2025

National Center for Health Statistics Research Data Center

Policy and Procedures Manual for Researchers

Contents

Introduction

The National Center for Health Statistics (NCHS) Research Data Center (RDC) provides approved researchers with access to restricted-use respondent-level data. The NCHS RDC policies and procedures are designed to maintain respondent confidentiality and prevent disclosure of confidential information. Researchers have important responsibilities in preventing disclosure by abiding by RDC policies and procedures. These policies and procedures are based on the laws that govern the collected data and apply to your role as a researcher in the RDC. Researchers at the RDC are legally obligated to maintain the confidentiality of respondent identities.

Your proposed research and the output you generate will be subject to an in-depth review. You will conduct your work in a secure computer environment with unique rules and procedures, which are outlined in this document.

Failure to comply with RDC policies and procedures will result in cancellation of your research project and potential disbarment from future access to restricted-use NCHS data and any other research activities related to the NCHS RDC.

RDC Procedures to Prevent Disclosure	Researcher Responsibilities
1. **Limited access**: Restricted-use data cannot be downloaded from the IT systems of the NCHS RDC or the Federal Statistical RDCs (FSRDC). Data must not be downloaded from the Virtual Data Enclave (VDE). a. RDC Analysts will email approved output or programs to you after a disclosure review.	• Do NOT attempt to remove data sets from RDC workstations or from the VDE. • Do NOT attempt to save output, files, or programs to transportable electronic media (for example, USB drives).
2. **Required research application**: The Review Committee carefully examines the variables requested, the analysis plan, and the desired output as specified in your application.	• Conduct ONLY the analyses that were approved in your application. • Any changes from your application must be approved by the review committee. Follow the instructions for amending your application.
3. **Required confidentiality training and completion of confidentiality paperwork**: Researchers must sign the "Data Access Form" and the "Designated Agent Form" prior to data access. These legal documents confer Designated Agent status to researchers and binds them to the laws that govern the restricted-use data. These forms are specific to the project and must be completed each time a researcher has a new approved application.	• Comply with all requirements, rules, and instructions outlined in the confidentiality training and those described in the legal documents you signed. • If there is an inadvertent identification of an individual or establishment, immediately bring it to the attention of RDC staff. Hold in strictest confidence the identity of any individual or establishment in any documents, data, discussion, or analysis.
4. **Data access policies, procedures, and rules:** The RDC must follow rules set by the Department of Health and Human Services (HHS) NCHS, and the data owner for data management and output release policies to prevent disclosure. The most common rules are highlighted in this document. However, special rules may be put in place upon approval of your application or as required by the data owner.	• Read data documentation for all years and cycles of data before you visit the RDC. • Follow all RDC data access rules, procedures, and policies. Violation of these rules can result in suspension of your research project.
5. **Analytic data sets for researchers:** RDC Analysts will merge restricted-use, public-use, and non-NCHS external data for the researcher. Whenever possible, restricted-use merge variables will be removed or may be randomized or coarsened versions of restricted-use variables.	• Follow the guidelines for providing the public-use and non-NCHS external data to your RDC Analyst. • Do NOT use ANY technique to learn the identity of any person, establishment, or sampling unit in any restricted-use file.
6. **Perform disclosure review**: A RDC Analyst must review all output before it can be released to the researcher.	• Follow the Output Review Policies (Part 4) and the Publishing Guidelines (Part 5) in this manual. • Only request an output disclosure review for the output listed in your approved application. • Do NOT remove any output from the RDC or VDE.

