

THE NATIONAL CHILDREN'S STUDY

Proposed Sampling Strategy:

Main Study

July 24, 2012

This document is intended to structure discussions about the future National Children's Study Main Study sampling design for the National Children's Study Federal Advisory Committee.

Introduction

The National Children's Study (NCS or the Study) is a longitudinal observational birth cohort study with a planned 21 years of follow-up for each enrolled child. The goal of the Study is to explore environmental exposure-health outcome relationships for children, where environment is broadly defined and interpreted in a genetic context. Health outcomes include but are not limited to acute and chronic diseases and conditions, and are intended to be assessed with quantitative and objective criteria for both improvement and decline.

The Study has a pilot or Vanguard Study that began field activities in 2009 with the goals of examining feasibility, acceptability, and costs of implementation. A Main Study with the goal of examining exposure-response relationships will be informed by the Vanguard Study results and is targeted to begin sometime in Fiscal Year 2013.

Initial recruitment in the Vanguard Study began with a two-stage geographic-based probability sample and door-to-door household recruitment. Analyses of the ongoing Vanguard, or pilot phase, of the NCS demonstrated that enrollment rates of pregnant women were lower than expected from recruitment assumptions. Therefore, continuation of the initial proposed sampling and recruitment strategy would not be affordable or sustainable and carry some scientific compromise, particularly with regard to the projected duration of the recruitment period to reach the target cohort size of 100,000 children. Subsequently, the NCS Program Office initiated additional recruitment strategies using a provider-based model and direct outreach to the public. Based on interim data analysis, a provider-based recruitment model has several advantages, including greater efficiency and reduced time of recruitment, with equivalent effectiveness compared to the other approaches.

Consequently, the NCS is proposing the use of a multi-layered cohort approach for the Main Study. This approach would involve as a first cohort geographically based probability area samples, then probability-based sample of birthing facilities and hospitals from within the areas, and then systematic selection of births. A second cohort would be pregnant women seeking prenatal care from providers associated with the selected birthing facilities, and a third cohort of preconception women using a broader list of providers from the same cooperating facilities as the basis for future recruitment activities. In addition, other cohorts may be utilized from outside the cooperating institutions and even outside the designated geographic area in order to target populations that may be underrepresented for any reason of scientific interest. These latter cohorts are intended to be analyzed independently of the core cohorts. Testing of significant aspects of this approach is currently underway at three NCS Vanguard Study locations.

Evolution of the Proposed Main Study Sampling and Recruitment Approach

In order to determine a recommended sampling approach for the Main Study, a series of meetings with various groups of sampling statisticians was planned—the first one with statisticians from other federal agencies was held on March 22, 2012. Topics discussed included dual sampling frame methodologies and the feasibility of these methodologies for the National Children's Study, use of research-ready health organizations including the advantages and disadvantages of using these organizations as sampling units in the NCS, and discussion of other alternative sampling methods. Sampling approach discussions were held with the Federal Consortium (on April 17, 2012), at the public meeting of the Federal Advisory Committee (on April 24, 2012), and with a group consisting of all participating Contractor organizations on May 26, 2012. A final open meeting was held with federal and non-federal statisticians on May 29, 2012. In addition, the NCS Program Office sought further insight through

multiple bilateral meetings with professional societies, advocacy groups, and individual statisticians in person, teleconference, and e-mail exchanges. All of these exchanges were instrumental in reaching the proposed approach.

We are proposing a multi-layered cohort approach for the Main Study design. In order to maintain consistency in language and understanding, we use the term cohort to describe a group of participants who share a common experience such as pregnancy or birth during a designated period and are enrolled in the Study within a defined time frame.

The rationale for using a layered cohort approach is our perception of differences among the characteristics of each cohort that have logistical, cost, or analytic implications and the difficulty of identifying and enrolling a single generalizable sample of women, spanning from preconception to birth, in a practical manner. We propose a set of layered cohorts that would comprise the NCS Main Study sample.

