned, merged, and harmonized.

The following deidentified data will be included in the data commons:



RESEARCH SUPPORT

Artificial Intelligence / Machine Learning Pilot Initiative

Initiative to support transformative research that leverages the power of artificial intelligence and machine learning

\$50,000 in direct costs for one year per awarded project

37 initial abstracts submitted in 2019

6 computational fields from applicants: *Imaging informatics, EHR Predictive analytics, Omics, Clinical decision system, Population health, Other*

55 peer reviewers, both academic and community

5 funded applications in basic science, clinical, and population health research in 2019

Secondary Analysis of Existing Datasets Pilot Initiative

Initiative to catalyze research & collaboration using nationally available or accessible datasets

\$15,000 in direct costs for one year per awarded project

29 applications in 2019

6 human health focus areas from applicants: *Cardiometabolic / Endocrine Cancer, Pediatrics, Neurological, Public Health, Digestive Disease*

61 reviewers

7 funded applications in basic science, clinical, and population health research in 2019

EDUCATION



Rethink, Reuse, Research: Secondary Analysis of Existing Data

New monthly seminar series to educate researchers about existing datasets and how they can be used for research:

- Focus on practical issues and challenges in using existing data for research.
- Led by biostatisticians and faculty with hands-on data experience.

PROPOSED TALK SCHEDULE:

