Data Sharing through a NIH Central Database Repository:

A cross-sectional survey of BioLINCC users

APPENDIX

Table of Contents

BioLINCC Opt-in/Opt-out E-mail

In an effort to better understand the benefits, experiences, issues, and barriers to investigators requesting data from data repositories, Yale University will be conducting a short survey of investigators that received data from the NHLBI Data Repository at any time between 2007 and 2014. The survey should take approximately 15 minutes to complete and will provide valuable information on the experiences of users of data repositories. Please note that your responses to the survey will be completely anonymous. NHLBI will not receive nor have any access to any of the survey responses. NHLBI may request specific tables to further the Institute's understanding of potential areas for improvements; however, your individual responses will remain anonymous.

At the conclusion of the survey and publication of findings, your contact information will be permanently removed from all Yale University systems.

Please respond with either a 'Yes' indicating that NHLBI has your permission to share your contact information only with Yale University and only for the purpose of carrying out the survey of NHLBI Data Repository investigators. No permission to share is implied for any other purpose with any other third party.

Or Respond with a 'No' indicating that you do not wish to participate in the survey.

Yale University Invitation E-mail

Subject: Yale Survey on Using Data from NHLBI's Data Repository (BioLINCC)

Yale University, in collaboration with the National Heart, Lung, and Blood Institute (NHLBI) of the NIH, is conducting a short survey of investigators that received data from the NHLBI Data Repository (BioLINCC) at any time between 2007 and 2014. This survey is intended to better understand the benefits, experiences, issues, and barriers to investigators requesting data from this repository.

After having the opportunity to use data from BioLINCC, we hope that you will consider providing feedback on this valuable resource so that efforts can be made to enhance its use by others.

Participation in the survey is voluntary and responses will be anonymized. This survey is expected to take no more than 15 minutes to complete and all participants will be automatically entered in a drawing to win one of five Amazon.com© gift certificates worth $100.

Please complete this survey by June 4, 2015.

Thank you for your participation!

Survey

Yale University is conducting a survey of investigators who have accessed clinical research data through the Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC), at the U.S. National Heart, Lung, and Blood Institute (NHLBI) of the NIH, from 2007 through 2014.

Specifically, we are interested in investigators' experiences with the data and perceptions of the value, importance, and challenges of data sharing initiatives such as this one.

Results from this project are intended to improve investigators' experience with BioLINCC, as well as to inform future clinical data sharing efforts, and are intended to be published in the biomedical literature.

There are no physical risks associated with this project. Participation is voluntary and responses will be anonymized. This survey is expected to take no more than 15 minutes to complete and all participants will be automatically entered in a drawing to win one of five Amazon.com© gift certificates worth $100. Completion of the survey indicates consent to participate.

Please contact us at jessica.ritchie@yale.edu or (203) 200-5346 if you have questions or concerns related to the survey.

I. Reasons for Data Request

1. For what primary research purpose(s) did you request data through NHLBI's Data Repository (BioLINCC)? (Please check all that apply)
   - To conduct an independent scientific study/studies
   - To conduct a pilot/preliminary analysis
   - To conduct an analysis with bio-specimens
   - To learn more about the BioLINCC data request process
   - Other (please specify): [Free text field]

2. Did any of the following influence your decision to request data through NHLBI's Data Repository (BioLINCC)? (Please check all that apply)
   - The study was closed and individual participant-level data were not available elsewhere
   - Original study investigators suggested BioLINCC as the appropriate data source
   - Difficulties in establishing a collaboration with original study investigators
   - Collecting data of similar size and scope was not feasible
   - Insufficient financial resources available for primary data collection
   - Insufficient time for primary data collection
   - Insufficient experience with primary data collection
   - Insufficient mentorship to support primary data collection
   - Unable to access similar data through home institution electronic medical records
   - Other (please specify): [Free text field]

3. How did you intend to use the data? (Please check all that apply) *[**BIFURCATION QUESTION**: If participants answer "To be combined", they will be presented with Question #4. If participants answer "As a standalone data source" or "Other", they will be presented with Question #6.]*
   - As a standalone data source

To be combined with other data sources from BioLINCC
To be combined with other data sources not from BioLINCC (i.e., non-BioLINCC studies, other public data)
Other (please specify): [Free text field]

4. When combining the requested data with other data sources, was the purpose of the project to conduct a meta-analysis? *[**NOTE:** Question #4 only asked of participants who answered "To be combined" to Question #3; **BIFURCATION QUESTION**: If participants answer "Yes", they will be presented with Question #5. If participants answer "No", they will be presented with Question #6.]*
   Yes
   No