The first layered cohort would be a multi-stage probability-sampled birth cohort. We would call this the core probability sample, as it would have the simplest recruitment strategy and probably the lowest cost compared with the other layered cohorts. This cohort would be comprised of women enrolled perinatally at hospitals or birth centers. The rationale is that the time of entry into the Study would be relatively uniform, and hospitals and birth centers are relatively easy to identify and enumerate for a sampling frame.

This multi-stage probability sample would start with a geographic frame, from which areas with approximately equal numbers of births would be probabilistically selected for the Study; these would be called Primary Sampling Units (PSUs). While it is possible to consider an alternative approach and generate a nationwide list of hospitals and birth centers and select facilities from that list, we feel that limiting the list of hospitals and birth centers to selected geographic areas is more likely to generate a complete and accurate listing. We also favor using geographic areas as the Primary Sampling Units to better control for field work costs and coverage of geographically based environmental exposures. The number, size, and locations of areas to form the geographic frame have yet to be determined.

Within the selected geographic areas (PSUs), selection of hospitals and birth centers would be from an enumerated list, with the probability of sampling proportionate to the number of births at the hospital or birth center. Women giving birth at the selected hospitals and birthing centers would be sampled systematically by an approach such as date or day of birth or 1 of n or some other method.

This cohort would have the following potential advantages:

- Probability-based sample that could be generalized to live births in the U.S.
- Participants would be enrolled with approximately the same starting point
- High expected rate of participation among selected institutions
- High expected rate of enrollment of newborns
- Broad demographic profile because most births occur in hospitals or birthing centers
- Cost effective based on data from prior studies
- Enhanced feasibility of collection of birth samples (cord blood and placental tissue) as participating hospitals will be known in advance, facilitating establishment of operational aspects of the collection.

The major disadvantage is that any prenatal data would be retrospective and based on recall and chart review with little or no opportunity for collection of prenatal environmental or biological samples.

A second layer cohort would be pregnant women who seek health care from prenatal care providers who are on the hospital privilege lists at the same selected facilities used to enroll the birth cohort described above. The women could be enrolled at any stage of pregnancy but the goal would be as early in pregnancy as possible to collect samples and document contemporaneous exposures with a target of 8 weeks of pregnancy. Health care providers would be randomly selected from hospital privilege lists provided by the participating facilities for provider lists above a threshold number yet to be determined. If the number of providers was small, then all providers would be contacted. All women who receive care from a selected provider would be eligible independent of domicile address. Pregnant women receiving care from the cooperating providers would be sampled using a systematic approach of one-in-n patients from a list, or a time interval sample.

This cohort would have the following potential advantages:

- Probability-based sample that could be generalized
- Leverage infrastructure and cooperation of institutions
- Ability to collect prenatal samples and document exposures prospectively
- Ability to document fetal loss
- Option to combine data with first layered cohort.

Potential disadvantages include:

- Variability among various demographic groups with regard to ability to receive prenatal care
- Variable start times within the pregnancy continuum with consequent greater spread among the cohort regarding contemporaneous data collection and a possible bias toward later exposures and events, unless inclusion criteria are adjusted to focus time of entry.

A third layer cohort would be preconception women using a broader list of providers than the prenatal providers from the same cooperating facilities as in the first two layered cohorts. The women would be followed for conversion to a pregnant state for up to 2 years. Once a woman becomes pregnant in this cohort, we intend to follow her and her child, if the pregnancy results in a live birth, using the same methods as the other two cohorts. The advantage of this cohort is the targeted ability to determine exposures during critical stages of early pregnancy, as well as exposures that may have occurred in the peri-conception period or those leading to infertility. These exposures, as well as early pregnancy outcomes such as fetal loss, may represent the tail end of a distribution that is truncated in the cohort of pregnant women. This cohort provides the opportunity to model such relationships, while making it logistically feasible to follow and recruit women. It is unlikely that this cohort is an unbiased sample but would favor women with access to health care and other demographic characteristics. Thus these women would bypass the systematic selection process for the pregnancy cohort or the birth cohort. We are interested in exploring technical methods to relate the data in this cohort to the other cohorts.

