

Tips on using NSFG data in the NCHS Research Data Center

The NCHS Research Data Center (RDC) is a physical space located at NCHS in Hyattsville, Maryland where researchers are allowed access to NCHS data files that cannot be released to the public. Applicants must submit a written proposal to use the RDC. The proposal does not have to be long. In addition to the cover letter, title, abstract, and Vita, 5-8 single-spaced pages is often sufficient. A proposal to use the RDC differs from a grant proposal. Less emphasis is placed on the literature review, and more emphasis is placed on the specific variables the researcher proposes to use. NCHS will review the proposal. Please allow a month for review.

The Research Proposal should contain the following:

- A. Cover letter, Project Title, and Abstract.
- B. Full personal identification and institutional affiliation. If a student, a letter from the department chair or advisor stating that the applicant is a student working under the direction of the department.
- C. Current resume or Curriculum Vita.
- D. Dates of proposed use of RDC or the remote access system.
- E. Source of funding.
- F. Summary of the research, including why publicly available data are insufficient.
- G. A complete list of the NCHS data requested - source, files, years, variables.
- H. Details concerning the user supplied data, if any, to be merged with the NCHS data. This includes documentation, file layout, number of records, and whether other researchers can use the user-supplied data.
- I. Software requirements. SAS, STATA, SPSS, SUDAAN, LIMDEP, HLM, "R", and Watcom Fortran are available in the NCHS RDC. State the specific version of the software you need (e.g. SAS Version 9.1 for Windows). More recent versions of software packages, and other software can be made available with sufficient lead time.

User Supplied Data

The NCHS RDC allows users to supply their own data to be merged with NCHS data sets. The user-supplied data may consist of proprietary data collected and 'owned' by the user or other publicly available data obtained by the user such as census data. Users *must* provide NCHS-RDC staff with complete documentation.

Cost of using the RDC

Researchers using the NCHS RDC are charged to defray the costs of space, equipment, staff time, etc. The cost per project is:

- \$200 per day for ONSITE use (in the RDC at NCHS);
- \$500 per year for REMOTE ACCESS to NSFG-Cycle 6 Contextual Data File.
- There are setup charges (usually \$500) for files requiring merging of additional data or creating custom file formats.

All applicants are required to sign the Agreement Regarding Conditions of Access to Confidential Data & the Researcher Affidavit of Confidentiality.

Special Notes to improve the progress of your project:

Once a proposal has been reviewed for disclosure risk and accepted, the researcher should use the publicly-available data files to create a data file containing only the public-use variables listed in his or her proposal. RDC staff will attach the requested contextual or user-provided data to this file. The researcher can then access the combined data file remotely through the remote access system or in-person at the RDC in Hyattsville.

NOTES:

- 1 Please run your program using publicly available variables before you submit it to the RDC, to be certain that you have removed errors. If such errors remain in your data set when you submit the program to the RDC, they will slow down the progress of your project.
- 2 When specifying variables to be included in your analysis file, please include all variables you think you might need. If variables are added, it may be necessary to perform another disclosure review, which may slow down the progress of your project.
- 3 A proposal expires after any 12-month period of inactivity from the proposal approval date.
- 4 When constructing the data file, researchers should not delete cases from the data sets. All observations (cases) should be retained, because some statistical packages (e.g. SUDAAN) require you to subset data only within the analytic step (e.g., the PROC statement) in order to estimate the variance correctly.

Your application package should be mailed to:

Peter Meyer, Director,
NCHS Research Data Center
3311 Toledo Road, Room 4113,
Hyattsville, Md. 20782

E-mail: RDCA@CDC.GOV

Phone: (301) 458-4277

Fax (301) 458-4039

Web page: <http://www.cdc.gov/nchs/r&d/rdc.htm>

Please also cc nsfg@cdc.gov when you submit your application, so that the NSFG staff can track the progress of your project.