Part 1: Confidentiality and the Research Data Center

Confidentiality and the Law

There are two laws that govern data collected by NCHS and hosted in the RDC:

- Section 308(d) of the Public Health Service Act (42 U.S.C. 242m(d)) states:

  ‘‘No information, if an establishment or person supplying the information or described in it is identifiable, obtained in the course of activities undertaken or supported under section 242b, 242k, or 242l of this title may be used for any purpose other than the purpose for which it was supplied unless such establishment or person has consented (as determined under regulations of the Secretary) to its use for such other purpose; and in the case of information obtained in the course of health statistical or epidemiological activities under section 242b or 242k of this title, such information may not be published or released in other form if the particular establishment or person supplying the information or described in it is identifiable unless such establishment or person has consented (as determined under regulations of the Secretary) to its publication or release in other form.’’ See: www.cdc.gov/nchs/data/misc/staffmanual2004.pdf (page 1).

- Confidential Information Protection and Statistical Efficiency Act (CIPSEA) (42 U.S.C. 3561-3583). CIPSEA states that the only people who can access confidential NCHS data are NCHS staff and its Designated Agents. Therefore, researchers wishing to access confidential data must become Designated Agents. CIPSEA stipulates that the penalty for willfully violating confidentiality is a class E felony with up to 5 years in prison or a $250,000 fine or both. The Freedom of Information Act does not apply to data collected under CIPSEA. Part of the CIPSEA statute states:

  ‘‘Whoever, being an officer, employee, or agent of an agency acquiring information for exclusively statistical purposes, . . . comes into possession of such information by reason of his or her being an officer, employee, or agent and, knowing that the disclosure of the specific information is prohibited under the provisions of this title, willfully discloses the information in any manner to a person or agency not entitled to receive it, shall be guilty of a class E felony and imprisoned for not more than 5 years, or fined not more than $250,000, or both.’’ See: www.cdc.gov/nchs/data/misc/staffmanual2004.pdf (page 2).

Sanctions for Violating Rules

Researchers who knowingly and willfully disclose confidential data will face the legal ramifications specified above within CIPSEA. Researchers who violate RDC policies and procedures will lose access to the RDC and their project data. The researcher’s employer and any sponsors will be notified of the violation.

Researchers who are suspected of violating RDC policies and procedures may be prevented from continuing their research. If this happens, access to the data is denied for all members of the research team. The NCHS RDC will conduct an investigation to determine if the violation applies to the whole research team or one or more team members. Research team members interested in continuing the project, and not involved in the violation, must submit a new application. If approved, a new data file is created, and the project is charged RDC fees.

Part 2: Approved Applications: Next Steps

You will have access to your analytic data file after the following requirements are met:

- Application is approved
- The RDC receives the confidentiality paperwork and training certificate(s) for each research team member
- The RDC receives your public-use or non-NCHS external data as approved in your application and your assigned RDC Analyst (RA) has completed the build of your analytic data file
- The RDC receives management and access fees, if applicable
- For FSRDC users – Special Sworn Status is approved for team member(s)
- For VDE users – the RDC receives your VDE Data Use Agreement and the RDC completes your secure room inspection

Your assigned RA will confirm receipt of delivered items. However, routine or regular project updates are not to be expected. Please review all correspondence as RAs will often communicate timelines for delivery via email.

Approved Applications

Approval of an application does not mean that NCHS endorses the merit of your research or its substantive, methodological, theoretical, or policy relevance.[1] NCHS approval reflects the judgment that the research, as described in the application, can proceed to data analysis and does not pose a significant potential disclosure risk. Approval of an application does not explicitly or implicitly guarantee that all output generated by the researcher will be released.

Confidentiality Training

All research team members must take the confidentiality training and complete the confidentiality forms. It should take about 20 minutes to complete the training. At the end of the training, you will take a quiz. You must score 100% on this quiz. Confidentiality Training is available at: www.train.org/cdctrain/course/1088489/.

[1] See: www.cdc.gov/rdc/publishing-guidelines/index.html#cdc_generic_section_6-required-disclaimer:~:text=Research%20Data%20Center.-,Prohibited%20statements

Confidentiality Paperwork

Confidentiality forms are specific to the application, so they must be completed every time a new RDC application is approved.

