Database and Survey Development for Clinical Research Studies

Panelists- **Joe Palmisano**, MPH, MS, Interim Director, BEDAC

Linda Rosen, MSEE, CDW Research Manager, OHRA

Tasha Coughlin, Data Manager, BU CTSI

Facilitator- **Howard Cabral**, PhD, Director,
Biostatistics and Research Design Program, BU CTSI

Objectives

To provide an overview of and answer questions about the:

- Biostatistics and Epidemiology Data Analytics Center (BEDAC) and the data system development, data management and analytic services offered
- data available from the BMC Clinical Data Warehouse for researchers and how to make a good data request.
- BUMC REDCap functionality, features, available end user support, and general tips

Biostatistics and Epidemiology Data Analytics Center (BEDAC)

A Full-Service Data Management Resource

Joseph Palmisano
Interim Director, BEDAC
Boston University School of Public Health

June 13, 2018

BOSTON UNIVERSITY

BEDAC Overview

- Resource established in 1984 by BUSPH
- Formerly the Data Coordinating Center (DCC)
- Support for hundreds of projects to date
- One-stop shop for all data needs
  - Data system development
  - Data management
  - Reporting
  - Statistical analysis

Data Systems, Management and Analysis Expertise

- Approximately 30 full-time data professionals
- Data Managers
  - REDCap, Qualtrics, StudyTrax, TeleForm
- Web/Database Developers
  - Web: ASP.Net, HTML, C++, C#, Java, PHP, Javascript, Perl
  - Database: SQL, MySQL, MS Access
- Statistical Programmers/Analysts
  - SAS, R
  - Support for SPSS, STATA, MPlus

Roles and Services

- Assessment/Questionnaire design
- Electronic Data Capture (EDC) system design:
  - EDC software: REDCap, Qualtrics, CASIC, ACASI, StudyTrax
  - Custom built EDC, reporting and tracking applications
  - Optical character recognition: TELEForm
- Data management:
  - Audit and monitor data
  - Data queries and cleaning
  - Analytic set creation
  - Data sharing
- Report development
- Statistical analysis
  - Support for abstracts, presentations and manuscripts

Project Support

- Investigators
  - BU, BMC, External Academic and Non-Academic Institutions
- Support for any/all data related project aspects
- Engage at any project stage
- Research Design:
  - Cohort, Cross-Sectional, Case-control, RCT, Survey, etc.
- Project types:
  - Registries, Large Research Centers, Consortia
  - Full project support
  - Modular support

Our Collaborators

- BUMC IT/ Information Security
  - Secure HIPAA compliant servers
  - Ensure data security and integrity
- BMC IT
  - BMC REDCap
- BUSPH Biostatistics
- CTSI
  - CTSI supported research software
  - BEDAC provided analytic support for CTSI approved projects
- Clinical Data Warehouse (CDW)
  - BMC EMR patient data

Working with BEDAC

- Services provided at hourly rates
- Quotes developed to meet project and budget needs
- Visit our website for more information
  - https://www.bu.edu/bedac/
- For questions or to discuss project needs:
  - Contact: Joe Palmisano
  - jpalm@bu.edu
  - (617) 638-5384
  - 85 East Newton Street, 900C (Fuller)

Using BMC and CHC data for research purposes

Clinical Data Access for Researchers

*The Clinical Data Warehouse.*In late 2005, Boston Medical Center embarked on a new project to bring data at the hospital into a single relational data warehouse. This means that clinical researchers who have IRB/HIPAA approval can now, with the help of a database developer:

- get data with a single query (rather than searching manually)
- cross-reference data from EPIC, Logician, SDK, SCM, OR and ED
- get simple patient counts defined by some user specified criteria
- relate local data with BMC data
- search text files (physician notes, pathology reports, etc.) for keywords
- and more, depending on the researcher's needs

Get Started

If your protocol has been approved by HIPAA and/or the IRB, then click here to access the data request form. Note that there is a charge for accessing the Data Warehouse for research purposes. To find out more about charges, click here. E-mail the Clinical Data Warehouse Research Manager, Linda Rosen with questions or a meeting request to discuss your particular data needs.

Linda Rosen lirosen@bu.edu

REDCap
Research Electronic Data Capture

Tasha Coughlin, REDCap Administrator
BU Clinical & Translational Science Institute (CTSI)

REDCap BACKGROUND

- Web-Based Software for Research Data Collection
- 2004 Developed at Vanderbilt
- 2012 BU CTSI joined REDCap Consortium
- 2018 Over 2k Users & Over 1,800 Projects at BU

CTSI: End User Support | **BUMC IT**: Maintains Server & Updates

ADVANTAGES

- **Web-Based**
- Fast Programming
- **Data Validation** & Data Quality Rules
- **HIPAA Compliant & Secure**
  - BUMC IT Server behind Firewalls
  - 2 Factor Login
  - Custom User Access & Audit Trail
- **FREE!**

FUNCTIONALITY

- Classic & **Longitudinal Design** (For Events & Multiple Arms)
- Design Data Collection Instruments (CRFs)
  - Data Dictionary & Online Designer
- **Create, Distribute & Track Participant Surveys**
- **Various Field Types**
- **Flagged Identifiers**
- Skip Logic (Branching)
- Reports & Data Exports to Stats Software