5. For the meta-analysis, which of the following was planned? (Please check all that apply) *[**NOTE:** Question #5 only asked of participants who answered "Yes" to Question #4.]*
   Summary-level data meta-analysis
   Participant-level data meta-analysis
   Other (please specify): [Free text field]

6. What was your primary research objective for the data requested through NHLBI's Data Repository (BioLINCC)? (Please check all that apply) [***BIFURCATION QUESTION**: If participants answer "New research", they will be presented with Question #7. If participants answer "Replication research" they will be presented with Question #8. If participants answer "Other", they will be presented with Question #9.]*
   New research
   Replication research
   Other

7. You indicated that your primary research objective was New research. Which of the following further describe this objective? (Please check all that apply) *[**NOTE:** Question #7 only asked of participants who answered "New research" to Question #6.]*
   To examine secondary endpoints
   To examine subgroup populations
   To leverage the data for a research question unrelated to the original research design (i.e., examine lost-to-follow-up rates or endpoints used in clinical trials)
   To leverage the data to create a cohort for comparison to another study
   Other (please specify): [Free text field]

8. You indicated that your primary research objective was Replication research. Which of the following further describe this objective? (Please check all that apply) *[**NOTE:** Question #8 only asked of participants who answered "Replication research" to Question #6.]*
   Replicate the main study primary endpoint findings
   Replicate the main study secondary endpoint findings
   Replicate the main study subgroup findings (for primary and/or secondary endpoints)

Other (please specify): [Free text field]

9. You indicated that your primary research objective was Other. Which of the following further describe this objective? (Please check all that apply) *[**NOTE:** Question #9 only asked of participants who answered "Other" to Question #6.]*
   Statistical methods research
   Epidemiological research
   Preliminary research to be used as part of a grant proposal
   Other (please specify): [Free text field]

10. Did your research focus on a medical product intervention (i.e., drug, biologic, medical device? *[**BIFURCATION QUESTION**: If participants answer "Yes", they will be presented with Question #11. If participants answer "No", they will be presented with Question #12.]*
    Yes
    No

11. What was the focus of your primary research question? (Please check all that apply) *[**NOTE:** Question #11 only asked of participants who answered "Yes" to Question #10.]*
    Efficacy
    Safety
    Pharmacodynamics
    Other (please specify): [Free text field]

II. Interactions with Original Study Investigators

12. Prior to or after requesting data through NHLBI's Data Repository (BioLINCC), did you contact the original study investigators to obtain the data? *[**BIFURCATION QUESTION**: If participants answer "Yes", they will be presented with Question #13. If participants answer "No", they will be presented with Question #16.]*
    Yes
    No

13. Did the original study investigators approve your data request? [***NOTE:** Question #13 only asked of participants who answered "Yes" to Question #12; **BIFURCATION QUESTION**: If participants answer "Yes", they will be presented with Question #14. If participants answer "No", they will be presented with Question #15.]*
    Yes
    No

14. If the original study investigators approved your data request, for what reason(s) did you also request data through BioLINCC? (Please check all that apply) *[**NOTE:** Question #14 only asked of participants who answered "Yes" to Question #13.]*
    Wanted to validate data made available by original study investigators

Original study investigators required co-authorship to make data available
Original study investigators required control of publication to make data available
Original study investigators required control of study design to make data available
Original study investigators required control of data analysis to make data available
Data made available by original study investigators had no or poor accompanying documentation
Data made available by original study investigators were poorly organized and could not be prepared for analysis
More straightforward to access data through BioLINCC
Other (please specify): [Free text field]

15. What reasons did the original study investigators provide for not approving your data request? (Please check all that apply) *[**NOTE:** Question #15 only asked of participants who answered "No" to Question #13.]*
    No interest in collaborating with external investigators
    Data cannot be made available to external investigators because original human subject consent forms do not allow
    Data cannot be made available to external investigators because of intellectual property issues
    Data cannot be made available to external investigators because of patient confidentiality issues
    BioLINCC was suggested as the appropriate data source
    A reason was not provided
    Other (please specify): [Free text field]

16. Prior to or after requesting data through BioLINCC, did you contact the original study investigators to request collaboration? *[**BIFURCATION QUESTION**: If participants answer "Yes", they will be presented with Question #17. If participants answer "No", they will be presented with Question #20.]*
    Yes
    No

17. For what reason(s) did you request collaboration with the original study investigators? (Please check all that apply) *[**NOTE:** Question #17 only asked of participants who answered "Yes" to Question #16.]*
    Needed additional data that were not included in the files provided by BioLINCC
    Needed additional statistical expertise due to data complexity
    Needed additional content expertise due to study design complexity
    Needed additional clinical expertise related to study question
    Wanted to work with original study investigators
    Other (please specify): [Free text field]

18. Did the original study investigators accept your request for collaboration? *[**NOTE:** Question #18 only asked of participants who answered "Yes" to Question #16; **BIFURCATION QUESTION**: If*

participants answer "No", they will be presented with Question #19. If participants answer "Yes", they will be presented with Question #20.]

