

The Cancer Prevention and Control Research Network (CPCRN) Data Sharing Guide for the Dataverse Repository

Updated 8/15/24

Introduction

CDC/NIH data sharing policy encourages that data collected in their funded programs are made available to the public to “enable validation of research results, provide accessibility to high-value datasets, and promote data reuse for future research studies.” Aligned with that, the Cancer Prevention and Control Research Network (CPCRN) has built a data management and sharing plan to maintain the collected data in a secure environment and facilitate data sharing in a timely and secure manner. For transparency and reproducibility, the CPCRN commits to making primary research data from all studies publicly available via the Dataverse Repository. This international data repository allows archiving, sharing, and citation of scientific data, promoting effective data management and collaboration. Datasets will be deposited in the CPCRN Dataverse upon publication of associated manuscripts. To protect participant privacy, all datasets shared publicly from all CPCRN studies will be de-identified following HHS Regulations for the Protection of Human Subjects and HIPAA standards. Data Use Agreements will be established, reviewed, and signed by all involved parties to ensure responsible data access and usage. The CPCRN Coordinating Center and Steering Committee will jointly review and update this data management plan annually to ensure it aligns with current regulations and best practices.

To ensure transparency and discoverability, CPCRN will support the development of comprehensive documentation for all datasets deposited in the Dataverse repository. This includes:

- 1) Study details: a clear description of the study design, settings, purpose, and data collection methods.
- 2) Sample characteristics: information about the sample size and any key population characteristics relevant to the data, such as focusing on a specific gender or demographic group.

3) Qualitative Research: Share: Interview guide and codebook (if applicable) in the Dataverse repository; Optional: Upload de-identified transcripts. If unavailable, provide contact information for access requests; Optional: Upload de-identified supporting documents like interviewer notes, if applicable,

4) Quantitative Research: Share: De-identified original dataset, codebook, and measurement tools in Dataverse; Upload complete surveys used in the study, if applicable; Provide instructions for measurement tools other than surveys, if applicable,

5) Multi-/Mixed Method Research: Share: All relevant files for both qualitative and quantitative components as outlined above.

CPCRN Data Sharing Instructions

When:

Start drafting or sharing with the Coordinating Center	After the completion of data collection
Publish	As soon as the manuscript arising from the dataset is published

All datasets from CPCRN5 should be uploaded to the Dataverse by Monday, December 16, 2024.

Where: Dataverse data repository: CPCRN Dataverse
<https://dataverse.unc.edu/dataverse/cpcrn>

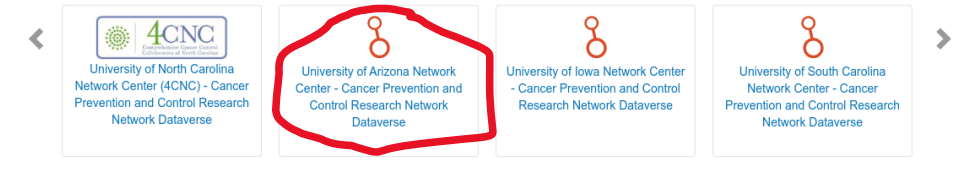
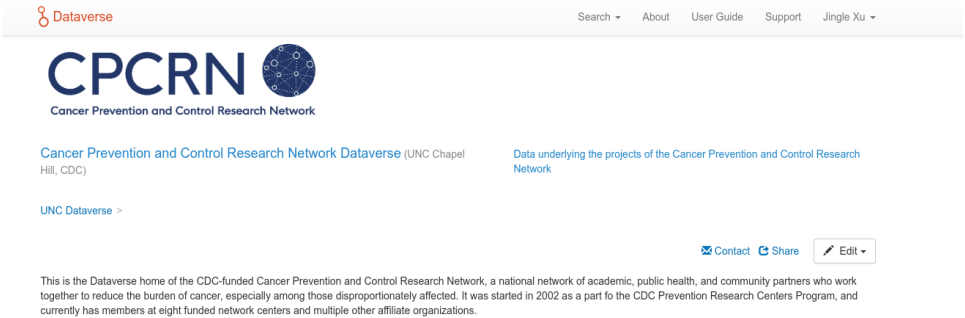
Note: if the dataset has been shared/uploaded elsewhere, a record of it can still be created on the Dataverse platform with a link to view the dataset.