Confidentiality Forms:

- Designated Agent Agreement Form (must be notarized)
- Data Use Agreement Form (must be notarized)

Provide the completed forms and your confidentiality training certificate to your RDC Analyst as soon as possible. Send forms via email (preferred method) or mail. Make sure the notary seal is visible. You may combine the forms for the same researcher, but do not combine all forms into a single document. You cannot begin your research until the RDC receives your completed forms. [Note: if federal employees use their government-issued PIV or ID card to sign these two forms, then they do not need to get the forms notarized.]

Creating Your Analytic Data File

The RDC tries to minimize access to raw restricted-use data. Therefore, RDC staff will merge the public, restricted-use, and any non-NCHS external data together to create your analytic data file.

Unless otherwise specified, provide data files to the RDC in SAS format. Similarly, your final analytic data file will be provided in SAS format unless other arrangements are made with your RDC Analyst.

Public-use and Non-NCHS Data

Researchers are responsible for providing the NCHS public-use data and any non-NCHS external data to their assigned RDC Analyst. Make sure these data are cleaned and complete before sending, for this will expedite the creation of your analytic data file. Follow these steps when providing your RDC Analyst with your public-use data set and any non-NCHS external data for merging:

1. Create a public-use data set that includes only the variables specified in your approved application. If you want to include additional variables not listed in your approved application, then you must amend your application. Discuss application amendments with your RDC Analyst. Requesting additional variables will result in your application going through the NCHS Review Committee process again.
2. If you would like to rename variables, include the original variable name in the variable description.
3. If you choose to create derived variables before working with the data, clearly define these variables in the application. The variable description should include the original variable name and an explanation of any arithmetic manipulation.

Save the code you used to create these variables, as your RDC Analyst may request it.

4. The non-NCHS external data file you send your RDC Analyst should only include the variables listed in your approved application.
5. Email the data files and a list of the variables to your RDC Analyst. If your data files are too large to email, ask your RDC Analyst for an alternative delivery option. RDC Analysts cannot use external, commercial data file share programs.

Important Notes About Submitting Public Data

- Any attempt to include variables that may lead to re-identification of study participants or establishments is a disclosure violation and will result in the cancellation of your project and possible legal actions.

- If you are requesting access to the restricted-use mortality files, you cannot include any public-use mortality variables, or any variables derived from the public-use mortality data.

Merging Data

After all the requirements have been completed as specified in section "Part 2: Approved Applications: Next Steps," the RDC Analyst will merge all the files based on the specifications you provided in the Data Linkages document or Merge Variables section of your approved application.[2] Discuss the merge with your RDC Analyst throughout the process to ensure that your analytic data file is created to your specifications.

After the merge is complete, and if requested in your approved application, the RDC Analyst will email the PROC CONTENTS of your analytic file for review. Check the number of observations, variables, and the list of variables to make sure everything has been included. Otherwise, you may verify the merge during your onsite visit. A RDC Analyst may not provide corrected files during your RDC visit onsite.

RDC Analysts will follow data owner policies to protect geographical, temporal, and perturbed or masked data. In the next section, we have summarized each policy for your reference.

Policies to Protect Geographical Information

Remove Geography: Geographical variables are removed when the variable is only needed to merge NCHS data to another data file.

[2] In rare cases, you may receive approval to merge NCHS public-use data and non-NCHS data (for example, Census data) to the restricted-use data. A justification for researcher-approved merges would have been included in your approved application.

Recode or Mask Specific Geography:

- Randomize Geography: A random version of the geographical variable will be provided for analysis.
- Coarsen Geography: When lower levels of geography can be grouped into larger areas, the RDC Analyst will create the coarsened variable and will not provide access to the underlying lower level of geography. RDC Analysts may request that the researcher write the code for how they want the variables created; researchers are always welcome to provide their own code and to review the code for created variables.
- Provide Geographical Identifiers: True geography will remain in the data file when used to make estimates at a lower level of geography if this has been approved in your application. However, most NCHS data are not representative at lower geographical levels (for example, state or county); therefore, this is the least common use of geography for researchers.