Yes
No

19. What reasons did the original study investigators provide for not accepting your request to collaborate? (Please check all that apply) [***NOTE:*** *Question #19 only asked of participants who answered "No" to Question #18.]*

    A reason was not provided
    Original study investigators were too busy
    Original study investigators felt the research question was low priority
    Original study investigators did not have funds to support collaboration
    Other (please specify): [Free text field]

III. Data Repository Experience

20. How did you learn about the NHLBI's Data Repository (BioLINCC)? (Please check all that apply)

    Internet search
    Communications with NHLBI
    Colleagues/other investigators
    Directed to BioLINCC by study investigators
    Other (please specify): [Free text field]

21. Is there anything you would have liked to have known prior to accessing the BioLINCC data?

    No
    Yes (please provide further details): [Free text field]

22. Were the data you received from BioLINCC suitable for your originally proposed project? [***BIFURCATION QUESTION****: If participants answer "Yes", they will be presented with Question #24. If participants answer "No", they will be presented with Question #23.*]

    Yes
    No

23. For what reason(s) was the data unsuitable for your proposed project? (Please check all that apply) [***NOTE:*** *Question #23 only asked of participants who answered "No" to Question #21.*]

    Data had no or poor accompanying documentation; could not determine if data were suitable
    Data were too complicated to use; could not determine if data were suitable
    Data were poorly organized; could not be adequately prepared for analysis
    Data had too many missing values; could not be adequately prepared for analysis
    Proposed main outcome variable was not available in data
    Main outcome variable was not available in data at the time points proposed for study
    Proposed main independent variable was not available in data
    Other (please specify): [Free text field]

24. Please consider your experience using the data you received from BioLINCC. Do you agree or disagree with the following statement: The documentation and data dictionaries (i.e., meta-data) received from BioLINCC were useful. [***BIFURCATION QUESTION:*** *If participants answer "Somewhat disagree" or "Strongly disagree", they will be presented with Question #25. Otherwise, they will be presented with Question #26.]*
    Strongly agree
    Somewhat agree
    Somewhat disagree
    Strongly disagree

25. You chose "Somewhat disagree" or "Strongly disagree." Please briefly explain your answer. [***NOTE:*** *Question #25 only asked of participants who answered "Somewhat disagree" or "Strongly disagree" to Question #24.]*
    [Free text field]

III. Project Details

26. Has your project been completed? *[**BIFURCATION QUESTION:*** If participants answer "No", they will be presented with Question #27. If participants answer "Yes", they will be presented with Question #31.]
    Yes
    No

27. For what reason(s) was the project not completed? (Please check all that apply) *[**NOTE:** Question #27 only asked of participants who answered "No" to Question #26.]*
    Project is in analysis/manuscript draft stage
    Data unsuitable for proposed project
    Realized investigator/team did not have sufficient expertise to analyze the data
    Lack of funding
    Lack of programming/statistical support
    Too busy with other responsibilities
    Research fellow or collaborator expected to lead project no longer affiliated with investigator
    Investigator no longer active in clinical research
    Others published same/similar work on same/similar data
    Other (please specify): [Free text field]

28. Do you plan to complete the project? *[**NOTE:** Question #28 only asked of participants who answered "No" to Question #26. **BIFURCATION QUESTION:*** If participants answer "No", they will be presented with Question #29. If participants answer "Yes", they will be presented with Question #31.]
    Yes
    No

29. Was there an issue with the BioLINCC data that prevented you from completing the project? *[**NOTE:** Question #29 only asked of participants who answered "No" to Question #28. **BIFURCATION QUESTION:** If participants answer "Yes", they will be presented with Question #30. If participants answer "No", they will be presented with Question #39.]*
    Yes
    No

30. Please briefly explain your answer. *[**NOTE:** Question #30 only asked of participants who answered "Yes" to Question #29.]*
    [Free text field]

31. Does your completed or anticipated final project differ from your pre-specified project? *[**NOTE:** Question #31 only asked of participants who answered "Yes" to either Question #26 or #28. **BIFURCATION QUESTION:** If participants answer "Yes", they will be presented with Question #32. If participants answer "No", they will be presented with Question #33. **NOTE:** If participants answer "No" to #26 and "No" to #31, they will be presented with #38]*
    Yes
    No

