

The National Children's Study: An Evolving System

Concept of Operations



Table of Contents

SECTION I	4
1 THE PURPOSE OF THIS DOCUMENT	4
2 AN INTRODUCTION TO THE NCS	4
3 THE NCS AS A SYSTEM	5
4 PLATFORMS IN SUPPORT OF THE NCS SYSTEM	6
5 A SNAPSHOT OF NCS ACTIVITIES	6
SECTION II	7
6 DATA APPROACH	7
6.1 NCS ORGANIZATIONAL STAKEHOLDERS	8
6.1.1 FEDERAL PARTNERS	8
6.1.2 NIH LEADERSHIP	8
6.1.3 OFFICE OF MANAGEMENT AND BUDGET (OMB)	8
6.1.4 INSTITUTIONAL REVIEW BOARDS (IRBs)	9
6.1.5 PROGRAM OFFICE	10
6.1.6 NCS ADVISORY GROUPS	11
6.1.7 INDEPENDENT STUDY MONITORING AND OVERSIGHT COMMITTEE (ISMOC)	11
6.1.8 INTERAGENCY COORDINATING COMMITTEE (ICC)	12
6.1.9 NCS FEDERAL ADVISORY COMMITTEE (NCSAC)	12
6.1.10 NCS STUDY CENTER STEERING COMMITTEE	12
6.1.11 EXECUTIVE STEERING COMMITTEE	12
6.1.12 STUDY CENTERS	13
6.2 STUDY METHODOLOGY DEVELOPMENT	13
6.2.1 SETTING STUDY GOALS	13
6.2.2 FORMATIVE RESEARCH	13
6.2.3 ALTERNATE RECRUITMENT SUBSTUDY	14
6.2.4 DESIGN BASED ON GOALS	15
6.2.5 STUDY CONTENT BUILT BY DESIGN	15
6.3 DATA APPROACH	16
7 DATA ACQUISITION	16
7.1 INFRASTRUCTURE	17
7.1.1 FACILITATED DECENTRALIZATION	17
7.1.2 INFORMATICS PLATFORM	17

7.1.3	OPERATIONAL DATA ELEMENTS	18
7.1.4	NCS CENTRAL DATABASE	18
7.2	PARTICIPANTS	19
7.2.1	IDENTIFICATION	19
7.2.2	ENUMERATION	19
7.2.3	RECRUITMENT AND RETENTION	20
7.2.4	PERMISSION AND CONSENT	20
7.2.5	ENVIRONMENT	20
7.3	DATA ACQUISITION	21
8	DATA ANALYSIS	21
8.1	DATA REPOSITORY	23
8.1.1	PRE-ANALYSIS ARCHIVES OVERVIEW	23
8.1.2	THE STAGING DATABASE	23
8.1.3	THE CENTRAL DATABASE	23
8.1.4	THE ANALYSIS DATABASE	24
9	DATA ACCESS AND USE	24
9.1	DATA CATALOG	25
9.1.1	COMPILE AND ANNOTATE	25
9.1.2	DISSEMINATE	27
9.2	DATA REQUEST	27
9.2.1	REQUEST	27
9.2.2	REVIEW AND RELEASE	28
9.3	PUBLICATIONS	30
9.4	APPENDIX A: NATIONAL CHILDREN'S STUDY GLOSSARY*	31
9.5	APPENDIX B: REFERENCES	35

SECTION I

1 The Purpose of This Document

The purpose of this document is to introduce the National Children’s Study (NCS) as an integrated system that supports the NCS data life cycle from the identification of elements to collect through the acquisition, processing, analysis, archiving, and dissemination of meaningful data sets.

This document is aimed at a general audience with familiarity with the NCS and is not intended as a comprehensive, end-to-end overview of the NCS, or as a manual of operations.

However, it will act as a foundation that will be augmented with supplemental documentation containing greater detail. These additional materials are intended to remain as technical internal documents that will elaborate on specific operations such as compliance, informatics, instrument development, analytics, sample repositories, and other themes.

2 An Introduction to the NCS

The National Children’s Study (NCS) is a national longitudinal study that will prospectively investigate the influence of biological, environmental, genetic, and social factors on the health and development of our nation’s children to ensure that future generations of children grow up strong and healthy. The NCS was mandated by the Children’s Health Act of 2000 (Public Law 106-310) and is being implemented by the National Institutes of Health (NIH) with input from the Centers for Disease Control and Prevention (CDC), the Environmental Protection Agency (EPA), and other federal departments and agencies. The NCS is a data-driven, evidence-based, community- and participant-informed study.

Within the NIH, the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) has provided the resources including space, personnel, expertise, funds, and a Program Office for the administration and implementation of the study. Also within the NIH, the National Institute of Environmental Health Sciences (NIEHS), along with other Institutes and Centers, is a resource for scientific advice and expertise.

The NCS is funded through a dedicated Congressional appropriation that is allocated to the Office of the Director at the NIH and subsequently distributed to the NICHD where the funds are managed by the NCS Program Office. The NCS funds both its support and scientific efforts through contract mechanisms.

The primary rationale for employing a contract mechanism is that the Federal government retains all rights to the data as it is collected by Federal contractors. In contrast, Federal law stipulates that data collected by a grant recipient is not owned by the Federal government and thus the responsibility of stewardship and access for the collected data remains with the

grantee. In addition, due to the number of parties involved, the greater control over budget and deliverables, and the length of the Study, a contract mechanism is the most efficient and effective means for fulfilling the mandate of the Children’s Health Act. Through the use of this contracting mechanism, the NCS began and continues to engage in a competitive process whereby contractors are sought for support of NCS activities and implementation of the Study.

To provide the study an unbiased, statistically-valid cohort of participants that can be generalized to the United States population, the NCS, through expert consultation and public discussion, developed a sampling strategy that identified and selected approximately 105 geographic areas from the approximately 3,000 counties within the United States. Based on this sampling frame, the NCS awarded and funded a set of contractors, referred to as NCS Study Centers, to implement the initial phases of the Study. Study Centers are responsible for the local implementation of the NCS and leverage guidance from the NCS Program Office, local Institutional Review Boards (IRBs) (when applicable), and local Community Advisory Boards (CABs).

In addition to the Study Centers, there are several support contracts that provide assistance in a variety of domains across the NCS including informatics, communications, operations, compliance, budgeting, instrument development, choreography, and technical help desk.

3 The NCS as a System

The NCS should be viewed as an integrated system comprised of many processes and components, not just a single study protocol. By framing the NCS as an integrated system, we can better understand the relationships of its many components, thus identifying interdependencies and priorities, which allows for effective resource allocation. A systems approach provides flexibility whereby activities can be reprioritized, reassigned, and reallocated in a dynamic fashion. The capacity to effect rapid changes provides benefits within the context of the larger system, even if a given operation may see a shift in resources and responsibilities. The flexibility of this approach also accommodates the funding environment of the NCS, allowing us to propose and evaluate multiple scenarios to accomplish our goals.

As an integrated system, the NCS has an opportunity to synthesize data coming from a variety of inputs simultaneously, thus enhancing its overall scientific contribution. This integrated system has several data acquisition strategies that highlight the scope and interdependencies of the NCS, including:

- Vanguard Study: A pilot study, which is an operational study aimed at understanding the feasibility, acceptability, and cost of recruitment, operations, and logistics, and visit assessments to inform the Main Study
- Main Study: The study activity that will focus on exposure-response relationships and be designed based on the data and evidence from the Vanguard Study

- Formative Research: Defined limited-scope and limited-duration research projects to address technical or methodological questions required to responsibly and effectively design components of the Vanguard Study or Main Study
- Substudies: Studies that are embedded within the Vanguard Study and, in the future and when applicable, the Main Study to address specific operational or scientific questions.

4 Platforms in Support of the NCS System

In designing, building, and maintaining the NCS system, there are several high-level areas of focus that can be generalized to other environments including business, scientific, and information technology (IT) infrastructure. To provide structure to these focus areas, characterized as areas of continual evolution and growth, the NCS has moved toward a set of distinct platforms that embody the best practices, standards, and resources for a given domain and, when integrated, can support a flexible and potentially optimally functioning system.

These platforms include but are not limited to:

- Compliance (for example, Federated IRB model)
- Operations (for example, Formal Project Management model)
- Informatics (for example, Secure, Open-Architecture, Standards-Based Systems)
- Research (for example, Operational/Logistical Pilot Study)

The ancillary benefits of these platforms lies in their ability to exist as stand-alone units that can be ported to other environments while retaining their benefits. For example, the informatics platform has provided a vision for the secure, open-architected, standards-based systems desired by the NCS. To ensure that these systems are designed in a compliant fashion, the informatics platform has incorporated into its set of standards a need for compliance with Federal security requirements. Using this platform, the NCS has managed to work collaboratively with the Office of the Chief Information Officer (OCIO) and the Study Centers to deliver 36 Federal Information Security Management Act (FISMA)-compliant academic institutions with NICHD ISSO Information Systems Security Officer -approved authority to operate. Similar examples can be drawn from the implementation of the Federated IRB model and the streamlined OMB submission process.