FAVES

- **Codebook**: quick reference for attributes of all field
- Bookmarks: Links to webpages, other REDCap projects & reports
- **Action Tags**: Unique Field Actions
- Designated Survey emails
- Twilio: 3rd party web service for texting survey links
- Messenger: Chat App
- **Custom Record Status Dashboards**
- Data Resolution workflow

TIPS

- **TEST** BEFORE MOVING TO PRODUCTION!!!
  - ADMIN IS ABLE TO MOVE PROJECT BACK TO DEVELOPMENT
  - **LIMIT USERS WITH PROJECT DESIGN & DELETE RECORD RIGHTS**
  - **INACTIVATE/ARCHIVE** EXPIRED SURVEYS & PROJECTS
  - CITE THE CTSI!

HELP

- REDCap Training Videos, Help & FAQ
- Email or Call Tasha
  - tawatson@bu.edu; rchelp@bu.edu; 617-358-7390
- Drop-In Sessions
  - 1st & 3rd Thursday of each Month, 1 – 3 PM
- Initial Consultations for New Users

QUESTIONS?

NEED REDCAP ACCESS? HTTPS://IS.GD/BUMC_REDCAP_ACCESS

NEED BERD OR REDCAP CONSULT? BU.EDU/CTSI/ABOUT/CONTACT-US/

CTSI RESOURCES: BU.EDU/CTSI

Support for Research →

Biomedical Bridge BUilders Initiative
Biostatistics, Data Management & Analysis
The Clinical Trials Network
Collaborative Research & Networking
Core Lab Support
Framingham Heart Study iPSC
Grant Writing & Editing Services
iPSC Culture Course
IRB Review
Pilot Awards
Protocol Implementation
Research Recruitment and Retention Program (R3)
Regulatory Support
Study Design
CRITC
BU Thrombotic Microangiopathy Collaborative (BUTMAC)
Voucher Program

Study Design

Biostatistics, Epidemiology & Research Design (BERD)

The CTSI's Biostatistics, Epidemiology & Research Design (BERD) program provides researchers with comprehensive support in designing and carrying out research at the design, implementation and analysis stages. As biostatistics assistance ranks among the top requests for CTSA services, BERD has generated a portfolio of solutions for the most common needs to ensure participation in the earliest stages of study development.

The BERD program aims to:

Collaborate and consult with research teams to help them design and carry out high-quality research. We offer free biostatistics and data management drop-in sessions on both the Medical and Charles River Campuses. For those investigators who cannot attend a drop-in, we can schedule a separate meeting through our request for consultations. If services requested go beyond what can be provided through CTSI consulting, we offer fee-for-services, as well as collaboration on grants and funded projects.

Educate and train research teams in proper approaches to design studies, collect and manage data, perform analyses and present results. We provide education, collaboration, and consultation on design and analysis through our regularly scheduled BERD seminars.

Identify novel methods for design, implementation, and analysis through collaboration with methods and experts. We invite you to attend one of our drop-in sessions, schedule a meeting, or attend a seminar to learn more about the novel methods provided and facilitated by our program.

BERD consulting services and collaboration are available to all BU investigators.

BERD faculty and staff also collaborate closely with the Clinical Research Informatics and Technology Consultation (CRITC) service, assisting with the use of information technology and informatics to facilitate clinical studies.

http://www.bu.edu/ctsi/support-for-research/study-design/ hjcab@bu.edu janicew@bu.edu

NIH Strategic Plan for Data Science

Additional Informational slides

NIH Strategic Plan for Data Science

- NIH defines data science as "the interdisciplinary field of inquiry in which quantitative and analytical approaches, processes, and systems are developed and used to extract knowledge and insights from increasingly large and/or complex sets of data."
- NIH supports the generation and analysis of substantial quantities of biomedical research data including numerous quantitative and qualitative datasets emanating from fundamental research using model organisms (such as mice, fruit flies, and zebrafish), clinical studies (including medical images), and observational and epidemiological studies (including data from electronic health records and wearable devices).

https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf ***6-4-18***

Genomic Data Growth

By 2025, the total amount of genomics data alone is expected to equal or exceed totals from the three other major producers of large amounts of data: astronomy, YouTube, and Twitter.

Indeed, next-generation sequencing data, stored at NIH's National Center for Biotechnology Information (NCBI), has been growing exponentially for many years and shows no signs of slowing.

