NIA
Data Management Plans - Help and Resources for Researchers

May 28, 2020
NACDA Director Dr. James McNally
Director of The Office of Research Resources (ORR) Dr. Partha Bhattacharyya
NACDA Project Manager Kathryn Lavender
Plan for Today

• NIA perspective and guidance from Dr. Partha Bhattacharyya

• Why data sharing matters

• Resources available from ICPSR and NACDA

• Preparing data for sharing
NIH Data Sharing Policy

- Consider Data Sharing Methods
- Ensure Timeliness
- Protect Human Subjects
- Understand Proprietary Data Constraints
- Provide Complete Data Documentation
- Consider Requesting Funds for Data Sharing

NIH Data Sharing Policy

Data sharing is essential for expedited translation of research results into knowledge, products and procedures to improve human health.

The Final NIH Statement on Sharing Research Data was published in the NIH Guide on February 26, 2003. This is an extension of NIH policy on sharing research resources, and reaffirms NIH support for the concept of data sharing. The new policy becomes effective with the October 1, 2003 receipt date for applications or proposals to NIH.

- Data Sharing Regulations/Policy/Guidance Chart for NIH Awards (08/30/2006) (MS Word - 58 KB) - This chart is designed as a quick guide only for the purpose of identifying various data sharing regulation/policy/guidance documents applicable to NIH funding.
- NIH Data Sharing Policy and Implementation Guidance (03/05/2003) - Guidance providing the NIH policy statement on data sharing and additional information on the implementation of this policy.
  - Frequently Asked Questions - Data Sharing (02/16/2004) - Listing of Frequently Asked Questions that will be updated as new questions are received. Please check back periodically for new questions and answers.

- Data Sharing Workbook (PDF - 75 KB) or (MS Word - 74 KB) (02/16/2004) - Workbook to show how investigators working in a variety of scientific areas have shared their data.
- NIH Data Sharing Brochure (PDF - 244 KB) (05/20/2003) - Printable brochure that summarizes main elements of the NIH Data Sharing Policy.

- Testimonials (MS Word - 22 KB) (03/05/2003) - First-hand accounts from researchers who have shared data.
- Other Data Sharing Documents and Resources (02/19/2004) - Additional resources relating to data sharing.
- For NIH Staff Use (02/16/2004)

Note: For help accessing PDF, RTF, MS Word, Excel, PowerPoint, Audio or Video files, see Help Downloading Files.
Frequently Asked Questions (FAQs)

Data Sharing

Got Questions? We’ve Got Answers! The NIH and its individual Institutes/Centers offer an array of Frequently Asked Questions (and Answers) in response to inquiries on policies and programs affecting the grants process. Bookmark this page for future reference and watch for updates.

1. Why should I share my final research data?
2. Who benefits from data sharing?
3. Is data sharing widely accepted as a good practice?

Go to: https://grants.nih.gov/faqs#/data-sharing.htm
NIA – Stay in the Know

https://www.nia.nih.gov/research/blog

Get weekly updates on NIA funding policies and research priorities.

Subscribe
WHY SHARE DATA?
• Behavioral and Social Research sciences (BSR) has a long history of making data available for secondary research.

• ICPSR was one of the first repositories established under this philosophy, beginning with Political Science data, but now we have been joined by other publicly available repositories.

• The US Government also has historically treated its data, particularly health data, as a public good, opening sharing the NHIS and other federally funded surveys and administrative data.

• The United States is somewhat unique in this, as data availability in other nations can be tightly controlled and difficult to use for secondary analysis.

• Over the past 20 years we have seen significant increases in data sharing among non-BSR disciplines such as biomedical studies, epidemiology and psychology. This has greatly enhanced our understanding in these fields.
Benefits of Data Sharing

• It fulfills your obligation to your funding agency.

• It promotes “Team Science,” the multidisciplinary use of research data.

• It can generate **new perspectives** and uses for your data to explore research questions that were not part of the original study design.