Policies to Protect Temporal Information

Coarsen Dates: The RDC Analyst will create the coarsened variable (for example, year, month, quarter) and will not provide access to the underlying exact date unless there is strong justification, and it was approved in your application. RDC Analysts may ask the researcher to write the code for how they want the variables created; researchers are always welcome to provide their own code and to review the code for created variables.

Creating Variables of Time from Dates: If exact dates are used to calculate time (for example, exact length of life calculated based on date of birth (DOB) and date of death (DOD)), the exact dates should only be used for variable creation and the resulting created variables (for example, length of time) should be used in the analysis. The exact dates would be deleted from the final analytic data file.

Policies to Protect Perturbed and Masked Information

Perturbation and masking are common disclosure limitation methods and are used on potentially sensitive information. These methods slightly alter the data so that they are allowed to be made public. If the methods were to be revealed, the public-use files could be compromised. The following are examples of variables that **cannot** be included together in the researcher's analytic data file:

- Public-use and restricted-use mortality variables
- Pseudo, masked, and true primary sampling units (PSUs) and Strata variables for most NCHS surveys (for example, the National Health and Nutrition Examination Survey)

Amending Your Approved Application

Research evolves and your analytic plan may change after you have submitted your application. Changes to your analytic plan may affect the disclosure risk of your project. Inform your RDC Analyst of any project changes that occur throughout the research

process. Update all relevant sections of your application to reflect these changes. Highlight all changes in the application, include a revision date, and submit the revised application to your RDC Analyst.

The following are examples of common situations which require an amended application and the process for each. Contact your RDC Analyst if you have specific questions about a possible application amendment.

New Variables: If you want to add new variables to your project, these new variables must be related to the original research question and will be reviewed by the RDC and NCHS. Add the variables to the data dictionary in your application and explain how they will be used.

- Adding public variables requires the RDC Analyst's approval and may require the entire NCHS Review Committee's approval.
- Adding restricted-use variables or non-NCHS variables or both always requires the RDC review and the NCHS Review Committee's approval. This process typically takes 4 weeks.
- Adding variables unrelated to the original research will most likely require a new application and will start a new 8 to 12-week review.

New Methods: If your analysis starts to take a different direction, inform your RDC Analyst, as this may significantly change the disclosure risk. Discuss these changes with your RDC Analyst. The NCHS Review Committee may need to review these changes. Highlight these analytic changes in your amended application.

New Types of Output: If you need new or different output compared to what was approved in your application, inform your RDC Analyst, as this may significantly change the disclosure risk of your project. Discuss output changes with your RDC Analyst. Output changes will involve review by the NCHS Review Committee. Highlight these changes in your amended application.

New Researchers: The Primary Investigator (PI) must inform their RDC Analyst about their departure or the departure of a team member. If the PI or team member who is using the data leaves the project and a new team member is needed, the new PI or team member must:

- Complete the Confidentiality Training and related confidentiality paperwork and send it to the RDC Analyst.
- The PI needs to update the application to reflect the change and include the new PI or team member's curriculum vitae or resume.

Change to the Location of Data Access: If the location of data access needs to change, discuss this with your RDC Analyst. This may incur additional costs. Follow these steps to update your location of access:

- Update the application to reflect the location change.
- Discuss the logistics with your RDC Analyst and any fees.

Part 3: Working with Restricted-use Data

Getting Ready for Your NCHS RDC Visit

1. NCHS RDC visits are **BY APPOINTMENT ONLY**. Regular RDC hours are 9 a.m. – 5 p.m. Please discuss alternative schedules with your RDC analyst.
2. Your RDC analyst will schedule and supervise all visits. However, when necessary, your RDC analyst may plan for another RDC analyst to supervise your visit.
3. If you miss an appointment, are late, or leave your RDC workstation, the RDC analyst may cancel your visit after two hours.
4. If you are not scheduled to visit the RDC, RDC staff will deny your entry into the RDC. Be prepared to show proof of scheduled visit (for example, an email from your RDC analyst).
5. Only authorized researchers are allowed into the RDC. Researchers not approved to visit the RDC will be denied entry. Authorized researchers are listed in the approved application and have completed required confidentiality training and forms.
6. At least 5 business days before your visit, email your statistical programming code and reference documentation (for example, codebooks, statistical reference materials) to your RDC Analyst. The RDC Analyst will review the documents and upload them to your account.
7. RDC computers do not have access to the Internet, so you must provide all your resources in advance of your visit. Electronic devices (for example USB drives) are not allowed in the RDC.

