

## **11. DATA USE AND CONFIDENTIALITY PROTECTIONS**

The National Children’s Study Publications Subcommittee of the NCS Steering Committee will oversee the orderly and timely presentation of pertinent findings and data from NCS to the scientific and medical communities as well as to the public. This will include scientific papers, abstracts, and presentations. The subcommittee will also assure fair and equitable participation in the analysis of the data set and in the presentation of the study results by all NCS investigators.

Press releases and media interviews; and presentations to lay and community groups are the responsibility of the Program Office and Study Centers.

### **11.1 Disclosure Controls**

The NCS Publications Subcommittee and the NCS Steering Committee will both have central roles in ensuring that study participants’ data are appropriately protected. Methods to be used include a broad suite of disclosure control tools, which balance minimizing risk to participants with the potential for societal benefit. The ultimate goal is to protect individuals while still making data accessible to those who might make valuable contributions based on those data. There are several ways in which the NCS will strive to ensure such protection. First, the NCS will employ secure treatment of identifying data through limiting the appearance of personal identifiers on distributed data sets. Second, the NCS will control access to sensitive information by identifying different levels of access to the data and customizing data access plans across levels to ensure adequate protections on all releases of data. Third, the NCS will utilize statistical disclosure control procedures to reduce the appearance of unique personal information in the data that could result in re-identification of a participant.

### **11.2 Public Use Data Sets**

Public use data sets for a given outcome and life stage will be developed for data sharing and made accessible to both the scientific/research community and the general public as soon as feasible, but no longer than within two years of the availability of a usable data set, and in accordance with NIH data sharing policy.<sup>1</sup> These types of public use data sets can be thought of as two levels of data.

#### **11.2.1 Data Sets for the Scientific/Research Community**

Data sets for professional researchers including academics, government workers, and others will be made available in compliance with the National Institutes of Health (NIH) data sharing policy:

“Data-use sharing agreements will put some limitations on who can use the data and how they are to be used. Such agreements will contain requirements, including those to protect the privacy of subjects and the confidentiality of the data. These agreements will incorporate confidentiality standards to ensure data security at the recipient site and prohibit manipulation of data for the purposes

---

<sup>1</sup> Relevant NIH policy and guidance on data sharing can be found at the following websites: NIH Data Sharing Policy [http://grants2.nih.gov/grants/policy/data\\_sharing/](http://grants2.nih.gov/grants/policy/data_sharing/); NIH Data Sharing Policy and Implementation Guidance (Updated: March 5, 2003) [http://grants2.nih.gov/grants/policy/data\\_sharing/data\\_sharing\\_guidance.htm](http://grants2.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm)

of identifying subjects. They will stipulate that the recipient not transfer the data to other users, that the data are only to be used for research purposes, that the proposed research using the data will be reviewed by an IRB [institutional review board], and the like.”

Data made available to the scientific and research community will be available in de-identified data sets. Data will be released only after a full and detailed analysis of risk of data disclosure is performed. All users will sign appropriate confidentiality agreements.

### **11.2.2 Public Use Data Sets**

Because there are no signed agreements or restrictions with regard to individual public data users, public use data files demand a very thorough initial review of the data for risk of disclosure. It is likely that a number of disclosure control techniques, jointly called perturbation, will be used on public data to ensure participants are fully protected from snoopers or inadvertent disclosures by those who are outside the research community.

### **11.3 Publication Policy**

Study-wide publications from data that are not yet released to the general public or to the broad scientific/research community will emanate from a de-identified, validated data set issued by the Coordinating Center to NCS investigators. The data set will be available for analysis by NCS investigators after the completion of a life stage (e.g., completion of the 1-year visit) using data from either the entire cohort or from a random replicate of the entire cohort (wave of data collection). The NCS Community of Investigators consists of investigators in the Program Office, the Interagency Coordinating Committee, the Steering Committee, the Coordinating Center, and all Study Center principal investigators (and their site investigators).

A series of derived variables based on raw data, validated by the NCS Coordinating Center and approved by the Steering Committee (e.g., standardized or normed growth measurements), will be included in the database. Both core publications and non-core publications are anticipated.

#### **11.3.1 Core Publications**

Core publications are study-wide publications that address study methodology, baseline cohort descriptions, and the priority exposures and outcomes of the NCS (as identified in the 28 core hypotheses and updated over time). The scope of the core publications will be specified by the Steering Committee in collaboration with the Program Office. The NCS Publications Subcommittee will announce pending availability of data for each core hypothesis, and all interested members of the NCS Community of Investigators will be invited to submit proposals for analyses. Once formed, writing groups will be assigned a Coordinating Center statistician and may begin analysis after receiving approval from the Program Office to expend funds on the effort.

### **11.3.2 Non-core Publications**

Non-core publications are study-wide publications not directly related to the Study's core hypotheses. Proposals for these publications will be generated by the NCS Community of Investigators as well as other government scientists (from lead agencies or otherwise). The Steering Committee and NCS Program Office may wish to advise or participate in the publication of non-core publications to ensure the maximal use of NCS data. Non-core analysis may be with or without collaboration with the Coordinating Center, resources allowing. The data access and publication proposal review process will be described in detail in the NCS Publications Subcommittee Policy Manual.

### **11.3.3 Approval Process**

All members of the NCS Community of Investigators may request permission to publish from the Publication Subcommittee. If the proposal is rejected, the decision may be appealed. The approval process will be outlined in detail in the NCS Publications Policy Manual.

