BDC Data Overview March 2024

Sweta Ladwa, MPH
NHLBI BDC Data Management Core Scientific Program Director

Prior to joining the NHLBI team, Mrs. Ladwa served as a consultant to various agencies within the Department of Health and Human Services (HHS) including the Centers for Medicare and Medicaid Services (CMS), The National Cancer Institute (NIH/NCI), and the Office of the National Coordinator for Health IT (ONC). She led numerous efforts relating to Health Information Technology, Epidemiology, and Big Data in areas such as data aggregation, coordination, and visualization for rare and common cancers, development of FHIR Implementation Specifications for PCOR and other use cases and curating de-novo and de-facto standards for biomedical research.

Mrs. Ladwa earned her master's degree in Public Health from George Washington University, and her bachelor's degree in Biology and Chemistry from Virginia Commonwealth University.

NHLBI's BDC Overview
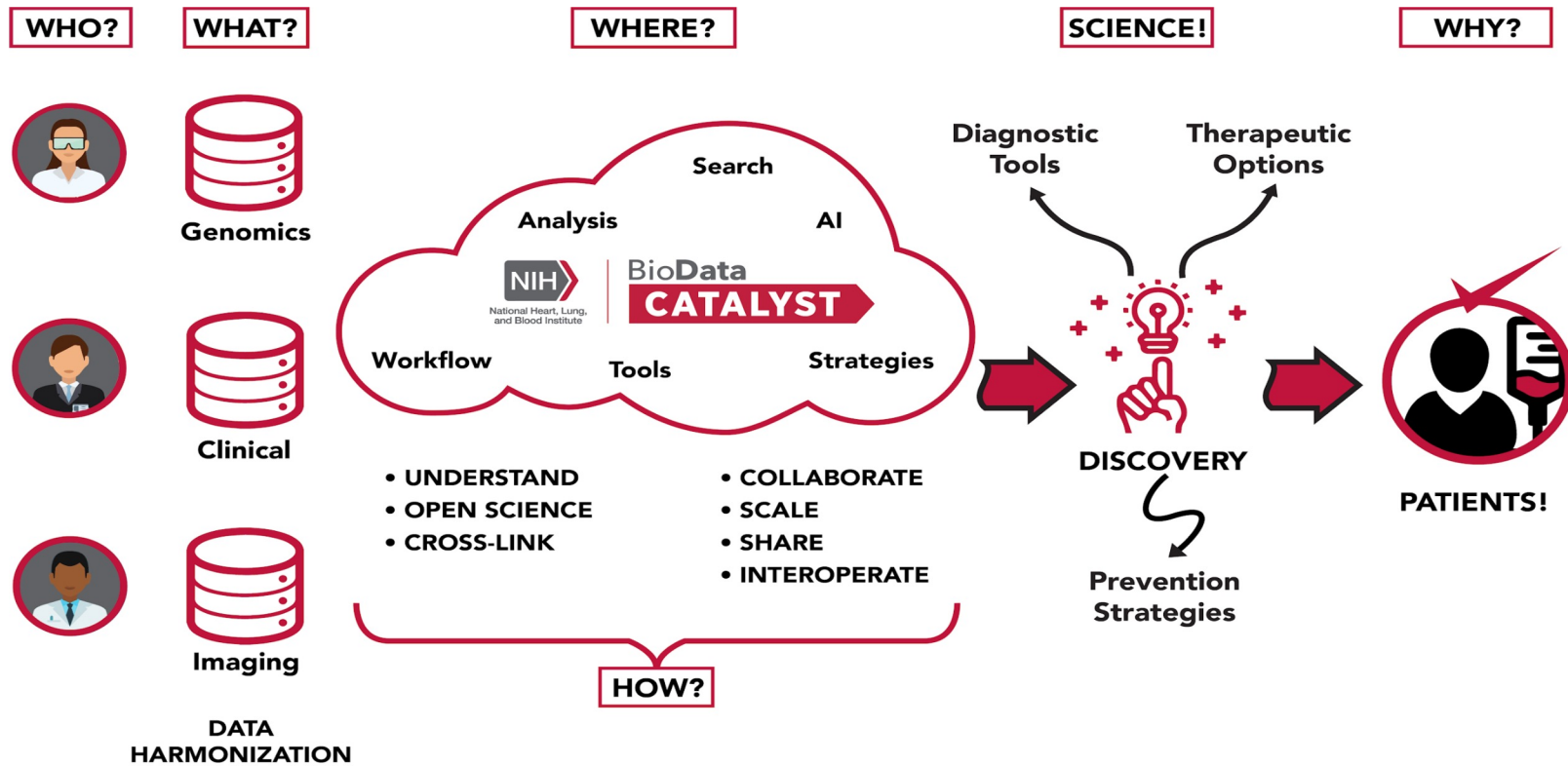
NHLBI's BDC Overview: Analytic Capabilities