NCHS RDC Rules

When you come onsite to a NCHS RDC, there are restrictions that the NCHS RDC requires to prevent the disclosure of confidential information.

1. An RDC Analyst will provide your analytic data file in your account that you will access on a RDC workstation.
2. If you have approval for multiple research projects, you are only allowed to work on one project, on one computer at a time.
3. Absolutely NO individual respondent-level data are allowed to leave the RDC facilities.
4. The RDC reserves the right to search all materials brought into the RDC facility and all materials that will leave the RDC facility.
5. Cell phones, laptops, and other electronic devices are not permitted in the RDC.
6. RDC workstations do not provide access to the local network or the Internet.
7. Researchers may derive variables from the analytic data set to create new variables, but they cannot introduce new data using their statistical code. Do NOT attempt to put any content in your code that would facilitate re-identification of a study participant. This is a disclosure violation and will result in the cancellation of your project and possible legal actions.

8. You are **not** allowed to bring items into the RDC that may enable you to identify study participants (for example, a national hospital directory, information from voter registration).
9. All notes generated in the RDC are subject to disclosure review by RDC Analysts before removal from the RDC.
10. Before you submit your request for output review, review it for potential disclosure concerns and unnecessary information. See "Part 4. Disclosure Review Policies and Procedures" below and review additional requirements for output review.
11. Allow for up to 3 weeks for your output to be reviewed by your RDC Analyst. If your output is approved, your RDC Analyst will email it to you.
12. If you would like a copy of your updated programming code after your visit, ask your RDC Analyst to review the code. If approved, the RDC Analyst will email the program code to you.
13. Create only data tables that appear in your approved application. Output that does not appear in your approved application will not be reviewed or released.
14. Submit your output in a plain text file (for example, .lst files from SAS) tab delimited text [.txt] files or comma-separated [.csv] files so files can be opened and read in Windows Notepad.

Working Onsite at a NCHS RDC

1. Your RDC analyst will escort you into the RDC office area. Do not enter the RDC facility unaccompanied.
2. There can be no more than two team members in the RDC at one time.
3. Always notify your RDC Analyst before you leave the RDC and when you expect to return.
4. Be respectful of other researchers and RDC staff while working in the RDC.
5. Store all personal belongings in RDC-provided lockers.
6. Do not bring any kind of electronic device (for example, cell phone, tablets, computer storage media, laptops) into the RDC.
7. Do not bring any documentation into the RDC unless it has been reviewed and approved by your RDC analyst.
8. The RDC provides paper and pens, by request only, to take notes during data analysis. All notes are subject to disclosure review before removal from the RDC. Your RDC analyst will conduct all disclosure reviews.
9. You must use all personal devices (for example, cell phone, tablets, laptops) outside of the RDC office area.
10. Discuss food and drink options with your RDC Analyst before taking these items into the RDC.
11. You can only work on one project at a time. You may not have different projects open simultaneously.
12. Save output files to your project folder. Do not save files to any other location on the computer (for example, the desktop or "My Documents"). Files saved to other locations will be automatically deleted.

13. Always follow RDC staff instructions when asked to leave the facility, change computers, or reschedule a visit.

Getting Ready for Your FSRDC Visit

NCHS RDC Analysts and FSRDC Administrators have unique roles and responsibilities. Contact your RDC Analyst and FSRDC Administrator before your first visit to ensure that everything will be ready before you go to your assigned FSRDC.

