NHLBI Data Repository (BioLINCC)

Sean Coady
National Heart, Lung, and Blood Institute

NIH Health Care Systems Research Collaboratory
Steering Committee Meeting

2/24/14

NIH National Heart, Lung, and Blood Institute

Outline

Studies currently in the data repository

Policies – Timing for release of data

Policies – Typical data modifications

De-identification notes

Submitting datasets

Requesting data from the repository

‘Issues’ associated with data requests

Oversight

Data repository activities

Data Repository Studies

- Data from more than 600,000 participants in over 100 observational and clinical trial studies.

- 32 Observational studies:

  General population studies (Framingham, ARIC, CHS, CARDIA, MESA) *Post-menopausal women* (WHI-OS); *African-Americans* (Jackson), *Puerto Ricans* (PRHHP*); Asian Americans* (HHP); Patient populations *Sarcoidosis* (ACCESS); *Sickle Cell* (CSSCD); *Iron Overload* (HEIRS); *Primary Pulmonary Hypertension* (PPH); *Human retrovirus infected blood donors* (REDS-HTLV, REDS2-MS); *Alpha$_1$ Antitrypsin deficiency patients* (Alpha$_1$ registry); *Women with suspected myocardial Ischemia* (WISE, 1024); *hemophilia (MHCS)*; Adolescents (Bogalusa, NGHS)

Data Repository Studies

- 77 Clinical Trials:

  Asthma Clinical Research Network Trials (BAGS, CIMA, SOCS/SLIC, MICE, DICE, BARGE, IMPACT, PRICE, MIA, LARGE, BASALT/TALC); Acute Respiratory Distress Network (ARMA/LARMA, ALVEOLI, FACTT, LaSRS); Community Interventions (REACT, PAD); Hypertension Treatment (TOHP I&II, DASH, DASH-Sodium, PREMIER, PATHS); Primary Prevention (MRFIT, ACT, LRC-CPPT); Secondary Prevention (BEST, BHAT, AMIS, CASS, MAGIC, WAVE, CAST, SOLVD, TIMI II & III, PEACE, ESCAPE, AFFIRM, OAT, ENRICHD, BARI 2D); Lung function/Respiratory (LHS, NETT, IPPB), Diagnostic (PIOPED); Other Treatment (MSH, WLM, RTS, VATS); Primary/Secondary Prevention (ALLHAT, HDFP, SHEP, WHI-CT, ACCORD);Emergency Medicine (ROC-HS, ROC-PRIMED, ROC-CPR); Adolescents (CAMP and CAMPCS, DISC, HIFI, TAAG)

Timing for release of data

- **Clinical Trials:**
  - ✓ **3 years after the final visit of the participants to their clinical trial sites, or**
  - ✓ **2 years after publication of primary outcome paper**

Timing for release of data

- **Clinical Trials:**
  - ✓ **3 years after the final visit of the participants to their clinical trial sites, or**
  - ✓ **2 years after publication of primary outcome paper**

- **Observational Epidemiology Studies:**
  - ✓ **3 years after the completion of each examination or follow-up cycle, or**
  - ✓ **2 years after the baseline, follow-up, genetic, ancillary study, or other data set is finalized within the study for analysis**