• It **increases use of a study** beyond what primary investigators can reasonably generate: HRS (6000 pubs), MIDUS (1292 pubs), NSHAP (1940 pubs).

• It **builds collaborations**, multidisciplinary opportunities, education and training uses, and increases the profile of primary researchers.

• **It makes you a better researcher.**
IT MAKES YOUR DATA DISCOVERABLE

Study of Women’s Health Across the Nation (SWAN): Cross-Sectional Screener Dataset, [United States], 1995-1997 (ICPSR 4368)

Citation


https://doi.org/10.3886/ICPSR04368.v5

Export Citation:

• RIS (generic format for RefWorks, EndNote, etc.)
• EndNote

NACDA – Data on Aging
Why have a data sharing plan that simplifies the use of your data?

A properly Designed Data Sharing Plan Can have a Significant Impact on the Use and Scientific Impact of Your Study
## LIFE - Lifestyle Interventions and Independence for Elders

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Last Updated</th>
<th>Study Period</th>
<th>NIA Division</th>
<th>Number of Subjects</th>
<th>Study Open Dates for Data</th>
<th>Clinical Trial URL</th>
<th>Primary Publication URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Trial</td>
<td>1/30/2019</td>
<td>2004 - Present</td>
<td>DGCG</td>
<td>1635</td>
<td>1/23/2018</td>
<td><a href="https://clinicaltrials">https://clinicaltrials</a>...</td>
<td><a href="https://www.ncbi.nlm">https://www.ncbi.nlm</a>...</td>
</tr>
</tbody>
</table>

### Resources Available

- Specimens and Study Datasets
- Study Publications (93)

### Materials Available

- DNA
- Plasma
- Serum
- Urine
### SWAN - Study of Women's Health Across the Nation

**Study Type:** Epidemiology Study  
**Last Updated:** 1/30/2019  
**Study Period:** 1994 - Present

**NIA Division:** DCCG  
**Number of Subjects:** 3302

**Study Open Dates:**  
**Specimen:** 1/23/2019  
**Clinical Trail URL:** N/A

**Primary Publication URL:** [https://www.swanst...](https://www.swanst...)

**Resources Available:**  
- Specimens and Study Datasets  
- Study Publications (448)

**Materials Available:**  
- DNA  
- Immortalized Cell Lines  
- Plasma Citrate  
- Plasma EDTA  
- Serum  
- Urine
Concerns About Data Sharing

Among the most common concerns mentioned in the “request for comments” on the new NIH data sharing requirements were made in regards to protecting respondent confidentiality.

• Data sharing does not require disclosure.
• A compliant data sharing plan will minimize disclosure risk while maximizing data access.
• All data can be shared safely if the appropriate safeguards are put in place.
• Confidentiality is not an excuse for not sharing data!
Developing a Data Management Plan

• NACDA and ICPSR have **decades of experience** in creating and developing data management plans to address best practices in data sharing.

• **We manage data** with a variety of structures, needs and requirements.

• We offer data management plans from public distribution to the physical enclaving of highly sensitive data.

• NACDA’s role is to advise and facilitate researchers in the **best practices** for data management and data sharing.

• NACDA can help you navigate the development of a **successful data sharing plan**.

• NACDA can help you distribute **NIA funded data** through our repository or through our open-source self-publishing site.

• NACDA can help you develop a data management plan for NIA data even if you chose to self distribute your study.

• **We want to help you be successful and productive.**
Data Management & Curation

ICPSR stores, curates, and provides access to scientific data so others can reuse the data and validate research findings. Curation, from the Latin “to care,” is the process that ICPSR uses to add value to data, maximize access, and ensure long-term preservation.

Data curation is akin to work performed by an art or museum curator. Through the curation process, data are organized, described, cleaned, enhanced, and preserved for public use, much like the work done on paintings or rare books to make the works accessible to the public now and in the future. With the modern Web, it’s increasingly easy to post and share data. Without curation, however, data can be difficult to find, use, and interpret. Through curation, ICPSR provides meaningful and enduring access to data.