NCHS RDC Analyst:

- Facilitates review of your application
- Creates your analytic data file
- Issues your invoice for fee payment
- Accepts and processes required confidentiality forms
- Transfers your analytic data file to Census
- Reviews your output for disclosure risk

FSRDC Administrator or Designee:

- Answers questions about the FSRDC RDC
- Ensures you have completed all Census Bureau security requirements for Special Sworn Status (SSS)
- Transfers your output to your assigned RDC Analyst

To work in an FSRDC you must become an NCHS Designated Agent and have Special Sworn Status (SSS) from Census. Contact your FSRDC Administrator for information about obtaining SSS. Keep your RDC Analyst up to date on where you are in the SSS approval process. After your SSS is approved and all requirements listed in "Part 2: Approved Applications: Next Steps" are completed, your Census account is created. At this point, your RDC Analyst will transfer your analytic data file to Census.

Working Onsite at an FSRDC

While working at a FSRDC, you are subject to **all** the same rules and restrictions as if you were working onsite at a NCHS RDC. An additional FSRDC rule is:

1. You must follow FSRDC procedures to transfer output files to your RDC Analyst for review. If you need help, ask your FSRDC Administrator.

Email all documents and reference materials to your RDC Analyst at least five business days before your visit to an FSRDC. No notes, papers, or hard copies of any kind are permitted inside a FSRDC.

Virtual Data Enclave (VDE) Procedures

While accessing data in the VDE, you are subject to all the same rules and restrictions as if you were working onsite at a NCHS RDCs.

Read and be familiar with the requirements and limitations imposed by the VDE Data Use Agreement (DUA) signed by your employer. The requirements and limitations in the VDE DUA must be followed. Failure to abide by the VDE DUA requirements will result in the cancellation of your project and possible legal actions.

VDE Procedures:

1. At least 5 business days before using the VDE, email your statistical programming code and reference documentation (for example, codebooks, statistical reference materials) to your RDC Analyst. The RDC Analyst will review the documents and upload them to your account. The VDE does not have access to the Internet.
2. A RDC Analyst will provide your analytic data file in your account that you will access via the VDE.
3. If you have approval for multiple research projects, you are only allowed to work on one project at a time.
4. Absolutely NO individual respondent-level data are allowed to leave the VDE.
5. Cell phones and other electronic devices are not permitted to be used when logged into an active VDE session.
6. Researchers may derive variables from the analytic data file to create new variables, but they cannot introduce new data using their statistical code. Do NOT attempt to input data into the VDE. This is a disclosure violation and will result in the cancellation of your project and possible legal actions.
7. Any notes you generated must not contain data or output that was displayed in the VDE. If you do this, it is a disclosure violation and will result in the cancellation of your project and possible legal actions.
8. Before you submit your request for output review, review it for potential disclosure concerns and unnecessary information.
9. Allow for up to 3 weeks for your output to be reviewed by your RDC Analyst. If your output is approved, your RDC Analyst will email it to you.
10. If you would like a copy of your updated programming code, ask your RDC Analyst to review the code. If approved, the RDC Analyst will email the program code to you.
11. Only create data tables that appear in your approved application. Output that does not appear in your approved application will not be reviewed or released.
12. Submit your output in a plain text file (for example, .lst files from SAS), tab-delimited text [.txt] files, or comma-separated [.csv] files so files can be opened and read in Windows Notepad.

Part 4: Disclosure Review Policies and Procedures

RDC disclosure review policies and procedures exist to protect the identity of study participants.