- Analytic workspaces (Jupyter, R, Python, SAS)

- GWAS (using GENESIS or HAIL)

- Variant Calling using GATK or TOPMed (WES/WGS Data)

- Image visualization and analysis including Machine Learning models

- App Development

- App Hosting

- Pipeline Development (CWL or WDL, Nexus coming soon)

- Bring Your Own Data (BYOD)

- Team Collaboration Workspaces

- Import and publish apps and workflows to/from dockerized apps (e.g. Dockstore)

- Search:

  - Search UI – Genotype and Phenotype

  - Search API

  - Synthetic Cohort Generation

  - Semantic meta-data search

NHLBI's BDC Overview: Data
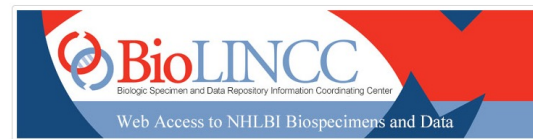
A few NHLBI Data Generating Programs

Trans-Omics for Precision Medicine (TOPMed)
Multi-omics data from NHLBI longitudinal cohort studies: WGS, sc-RNA, metabolomics, proteomics, methylation, DNA, etc.

HeartShare/AMP (NHLBI)
Retrospective NHLBI cohort data + Prospective Data Collection
Retrospective: Multi-modal imaging (echos, ECGs, etc.), clinical, multi-omics, etc.
Prospective: Clinical, Omics, Imaging, EHR

National Sleep Research Resource (NSRR)
Multi-modal sleep data: clinical, polysomnography, actigraphy, etc.

Pediatric Cardiac Genomics Commons (PCGC)
Multi-modal data on pediatric populations: clinical, omics, EHR, imaging

CONNECTS
COVID-19 related observational studies and clinical trials data

Trans-NIH Data Generating Programs and BDC

RECOVER
- Longitudinal Epidemiological Study
- Multiple modalities: clinical, genomic, imaging (multiple modalities), clinical trials, censor, mobile health, biospecimen, etc.

Regenerative Medicine Innovation Program (RMIP)
- Multi-modal cell characterization data
  - Single-Cell RNA, cyTOF, WGS/WGBS, etc.

Pause for questions and comments

BDC Hosted Data Released December 2023 to March 2024

Sweta Ladwa, MPH
NHLBI BDC Data Management Core
Scientific Program Director

Prior to joining the NHLBI team, Mrs. Ladwa served as a consultant to various agencies within the Department of Health and Human Services (HHS) including the Centers for Medicare and Medicaid Services (CMS), The National Cancer Institute (NIH/NCI), and the Office of the National Coordinator for Health IT (ONC). She led numerous efforts relating to Health Information Technology, Epidemiology, and Big Data in areas such as data aggregation, coordination, and visualization for rare and common cancers, development of FHIR Implementation Specifications for PCOR and other use cases and curating de-novo and de-facto standards for biomedical research.

Mrs. Ladwa earned her master's degree in Public Health from George Washington University, and her bachelor's degree in Biology and Chemistry from Virginia Commonwealth University.

December 2023-March 2024

NSRR	National Sleep Research Resource (NSRR) - CFS
CONNECTS	CONNECTS_C3PO_Assay
CONNECTS	CONNECTS_ACTIV4C_DataDictionaryFile
CONNECTS	CONNECTS_ACTIV4A
TOPMed	TOPMed_SPIROMICS
TOPMed	TOPMed_BostonBrazil_SCD
PCGC	PCGC_SRA
LungMAP	LungMAP