**Quality**

Data at ICPSR are enhanced with meaningful information to make it complete, self-explanatory, and usable for future researchers. As a repository, ICPSR adheres to standards that demonstrate it is organizationally, procedurally, and technologically sound as a trustworthy data custodian.

- Preparing Data for Archiving
- Guide to Archiving Social Science Data for Institutional Repositories (pdf)
- Curating Data at ICPSR

**Preservation**

Digital preservation is the proactive and ongoing management of digital content to lengthen the lifespan and mitigate against loss. ICPSR preserves its data resources for the long-term, guarding against deterioration, accidental loss, and digital obsolescence. ICPSR has a 50-year track record of reliably storing research data.

- Trusted Digital Repositories
- Digital Preservation Policies and Planning at ICPSR

**Access**

ICPSR hosts data in a repository with powerful search capabilities. Indexed by all the major search engines, ICPSR data are easily discoverable and widely accessible to the public.

- Data Management Plans
- Guidelines for OSTP Data Access Plan

**Confidentiality**

Data at ICPSR are screened for confidentiality and privacy concerns. Stringent protections are in place for securing and distributing sensitive data.
ICPSR Data Management Plan Resources

Guidelines for Effective Data Management Plans

Many federal funding agencies, including NIH and most recently NSF, are requiring that grant applications contain data management plans for projects involving data collection. To support researchers in meeting this requirement, ICPSR is providing guidance on creating such plans.

These guidelines can also be downloaded as a single PDF.

Go to the link below to watch ICPSR’s Data Management Plans Webinar:

https://www.icpsr.umich.edu/web/pages/datamanagement/dmp/index.html
ICPSR FAQs on Data Management Plans

- Are letters of support from ICPSR necessary when applying for National Institute of Justice (NIJ) sponsored research?
- Do I need to pay ICPSR to make data from my NSF grant available?
- Does ICPSR have people who can assist in the preparation of metadata?
- Can research data be copyrighted?
- Do you have recommendations for intellectual property language for proposals that are collaborative and involve two or more PIs at two or more institutions who will be generating data together?
- Have proposals been rejected because of an unacceptable data management plan? Or do agencies work with the PI to improve a data management plan if they find the proposal worth funding?
- How do data management plans relate to institutional statements of capabilities?
- How much time is needed prior to proposal submission to obtain a letter of support from ICPSR?
- Is there a list of generic data formats for those data types that are non-ASCII, i.e., images and sounds, etc.?
- Is videotaped data included as part of the data-sharing requirement? How do you protect the confidentiality of your participants?
- Should a researcher obtain permission from a data archive before naming it in a data management plan?
- The generic examples that ICPSR provides are great, but they are boilerplate. Would that be a problem during agency review?
- What if the data I use come from a confidential source?
- Which of your data management plan examples are for NSF and which are for NIH?
- Will NIH and NSF allow investigators to request funds to pay for data archiving? Assistance in preparing data for archiving in the final year of a grant?
ICPSR Data Management Plan Resources

Guide to Social Science Data Preparation and Archiving

Best Practice Throughout the Data Life Cycle: 6th Edition

Inside the Guide

About ICPSR
- Why Should I Archive Data?
- How Do I Deposit Data?
- What Services Does ICPSR Offer?
- Using the Guide

Importance of Data Sharing and Archiving
- Planning Ahead for Archiving and Preservation of Data
- The Data Life Cycle

Proposal Development and Data Management Plans
- Recommended Elements for Data Management Plans
- Other Considerations
- Summary Table: Elements of a Data Management Plan
- An Example Data Management Plan for Depositing Data with ICPSR

Project Start-Up
- Importance of Good Data Management

https://www.icpsr.umich.edu/web/pages/deposit/guide/
We preserve and share data on aging.
The NACDA data archive has spent the past 40 years identifying, gathering and organizing data on aging and the aging life course so we can share these resources with the research community.