General Output Policies[3]

1. The creation of a data set is not appropriate output and will not be released.
2. Absolutely **no** output will leave the RDC without first being reviewed by a RDC Analyst for possible disclosure of confidential information.
3. Output for review **must** match the research questions and the output described in your approved application.
4. Review and release of intermediate output is **not** allowed. If the output described in your approved application is later determined to be intermediate output, the RDC will not release it. You **must** constrain your output to what you need for a final research paper or journal article.
5. RDC Analysts may apply cell suppression criteria: Guidelines may differ by data system and possibly by survey year because of sample size, sample design, and content. Sometimes specific projects pose additional disclosure concerns, and additional or more stringent cell suppression may be applied.
6. Researchers must plan ahead to allow the amount of time needed by the RDC to review and approve output. The RDC will usually return your approved output via email within 3 weeks of your request. Lengthy output not intended for a standard journal article or presentation may take more than 3 weeks to review.
7. Researchers may only use their approved output and statistics in a way that does not pose any additional disclosure risks to study participants.
8. Submit your output in a plain text file (for example, .lst files from SAS), tab-delimited text [.txt] files, or comma-separated [.csv] files so files can be opened and read in Windows Notepad.
9. You **must** constrain your output to what you need for a final research paper or journal article.

Output Policies Specific to the NCHS RDC, FSRDCs and VDE

All output must match what was approved in your application. Follow the steps outlined in this manual for creating and submitting output for review. Otherwise, you may be asked to return to the RDC to redo your output so that it conforms to RDC policy and your approved application.

Before submitting output to your RDC Analyst for review, you must:

- Make sure it is in a form that can be released by the RDC. Output will be released in a plain text file (for example, files that can be opened and are readable in Windows Notepad, such as tab-delimited text [.txt] files or comma-separated values [.csv]).
- Create data tables that will appear in your paper or journal article. Data tables must match what is in your approved application.

[3] Also see RDC output policies at: www.cdc.gov/rdc/output/index.html

- Remove any output that you feel could lead to the identification of an individual or institution. If you have questions, ask your RDC Analyst.
- Remove any individual-level data from your output. Output that has individual record-level information is not permitted.
- Extreme values or values representing an individual must be removed. Examples include minima, maxima, medians, and modes. If a procedure such as Proc Univariate creates extreme observations, (for example, 0, 1, 99, and 100 percentiles) those extreme values must also be removed.
- Collapse variables categories if any category has a frequency less than five. If you are unable to collapse, then all cells with a frequency less than five should be asterisked. Complete this action before you submit your output to your RDC Analyst.

Make sure output submitted for review includes:

- A description of the output (for example, A regression of…).
- Descriptions of (sub)samples in the analysis and output (for example, Black males ages 20–29).

Intermediate output[4] (output that will not appear in publication) can be created and used onsite at the RDC, but the RDC will not review or release it.

- Examples of intermediate output include: unweighted n's, tables of preliminary descriptive statistics, large volumes of numbers, and preliminary regression models.
- Intermediate output increases disclosure risk: Similar tables based on different subsamples may cause complementary disclosure problems because comparison of tables could reveal information about the sample and individual characteristics.
- You **must** constrain your output to what you need for a final research paper or journal article.

Stage Output for Review

- Do NOT include any hard-coding of data in your code intended for release. Any comments or notes should not facilitate re-identification of a subject/establishment/geographic unit or disclose any information you learned while in the RDC.

- Put the output file(s) that you want reviewed in a separate "Review" folder and add the date to the folder name (for example, Z:\Review_MMDDYYYY).

[4] See: www.cdc.gov/rdc/output/index.html#cdc_generic_section_3-intermediate-output

- Before you leave the RDC, let your RDC Analyst know if you have output for review. Via email, provide the complete path to the output (for example, Z:\Review_11302019) for your RDC Analyst.

- FSRDC users must follow output staging procedures provided by Census.

Output Review

- Within 3 weeks of your request, your output will be reviewed and returned to you or your RDC analyst may have questions for you about the output.

- Output review will exceed 3 weeks if the amount output is large.

Part 5: Publishing Your Research

Your Disclosure Responsibilities Prior to Publication

RDC researchers must send their near final draft information products (for example, journal article, manuscript, white paper, slides, conference presentations, posters, and other products) to the RDC for review before submitting the information product for publication. An additional review may be warranted if substantive changes are made to how the output are reported in the tables or text if subsequent versions of your information product are made (for example, changes made based on a revise and resubmit for a journal article).