This cohort would have the following potential advantages:

- Leverage infrastructure and cooperation of institutions
- Ability to collect preconception samples and document exposures prospectively increasing reliability of exposure assessment
- Ability to document time to pregnancy, infertility, and early fetal loss.

Potential disadvantages include:

- Cooperation rate among individual providers with refusals potentially introducing bias
- Screening of women for intention to become pregnant has been unreliable and costly in previous arms of the Vanguard Study, therefore this sample might be either highly targeted and therefore not generalizable or inefficient and subject to cost constraints.
- Variability among various demographic groups with regard to ability to receive routine medical care
- Due to variability and potential bias, may not be able to combine data with the two other probability-based layered cohorts

Additional cohorts could be outside the cooperating institutions and even outside the designated geographic area and would target populations that may be underrepresented for any reason of scientific interest. An example of one of these cohorts would be a small sample of pregnant women residing in a community where fracking is taking place, where the scientific interest lies in the environmental exposure, but the area or number of births may be so small that the probability of selection into any other cohort is low. These cohorts could be part of ancillary studies that would leverage the resources of the NCS. These targeted cohorts are not expected to be part of the larger probability samples described above although probability based approaches may be used. These cohorts are intended to be analyzed independently of the core cohorts. We propose a scientific review process to screen proposals for targeted cohorts for alignment with the Study goals and prioritization with available resources.

Relevant Questions for the Proposed Approach

- **How large would each of the cohorts be?**
Cohorts one and two (the birth and prenatal cohorts) would comprise 90 percent of the total sample size. Cohorts three and four (the preconception and supplemental cohorts) combined would be up to 10 percent of the total sample size of 100,000.
- **What proportion of all births in the United States occurs in hospitals and birthing centers?**
Based on data from 2006, about 99 percent of births occur in hospitals and birthing centers. The proportion of at home births is estimated to be about 0.6 percent overall with some rural states such as Montana, Oregon, and Vermont around 2 percent.
- **Why use an area frame to determine Primary Sampling Units? Why not a list of hospitals with birthing centers?**
An area frame has two advantages. The first is that the number of hospitals with birthing centers within an area is limited so assessing coverage and generating a list should be straight forward. The second is that logistically we would like to leverage geographic clustering to control the number of field offices and field personnel.

A list of all hospitals and birthing centers can be generated from the universe of licensed hospitals in the United States (about 6,000). However, birth data is generally available from about 80 percent of hospitals from national databases with variability in detail and quality. We anticipate it would be feasible to obtain the relevant data with consistent quality and completeness from close to 100 percent of hospitals in a defined geographic area.

- **Will using a birth cohort approach as the first layer bypass the expectation for assessing prenatal exposures?**

For a cohort enrolled perinatally, the prenatal exposure data will not be prospectively collected and any prenatal biological samples will be serendipitous. Through cooperation with prenatal care providers we hope to generate a prenatal health history via medical chart abstraction, which can serve as a partial exposure history, but will not include the specific examination of environmental exposures of interest. Some environmental scientists argue that a local environment is relatively stable and that sampling can reflect chronic exposures that represent the environment several months earlier. However, we remain uncertain about the reliability of such sampling. By using a proportion of perhaps 40 percent of the overall Study population as a birth cohort, we believe we can obtain a useful sample that can generate accurate generalizable data. Further, while we are missing the individual household exposures, we can still combine these data with general exposure data collected at the municipal or neighborhood level (water quality, air quality, known industrial pollution) to achieve additional retrospective exposure information.