NACDA currently supports approximately 1,600 individual studies on ageing and health and adds new studies to our collections each year.

NACDA is part of ICPSR
Twitter @NACDA_Aging
NACDA Data Management Plan Resources

Need a Letter of Support?
Contact us:
icpsr-nacda@umich.edu
So You Want to Share Your Data…

Why deposit your data with us?

• To fulfill grant obligations
• To broaden research potential of the data
• It’s free (for you and secondary users)
• It will be preserved
• It’s easy to track usage
Preparing Data for Sharing

1. Remediate risk of respondent confidentiality concerns
2. Provide complete collection metadata (sampling info, etc.)
3. Provide complete variable metadata (variable and value labels)
4. Provide documentation (codebook, questionnaire, user guide)

Here is a sort of painful and funny video that accurately illustrates the need for core work that we do: https://www.youtube.com/watch?v=N2zK3sAtr-4
(No need to watch the whole thing - the message is clear pretty quickly.)
Preparing Data for Sharing

How long does it take?

- How many files do you have?
- How many variables are there?
- Are the labels included in the data file?
- Do you have a release date in mind?

Reach out to us – the timeframe for release depends on many factors.
Preparing Data for Sharing

Is your data **longitudinal**?

If so, make sure to highlight variables users should use to merge, share syntax, and consider providing a merged file.

https://giphy.com/gifs/ioelondon-life-closer-lifelong-vRGJJ4d4bcXPUNy1eH
Preparing Data for Sharing

Check out our “Data Tools/Resources” pages
Preparing Data for Sharing

Longitudinal data is collected from the same sample at different points in time. The sample can consist of individuals, households, establishments, and other units of observation and/or analysis. Using longitudinal data is a great way to measure change.

NACDA has longitudinal data organized by series and study, and even dataset within study.

For example, data organized by series means that we have several studies (usually 2-3 or more) that can be used together (and/or were intended to be used together) because they have the same questions across years, or because the studies have the same sample of respondents. Therefore, users can see all of the studies that are intended to be analyzed together by the principal investigator and have the components to do so (such as a consistent ID variable to sort and merge by). This also means that users will often need to download files from each study page in order to merge them, as there may not be a merged file already created/provided.

The SWAN series is an example of multiple waves by study within series, in addition to MIDUS and MIDIA, and NSHAP.

Data that are organized by dataset within study means that a single study was created and all of the waves and/or components of the whole study are downloadable from that same study page. The datasets are clearly meant to be used together, and there should be consistent variables to sort and merge by. Users may still need to download all of the study files or multiple files, however, they will only need to do this from a single study page.
Preparing Data for Sharing

Contact us!
Reach out to us at icpsr-nacda@umich.edu

We will help you figure it out.
Depositing Data with NACDA

Visit our site: https://www.icpsr.umich.edu/web/pages/NACDA/index.html

Select the “Deposit” tab

About Us

The National Archive of Computerized Data on Aging (NACDA), located within ICPSR, is funded by the National Institute on Aging. NACDA’s mission is to advance research on aging by helping researchers to profit from the under-exploited potential of a broad range of datasets.

NACDA acquires and preserves data relevant to gerontological research, processing as needed to promote effective research use, disseminates them to researchers, and facilitates their use. By preserving and making available the largest library of electronic data on aging in the United States, NACDA offers opportunities for secondary analysis on major issues of scientific and policy relevance.

The NACDA staff represents a team of professional researchers, archivists and technicians who work together to obtain, process, distribute, and promote data relevant to aging research.
Depositing Data with NACDA

Depositors should review this list prior to depositing. If you have questions or concerns, please email icpsr-nacda@umich.edu and we will work with you to resolve any issues. Not all of these items are required, however, the more complete information provided the more likely the materials will be easily replicated and reused.