5 A Snapshot of NCS Activities

The NCS system and its supporting platforms have been an ever-evolving entity. During this evolution, individual activities have changed in nature and specific objectives. The NCS Vanguard Study, which began in January 2009, will continue until each enrolled child has been followed for 21 years. Enrollment for the Vanguard Study is currently estimated to be complete sometime late in calendar year 2011, although that date will be empirically determined based on progress respective to Study goals. The Vanguard Study is intended to remain dynamic in nature and content in order to evaluate multiple procedural and operational scenarios in terms of feasibility, acceptability, and cost.

The NCS Main Study will be informed by the NCS Vanguard Study and once sufficient data from the Alternate Recruitment Substudy and the formative research projects is available, the design and operations strategies for the Main Study will be framed. These strategies will need to have the flexibility and scalability to address the dynamic nature of the NCS in terms of scope, schedule, risks, and cost. The Main Study, like the Vanguard Study, will continue until each enrolled child has been followed for 21 years. The Main Study will examine exposure-response relationships and while it will remain dynamic, it is expected to be far less variable than the Vanguard Study.

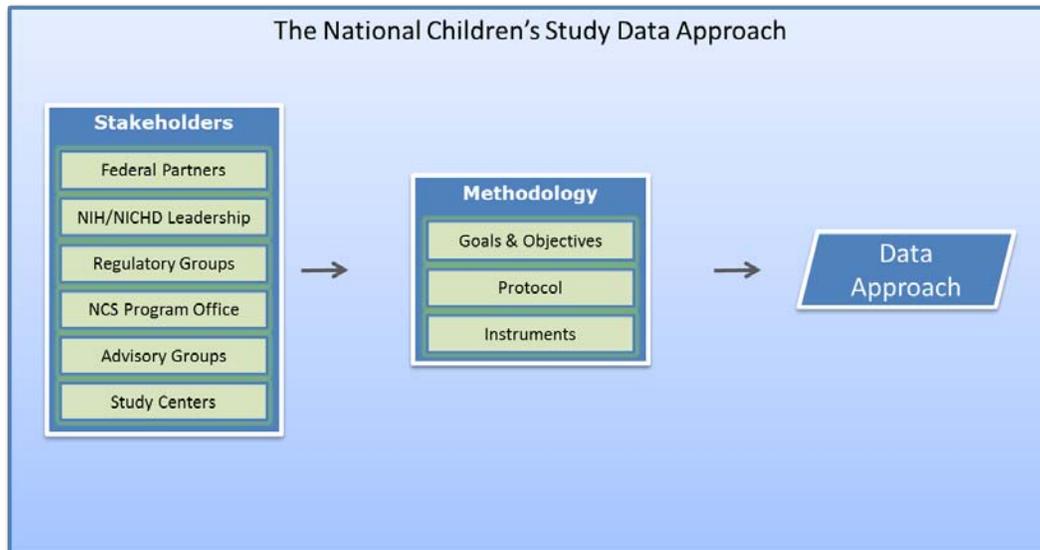
The Vanguard Study and the Main Study will operate in parallel, with the Vanguard Study cohort always preceding the Main Study cohort. This will allow the procedures, processes, and policies developed and vetted in the Vanguard Study to be optimized and scaled for inclusion in the Main Study through an ongoing, iterative process. By recognizing the NCS as an integrated system, the Vanguard Study and the Main Study can offer different combinations of study visit content and variations in infrastructure while successfully addressing a broader range of scientific topics.

SECTION II

6 Data Approach

The approach to data collection in a clinical study is driven by an understanding of what data will be collected, why it is being collected, and how it will be used. This understanding is derived from a complex network of inputs that interact at many levels and ultimately must be integrated to provide a foundation from which data collection can occur.

In the following section, we will walk through the high-level schematic shown below. By highlighting the organizational, methodological, and infrastructural components, we are able to illustrate the inputs used to shape the NCS approach to data. The schematic displays entities in rectangles and processes in trapezoids.



6.1 NCS Organizational Stakeholders

The NCS has many organizational inputs that are incorporated and leveraged in creating its evolving approach to data. These stakeholders provide guidance on a broad scope of issues ranging from governance to compliance to scientific rationale. To recognize the roles and responsibilities of these stakeholders their relationships with the NCS are described below.

6.1.1 Federal Partners

The NCS was mandated by the Children's Health Act of 2000 (Public Law 106-310) and listed NIH, CDC, and the EPA as participants in fulfilling the mandate under the leadership and guidance of the Director of the NICHD.

6.1.2 NIH Leadership

Currently, the NCS is included in the Federal appropriations bill as a separate entity with its own designated funding allocation. This appropriation is distributed directly to the NIH Director who, in coordination with the NICHD Director and the NCS Director, allocates the funds while continuing to provide oversight of their use. To fulfill its part in meeting the mandate, the NICHD established an NCS Program Office and provided needed resources including space, personnel, expertise, and funds for the administration and conduct of the study. The daily operational and scientific guidance for the NCS is provided by the NCS Director and the NCS Program Office staff.

Also within the NIH, the NIEHS, along with other Institutes and Centers, provides the NCS with scientific advice.

6.1.3 Office of Management and Budget (OMB)

The OMB is the largest component of the Executive Office of the President and aims to implement the commitments and priorities of the President. To achieve this mission, the OMB leverages five distinct processes including management, which involves oversight of agency

performance, Federal procurement, financial management, and information/IT (including paperwork reduction, privacy, and security). Specifically, the OMB is charged with regardless of form or format, whether numerical, graphic or narrative in form, and whether oral or maintained on paper, electronic or other media.” Therefore, any information requested of an individual, whether an existing public record or not, is considered “information” and subject to the requirements of the Paperwork Reduction Act of 1995 (PRA). The PRA was developed to: “ensure the greatest possible public benefit from and maximize the utility of information implementing the Privacy Act of 1974 and the E-Government Act of 2002, of which the FISMA of 2002 is one component.

The OMB regulations define “information” as “any statement or estimate of fact or opinion, created, collected, maintained, used, shared and disseminated by or for the Federal Government.” The PRA requires Federal agencies to: (1) seek public comment on proposed collections, and (2) submit proposed collections for review and approval by the OMB.

The NCS fulfills the requirements of the PRA as prescribed by the OMB by submitting all data collection instruments and respective burden-hour calculations to the Office of Information and Regulatory Affairs (OIRA) of the OMB for review and approval. The OIRA evaluates whether the collection of information by the agency is necessary for the proper performance of the agency, including whether the information has practical utility, and minimizes the Federal information collection burden. The OIRA also reviews the extent to which the information collection is consistent with applicable laws, regulations, and policies related to privacy, confidentiality, security, information quality, and statistical standards. In addition, OIRA provides scientific and statistical advice based on the expertise of its staff. The only exception is those cases where the total participant population will not exceed nine individuals, as this is currently outside the requirements of the PRA.

As part of its overall compliance platform, the NCS has continued to evolve its approach to communication, submission, and collaboration to comply with the guidance provided in the OMB references, *Standards and Guidelines for Statistical Surveys* (September 2006) and *Questions and Answers When Designing Surveys for Information Collections* (January 2006).

6.1.4 Institutional Review Boards (IRBs)

The Code of Federal Regulations Title 45 Part 46 requires that federally funded research, such as the NCS, be approved by an Institutional Review Board (IRB). Oversight for IRBs is provided by the Department of Health and Human Services (DHHS) through the Office for Human Research Protections (OHRP), which provides oversight and monitoring of the IRBs through a program of assurances. The IRBs are required to apply for and obtain a Federal-wide Assurance number and report to OHRP.

The role of an IRB is to approve and monitor human subject research by making an assessment of the potential risks and benefits to participants and stipulating changes to the study if the risks or risk management is, in the view of the IRB, not acceptable. IRBs evaluate the informed consent process as a component of risk management to inform potential study participants

about the study goals and content. Informed participants and vigilant investigators are considered essential for risk management. IRBs may also not approve a study at all.

The Federal system is designed so that local IRBs have the authority to make the relevant determinations and there is no formal appeal mechanism at a state or Federal level, although in some cases an IRB may refer evaluation and determination to a Federal panel. The Federal regulations describe a system of risk classification beginning with minimal risk, generally defined as the risks encountered in daily living or those associated with routine medical care. The NCS is currently classified as a minimal risk Study.

The broad geographic scope of the NCS involves multiple institutions and IRBs. To manage the large number of IRBs, the NCS is implementing a Federated IRB model. The Federated model is a mechanism for establishing a shared set of principles for the review of NCS documents, for sharing information across all IRBs reviewing the Study and for providing the opportunity to streamline local IRB review by allowing the option of reliance on a lead IRB as the IRB of record. For the NCS and those contractors who choose to reassign reliance, the lead IRB is the NICHD Intramural IRB.