- **How will you collect biological specimens such as cord blood and placentas when the women are identified at birth?**

In most cases we expect to enroll women based on systematic sampling and would request that all protocol specified specimens be collected on all births at participating hospitals during the enrollment window. Those women that do not enroll and for which there is no other reason to retain the specimens would have the specimens discarded. For women that enroll, the Study would receive the specimens. Some hospitals routinely collect blood on pregnant women for type and cross and may collect placentas and cord blood on all births. We would plan to leverage those specimens from facilities that collect them.

- **How do you expect to enroll pregnant women, particularly women who are early (less than 12 weeks) in pregnancy?**

We will use a list of prenatal care providers from the selected hospitals and birth centers as a first step and then, if the number is manageable and the staff cooperative, attempt to use all the providers. If the list is large we will take a random sample of the providers guided by a measure of size based on the number of annual deliveries. We would leverage the cooperation of the hospital or birthing center participation to support participation of prenatal care providers. We will enroll women at any stage of pregnancy but would encourage early enrollment.

- **How will the pregnancy cohort be related to the birth cohort?**

We plan that both cohorts will be probability samples and through the use of the same geographic area and same facilities we can align the two stages of the sampling strategy. The use of the prenatal care providers and selection of pregnant women add stages to the prenatal cohort sampling. We can analyze the demographic and health profiles of both cohorts to confirm the characteristics and detect possible bias in the population recruited compared to community data.

- **If you are enrolling both a birth cohort and a prenatal cohort at the same facility, will you not bias one or the other if they have to compete for the same pool of pregnant women? Would not the birth cohort favor women who did not seek prenatal care, which in general is less than 5 percent of all pregnancies?**

The overall strategy is to use systematic methods and track when a woman is offered enrollment. One approach could be to enroll each cohort at different times with the expectation that the birth cohort may be easier to fill. Subsequently the prenatal cohort would enroll. We are exploring various options and will do modeling to help guide a selection.

- **Would it not be easier to just enroll a prenatal cohort across a continuum of pregnancy lengths and not have a separate birth cohort?**

We believe starting with a birth cohort would be more advantageous mainly because it is easier to build sampling frames of birth hospitals and birthing centers than to build sampling frames of prenatal care providers. An additional efficiency of the birth cohort is the hospital engagement for birth biospecimen collection. This hospital based recruitment would be leveraged for the prenatal cohort with recruitment limited to only the privilege lists of sampled hospitals, again to make enumeration of the provider population simpler, and to use the relationship with the birth hospital to facilitate potential prenatal provider cooperation as well as biospecimen collection. Furthermore, we anticipate the costs of the birth cohort will be lower because it will be easier to recruit and we will not incur expenses for the prenatal visits. We can use the data from the prenatal cohort to calibrate the reliability of the retrospective recall and chart review approach for prenatal exposures that we will use for the birth cohort. As we have seen in our current recruitment substudies, having a large variation in entry points to the study actually creates several cohorts of women for analysis, each with different sets of data. In essence, what we have done is separate these cohorts at the outset, so that the data collection is more uniform within the cohort, which will lead to better sample sizes for analysis.

- **How will the preconception cohort be enrolled and how many preconception women do you target?**

We plan to use the same hospitals and birthing centers as in the other two cohorts, but expand the staff listing to include all providers that provide health care to age eligible women. We would encourage broad outreach and enrollment. We estimate that for every woman enrolled that would become pregnant, we would have to follow at least 6 women for about 2 years.

- **What is the rationale for the preconception cohort?**

We would like to collect data on exposure around the periconceptional period that would have an impact on early fetal development, especially organogenesis, allowing for investigations on fertility, fetal loss and malformation. We would also like to enroll women as early as feasible and beginning with a preconception cohort may enrich for identifying women early in pregnancy.

Provider-Based Sampling in the Vanguard Study

The Alternative Recruitment Strategy (ARS) portion of the Vanguard Study resulted in the identification of a provider-based recruitment approach as a strong candidate for the major recruitment mechanism for the Main Study. The NCS Program Office has been working on a substudy testing the feasibility, acceptability, and cost of a Provider-Based Sampling strategy (PBS) in the Vanguard Study. This strategy involves establishment of a sampling frame of providers who provide prenatal care as a secondary sampling unit to women residing in a geographically based primary sampling unit. From this list, a sample of provider locations is selected with probabilities of selection proportional to the size of practice. Eligible participants are then recruited from selected provider locations.