Ease of Understanding:
- File and folder structure named with unique, descriptive titles; if providing multiple datasets, include dataset to documentation legend or name correspondingly
- Any publication or promotion deadlines related to the plan for the release of the data and materials
- Intended access type; public or restricted access
  - If restricted access is needed, plan to discuss dissemination with archive and collaborate on terms of use agreement
- Do you already distribute data with us? If so, is this deposit an update to the data you supplied in the past? Please provide ICPSR study number and dataset relation.

Complete Documentation:
- Original questionnaire
- Codebook; with unweighted frequencies is best for data comparison
- Identify and describe computed and derived variables
- Interviewer instructions
- User Guide
- Final Report
- Citations of related publications the data were used in
- Question text, DDI, or another text-source for the question
- Variable groupings (Especially for large datasets, it is useful to categorize variables into conceptual groupings)

Complete Study-Level Metadata:
- Clear and consistent title across files and metadata. **NOTE:** to conform to metadata standards, the title will appear on the ICPSR (NACDA) site as Title, Location, Years, such as Aging in Women (United States), 2005-2009.
- Description of the project; goals, main topics, etc.
Depositing Data with NACDA

NACDA welcomes and encourages deposits of digital data. Deposits are made using a secure form to describe the data collection and upload content.

Users must have or create a MyData, Facebook, or Google account to sign in to the deposit form.

For a discussion of best practices in preparing data for sharing, please refer to ICPSR’s Guide to Social Science Data Preparation and Archiving.

Are you planning your research project, but not sure about how you plan to archive your data afterwards? Check out our Guidelines for Effective Data Management Plans.

NACDA has two options for sharing your data!

NACDA offers curation services as staff obtain, process, preserve, and disseminate electronic data to researchers for secondary analysis in aging research. After a data collection is deposited with NACDA, the data and documentation will be reviewed, enhanced, and made available in several widely used formats, including ASCII, tab-delimited (for use with Excel), SAS, SPSS, and Stata. Documentation is provided as PDF.

The NACDA-Open Aging Repository (NACDA-OAR) offers free self-publishing options for gerontological researchers to share their research projects. Data collections published in NACDA-OAR are immediately distributed as-is for research use in the same condition and format submitted by the depositor.

2 ways to deposit
Midlife in the United States (MIDUS 3), 2013-2014 (ICPSR 36346)

Version Date: Apr 30, 2019

Principal Investigator(s):
Carol Ryff, University of Wisconsin-Madison; David Almeida, Pennsylvania State University; John Ayanian, University of Michigan; Neil Brinkley, University of Wisconsin-Madison; Deborah S. Caru, Rutgers University; Christopher Coe, University of Wisconsin-Madison; Richard Davidson, University of Wisconsin-Madison; Joseph Grigsby, Florida State University; Amin Karanam, University of California-Los Angeles; Robert Kuepper, University of Minnesota; Marjorie Lachman, Brandeis University; Gayle Lovin, University of Wisconsin-Madison; Marsha Malick, University of Wisconsin-Madison; Daniel Mroczek, Northwestern University; Barry Radio, University of Wisconsin-Madison; Teresa Seeman, University of California-Los Angeles; Richard Sloan, Columbia University; Duncan Thomas, Duke University; Maxine Weinstein, Georgetown University; David Williams, Harvard University

Series:
- Midlife in the United States (MIDUS) Series

https://doi.org/10.3886/ICPSR36346.v7

Version V7 (see more versions)
Depositing Data with NACDA-OAR

https://www.openicpsr.org/openicpsr/nacda-oar

About the NACDA - Open Aging Repository

The NACDA - Open Aging Repository (NACDA-OAR) from the National Archive of Computerized Data on Aging (NACDA) offers free self-publishing options for gerontological researchers to share their research projects. Data collections published in the NACDA-OAR are immediately distributed as-is for immediate research use. NACDA-OAR services are primarily meant to be a resource for NIA-funded projects, although all aging-related data may use the service.

NACDA-OAR self-publications should include all data, annotated program code, command files, and documentation that are needed to replicate research findings. For general instructions for publishing data, see the NACDA-OAR Deposit Instructions. For best practice in preparing data for sharing, please refer to ICPSR’s Guide to Social Science Data Preparation and Archiving.