As part of its overall compliance platform, the NCS has continued to evolve its Federated IRB model. All Institutions may select from various tiers of participation in the Federated model: reliance on the NICHD IRB, shared responsibility, or maintaining independence. At any time, an IRB may shift from one tier of participation to another with adjustment of a Memorandum of Understanding.

6.1.5 Program Office

The NCS Program Office (PO) directs the implementation of the day-to-day operations of the Study. The Program Office houses scientists, staff on detail assignment from other organizations, program analysts, and support contractors. The NCS PO is organized around functional domains including: Administrative, Communications, Planning, Operations, and Analysis. Each functional domain is supported by a team with responsibilities that align with the larger integrated needs of the NCS PO and include:

- **Administration:** ensures the operational support of the NCS Program Office and is responsible for records management and document archiving.
- **Communications:** in liaison with the NICHD Public Information and Communications Branch and the NIH Communications Office, is responsible for supporting effective communications regarding NCS activities and plans with contract investigators, participant communities, Study Centers, and the general public. The Communications Team coordinates the NCS Steering Committee and Public Advisory Committee meetings, organizes presentation materials, and prepares an external newsletter and an internal daily bulletin and quarterly report. The NCS is complex and requires effective communications to coordinate activities. The NCS Program Office recognizes the essential need and priority for effective communications and the fact that people

absorb information differentially. The communications strategy is intended to be multi-modal and redundant coupled with archiving of essential information. Therefore, the Communications Team uses several communications channels to relay NCS messages, including:

- Internal daily bulletins
 - Weekly newsletters to the NCS contractors
 - Directed e-mails
 - Teleconferences, including a general weekly teleconference
 - In-person meetings
 - Documents circulated electronically and posted to a secure NCS portal
- **Planning Team:** gathers information, develops general specifications, determines resource needs, identifies tactical goals, establishes priorities, and develops recommendations for NCS programmatic activities, including protocol development needs and analysis plans. The team conducts gap analyses, identifies risks, and locates additional external resources—such as projects, programs, and organizations—that can be engaged to support the mission of the NCS.
 - **Operations Team:** guides and monitors the technical implementation of the Study field operations. The Operations Team also includes the Contracting Officer’s Technical Representatives (COTRs) that provide the technical guidance to the Contracting Officers on the NCS Study Center contracts.
 - **Analysis Team:** evaluates the integrity (data quality), feasibility (scientific robustness), acceptability (burden on participants and study infrastructure), and cost of Study data. The team recommends changes in study protocol and operations based on these evaluations.

6.1.6 NCS Advisory Groups

The NCS receives scheduled and structured input through several standing advisory groups. The NCS Program Office has a responsibility to integrate input from all stakeholders to develop policies and procedures in conformance with all applicable federal laws, regulations, and policies. All NCS policies and procedures are vetted through all relevant oversight committees and individuals.

6.1.7 Independent Study Monitoring and Oversight Committee (iSMOC)

The independent Study Monitoring and Oversight Committee (iSMOC) monitors the NCS data integrity and usage and the protection of Study participants. The committee reports to the Study Director and its responsibilities include:

- Monitoring human subject safety through review and evaluation of accumulated Study data
- Reviewing Study conduct and progress
- Making recommendations concerning continuation or modification of the Study

The iSMOC considers Study-specific data and relevant non-Study background information when reviewing the Study protocol and identifies any major monitoring concerns prior to or shortly after implementation. During the Study, the iSMOC reviews data regarding procedure-related adverse events, unanticipated problems involving risks to subjects or others, adherence to the protocol, factors that might affect Study outcome or compromise Study data (for example, protocol violations, losses to follow-up, or breach of subject confidentiality), and barriers to Study progress or completion (such as slow enrollment, new data or findings, other milestones, change in resources, or rate of data accumulation). The iSMOC also recommends appropriateness of notification and referral of individual participants or communities for significant findings that may trigger a medical or public health action. Confidentiality is maintained during all phases of iSMOC review and deliberations; therefore, the meetings or formal meeting records are not public. Summaries of iSMOC meetings may be selectively shared at the discretion of the Committee, provided that they will not compromise Study confidentiality or integrity.

6.1.8 Interagency Coordinating Committee (ICC)

The Interagency Coordinating Committee (ICC) represents selected federal agencies with an interest in the Study, oversees broad Study issues, and ensures interagency collaboration. The representatives ensure, at a high level, that the mission and goals of the NCS reflect the scientific interests of the participating federal agencies. The committee is currently composed of staff from DHHS and the EPA.

6.1.9 NCS Federal Advisory Committee (NCSAC)

The NCS Federal Advisory Committee (NCSAC), constituted under the Federal Advisory Committee Act, provides strategic advice and recommendations regarding the Study to the Director of the NIH, the Director of the NICHD, and the Director of the NCS. The NCSAC meets quarterly and meetings are open to the public, with summaries posted on the NCS public Web site.

6.1.10 NCS Study Center Steering Committee

The NCS Study Center Steering Committee membership consists of representatives from the contracted NCS Study Centers. The Steering Committee provides a forum to discuss issues forwarded by the NCS Program Office, the Study Center Principal Investigators, and the Executive Steering Committee. The Steering Committee can recommend tactical protocol modifications that do not change the direction or overall cost of the Study, subject to confirmation by the Program Office. The expanded Steering Committee may refer management and operational issues to the Executive Steering Committee.

6.1.11 Executive Steering Committee

The Executive Steering Committee is a subset of the NCS Study Center Steering Committee composed of Study Center Principal Investigators (PIs) having regional representation. The committee meets monthly by teleconference and is tasked to discuss protocol and operational

topics. This committee addresses issues that are budget-neutral and includes topics that do not change the Study's mission.

6.1.11.1 NCS Study Center Community Advisory Boards (CAB)

Each of the geographic locations participating in the NCS has a local Community Advisory Board (CAB) that is supported by an NCS Study Center. CAB members are community representatives who provide Study Center staff with guidance regarding implementing the Study in their communities.

6.1.12 Study Centers

The NCS Study Centers are the entities that are under contract with NICHD to implement the Study in selected locations throughout the United States. The Study Centers work within their designated location(s) to recruit participants and collect and process data.

The success of the Study in each location requires the collaboration of researchers, government officials, health care workers, social service agencies, and community groups, such as schools, churches, and local governments.

Each NCS Study Center has both primary and secondary Project Officers within the NCS Program Office who function as Contract Officer Technical Representatives (COTR).

6.2 Study Methodology Development

6.2.1 Setting Study Goals

The Vanguard Study is a pilot study designed to help provide data for the detailed preparation needed to design and implement the NCS Main Study.

The Vanguard Study is evaluating the feasibility, acceptability, and cost of three different recruitment strategies, as well as Study procedures and outcome assessments that are to be used in the Main Study.

The Main Study will focus on exposure outcome relationships with a data driven, evidence-based approach.

6.2.2 Formative Research

The NCS has initiated formative research projects that are limited in scope and duration, are intended to augment the Vanguard Study by addressing specific technical questions, and provide information on the acceptability, feasibility, and cost of the research. Each formative research project must comply with OMB, OHRP, FISMA, and local IRB requirements.

These formative research projects will provide data to explore new and potentially cost-effective approaches in many areas, including genetic, cognitive, and environmental assessments that have not previously been evaluated from an operational perspective.

Current NCS formative research projects are examining outcomes in the following areas:

- Real-time or near-time analysis of Study samples, specimens, and measurements
- Study logistical analysis and improvements
- Biospecimen collection and processing
- Environmental sample collection and processing
- Physical measures
- Questionnaire development and validation
- Study infrastructure development

Based on the results of these formative research projects, the Vanguard Study can evaluate the types of research questions that will be feasible for the Main Study.

6.2.3 Alternate Recruitment Substudy

The NCS is currently conducting a participant recruitment substudy to determine the most efficient and successful way to recruit participants into the NCS.

Study Centers are engaged in three different recruitment strategies:

- **Provider-based Recruitment:** In this schema, the Study will be introduced to potential participants through their existing health care system. Women will receive information about the Study from their health care providers (such as a doctor, midwife, or public health nurse) in a familiar and trusted environment. These efforts will be supplemented with general outreach activities aimed at raising awareness and recognition of the Study at a broader level.
- **Enhanced Household-Based Enumeration:** In this schema, the Study will be introduced to potential participants through NCS field workers performing door-to-door enrollment in the selected Study segments. These household visits will complement information the families receive about the Study through community engagement and media activities.
- **Two-Tiered Recruitment (High-Intensity/Low-Intensity):** Public media will be the primary channel through which potential participants are introduced to the Study. Participants will be enrolled in the low intensity data collection effort, with some participants being invited to join a higher intensity data collection.