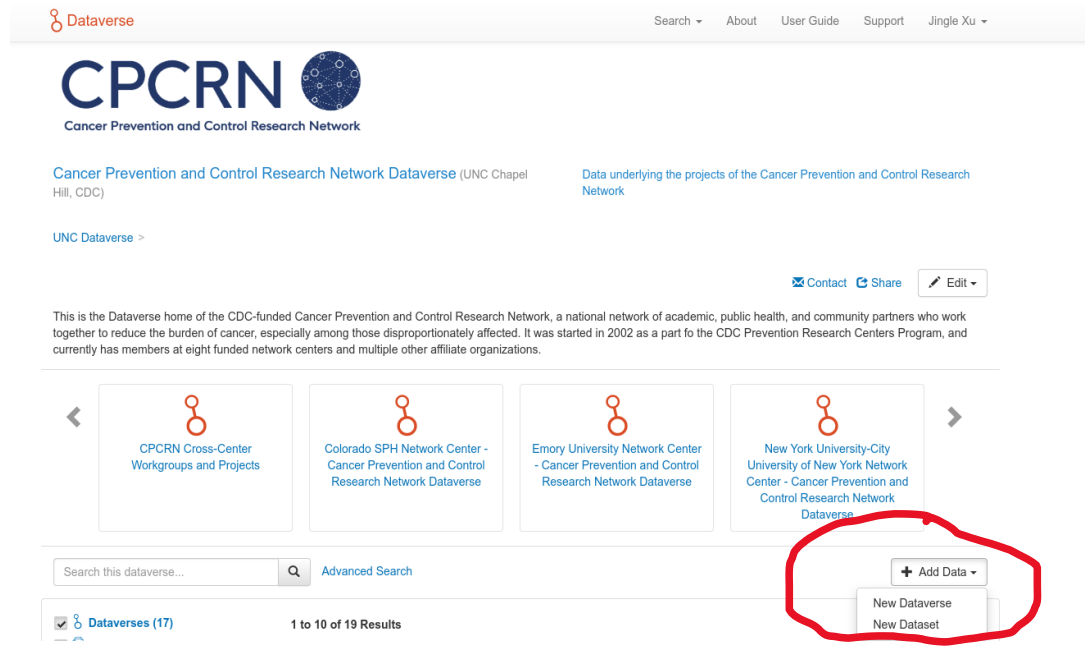
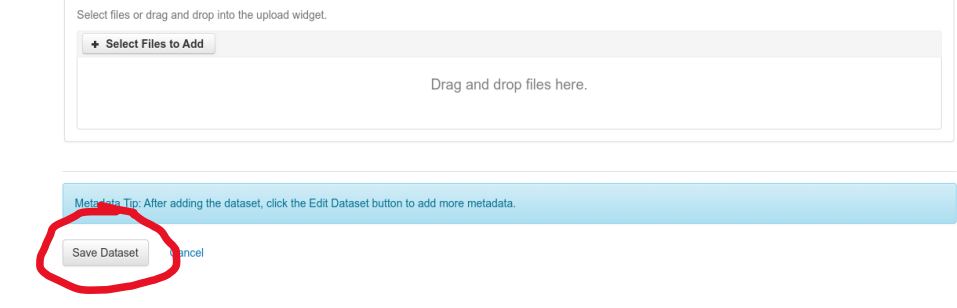
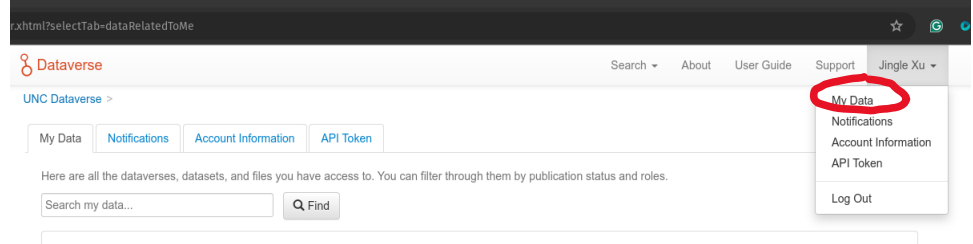
If your institution works with Dataverse, you may upload your dataset to the Dataverse of your institution that is linked to the CPCRN Dataverse.