Start Your Deposit

NACDA's Role

NACDA does not alter datasets deposited in NACDA-OAR in any way. Data are distributed in the same condition and format submitted by the depositor.

NACDA-OAR is one of many services provided by the NACDA Program on Aging to archive and promote gerontological research, aging health and a better understanding of the life course. Please contact NACDA at icpsr-nacda@umich.edu with any questions you might have about your data and archiving options.

NACDA staff may review data deposited and published to the NACDA - Open Aging Repository at random. If any issues are found during these random reviews, NACDA will contact depositors to remediate the concerns. In the event any data contain direct identifiers or pose a risk to respondent anonymity, NACDA reserves the right to make the data unavailable.
Why Self Publish with NACDA-OAR?

- You control the content and versions
- Quick availability and DOI
- Discoverable through ICPSR and NACDA site searches
- Fulfills your NIH Data Sharing Requirements
- Rapid turnover and release
- We track the use of your data
# National Poll on Healthy Aging (NPHA), [United States], October 2017

**Principal Investigator(s):** Preeti Malani, University of Michigan; Jeffrey Kullgren, University of Michigan; Erica Solway, University of Michigan

**Version:** V2

<table>
<thead>
<tr>
<th>Name</th>
<th>File Type</th>
<th>Size</th>
<th>Last Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPHA_wave2_Oct2017.dat</td>
<td>text/plain</td>
<td>830.2 KB</td>
<td>04/28/2020 10:57:AM</td>
</tr>
<tr>
<td>NPHA_wave2_Oct2017.sas</td>
<td>application/x-sas</td>
<td>50.5 KB</td>
<td>04/28/2020 10:17:AM</td>
</tr>
<tr>
<td>NPHA_wave2_Oct2017.sav</td>
<td>application/x-spss-sav</td>
<td>588.6 KB</td>
<td>04/28/2020 10:37:AM</td>
</tr>
<tr>
<td>NPHA_wave2_Oct2017.sps</td>
<td>text/x-spss-syntax</td>
<td>37.5 KB</td>
<td>04/28/2020 10:37:AM</td>
</tr>
</tbody>
</table>

**Usage Metrics**

- **Overall Project Metrics**
  - **Views:** 153
  - **Downloads:** 6
  - **Publications:** 5

**Download Detailed Metrics**

**Published Versions**

- V2 [2020-04-28]
NACDA or NACDA-OAR?

**NACDA**
- We publish and prioritize NIA-funded research
- Time to curate and convert files
- We perform confidentiality review and create page content and final file versions based on info you provided
- We manage versioning

**NACDA-OAR**
- Self-publish gerontological content
- Quick turnaround and DOI
- You perform confidentiality review, create page content, and upload final file versions
- You control versioning
About Us

The National Archive of Computerized Data on Aging (NACDA), located within ICPSR, is funded by the National Institute on Aging. NACDA’s mission is to advance research on aging by helping researchers to profit from the under-exploited potential of a broad range of datasets.

NACDA acquires and preserves data relevant to gerontological research, processing as needed to promote effective research use, disseminates them to researchers, and facilitates their use. By preserving and making available the largest library of electronic data on aging in the United States, NACDA offers opportunities for secondary analysis on major issues of scientific and policy relevance.

The NACDA staff represents a team of professional researchers, archivists and technicians who work together to obtain, process, distribute, and promote data relevant to aging research.
Contact Us and Stay Informed

Web Address - https://www.icpsr.umich.edu/icpsrweb/NACDA/

Twitter handle - @NACDA_Aging

Facebook - NACDA Program on Aging - @NACDA.Aging.Program

LinkedIn - linkedin.com/company/nacda-aging

Email icpsr-nacda@umich.edu
Thank you!!

Email us!

Look for these slides on the NACDA announcements page!

jmcnally@umich.edu
kfrania@umich.edu

NACDA – Data on Aging