After enrollment, the protocol is identical for each Study location with the exception of the broader low-intensity catchment area around the high-intensity area in the two-tiered recruitment schema. Each schema is being implemented in 10 locations that are geographically and demographically diverse. There will be no attempt to have a population that can be

generalized to the U.S. for any of the schema. Nonetheless, the NCS has a specific interest in examining potential enrollment bias using U.S. Census data as a reference frame. The groups of 10 locations are approximately equally resourced and each has a specific communications theme. In sum, 30 locations are engaged in the Alternate Recruitment Substudy. Together with the initial seven locations that piloted the Household-based Enumeration recruitment strategy, the NCS has 37 active Study locations.

The primary outcome measures for the alternate recruitment substudy are the recruitment and retention rates of the three alternate recruitment strategies. No specific numerical target is prospectively identified. Recruitment will end for any given location when a steady state is reached. The definition of steady state is when a Study location maintains about the same recruitment rate for 3 consecutive months.

The goal of the alternate recruitment substudy is to define quantitatively the effectiveness and costs of different recruitment strategies in different locations. Analysis is not intended to select a single recruitment strategy for the NCS Main Study, but rather to provide data to construct a hierarchy of approaches that may be applied in different settings.

6.2.4 Design Based on Goals

6.2.4.1 Feasibility, Acceptability, and Cost

The feasibility of the NCS will be based on the technical performance of Study procedures and operations. Acceptability will look at the impact of Study procedures and operations on Study participants and Study infrastructure. The cost merit of the NCS will be ascertained through qualitative and quantitative assessments of the level of effort, material, logistics, and funds required to implement Study procedures and operations.

6.2.5 Study Content Built by Design

The goals of the NCS Vanguard Study require a research study design that systematically examines the scalability of all the operations associated with Study implementation and operations, allowing an accurate and fiscally responsible design of the much larger Main Study.

The NCS needed to develop a framework based on operations in survey research and then extend the concepts to other fields due to the fact that no formal consensus design or data standards exist for pilot studies and operational data elements. The development process is iterative beginning with the NCS Program Office. The Program Office circulates drafts for comment to relevant advisors and stakeholders and then tracks the changes and versions until field testing. Additional development is informed by the results of the field testing and parallel developments in other fields. A multitude of survey instruments are required to support the capture of the required interview data, sample collection information, and measurements for the NCS.

Study instruments can be administered in different modes, including: Computer Assisted Personal Interview (CAPI), Computer Assisted Telephone Interview (CATI), Automated

Computer Assisted Self Interview (ACASI), Self- Administered Questionnaires (SAQ), and Computer Assisted Web Interviews (CAWI). Certain instruments will be administered in a single mode while others will have a multi-mode capability (for example, administration either in person or over the phone).

The overall data collection protocol for the NCS is varied and encompasses questionnaires, biospecimen and environmental sample collection, physical measurements, neurological assessments, and other types of assessments. Many of the instruments are used across several Study visits or may be administered in different modes (for example, CAPI or CATI). A robust survey design tool to support the definition of these instruments, their modes of administration, and other relevant information is under discussion.

The NCS Program Office has a structured work flow process and team for instrument development. Details of the instrument development process will be included in an Appendix to this document.

6.3 Data Approach

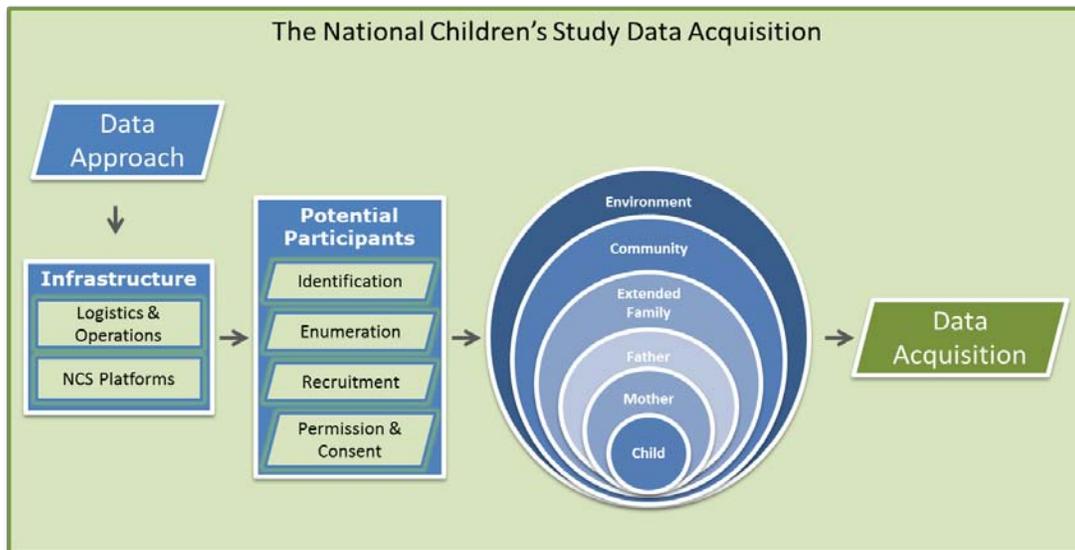
The overall approach of the NCS is based on the convergence of stakeholder input, the Study methodology, and the infrastructure. Study planning is followed by identification of volunteers to donate data to the Study. The data sources for the NCS are the enrolled participants and their environment.

After receiving permission from the Study participants to collect their data, a systematic structured data acquisition process that is centrally specified but locally implemented occurs. Data acquisition is followed by data transmission using specified standards to either a central data repository or to a data storage facility (such as a biobank) or to a data processing facility (such as a regional laboratory) and then subsequently to a central data repository.

Within the central data repository, the data are processed and quality controlled, transformed, and catalogued. The catalogued data are then available for analysis with additional cataloguing of the resulting analytic datasets.

7 Data Acquisition

The acquisition of data in the NCS is shaped by several factors, including how the data will be collected, the sources of the data, and the overall environment in which the collection process will occur. This section describes the high-level schematic shown below, which highlights the infrastructure, participants, and overall environment that shape the acquisition of data within the NCS. Rectangles represent entities, trapezoids represent processes, and circles represent data sources.



7.1 Infrastructure

7.1.1 Facilitated Decentralization

The initial phase of the NCS Vanguard Study utilized a centralized model of operations that applied to most activities from the management of instruments to warehousing and distribution to the informatics platform. Based on these early experiences, the centralized model was found to have multiple technical and logistical challenges. In response to these challenges, the NCS Program Office implemented a revised operations model, termed *facilitated decentralization*, that is focused on the consistency and quality that comes from centrally provided specifications while leveraging the flexibility and innovation of local selection and implementation.

7.1.2 Informatics Platform

As part of the facilitated decentralization model, the NCS has evolved its approach from a centralized informatics platform to a request that Study Centers provide non-proprietary, open-architected, modular, scalable, and standards-based informatics solutions. The content, format, and security functions of those solutions must meet those specified by the NCS Program Office. Study Centers investigated, selected, implemented, and began deployment of case management, data acquisition, collection, and storage platforms, and practices and policies that ensured that the content, format, and handling of the data complied with the NCS PO specifications. Developed through a consortium of partners, these specifications include requirements for data fields, tables, and relationships; formatting and transmission standards; a central data archive; and specifications and guidelines for data security, participant confidentiality, and regulatory compliance.

While the Alternate Recruitment Substudy (ARS) phase of the Vanguard Study will follow the same standards of confidentiality applicable to the Initial Vanguard Study, the decentralized

model puts the primary responsibility on the individual Study Centers to implement, assess, and maintain the security controls that are necessary to assure the confidentiality and integrity of participant data. To ensure that the Study Centers are ready to meet the specifications surrounding security, each Study Center must undergo the Federal Information Security Management Act of 2002 (FISMA) certification process, which involves review by the NICHD Information Systems Security Officer (ISSO) and a final Authority to Operate (ATO) designation granted by the NICHD Chief Information Officer (CIO). As part of the certification process, each Study Center and the NCS Program Office Data Warehouse Center is required to author several formal documents, including a Security Plan, a Privacy Impact Assessment (PIA), and a final risk assessment report. With all necessary documents submitted and approved, the ATO can be granted, which authorizes the initiation of data collection.

Throughout the ARS implementation process, the various data platforms and systems that are deployed in the NCS Vanguard Study will be systematically evaluated to inform the design of the information management system for the Main Study.

7.1.3 Operational Data Elements

The NCS ARS Vanguard Study aims to evaluate the feasibility, acceptability, and cost of study operations and therefore requires the collection of data elements characterizing those parameters. The NCS Program Office has defined a catalog of these data elements, termed Operational Data Elements (ODEs), to accomplish this goal.

