

## RDC Research Proposal

<b>General Information</b>	
<b>Date:</b>	
<b>Title of Project:</b>	
<b>Name of Data System and Years:</b>	
<b>Other External Data Files:</b>	
<b>Mode of Access:</b>	<input type="checkbox"/> NCHS RDC, Hyattsville, MD <input type="checkbox"/> NCHS RDC, Rockville, MD <input type="checkbox"/> NCHS RDC, Washington, DC (Government Only) <input type="checkbox"/> NCHS RDC, Atlanta, GA <input type="checkbox"/> Federal Statistical RDC, specify: _____
<b>Statistical Software:</b> (Check all that apply)	<input type="checkbox"/> SAS/SUDAAN <input type="checkbox"/> Stata <input type="checkbox"/> Other, specify: _____
<b>Proposed Start Date:</b>	
<b>Funding Source:</b>	
<b>Billing Address:</b> (include contact person)	

List the name, institution, contact information, and role for anyone who will contribute to publications resulting from this project. Everyone listed must submit a C.V or resume. Add sections as needed.

<b>Research Team</b>		
<b>Role:</b>	<b>Principal Investigator</b>	<b>Co-Investigator</b>
<b>Name</b>		
<b>Email</b>		
<b>Phone</b>		
<b>Institution</b>		
<b>Mailing Address</b>		
<b>SSS Status? Y or N</b>		
<b>US Citizen? Y or N</b>		
	<b>Programmer</b>	<b>Programmer</b>
<b>Name</b>		
<b>Email</b>		
<b>Phone</b>		
<b>Institution</b>		
<b>Mailing Address</b>		
<b>SSS Status? Y or N</b>		
<b>US Citizen? Y or N</b>		
	<b>Advisor (For Students and Post-Docs)</b> [ <input type="checkbox"/> ] <a href="#">RDC-Student-Advisor Form</a>	<b>Other, specify:</b>
<b>Name</b>		
<b>Email</b>		
<b>Phone</b>		
<b>Institution</b>		
<b>Mailing Address</b>		
<b>SSS Status? Y or N</b>		
<b>US Citizen? Y or N</b>		

Complete as applicable for your project. Address any "Yes" responses in the body of the proposal.

<b>RDC Proposal Summary Information</b>		<b>YES</b>	<b>NO</b>
<b>Geographic variables</b>			
Level of geography to be shown in <b>output</b> (check all that apply)			
National			
Regional			
State			
MSA			
County			
Urban/rural classification			
Census tract			
Latitude/Longitude			
Other			
Will geographic identifier(s) be removed after merge			
If yes, can true geographic identifiers be replaced with masked versions of these variables			
Is GIS or mapping proposed			
<b>Dates and Temporal information</b>			
Are exact dates requested other than to calculate time of follow-up			
If more than 1 year/cycle, will years/cycles be presented separately			
<b>Merging of data with restricted data</b>			
Are external data being merged with EPA data			
<b>Linked Data Products</b>			
Are you using public-use NCHS mortality data			

**A. Abstract:** Please limit the project description abstract to 300 words.

**B. Research Question:** Describe study purpose, hypotheses, goals, or research questions.

**C. Background:** Include a short literature review, no more than 2 pages, focusing on papers that discuss your topic and address the methodology that you plan to use. Please limit your reference list to 10 items or less.

**D. Public Health Benefit:** In one paragraph, describe how does your research is beneficial and applicable to EPA.

**E. Data Requirements:**

Please address the four items below and provide an explanation for “yes/checked” responses from the RDC Proposal Summary Information table (above).

**1. Survey, Years, Files:**

List survey name(s) and survey years you would like to access. For examples, NHIS 2005-2007 Household, Person and Sample Adult Files, NAMCS 2005-2006 Provider and Patient Visit Files, or NHANES 2005-2006 Examination and Demographic Files.

**2. Restricted-use Data:**

List and describe the restricted-use variables that you will need. These variables must be listed in the Data Dictionary section of this proposal. Explain why each variable is needed and how you will include them in your analysis. Specify how geographic variables, if applicable, will be used to merge files, analyze the data and/or be presented in output.

**3. External Data:**

If you plan to provide data from another source (such as Census or EPA) to merge to the restricted-use data, please describe the source, list the filename(s), and provide a general description of the non-NCHS data. The variables from the non-NCHS data must be listed in the Data Dictionary section.

**4. Merge Variables:**

In detail, please describe the merge procedures used to produce your analytic dataset(s). Highlight the variables used in the merge routine in the Data Dictionary. Leave blank if not applicable.

**F. Methodology:**

We highly recommend you familiarize yourself with the analytic guidelines of the data you intend to use. Any deviations from the methodology suggested in the guidelines will require explanation as it may pose a disclosure risk. Please address these three items in your proposal:

**1. Unit or Level of Analysis and Subpopulation(s):**

There can be many levels of analysis: be as detailed as possible in your description. A common example for an analysis using NHANES data is where the unit of analysis is the person while the subpopulation is adults ages 18-64. A common example involving geography is when the researcher aggregates persons to the state level to make comparisons between state policies.

**2. Analysis Plan:** Provide an overall analysis plan that specifies what analytic procedures or models you will use, such as prevalence estimates, logistic regression, or log-linear modeling. Also, list the specific statistical package procedures you will use.

**3. Complex Survey Design:** Indicate how you will address sample weights, design variables, and other adjustments for the use of complex survey data, if applicable, using the statistical software listed in the General Information area. A detailed description per weight, design variables, and other adjustments are required and central to understanding the limitation of the data. This is a critical element to describe and is central to the proposal review process.

#### **G. Output:**

Please describe in full detail all of the output you would like to take out of the RDC that is needed for the final product you wish to create. Please be very specific in describing your output needs (i.e., state groupings, levels of output and how restricted-use data will be displayed) as this section is necessary for the Review Committee to assess disclosure risk of your output. Your examples should reflect the geographic variables you will use in your output. **Note: Any output produced in the RDC that is not listed in your approved proposal will not be allowed out of the RDC.**

- 1. Table Shells:** Include detailed examples of all table shells, and/or models and all statistical estimates/parameters needed with titles/labels. Indicate the subsample and unit of analysis used in each type of table and/or model. Your proposal will not be approved without this information.
- 2. Presentation of Results:** Describe how you will present the results (in a report, publication in a peer-reviewed journal, presentation at a scientific meeting, or used for internal policy analysis).

#### **H. Data Dictionary:**

Include a data dictionary for each data source. Provide a public-use, restricted-use data, and/or external data dictionaries of the data you need. This should be a listing of variables you would like in your dataset. See instructions and examples for [creating the data dictionary](#). When asking for multiple years of data, make sure to reflect the public-use file layout for each year as variable names can change over years. Include all explanations in Section E. Data Requirements.

**Note:** Provide a single data dictionary that includes all the variables you need in your analytic data set.

I. Literature References: Please provide no more than 10 references that are relevant to your project.

J. Resumes/C.V.: Please include a 2-page C.V. for each member of the research team (not as attachments).