Browse Data in BDC

https://gen3.biodatacatalyst.nhlbi.nih.gov/discovery

Browse and Build Cohorts without Login

https://openpicsure.biodatacatalyst.nhlbi.nih.gov/picsureui/

Pause for questions and comments

Future Data-Related Updates in BDC

Sweta Ladwa, MPH
NHLBI BDC Data Management Core
Scientific Program Director

Prior to joining the NHLBI team, Mrs. Ladwa served as a consultant to various agencies within the Department of Health and Human Services (HHS) including the Centers for Medicare and Medicaid Services (CMS), The National Cancer Institute (NIH/NCI), and the Office of the National Coordinator for Health IT (ONC). She led numerous efforts relating to Health Information Technology, Epidemiology, and Big Data in areas such as data aggregation, coordination, and visualization for rare and common cancers, development of FHIR Implementation Specifications for PCOR and other use cases and curating de-novo and de-facto standards for biomedical research.

Mrs. Ladwa earned her master's degree in Public Health from George Washington University, and her bachelor's degree in Biology and Chemistry from Virginia Commonwealth University.

Data-Related Updates

BDC Data Model in development
- A common, LinkML-based model to facilitate a consistent representation of data rooted in standards and ontologies to facilitate search and use case-based harmonization

FHIR Server in development
- Phase 1 focused on supporting metadata for discoverability and EHR data exchange
- Phase 2 focused on supporting SDOH data

Implementing DataCite DOIs
- Increases FAIR-ness of data and tools/workflows
- Benefits include ease of reference for publications

NIH Researcher Auth Service (RAS)
- Allows researchers' data access to follow the researcher to the NIH computational environment of their choice to facilitate interoperability

Supporting HLBS researchers
- Adding additional data and metadata to support HLBS researcher needs
- Supporting the NHLBI Supplement to the NIH Final Policy for Data Management and Sharing
- Continuous improvement based on feedback from BDC stakeholders

BDC and Interoperability
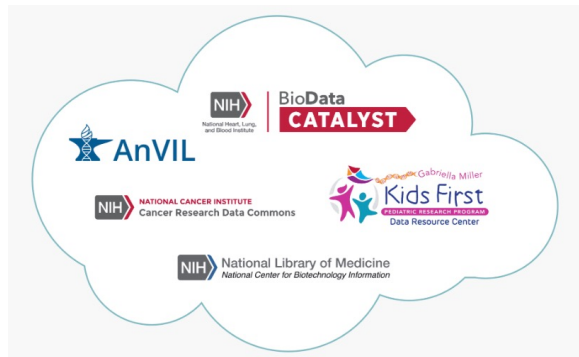
NIH Cloud Platform Interoperability (NCPI)
Federate data analysis (bring compute to data with NIH Partner Systems)

ARPA–H Biomedical Data Fabric Toolbox
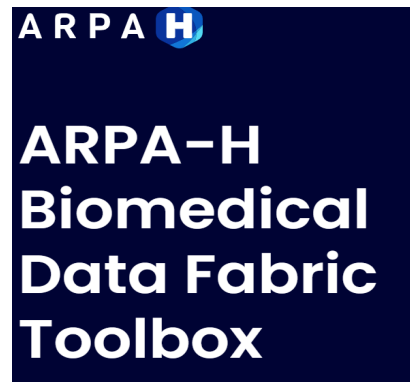
Advanced Research Projects Agency for Health (ARPA-H) Biomedical Data Fabric Toolbox
Support a national, unified consistent layer of data services to work across many different systems and environments

GA4GH DRIVER PROJECT

Global Alliance for Genomics and Health
Work to interoperate with international data sources, including UK Biobank and the NIH All of Us Research Program

Acknowledgements

- NHLBI Director: Dr. Gary Gibbons
- BDC Executive Sponsor: Dr. David Goff
- BDC Scientific Program Director: Dr. Regina Bures
- BDC DMC Scientific Program Director: Mrs. Sweta Ladwa
- BDC Steering Committee: Dr. Ingrid Borecki, PIs, Consortia
- BDC EEP
- BDC Coordinating Center: Dr. Ashok Krishnamurthy/Dr. Stan Ahalt
- RTI International: Mrs. Becky Boyles
- BDC Data Management Core: Mr. Chris Siege and Dr. Anne Thessen
- University of Chicago: Dr. Bob Grossman
- Velsera: Dr. Jack DiGiovanna
- Harvard Medical School: Dr. Paul Avillach
- Broad Institute: Dr. Alisa Manning