The ODEs were developed by identifying specific logistical questions, then constructing data elements necessary to address the questions. The ODEs were subsequently grouped into logical tables and the tables were related to one another to form a data framework. The ODEs and framework were vetted through the various advisory groups and Study Centers and continue to evolve, with a refresh cycle currently planned for every 90 days. Examples of ODEs are the time to perform operations such as travel, interviews, logistics set-up, types of equipment, personnel categories and level of effort, monetary costs, and compliance with planned activities. As no consensus standards exist for ODEs, the NCS is in discussion with several standards organizations to establish a general framework for ODEs as a formative research project.

The ODEs are a subset of the larger Master Data Elements Specification (MDES) that includes the specifications for all operational and study data, collected during the ARS Vanguard phase of the NCS.

7.1.4 NCS Central Database

The data acquired at the Study Centers, whether a biological or environmental sample or an electronic file, is transferred to the custodianship of the NCS Program Office on a regular basis. This transmission process is also dictated by a set of specifications from the NCS Program Office and undergoes additional quality control measures to assess its integrity upon receipt. Regardless of the type or route data takes, it is tracked and recorded. In cases of electronic data transmission, there are specific procedures used to mask or remove personally identifiable information (PII) prior to submission to the NCS Program Office. Described in greater detail in a

future section, the NCS Central Database, also known as the Vanguard Data Repository (VDR), is a family of databases that accepts validated Extended Markup Language (XML) transmission data files submitted by Study Centers for centralized storage of NCS data.

7.2 Participants

The NCS is dependent upon volunteer participants who enroll in the Study and in the process contribute the data needed for the Study. Under the current ARS sampling strategy, potential participants are recruited from 30 geographic locations throughout the United States. These locations were selected from a larger pool of 105 locations using a scientifically-based method aimed at ensuring that children and families across the nation—from diverse ethnic, racial, economic, religious, geographic, and social groups—have the opportunity to participate in the Study.

7.2.1 Identification

In the NCS system, anyone providing information is considered a “Person.” A person can be introduced to the Study through direct contact with NCS personnel or affiliates, referral by others, or self-referral. If that person meets eligibility criteria, she or he is a Potential Participant. A potential participant can give permission through the informed consent process to become an Enrolled Participant.

Women are eligible for enrollment in the NCS if they reside in one of the designated geographic Study segments AND are either pregnant or between the ages of 18 and 49 and could become pregnant.

In the current protocol, enrolled participants can be:

- Women who anticipate becoming pregnant
- Pregnant women
- The children born to pregnant women enrolled in the study
- The father of the child from an enrolled pregnant woman

Enrolled Participants and their physical, community, social, and familial environments are the sources for the data collected by the informatics platforms used in the NCS. Sources include the child, mother, father, extended family, community, and social and physical environment to which enrolled participants are exposed. The NCS will collect data from living relatives other than the parents when feasible, including other caregivers and people that influence family dynamics and child rearing practices.

7.2.2 Enumeration

Within each identified geographic segment, eligible women will be identified through the different ARS strategies and invited to participate. Information about household composition and living arrangements, as well as other aspects of a woman’s social network, will also be collected. The NCS is interested in information flow and influences on both NCS participation and the child’s developmental environment.

7.2.3 Recruitment and Retention

Study recruitment and retention activities will occur at multiple levels (national, regional, state, and local) with the goal of fostering mutually beneficial partnerships that cultivate a sense of personal ownership for the Study. Participation in the NCS should strive to build a sense of trust and pride in the valuable contributions of Study partners and participants to the future health of children in the U.S.

While the Study and the Study Centers ensure the protocol is implemented and information is provided to partners on an ongoing basis, communities provide invaluable advice and consultation on ways to successfully implement the Study and sustain broad support throughout its course. Community engagement and outreach approaches that successfully support recruitment and retention at one Study Center need to be empirically tested to determine their applicability to another Study Center location. A continuum of approaches to community engagement, recruitment, and retention is needed to reflect the varying characteristics of the communities and Study participants involved.

7.2.4 Permission and Consent

Informed consent refers to the regulated process in which eligible participants are presented with the relevant information needed for them to fully understand the opportunities, responsibilities, and potential risks of the Study so they can make an informed choice regarding participation in the NCS. Study personnel will use a variety of delivery modes to provide participants with the needed information and must obtain consent before progressing further and enrolling the participant in the Study.

While participants of legal age indicate agreement to enroll in the Study by signing a consent form, parents and legal guardians provide permission for minor children to participate. When a child is considered to be able to understand Study procedures, assent of the child will be sought. When a child reaches the age of majority in the area in which they live, the child as a Study participant will have the opportunity to consent.

The upper age limit for the NCS follows the Food and Drug Administration Amendments Act of 2007, which states the upper limit of a child for research purposes is 21 years. The rationale for this law is based on data demonstrating physical, neurologic, and endocrine growth and changes that continue into the early twenties. As emerging data continues to emphasize the critical impact and lifelong implications of early events in a person's development, the NCS has favored a strategy of early data acquisition to capture as much information about early events as feasible and affordable. The NCS terms this approach "front loading" of data acquisition and applies it both to pregnancy and early childhood, as feasible.

7.2.5 Environment

The NCS aims to support improved understanding of the ways the social and physical environment, in combination with other exposures, affects children's development. Examples include, but are not limited to language, culture, change in location, travel, access to health

care, participation in organized activities, recreational opportunities, structured learning, out of home child care, and community involvement.

Women with a high likelihood of pregnancy will be monitored to assess their exposures to food and supplements, medicinal products and devices, environmental chemicals and other potential toxins, noise, and natural events prior to and during pregnancy.

7.3 Data Acquisition

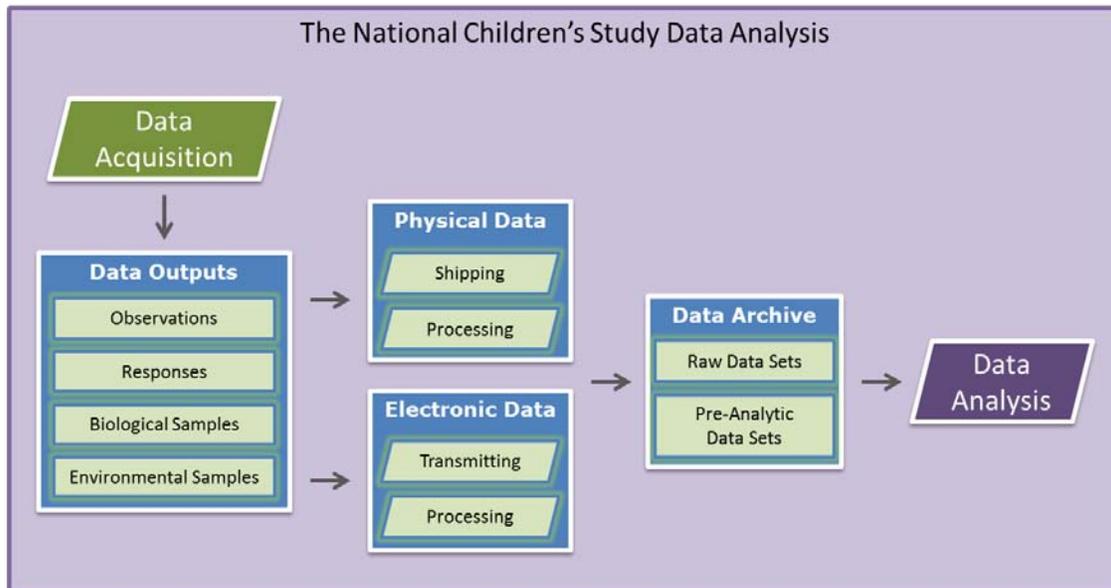
Once enrolled in the NCS, repeat visits are planned for mothers, fathers, and children. At each visit, a combination of operational data, questionnaire data, biological and environmental samples, observational data, and data from physical examinations of Study participants will be collected according to the current protocol. The protocol, particularly the Vanguard Study protocol, is dynamic and will continue to evolve based on scientific and logistical needs. In addition, enrolled participants in both the Vanguard Study and the Main Study may have different visit schedules and assessments to leverage the NCS system for a larger scope of questions and data than if every participant had the same schedule and content.

The modes and tools for data acquisition are numerous, reflective of the scope of data types and scientific questions the Study may pose and of encouragement by the NCS Program Office to be exploratory and innovative as long as quality and security standards are met. Modes may include direct in-person interviews, telephone interviews, computer assisted interviews, self-administered questionnaires, and real time and recall observations.

8 Data Analysis

The analysis of data in the NCS is driven by an understanding of the types of data available, the format of the data, and the organization and location of the data.

In this section, we will discuss the high-level schematic shown below, which highlights the data types, formats, and locations that shape the analysis of data within the NCS. Entities are represented by rectangles and processes are represented by trapezoids.



After the data has been acquired from Study participants and their environments and quality has been checked and archived, the data must be processed and analyzed.

For this discussion, data will be considered generically and independent of type. Some data that is captured in digital form or can be readily converted to digital form can be processed for quality control and any mathematical operations or transformations immediately.