How:

Option 1	<p>Share your dataset with the Coordinating Center and the Coordinating Center will upload your datasets to the CPCRN Dataverse.</p> <p><i><u>If you choose this way, please fill out the form via this link (https://forms.office.com/r/j5A3s3bSUK) and share your dataset files via Google Drive and notify Jingle Xu as instructed below.</u></i></p>
Option 2	<p>Upload your datasets to the CPCRN Dataverse directly.</p> <p><i>You don't have to complete the form if you choose this way.</i></p>

Option 1. Share your dataset with the CPCRN Coordinating Center	
Step 1	Learn about your institution's rules on sharing de-identified data. Act accordingly.
Step 2	<p>Prepare the study information and dataset files according to the tables in the next section via this link (https://forms.office.com/r/j5A3s3bSUK). De-identify any information or files that you plan to provide.</p> <p>Note: you can selectively make part of the uploaded files viewable or downloadable to the public</p>
Step 3	Upload the dataset metadata and required dataset files to your Center's Dataverse Google Drive folder. Make sure only authorized personnel have access to that folder. Notify Jingle Xu via email at jxu2@unc.edu .
Step 4	Delete your dataset files from your Center's Dataverse Google Drive folder after you receive notification from Jingle Xu that the files have been received.
Step 5	In 1-2 weeks, the Coordinating Center will share a ZIP file of the built dataset exported from the Dataverse data repository in your Center's Dataverse Google Drive folder and send an email notification. You will have 2 weeks to review the file and return an approval or any corrections that ought to be made to Jingle Xu at jxu2@unc.edu .
Step 6	Notify Jingle Xu via email at jxu2@unc.edu when the dataset is ready to be published (i.e., the publication arising from the dataset is published). <i>If the publication based on the dataset has been published before you share the dataset with the CPCRN team, please mention that in the approval email in step 5.</i>
Step 7	You will receive confirmation and a link to access the published dataset from Jingle Xu once the dataset is published.

Option 2. Guidance on uploading your dataset to Dataverse	
<p>Step 1</p>	<p>Prepare the study information and dataset files according to the tables in the next section. De-identify any information or files that you plan to provide.</p> <p>Note: you can selectively make part of the uploaded files viewable or downloadable to the public</p>
<p>Step 2</p>	<p>For external users to UNC: if your institution has access to the Dataverse service, upload your dataset to the Dataverse of your institution that is linked to the CPCRN Dataverse (e.g., University of Arizona Network Center – CPCRN Dataverse).</p> <p>This is the Dataverse home of the CDC-funded Cancer Prevention and Control Research Network, a national network of academic, public health, and community partners who work together to reduce the burden of cancer, especially among those disproportionately affected. It was started in 2002 as a part of the CDC Prevention Research Centers Program, and currently has members at eight funded network centers and multiple other affiliate organizations.</p>  <p>If your institution does not have access to the Dataverse service, <i>please select option 1 for now.</i></p>
<p>Step 3</p>	<p>Open the host Dataverse where you would like to upload the dataset to (e.g., CPCRN Dataverse)</p>  <p>This is the Dataverse home of the CDC-funded Cancer Prevention and Control Research Network, a national network of academic, public health, and community partners who work together to reduce the burden of cancer, especially among those disproportionately affected. It was started in 2002 as a part of the CDC Prevention Research Centers Program, and currently has members at eight funded network centers and multiple other affiliate organizations.</p>