Email your information product to your RDC Analyst to start this process. Place "Product Review" in the subject line of your email.

When conducting these reviews, the RDC does not comment on scientific merit or impose any merit-based publishing guidelines.

When writing your information product (for example, journal article, manuscript, white paper, slides, conference presentations, posters, and other products), follow these requirements:

1. Adhere to any specific requirements in your approved application and any requirements in emails received from your RDC Analyst regarding your output.
2. Journals can ask that the data be provided that is referenced in your article. This is not allowed. We recommend that you share the data dictionary and merge procedures that were included in your approved application.
3. Do not place any information in your information product:
    a. That could identify an individual or establishment.
    b. That could be used to identify geographical areas where respondents live or were sampled unless you specifically have permission to make estimates for those areas in your approved application.
    c. That provides exact dates.
    d. That could unintentionally reveal information about any of the above items. For example, publishing information on external sources of data after they have been merged with NCHS data could facilitate the identification of areas, dates, or individuals.
4. Follow NCHS citation requirements for data.
5. Include the disclaimer (shown below) in your information product.
6. Ask the RDC Analyst for help or clarification any time you have a disclosure concern or question.

Disclosure Concerns to Consider

Ask yourself the following questions as you write your publication based on restricted-use data:

Question	Action
1. Are you mentioning anything in the publication that you learned while looking at the restricted-use data in the RDC?	• If you intend to discuss something that you observed while in the RDC but this information did not undergo disclosure review, share this information with your RDC Analyst to assess if it poses a disclosure risk.
2. Were restricted-use geographical variables used in your project?	• Do not mention geographical information not available in the public files (for example, specific states, counties), unless it was part of your approved application. • Do not discuss the number of geographical units represented in your output without written permission. • Do not present unweighted sample sizes related to lower levels of geography.
3. Did your project use any restricted-use temporal information/variables?	• Do not mention exact dates. • Do not mention coarsened dates (for example, season, year) without written permission. • Do not present unweighted sample sizes related to temporal components (for example, dates).
4. Did your final analytic data set include data from another data source that was merged on geographical or temporal components?	• Do not reveal any information that could facilitate the identification of areas, dates, or individuals who participated in NCHS or non-NCHS surveys.

5. Could an individual or establishment possibly be identified by any information in your publication?	• Discuss inclusion and exclusion criteria for subpopulation analysis in a way that does not inadvertently identify small cells or extreme cases. This includes information on individual(s) or establishment(s). Review for:
	  ○ Fewer cases than those specified in the required cell suppression criteria.
	  ○ Extreme cases or outliers that were identified during analysis.
	  ○ Unweighted counts that could be subtracted from totals to reveal a sample size less than that specified in the required cell suppression criteria.

How to Cite the RDC in Your Publication

When describing where the restricted-use data for your study came from, specify that restricted-use data were accessed through a Census or NCHS Research Data Center. Also, indicate why these data were essential to your research study. Here are two examples:

- Geographical variables including state, county, and tract were used to merge Census variables that provided neighborhood contextual information. State, county, and tract are restricted-use variables, and these data were accessed through a (Census or NCHS) Research Data Center.
- To merge the patient and agency files, we needed the restricted facility identification variable. Because the analyses required restricted-use data, the data were accessed through a (Census or NCHS) Research Data Center.

Do not place these kinds of statements in your products:

- Analyses of these data were approved by the NCHS Research Data Center.
- The NCHS Research Data Center reviewed and approved our results.

When displaying output in your information product, provide the Data Source. The content of a Data Source citation is usually provided by the data owner in their data documentation.

Include the following disclaimer in your product:

- The findings and conclusions in this paper are those of the author(s) and do not necessarily represent the views of the Research Data Center, the National Center for Health Statistics, or the Centers for Disease Control and Prevention.