The goal of this work is to investigate the feasibility of using a probability sample of providers as an alternative to the previous geographic secondary sampling frame. Three Study Centers—Baylor College of Medicine (Harris County, TX), the University of Massachusetts (Worcester County, MA), and the University of Louisville (Jefferson County, KY)—were selected to participate in this participant recruitment pilot as primary sampling units.

These sites are using secondary sampling units of birth centers and prenatal care providers located within the respective counties. Recruiting of children at birth and pregnant women any time prior to birth will utilize this source of potential participants without restriction as to whether or not the providers' locations fall within the geographic boundaries of the primary sampling unit. The women recruited will be required to reside within the primary sampling unit, without further geographic restrictions, as there were in the Provider-Based Recruitment strategy. We are aware of the difference between the current Provider-Based Sampling substudy, which limits the sample of eligible women to a geographic area based on address but does not limit the location of prenatal care providers and the proposal for the Main Study, which would limit the location of birth centers to a geographic area but not place restrictions on eligible women with regard to domicile. This difference arose due to the timing of the development of the two proposals, but nonetheless we believe the Provider-Based Sampling substudy will generate essential data to inform the Main Study proposal.

The implementation of the Provider-Based Sampling pilot has been designed to determine the feasibility, acceptability, and cost of this sampling approach and recruitment method, along with identification of logistical issues and methods of importance to fielding the Study, and the kinetics of listing and recruitment.

Value of the Provider-Based Sample (PBS) Pilot for the Design of the NCS Main Study

The use of a three-site Provider-Based Sample pilot is an essential and critical step in establishing critical parameters for scale-up of the proposed design of the Main Study for the NCS. The ARS Vanguard effort has shown that specific values of operational design elements may have a high degree of leverage on overall cost, data quality, efficiency, feasibility, acceptability, and general utility of the Study and its results. As part of our data-driven approach for the National Children's Study, it is important to identify those critical elements and their range of effective operation for a provider-based sample approach.

As such, the design of the Provider-Based Sample pilot and its implementation methods have been developed and adjusted to provide the necessary data elements for the Main Study.

The PBS design elements have important relationships with respect to the recommended Main Study approach. These include:

- Geographic basis for Primary Sampling Units
- Prenatal care providers perform recruitment
- Systematic sampling of women within provider locations
- Efficiencies involved with data collection when recruiting at hospital or birth center locations
- Enrollment of pregnant women and babies at birth
- Generation and validation of provider lists
- Collection of critical operational data elements.

The results of the PBS will yield important analytic insights that will inform the design of the Main Study. These include:

- Analyses of operational data elements to inform costs and resources required for the Main Study
- Facility cooperation rates
- Provider cooperation rates and refusal characteristics
- Participant cooperation rates and refusal characteristics
- Frame coverage
- Effectiveness of methods to screen women so they only have a single chance to be selected
- Kinetics of recruitment
- Gestational Age distribution of enrolled women
- Demographics of enrolled women and children including race, ethnicity, and socioeconomic status
- Measurement of bias in the recruited population
- Acceptability of the recruitment method and operational aspects at all stages of sampling
- Efficiency of recruiting women
- Retention between enrollment and birth
- Cost and quality of frame development
- Effectiveness of frame construction methods.

THE NATIONAL CHILDREN'S STUDY RECRUITMENT UPDATE

This section of the document provides an update on National Children's Study recruitment into the Vanguard Study, as of June 14, 2012.