<p>Step 4</p>	<p>Click on “Add Data”, select “New Dataset”</p>  <p>The screenshot shows the Dataverse homepage for CPCRN. At the top right, there is a navigation menu with 'Search', 'About', 'User Guide', 'Support', and 'Jingle Xu'. Below the CPCRN logo, there are links for 'Cancer Prevention and Control Research Network Dataverse (UNC Chapel Hill, CDC)' and 'Data underlying the projects of the Cancer Prevention and Control Research Network'. A 'Contact' button is visible. A paragraph describes the network's mission. Below this are four network center cards. At the bottom right, a '+ Add Data' button is circled in red, with a dropdown menu showing 'New Dataverse' and 'New Dataset'.</p>
<p>Step 5</p>	<p>Fill out the dataset metadata and upload dataset files accordingly. (Note: more items of metadata are available after you add the dataset and edit it.)</p>
<p>Step 6</p>	<p>Click on “Save Dataset” in the bottom left corner to add the dataset.</p>  <p>The screenshot shows a file upload interface. At the top, it says 'Select files or drag and drop into the upload widget.' Below is a 'Select Files to Add' button and a large area for dragging files. A blue tip box says 'Metadata Tip: After adding the dataset, click the Edit Dataset button to add more metadata.' At the bottom left, the 'Save Dataset' button is circled in red.</p>
<p>Step 7</p>	<p>After you add the dataset, the dataset is saved under “Draft”. You can access the dataset in draft by visiting “My Data” under your profile.</p>  <p>The screenshot shows the user profile dropdown menu. The 'My Data' option is circled in red. Other options include 'Notifications', 'Account Information', 'API Token', and 'Log Out'.</p>

Dataverse

UNC Dataverse >

My Data | Notifications | Account Information | API Token

Here are all the dataverses, datasets, and files you have access to. You can filter through the

Search my data...

Dataverses (4)

Datasets (200)

Files

Publication Status

Published (203)

Unpublished (1)

Draft (1)

In Review (0)

Deaccessioned (0)

Roles

Admin

File Downloader

Dataverse Creator

Contributor

1 to 10 of 204 Results

Replication Data for: The special sauce of th
learned in developing the evidence base, bu

Unpublished Admin Contributor

Feb 14, 2024 - University of North Car
Dataverse

Wheeler, Stephanie; Lee, Rebecca J;
Prajakta; Birken, Sarah A; Fernández
Brandt, Heather M; Williams, Rebecc
Control Research Network: 20 years |
translating research into practice", htt

Purpose: To clarify the ingredients of the "special
and loosely connected network of individuals bal

University of North Carolina Network Center

Published Admin

Mar 9, 2023 - Cancer Prevention and


This is the Dataverse for the Universi
and Control Research Network, called

Step 8

You are able to edit the dataset by visiting the dataset page.

Dataverse

Search ▾ About User Guide Support Jingle Xu ▾



University of North Carolina Network Center (4CNC) - Cancer Prevention and Control Research Network Dataverse (UNC Chapel Hill, CDC)

UNC Dataverse > Cancer Prevention and Control Research Network Dataverse >
University of North Carolina Network Center (4CNC) - Cancer Prevention and Control Research Network Dataverse >

Replication Data for: The special sauce of the Cancer Prevention and Control Research Network: 20 years of lessons learned in developing the evidence base, building community capacity, and translating research into practice

Draft Unpublished

Wheeler, Stephanie; Lee, Rebecca J; Young, Alexa L; Dodd, Adam; Ellis, Charlotte; Weiner, Bryan J; Ribisl, Kurt M; Adsul, Prajakta; Birken, Sarah A; Fernández, María E; Hannon, Peggy A; Hébert, James R; Ko, Linda K; Selman, Aaron; Vu, Thuy; Brandt, Heather M; Williams, Rebecca S, 2024, "Replication Data for: The special sauce of the Cancer Prevention and Control Research Network: 20 years of lessons learned in developing the evidence base, building community capacity, and translating research into practice", <https://doi.org/10.15139/S3/M6TE2F>, UNC Dataverse, DRAFT VERSION, UNF:6:mzJf50yJM9lv+yamr8rsQ== [fileUNF]

[Cite Dataset](#) [Learn about Data Citation Standards.](#)

Access Dataset ▾

Publish Dataset

Edit Dataset ▾

Contact Owner Share

Dataset Metrics ⓘ

0 Downloads ⓘ

Step 9	Click on the “Publish Dataset” button after you complete editing and are ready to share the dataset to the public.
Step 10	Select the Terms of Use
Step 11	Notify Jingle Xu at jxu2@unc.edu of the successful publication of the dataset and provide the link to the published dataset.