Data in the form of biospecimens or environmental samples is considered for this discussion latent data in that the specimen or sample must be processed and then generally assayed before statistically analyzable data emerges.

As shown in the above work process flow diagram, all data can follow one of several pathways:

- **Immediate analysis at the time of collection:** Examples are physical measures that can be readily entered into a database that has parameters for accepting only plausible values or a biofluid that is analyzed at the time of collection using point of care technology.
- **Data acquisition and subsequent transfer of the data to be processed at another location and time:** Examples are biospecimens such as fluids that are shipped to a clinical laboratory for processing and analysis or data from a cognitive assessment instrument that are processed for quality control and calculation of an integrated or composite score.
- **Data acquisition and subsequent transfer to an archive, either digital or physical, for subsequent processing and analysis:** Examples are environmental samples that are shipped and stored or responses to questions that are transferred to a digital archive

and stored prior to any quality control processing. Data can be retrieved from the archive and, depending upon its format and other factors, analyzed immediately or processed and then analyzed.

All pathways eventually lead to analysis and production of analytic data sets. No matter which pathway the data follow, the individual data elements will be tagged and catalogued on the basis of numerous codes, allowing the NCS to track the life cycle of each data point. Some of the types of codes attached to data are:

- Participant code (a code generated without any personally identifiable information)
- Date and time stamp
- Study site code
- Processing or shipping data
- Type of content
- Relationship to other data elements

Collectively, all the information for the data elements, whether digital or physical, will be available in a catalog.

8.1 Data Repository

8.1.1 Pre-Analysis Archives Overview

The Vanguard Data Repository (VDR) is a family of databases joined in a workflow. The origin of the data into the VDR is Extended Markup Language (XML) transmission data submitted by the NCS Study Centers. The data are submitted without Personally Identifiable Information (PII) based on a list of pre-identified data fields. Although the PII fields are collected, they will be retained at the local Study Center and not transmitted to the central Vanguard Data Repository. Across the workflow a series of transformations take place.

8.1.2 The Staging Database

First in the family of VDR databases is a Staging Database. The Staging Database mirrors the transmission tables specified in the Master Data Elements Specification (MDES). Data from the Study Centers in the specified format and tagged with XML metadata that passes a schema validation tool process (termed XSD) is loaded into the Staging Database.

XSD validation performs several checks on the transmission data set before entering the Staging Database. If a data submission fails any check, the transmission is rejected and NCS Study Centers receive instructions to fix the errors prior to resubmitting.

8.1.3 The Central Database

The Central Database has many more tables than the Staging Database because it includes the lookup tables corresponding to each code list that is associated with the data elements in the MDES.

In the central database, coded responses in the operational and instrument tables are linked to these lookup tables with a logical device termed a foreign key constraint. An additional data schema validator (XSD) tests whether each coded response is a valid response in the associated code list and whether links to records in other tables are valid. The testing of data links ensures that data fields can be properly interpreted and no potentially linked data fields are orphaned.

Study Center data is sent coded with identification keys. Before a submission moves from the Staging Database and is co-mingled with data from other Study Centers, the data has to be keyed again because the generation of the identification keys across all the data collection points is locally controlled.

To ensure unique and consistent identification across the entire system, the identification keys are checked again and if necessary adjusted so data dependent upon other data, termed “children,” uniquely point to the correct data they are dependent on, termed “parents,” and the “parents” can be uniquely identified. The original keys are kept as additional data fields and new primary and foreign keys are added as needed.

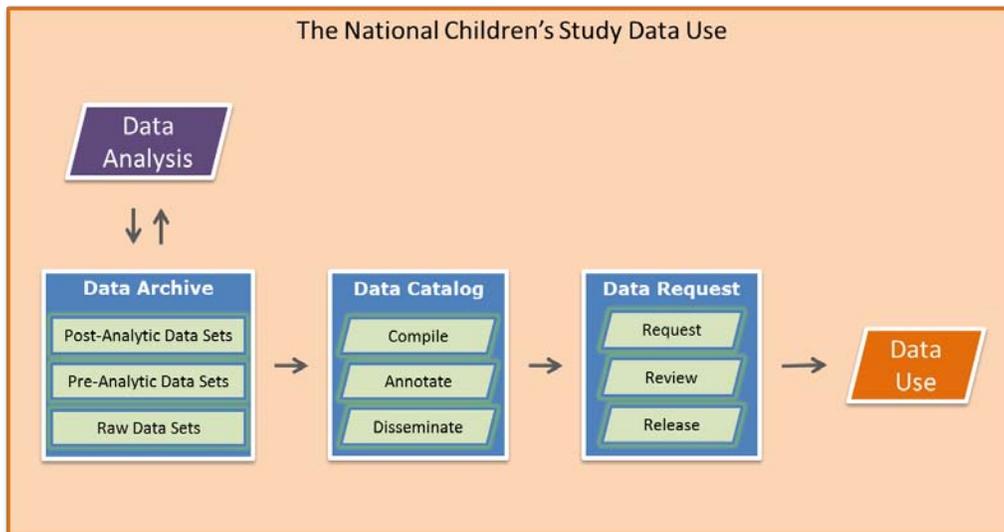
8.1.4 The Analysis Database

The last stop in the VDR workflow is the analysis database. In the current structure the analysis database exactly mirrors the Central Database. Its content is different, however, because as data moves from the Central Database to the Analysis Database, specific fields that may still contain PII that was inadvertently transmitted are removed. All current logical and technology practices for removing PII are supported. The analysis database is considered the data archive and is under the management of an NCS contractor.

9 Data Access and Use

The access and use of data in the NCS is driven by an understanding of the data available, the processes to search and request the data, and the review and approval processes needed to release the data.

This section provides an overview of the high-level schematic shown below, which highlights the processes involved and discusses the potential outcome of publication. Entities are represented by rectangles and processes are represented by trapezoids.



9.1 Data Catalog

9.1.1 Compile and Annotate

As defined by the Federal Information Processing Standards (FIPS Publication 11-3), “data is a representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means.”

To adhere to this definition, the NCS has to ensure the information it collects is presented in a form that it is not only accessible but usable. The *M-06-02 memorandum (Improving Public Access to and Dissemination of Government Information and Using the Federal Enterprise Architecture Data Reference Model)* requires Federal agencies to “organize and categorize agency information” and “make data searchable across agencies.” Despite the unique nature of research data, the NCS still aims to use formal information models so NCS data is cataloged and organized such that access, search, request, and analysis are maximally supported.

Similar to *Data.gov*, the NCS Data Lifecycle Management process will facilitate the identification of information by offering a catalog of datasets while supporting the principles put forth by the E-Government Act of 2002 requiring “the adoption of standards, which are open to the maximum extent feasible, to enable the organization and categorization of Government information . . . in a way that is searchable electronically, including by searchable identifiers.”

This approach to accessibility also provides the research community opportunities to more easily extend the collection and discovery platforms by integrating extant data sources. Within the Catalog, data is designated (or tagged) as either primary or secondary (that is unanalyzed or latent) and ultimately resides in a repository or archive:

- NCS Vanguard Data Repository (VDR)—The VDR stores all of the operational and instrument data element specifications, code lists, and disposition codes as described in the *NCS Vanguard Data Repository (VDR) System Specifications* document. It sets the central protocol for the local execution of the NCS Recruitment Substudy for participating Study Centers. Below are the general categories of data collected, tagged and organized in the catalog.
 - Community outreach and engagement
 - Hospital outreach and engagement
 - Field operations data
 - Study participant data

- Biological and Environmental Sample Repository (BESR)— When the Vanguard Study moves beyond the minimal data collection phase and institutes biospecimen and sample collection, the BESR will store the data locally or regionally at a shared facility or transmit them to a national repository. (Because the collection of specimens (both biological and environmental) is conducted by Study Centers under a contract to the NCS, all specimens from the NCS participants collected at any site or by any individual or a group will be the sole property of the NCS.)
 - The current NCS Repository contractor receives, processes, stores, and distributes biological and environmental samples. Biological and environmental sample archiving tasks are highly technical and require extensive experience, expertise, specialized facilities, and long-term commitment. Each Study participant’s samples are unique and cannot be replaced if lost, damaged, or contaminated. It is essential, therefore, that samples be stored under optimal conditions, which necessitate a wide range of storage temperatures, avoiding the potential of cross-contamination with respect to environmental contaminants or degradation, and avoiding loss of or damage to samples due to a natural or manmade disaster.
 - The NCS continues to explore options in building capacity to continually process, store, and retrieve samples and data in a form suitable for future scientific research that is simple, reliable, cost-effective, and pragmatic for both the NCS research community and future researchers. Details regarding an NCS biobank will be developed in the future.

- Analytic Data Sets—Once data elements have been processed and analyzed, the resulting analytic data sets will be digitally archived. The analytic data sets will be tagged and catalogued on the basis of data element types, subsequent processing and analysis, and additional relevant information. Analytic data sets can be cataloged on the basis of domains and content.