Data De-identification

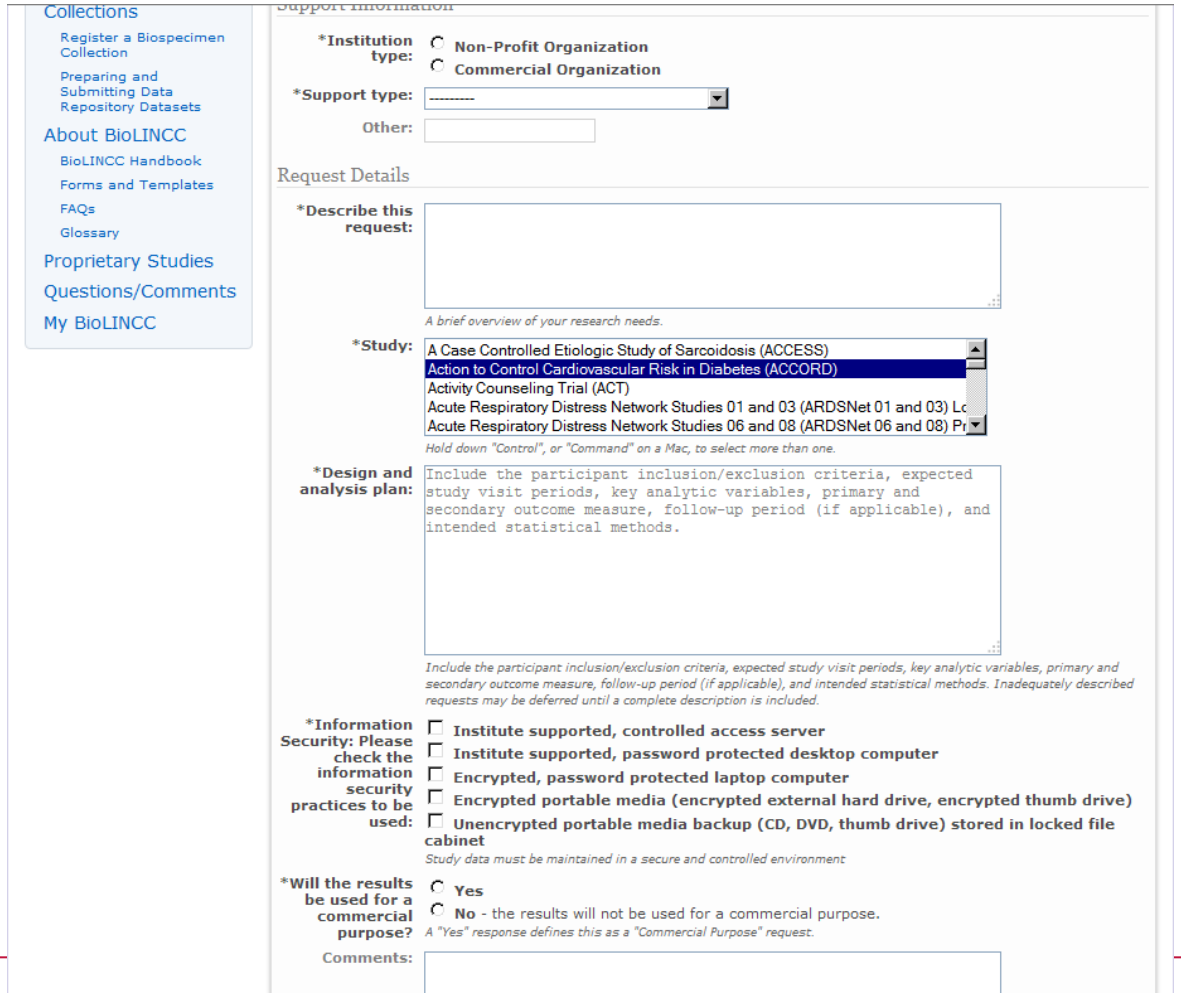
1. Be prepared in a manner consistent with the informed consent (eg. Commercial use restrictions, sharing restrictions, research use restrictions)
2. Obvious identifiers removed (i.e. HIPAA identifiers).
3. Geography is generally removed.
4. Dates are recoded as a time interval.
5. Sensitive data removed when not integral to the study
6. Other data may be recoded to minimize the possibility of re-identification of a participant.
    a. Top/bottom coding for height/weight
    b. Grouping of race/ethnicity
    c. Grouping of marital status, education, income, etc.

Submitting Data

- **Data:**
  - ✓ **Individual level data, preference is for SAS v9 format, or text files with appropriate coding manuals.**
- **Documentation**
  - ✓ **Case reporting forms, protocols, manual of operations, data dictionaries or coding manuals in MS Word format for 508 compliance (select documentation will be posted on the website)**
  - ✓**Basic readme file with a description of the data, de-identification notes, or data 'tips'**
- **Data and documentation uploaded to a secure FTP site**
- **Data and documentation reviewed for compatibility with guidelines and usability**

Requesting data from the Repository

1. Register on website and complete data request form



Requesting data from the Repository

1. Register on website and complete data request form

2. Obtain IRB approval or Institutional certification of exemption from IRB review (Framingham and Jackson require IRB approval)

Requesting data from the Repository

1. Register on website and complete data request form

2. Obtain IRB approval or Institutional certification of exemption from IRB review (Framingham and Jackson require IRB approval)

3. Sign a Research Materials Distribution Agreement that:
   - ✓ Prohibits the transfer of data
   - ✓ Limits use of the data to three years
   - ✓ Specifies a required acknowledgment in manuscripts/abstracts
   - ✓ Requires notification of any new projects

4. Obtain institutional sign-off on the research materials distribution agreement.

Data Requests - Issues

- There is no scientific merit review of requests for data or screening of manuscripts prior to Journal submission
  - “Head scratching” data requests
    - Pre-hypertension & risk for CVD (sickle-cell, DISC, Alpha1)
  - Possibility of poor or incorrect analyses getting published
    - “Embarrassing presentation”
    - “very sloppy work”
- Ask for a progress report once a year, but response rate is low.
- Manually search for manuscripts based on investigator name and institution

Oversight

Data repository operates under NHLBI IRB approved protocol

- Policy and procedures are somewhat insulated from arbitrary changes that may be encouraged by NHLBI staff or leadership
- Protocol is not strictly rigid. There is some flexibility in the protocol (there are both ‘rules’ and ‘guidelines’)
- Management of the data repository is also insulated from arbitrary demands by those requesting access.

Data Repository Activities

- Approximately 640 investigators from the US and 20 other countries have received data
- Nearly 35% of requested datasets include data from clinical trials
- More than 450 publications in 179 different Journals
- Several data repository investigators elected to collaborate with study investigators

Participant Privacy