Initial enrollment began at 2 locations in January 2009 and expanded to 7 locations during the spring of 2009 using a household-based door-to-door recruitment approach in selected geographic areas within selected counties across the United States. The household-based recruitment continued in the initial 7 locations until late summer 2010. Subsequently in the autumn of 2010 recruitment was expanded to an additional 30 locations using 3 different methods based on method of initial contact. The first was enhancements to the door-to-door household-based approach in 10 locations. The second was direct outreach to participants through media and local events at another 10 locations. The third was referral by a trusted individual or organization using primarily health care providers at an additional 10 locations. The aggregate of all these activities was known as the Alternate Recruitment Substudy. Active recruitment continued at these 30 locations from autumn 2010 through early 2012. Both pregnant and non-pregnant women were enrolled. Non-pregnant women were followed for up to 2 years if they did not become pregnant. If a woman became pregnant, she was enrolled into the pregnant woman cohort.

The following tables and figures summarize the recruitment experience at all 37 locations. The first table is an overall summary. The first figure provides further details on enrollment at the initial 7 locations. The second figure provides further details on enrollment at the 30 locations in the Alternate Recruitment Substudy grouped by recruitment strategy. The final 2 tables show data on recruitment efficiency and proportions of pregnant women enrolled in the different arms of the Alternate Recruitment Substudy.

Table 1: Recruitment status of Vanguard Study participants, as of 6/14/2012.

	Initial Household (2009 cohort)	Alternate Recruitment (2010 cohort)	All Vanguard to Date
Locations	7	30	37
Recruitment duration, months	18 active+ 19 follow up	14 active+ 3 follow up	
A. Women eligible for contact	35000	50700	85650
B. Contacted for pregnancy screen (% of eligible)	34350 (98%)	44600 (88%)	78950 (93%)
C. Completed screen (% of contacted)	30900 (90%)	38350 (86%)	69250 (88%)
D. Pregnant or trying (% of screened)	3100 (10%)	7000 (18%)	10100 (15%)
E. Enrolled (% of pregnant or trying)	2000 (64%)	5100 (73%)	7100 (71%)
F. Babies enrolled	1200	2450	3650

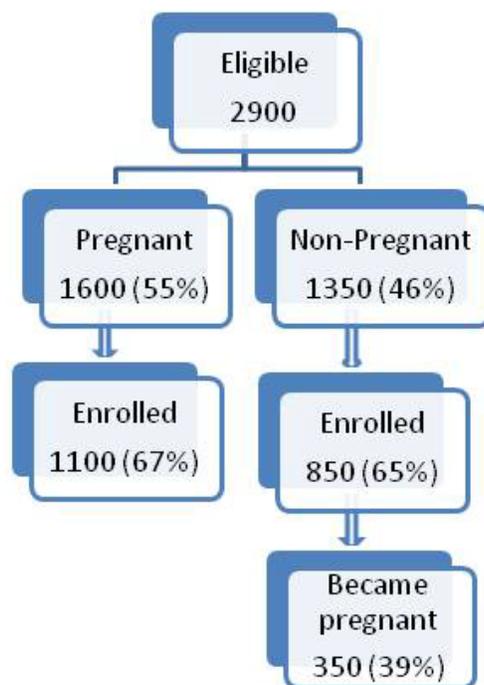


Figure 1: The recruitment experience in the initial household enumeration cohort of 2009, separating the pregnant eligible cohort from the pre-conception cohort. Based on data acquired prior to 6/14/2012, with a total recruitment period of approximately 34 months and a 25-month enrollment period for non-pregnant women.