32. In what ways does your completed or anticipated final project differ from your pre-specified project? (Please check all that apply) *[**NOTE:** Question #32 only asked of participants who answered "Yes" to Question #31.]*
    Modified planned data source by combining data received from BioLINCC with other data sources
    Modified planned data source by not combining data received from BioLINCC with other data sources
    Modified study sample
    Modified primary endpoints
    Modified secondary endpoints
    Modified selection of main independent variables
    Modified statistical analysis plan
    Other (please specify): [Free text field]

33. Was your research project published? [***NOTE:** Question #33 only asked of participants who answered "Yes" to Question #26*. ***BIFURCATION QUESTION:** If participants answer "Yes", they will be presented with Question #34. If participants answer "No", they will be presented with Question #36.]*
    Yes
    No

34. In what format was your research project published? (Please check all that apply) [***NOTE:** Question #34 only asked of participants who answered "Yes" to Question #33*.]
    Original research article in a peer-reviewed biomedical journal
    Systematic review/meta-analysis in a peer-reviewed biomedical journal
    Non-systematic review article in a peer-reviewed biomedical journal
    Commentary / viewpoint / editorial in a peer-reviewed biomedical journal
    Letter in correspondence in a peer-reviewed biomedical journal
    Weblog post or other on-line forum

Self-published
Other (please specify): [Free text field]

35. Please provide the publication citation and PubMed ID (if applicable) or other citation (such as web address). [***NOTE***: *Question #35 only asked of participants who answered "Yes" to Question #33.*]
    [Free text field]

36. When attempting to publish your research, were any of the following concerns raised by editors or peer reviewers during the peer-review process? (Please check all that apply)
    I did not attempt to publish
    No substantive concerns were raised beyond minor comments and suggestions to clarify/improve the research
    Concern that I was not one of the original study investigators
    Concern about the original study design that I could not address
    Concern about my research project design that I could not address without additional data
    Concern about my research project design unrelated to the data
    Concern about my research methodology and analysis
    Concern about the importance of my research
    Other (please specify): [Free text field]

37. How many research projects did you complete, or do you plan to complete, through your single request? [***NOTE:*** *Question #38 only asked of participants who answered "Yes" to either Question #26 or #28.*]
    One
    Two
    Three
    Four or more

38. Did using data from BioLINCC aid in any future grant applications? (Please check all that apply)
    Yes, use of the BioLINCC data established a publication record that was then included in a grant application
    Yes, use of the BioLINCC data furthered the understanding of questions of particular interest/identified gaps that then served as the basis of a grant application
    Yes, other (please specify): [Free text field]
    No

39. What was the primary funding source used to support this project?
    Self-funded
    NIH
    Non-NIH Federal
    Non-Profit Organization or Foundation in US
    Industry
    Non-US Government or Organization
    Other (please specify): [Free text field]

40. What additional funding source(s) were used to support this project? (Please check all that apply)
    - Self-funded
    - NIH
    - Non-NIH Federal
    - Non-Profit Organization or Foundation in US
    - Industry
    - Non-US Government or Organization
    - Other (please specify): [Free text field]
    - None

IV. Requestor Demographics

41. Which of the following best classifies your primary employer at the time when you requested data through BioLINCC?
    - Academic Institution
    - Private Industry
    - Non-Profit Organization
    - For-Profit Hospital
    - Government
    - Other (please specify): [Free text field]

42. How would you classify your career status with respect to clinical or epidemiological research at the time when you requested data through BioLINCC?
    - In training (< 3 years of active engagement in clinical research, still receiving formative training in research methods)
    - Early stage career (3-10 years of active engagement in clinical research)
    - Established in the field (> 10 years of active engagement in clinical research)

43. Have you ever been closely involved (as Principal or Co-Investigator) in the conduct of a randomized controlled trial?
    - Yes
    - No

44. Have you ever deposited clinical trial data in the BioLINCC repository?
    - Yes
    - No

45. Please indicate your age range.
    - 34 years or younger
    - 35-49 years
    - 50-64 years
    - 65 years or older
    - Prefer not to answer

46. Please indicate your gender.
    - Male
    - Female

Prefer not to answer

47. Please indicate your ethnicity.
    Hispanic or Latino
    Not Hispanic or Latino
    Prefer not to answer

48. Please indicate your race.
    American Indian or Alaska Native
    Asian
    Black or African American
    Native Hawaiian or Other Pacific Islander
    White
    Other (please specify): [Free text field]
    Prefer not to answer

49. Thank you for completing this survey. To be eligible for entry into the Amazon.com gift certificate drawing, please provide your email address. Your email address will be kept separate from your survey responses in order to ensure anonymity.

50. Would you like to receive a notification when the results of this study are published? If so, please re-enter your email address. Your email address will be kept separate from all other responses to ensure confidentiality.