What to upload/provide:

The following table maps out what information and dataset files to provide or prepare for the sharing of data collected in qualitative, quantitative, and mixed/multi-method studies, respectively, via the Dataverse Repository. If you choose Option 1 (to share data via the CPCRN Coordinating Center), you are supposed to provide the following information and dataset files to the Coordinating Center via this link

(<https://forms.office.com/r/j5A3s3bSUK>) and your Center’s Dataverse [Google Drive](#) folder, separately. Please note the Dataverse Repository provides other dataset metadata items that are not included in this table. If you choose Option 2 (i.e., to upload datasets to the Dataverse Repository yourself), besides the required items, you can provide any other dataset metadata if applicable.

*required; †if applicable; others are completely optional

Multi-/Mixed Method Research	
<i>Study Information (Dataset Metadata)</i>	
*Mixed-method design	Design of the mixed-method study.
*Report the items of qualitative and quantitative research listed below	
<i>Dataset Files</i>	
*Upload the quantitative and qualitative dataset files listed below.	
Qualitative Research	
<i>Study Information (Dataset Metadata)</i>	
*Title	Full title of the project or dataset
Alternative title	The common name of the project or dataset (i.e., abbreviations of the title)
*Author(s)	Authors and affiliations of the publication or grant.
Alternative URL	If the dataset is also available elsewhere, put the link to other sources here.
Other ID	If the dataset is also available in another repository, put information here.
*Contact	Contact information to whom the interested ones can request more about the data from.
*Subject	Medicine, Health, and Life Sciences

*Language	Language of the dataset
*Keywords	Implementation Science, Cancer Prevention and Control, etc.
*Description	Write a summary including the following information about the dataset (maximum 500 words)
Design	Design of the study (e.g., Grounded Theory, Ethnography, Case Study, Qualitative Descriptive, etc.)
Sample Size	Total number of the participants in the dataset
Purpose	Purpose of the study
Setting	Country? Community? Hospital?
Participants	Who? Relevant to/diagnosed with what cancer?
†Related Publications	You can add multiple publications arising from this dataset or the most relevant one.
Notes	Add any additional information about this dataset
Time period covered	This refers to the time period that the dataset covers, <i>not the time period of data collection</i> .
†Date of collection	From [date] to [date] the data was collected.
*Kind of data	Include the type of data (interview transcripts?) and the data collection (focus group? One-to-one in-depth interviews?)
*Dataset Files	
*Interview Guide	Semi-structured interview guide or focus group guide
*Codebook	Final version of the codebook
†De-identified transcripts	Interview/focus group transcripts in which any identifiable information has been removed
†Interviewer Notes	If the interviewer notes are informative, please upload them as well.
Other	Any other documents that provide useful information about this dataset.
Quantitative Research	
<i>Study Information (Dataset Metadata)</i>	
*Title	Full title of the project or dataset
Alternative Title	The common name of the project or dataset (i.e., abbreviations of the title)
*Author(s)	
Alternative URL	If the dataset is also available elsewhere, put the link to other sources here.
Other ID	If the dataset is also available in another repository, put information here.

Contact	Contact information to whom the interested ones can request more about the data from.
*Subject	Medicine, Health, and Life Sciences
*Language	Language of the dataset
*Keywords	
*Description	Write a paragraph including the following information about the dataset (maximum 500 words)
Design	Design of the study (e.g., RCT, cross-sectional, cohort study, case-control study, etc.)
Sample Size	Total number of the participants in the dataset
Purpose	Purpose of the study
Setting	Country? Community? Hospital?
Response rate/attrition	
†Related Publications	You can add multiple publications arising from this dataset or the most relevant one.
Notes	Add any additional information about this dataset
Time period covered	This refers to the time period that the dataset covers, <i>not the time period of data collection</i> .
†Date of collection	From [date] to [date] the data was collected.
*Kind of data	Include the type of data (survey?)
*Dataset Files	
*The original de-identified dataset	For quantitative data, you must upload the complete de-identified dataset
*Codebook	Codebook of the dataset
*Survey (If applicable)	If your study uses surveys to collect the data, please share the surveys as well.
*Tool (If applicable)	If your study uses any other tools to collect the data, please include introductions of the tools and/or where the interested researchers can access the tools.

If you have any questions, please reach out to Jingle Xu (jxu2@unc.edu) or Becky Lee (beckylee@unc.edu) at the CPCRN Coordinating Center.