A composite data catalog will provide the research community with a readily available resource for identifying desired data sets and samples that can be requested for use. The NCS catalog of data sets will leverage several design principles, including:

- Standardized harvestable metadata and data sets
- Centralized, secure access points
- Processes, policies, and practices that leverage the Federal Enterprise Architecture (FEA) to ensure alignment and coordination of data access across Federal agencies

To achieve this objective and establish the basis for the use of NCS data by the larger research community, the NCS Program Office must:

- Maximize the accessibility, both in format and timely release, of collected data (i.e., raw data, analytic data sets, and latent data [biological/environmental samples]);
- Ensure that the data provided meets required quality, accessibility, and interoperability standards; and
- Continue to protect participants and their communities against risk of intentional or inadvertent disclosures of sensitive personal information. Therefore, the Data Access & Use process is firmly rooted in the use of the comprehensive data catalog described in the previous section.

9.1.2 Disseminate

The NCS has an expectation and an obligation to disseminate data in the form of data sets that are compliant with Federal regulations and policies, of impeccable quality, well cataloged, and conforming to international standards to allow meta-analyses and comparisons with other data sets. In addition, the NCS will perform and support analyses to interpret the data and share those analyses through presentations, publications, and, if applicable, press announcements.

Data dissemination will occur as rapidly as feasible to maintain data quality and analytic accuracy. The NCS data structures and formats are designed to maximize compatibility and interoperability with other data resources and will continue to evolve to meet those goals.

9.2 Data Request

9.2.1 Request

All information that has been gathered in the NCS is to be kept private to ensure compliance with the Public Health Service Act (42 U.S.C. 241(d)). To ensure this privacy, researchers requesting access to any NCS data products not released as public use files must complete an application for a Data Use Agreement (DUA). This application includes:

- A data request that describes the reason public use data is not be sufficient for the purposes of an intended scientific inquiry
- A data security plan
- Notarized affidavits of nondisclosure for all persons accessing the data
- A license that legally binds the researchers to the terms of the agreement

The current Study Center PIs are responsible for tracking the use of NCS data at his or her location as well as ensuring that every with access to the data at the location reads, understands, signs, and adheres to the DUA. The DUA is reviewed by the Data Access Confidentiality Committee (DACC), which makes the decision concerning approval of the DUA.

Once they have identified a desired data set, users may make a request to access to the data. The process for accessing NCS data depends on the type of data requested. Following are the categories of data sets:

- **Administrative:** NCS Study Center investigators (and their delegates) will have ongoing access to administrative data (that is, exact address, household, and tracing data) related to their local respondents.
- **Crosswalk:** Study Center and the Coordinating Center staffs need access to this highly identifiable data tool for administrative purposes. Access to this form of data will be strictly limited to those whose responsibilities to the NCS require it.
- **Raw:** Non-finalized data files and documentation from the Vanguard and Main Studies will be available to affiliated scientists for data review and evaluation, remediation, post-processing, and variable creation.
- **Interim local:** Cleaned local case data files will be available to local Study Center-affiliated scientists while final national data sets and supporting data use documentation are in development.
- **Public use:** These data have undergone disclosure risk analysis and have been determined not to present disclosure risk.
- **Restricted-use:** These are final data with known disclosure risk.
- **Controlled-use data:** These data require additional security measures given its sensitive and/or finite nature.
- **Supplemental Methodological Study and Adjunct Study:** Produced through either a Supplemental Methodological Study or an Adjunct Study, these data will be subject to the DACC-approved data sharing plan as part of the Adjunct Study review process prior to final Adjunct Study approval.

9.2.2 Review and Release

The NCS operates under a Certificate of Confidentiality issued by DHHS through delegation to the NIH. The Certificate of Confidentiality protects the privacy of the research participants by withholding, from all persons not connected with the Study, any personally identifying characteristics of the research participants except in cases of real or threatened harm as defined by law. The NCS retains responsibility and accountability for establishing the needed governance infrastructure around data access and use. When considering the security and role-based access of NCS data, there are several categories of data access to be considered:

- **Public access:** NCS final analytical data and supporting documentation shall be made available to the research community for the intended scientific purpose in a timely manner regardless of affiliation (see NCS interpretation, 2003 NIH Data Sharing Policy, at http://grants1.nih.gov/grants/policy/data_sharing/index.htm).

- **Access related to task:** Access to non-final NCS data will be guided by roles in the data collection and finalization process.
- **Equal access within responsibility:** Within each role, access will be equal among researchers, regardless of scientific institution or affiliation. This approach maximizes data access for intended scientific purposes and minimizes unmitigated disclosure risk.
- **Variable access using versioning of data sets:** Some data sets will require greater level of detail in personally or community identifiable data for scientific analysis (i.e., geographic information), or in data that, by their nature, are particularly sensitive (i.e., biospecimen and environmental samples, tribal affiliations) or naturally identifiable (i.e., facial images). The NCS will provide versions of data sets for different scientific purposes.

Several governing bodies and channels of communication are leveraged in defining the policies, processes, and approval of data access and use. While each committee has unique missions, membership, and status respective of the type of data requested, they all share the understanding that data security is a fundamental aspect of the NCS and that effective security relies on strong and frequent communication among the key constituencies. The level of involvement of these different entities varies and ranges from high-level oversight as with the Independent Study Monitoring and Oversight Committee (iSMOC) to the regular operational involvement of the Data Access and Confidentiality Committee (DACC). The iSMOC makes recommendations to the NICHD Director, the NCS Director, and the Ethics Advisory Committee, Subcommittee of Federal Advisory Committee, concerning protocol and operational issues to ensure the safety of participants as well as the validity and integrity of the data.

The DACC is a federal interagency committee that establishes policies regarding data access and confidentiality for the NCS. The DACC is responsible for establishing the policies of data security related to who has access to Study data, which data may be accessed, and when data may be accessed, including after the Study's completion. These policies are established in accordance with NICHD/NIH data sharing and confidentiality principles. The DACC reviews manuscripts and presentations to assess and reduce the risk of disclosure, that is, the direct or indirect identification of Study participants and their families. The DACC policies are required to align with the policies of the NICHD, the NIH, and other organizations.

A primary function of the DACC is the review and approval of DUAs. The DACC Chair may solicit input from DACC members, the NICHD CIO and the NCS Director before making a decision on approval. Approval is conditioned on the terms of access granted; the agreement is in force for a named set of investigators at a particular location and security plan to pursue a specified line of inquiry for a period of time. Alterations to these elements require an approved amendment to the agreement prior to implementation. Federal regulations will be enforced upon recommendation by the DACC and at the direction of the NCS Director, including but not limited to revocation of data, termination of data agreements, fines, and contractual repercussions as appropriate. Researchers using public use data only are not required to submit an application for data use agreement.

The following are considered as part of the review of DUA requests:

- **Minimal alteration:** Disclosure risk limitation techniques will be carefully selected to minimize purposeful distortions of data to the extent possible for the purposes of a given dataset.
- **Transparency of strategy:** Disclosure limitation decisions will be mindful of the integrity of the NCS and the quality of its data files.
- **Equality of responsibility:** All persons in contact with personally and/or community identifiable data used by or collected by the NCS are required to adhere to NCS data security guidelines.

Prior to the release of public NCS data, the DACC will develop a Disclosure Analysis Plan (DAP). This plan will identify known disclosure risks to Study participants from the NCS data file proposed to be released, risks in conjunction with previously released NCS data files, and risks in conjunction with other publically accessible data. The plan will identify acceptable levels of disclosure risk for the purpose and intended access of a given data set and propose a method to limit known disclosure risk within the acceptable risk level. The DACC will submit the DAP to the Disclosure Review Board (DRB).

The DRB is an independent advisory board comprised of statisticians specially trained to estimate statistical disclosure risk from data files. The DRB will be closely integrated with the charge and activities of the DACC. The DRB will review the DAP created by the DACC for completeness. This will include a review of the data description and variables identified as identifiable to individuals or communities. The DRB will make recommendations to the DACC in relation to the proposed methods of disclosure limitation and then the DACC will either accept or reject these recommendations in any revision of the DAP.

9.3 Publications

The NCS Program Office encourages publication and presentation of results at all stages of the Study. While formal review and approval of certain publication materials is not required, the NCS Program Office does require notice of any intended publication submissions. If any restricted-use NCS data are intended for publication, including presentations, abstracts, and manuscripts, the material must be provided to the NCS Confidentiality Officer for disclosure review prior to submission for view by persons not party to an NCS Data Use Agreement including the general public.

The NCS Program Office has established a Publications Committee consisting of currently active NCS investigators and Program Office staff for identifying, prioritizing, writing, and reviewing primary NCS publications; that is, all manuscripts regarding centrally collected and integrated data. The Publications Committee will not serve as a formal review mechanism for such publications that are based upon public use data, single Center (“local”) data, formative research, Supplemental Methodological Studies, or Adjunct Studies. Investigators are encouraged to prepare publications based upon those data; however, the Publications Committee will not oversee those publications.