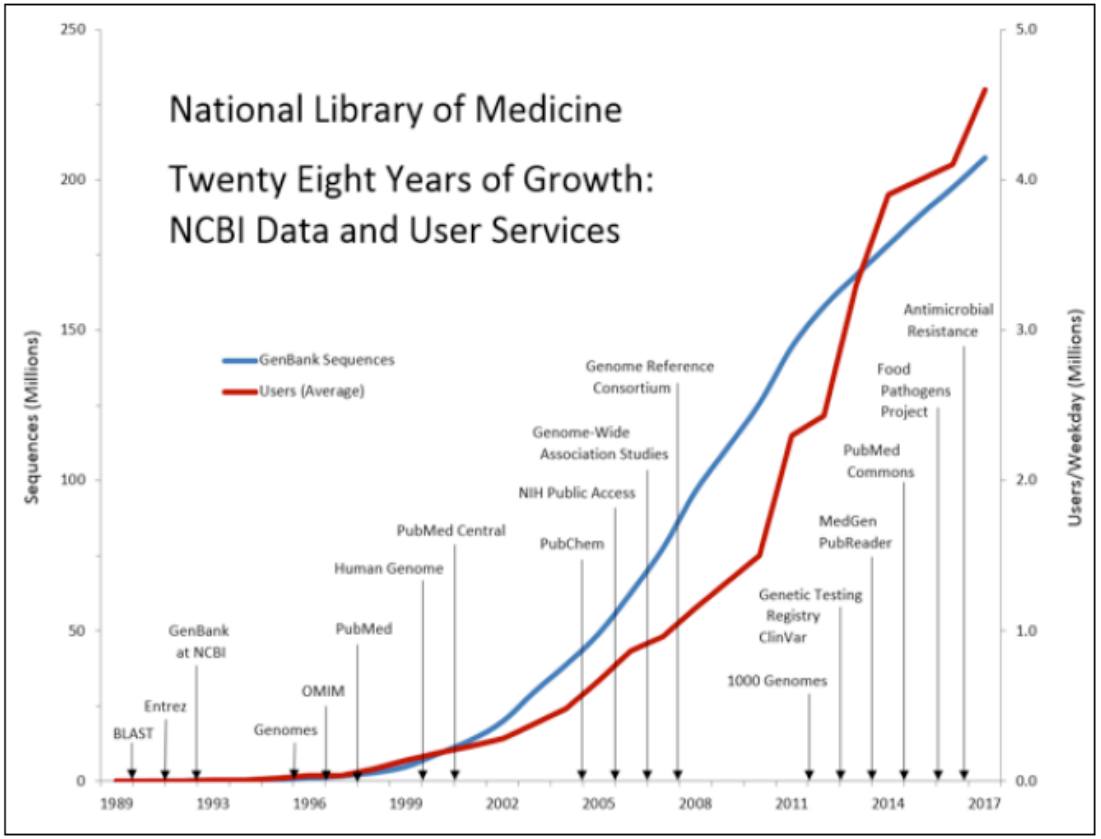

Figure 1. Growth of NCBI Data and Services, 1989-2017 Credit: NCBI

New Research Methods and Tools Needed

- Advances in storage, communications, and processing have led to new research methods and tools that were simply not possible just a decade ago.
- Machine learning, deep learning, artificial intelligence, and virtual-reality technologies are examples of data-related innovations that may yield transformative changes for biomedical research over the coming decade.
- The fastest supercomputers in the world today perform a quadrillion (10^{15}) calculations each second: known as the petascale level. The next frontier is exascale computing (which is 1,000 times faster than petascale, or a quintillion (10^{18}) calculations each second).

Clinical Data and Information Security

- Throughout the research enterprise, NIH must continue to balance the need for maximizing opportunities to advance biomedical research with responsible strategies for sustaining public trust, participant safety, and data security.
- NIH must develop, promote, and practice robust and proactive information-security approaches to ensure appropriate stewardship of patient data and to enable scientific advances stemming from authentic, trusted data sources.
- NIH will ensure that clinical-data collection, storage, and use adheres to stringent security requirements and applicable law, to protect against data compromise or loss.
- NIH will also comply with the health-information security standards.
- Data quality and integrity must be maintained at all stages of the research life cycle—from collection through curation, use, dissemination, and retirement.

What is FAIR?

Biomedical research data should adhere to FAIR principles, meaning that it should be Findable, Accessible, Interoperable, and Reusable.

- To be **Findable**, data must have unique identifiers, effectively labeling it within searchable resources.
- To be **Accessible**, data must be easily retrievable via open systems and effective and secure authentication and authorization procedures.
- To be **Interoperable**, data should "use and speak the same language" via use of standardized vocabularies.
- To be **Reusable**, data must be adequately described to a new user, have clear information about data-usage licenses, and have a traceable "owner's manual," or provenance.

https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf 6-4-18

NIH Strategic Plan for Data Science: Overview of Goals and Objectives

The central aim is to modernize the data-resource ecosystem to increase its utility for researchers and other stakeholders, as well as to optimize its operational efficiency.

Data Infrastructure	Modernized Data Ecosystem	Data Management, Analytics, and Tools	Workforce Development	Stewardship and Sustainability
•Optimize data storage and security •Connect NIH data systems	•Modernize data repository ecosystem •Support storage and sharing of individual datasets •Better integrate clinical and observational data into biomedical data science	•Support useful, generalizable, and accessible tools and workflows •Broaden utility of and access to specialized tools •Improve discovery and cataloging resources	•Enhance the NIH data-science workforce •Expand the national research workforce •Engage a broader community	•Develop policies for a FAIR data ecosystem •Enhance stewardship

https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf ***Released 6-4-18***