9.4 Appendix A: National Children’s Study Glossary*

**Terms as used in the National Children’s Study*

Acceptability: The impact of each visit and each assessment on the participants, study personnel, and study infrastructure.

Adjunct Studies: Modular, focused studies that leverage the NCS infrastructure, and occur during and integrate with the Main Study. They are initiated, developed, and implemented from outside the NCS Program Office and funded by other than the NCS appropriation. They may be initiated by a variety of sources, such as Federal agencies, Study Centers, independent investigators, industry, and research advocacy organizations.

Address Look Up Tool/ALT: An information application that allows the user to examine whether a potential participant is geographically eligible to participate in the NCS. The user does not have direct access to address lists; the response generated is match or no match.

Alternate Recruitment Strategy/Schema: See Recruitment Strategy.

Ancillary Study: NCS terminology for these during the Vanguard phase is Supplemental Methodological Study and during the Main Study is Adjunct Study.

Authority to Operate/ATO: An Authority to Operate indicates that an information capture and management system has been determined, by the NICHD Chief Information Officer, as compliant with the Federal Information and Security Management Act of 2002. Authority to Operate is one of the conditions to be met before information collection may begin.

Cost: The time, money, number/type of personnel, and effort required to perform an aspect of the Study.

Facilitated decentralization: Centralized development of specifications, quality controls, and policies with the flexibility of local implementation allowing for adaptation to local skills, resources, and other factors.

Feasibility: Assessment of the technical performance of an aspect of the study visits or infrastructure.

Federal Advisory Committee (NCSAC): A Federally chartered committee operating under the Federal Advisory Committee Act that provides advice and recommendations to the Directors of NIH, NICHD, and the NCS on scientific and policy matters regarding the NCS.

Federated IRB: The Federation of NCS IRBs is an IRB model that provides a mechanism for creating a shared set of principles and processes for review of the NCS and for sharing information across all IRBs reviewing the NCS protocol. Institutions can choose to review submissions independently and share information with other members of the federation or

streamline local IRB review by utilizing a facilitated review mechanism or by relying on the review conducted by the NICHD IRB.

FISMA: Federal Information and Security Management Act of 2002/FISMA: The Federal Information Security Management Act of 2002 ("FISMA", 44 U.S.C. § 3541, *et seq.*) is a United States federal law enacted in 2002 as Title III of the E-Government Act of 2002 (Pub.L. 107-347, 116 Stat. 2899). The act recognized the importance of information security to the economic and national security interests of the United States. The act requires each federal agency to develop, document, and implement an agency-wide program to provide information security for the information and information systems that support the operations and assets of the agency, including those provided or managed by another agency, contractor, or other source.

Formative Research: Research addressing acceptability, feasibility, and/or cost related to methods and process, performed during the Vanguard Study. May or may not be “integrated” with the Vanguard Study—that is, may or may not utilize NCS participants and/or NCS laboratory materials. If it is integrated it is termed a Substudy. Formative research is initiated by the NCS Program Office (for example, via request for Letters of Interest) and funded by the NCS appropriation. Its purpose is to inform the Main Study.

Information Management System (IMS): An integrated system of hardware, software, and connectivity that provides administrative, computational, and telecommunications support for the Study. The IMS securely connects all Study Centers and associated locations, the NCS Program Office, various support contractors, and participating laboratories and repositories to support all NCS operational activities and protect the privacy and integrity of the data.

Independent Study Monitoring and Oversight Committee: A committee of independent experts that provides advice to the Director of NICHD and to the Study Director regarding participant and data safety and integrity, and matters related to confidentiality and privacy.

Interagency Coordinating Committee (ICC): A committee of staff and scientists from Federal agencies that support the development and planning of the NCS. The ICC provides strategic advice and advises regarding Study alignment with the missions of the respective Federal agencies.

Letters of Interest: Documents acknowledging interest by a current contractor to perform a specific activity in response to invitation by the Program Office, such as regards formative research projects.

Main Study: The longitudinal NCS, measuring exposures and outcomes from before birth to age 21, to assess effects of the environment on child health and development. Environment is defined broadly, to include the biological, physical, chemical, and psycho-social cultural environments. Gene-environment interaction is a key factor as well.

Manual of Operating Procedures (MOP): Details all procedures to be used throughout the conduct of the Study. In addition to detailing what is collected in greater depth than the Study

Protocol, it also details the multiple aspects of how samples and information are gathered and submitted for processing and storage.

Master Data Elements/MDES: A set of data specifications describing the format of information files Study Centers transmit to the Vanguard Data Repository. The specifications for the MDES are updated as instruments are added to the protocol and questions are revised. The MDES includes operational and study visit measure data specifications. The MDES is a method to establish uniform data formats for information collected using different capture systems.

Primary Sampling Units (PSUs): Individual components into which the target population is divided for the first sampling stage. In the NCS, the PSUs roughly correspond to U.S. counties.

Project Officer: Program Office staff who serve as liaisons to the Study Centers and advise Contracting Officers regarding technical aspects of the Study; they are also referred to as the Contracting Officer Technical Representative (COTR).

Recruitment Strategy/Recruitment Schema: A description of a recruitment methodology that will be subjected to analysis regarding feasibility, acceptability, and cost. The three alternate recruitment strategies undergoing assessment are:

- Provider Based Recruitment Strategy
- Enhanced Household Based Recruitment Strategy
- Two-Tiered Recruitment Strategy (High Intensity/Low Intensity)

Segments: In a multi-stage probability sample, primary sampling units (PSUs) are divided into smaller geographic areas called secondary sampling units or segments. In the NCS, the boundaries for these segments correspond to “neighborhoods or communities.”

Statistically valid sample: A participant population whose results are generalizable to the U.S. population.

Steering Committee: A committee with representation from the NCS-contracted Study Centers to provide, to the NCS Study Director, advice regarding Study implementation.

Study Centers: Contracted organizations responsible for participant recruitment and data collection within the given Study Locations.

Study Locations: The counties (or in sparsely populated areas, groups of contiguous counties) that were selected as primary sampling units in the first stage of sampling.

Study Sites: The segments that were selected in the second stage of sampling, within the Study Locations. The Study Sites correspond roughly to neighborhoods, and will be the areas from which Study participants are recruited and in which most of the community-level data collections will occur.

Substudies: Focused clinical research studies, on a portion of the cohort, which “integrate” with the Vanguard (and later the Main) Study; that is, they involve a portion of the cohort and/or laboratory materials. During the Vanguard phase, a type of formative research, pertaining to

scientific, operational, or methodological questions of limited scope. Initiated by the NCS Program Office and funded through the NCS appropriation.

Supplemental Methodological Studies: Focused methodological studies initiated, developed, and implemented from outside the NCS Program Office, and funded with other than NCS funds. Supplemental Methodological Studies occur during and integrate with the Vanguard phase (that is, they utilize NCS participants and laboratory materials). Their purpose is to inform the Main Study by assessing feasibility, acceptability, and/or cost of logistic, methodological, and operational activities, recruitment-related activities, or specific study visit assessments.

Vanguard Data Repository/VDR: A central information system that receives information transmitted from Study Centers active in the Vanguard Study, including data collected in the Recruitment Substudy. Data received through the Vanguard Data Repository are transferred securely to an NCS Support Contractor for data cleaning, transformation, and further distribution to analysts and organizations upon Program Office direction.

Vanguard Locations: The Study Locations actively recruiting for and participating in the Vanguard Study.

Vanguard Study/Vanguard Phase: The pilot phase of the NCS that is designed to study the feasibility, acceptability, and cost of methodological, operational, and logistic activities, recruitment-related activities, and study visit assessment measurements, for the purpose of informing the Main Study.

Vanguard Recruitment Pilot: A substudy of the Vanguard Study, assessing alternate recruitment strategies.

Visit Assessments: The various functions performed at Study visits, including tests, measurements, interviews, questionnaires, and collection of laboratory materials (biospecimens and environmental samples).

Visit choreography: The process of designing the order and flow of individual components within Study visits.

Revised March 5, 2011

9.5 Appendix B: References

- Privacy Act of 1974, revised 2004 (http://www.usdoj.gov/oip/04_7_1.html)
- NIH Data Sharing Policy, revised 2003 (http://grants.nih.gov/grants/policy/data_sharing/)
- Report on Statistical Disclosure Limitation Methodology, revised 2005 (<http://www.fcsn.gov/working-papers/spwp22.html>)
- Office of Management and Budget (OMB)'s Guidance on Privacy (http://www.whitehouse.gov/omb/privacy_default/)
- OMB's Standards and Guidelines for Statistical Surveys (http://www.whitehouse.gov/omb/assets/omb/inforeg/statpolicy/standards_stat_surveys.pdf)