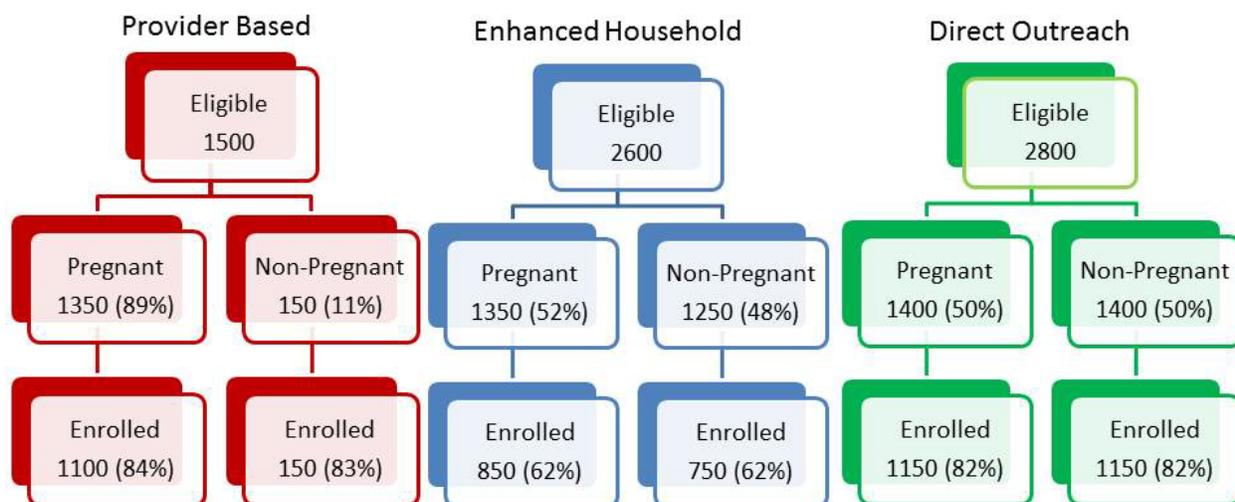


Figure 2: Comparisons of the recruitment strategies, as of 6/14/2012, with a recruitment period of about one year.

Table 2: Screening efficiency of the recruitment strategies, as of 6/14/2012.

	Provider Based	Enhanced Household	Direct Outreach
Number of locations x Weeks in field	701	747	719
Mean number of women enrolled per week	1.5	2.1	3.1
Mean number of women screened per woman enrolled	2.9	13.8	8.6

Table 3: The pregnancy related characteristics of women enrolled in the NCS, as of 6/14/2012.

	Provider Based	Enhanced Household	Direct Outreach
Women enrolled	1250	1600	2250
Proportion of enrolled who were pregnant vs. trying at enrollment	89%/ 11%	52%/ 48%	50%/ 50%
Proportion of pregnant enrollees whose pregnancy was <14 weeks of gestational age at enrollment	23%	23%	22%

VANGUARD STUDY NEXT STEPS

Based on the preliminary observations of the efficiencies of provider-based recruitment, the next step in the Vanguard Study is to adjust the sampling frame to further examine the potential of the provider-based approach. In the Provider-Based Recruitment arm of the Alternate Recruitment Substudy, eligible women were limited to those that had an address in the smaller geographic segments that formed the Secondary Sampling Units within the larger Primary Sampling Units.

In the arm of the Alternate Recruitment Substudy that will go into the field in summer 2012, the geographically based Secondary Sampling Units have been replaced with a list of health care providers that serve women who reside in the Primary Sampling Units. The health care providers do not need to be located in the Primary Sampling Unit, but only their patients that reside in the Primary Sampling Unit are eligible in this design. The development of a list of health care providers as a Secondary Sampling Unit is called Provider-Based Sampling. The rationale is that the geographic-based Secondary Sampling Units of the Provider-Based Recruitment strategy were so limiting that many pregnant women that visited a selected health care provider could not enroll in the Study because their home address was outside one of the Secondary Sampling Units. In order for the National Children's Study to more fully assess the feasibility, acceptability, and cost of enrollment, we chose to eliminate the geographic restriction of the Secondary Sampling Units. Thus all women who reside in a Primary Sampling Unit, which is usually a county, and receive health care at a selected provider, are eligible for the Study.

THE NATIONAL CHILDREN'S STUDY **ADMINISTRATION**

This part of the document provides an overview of National Children's Study operations, administration, and organization.

Administration of the National Children's Study

The National Children's Study (NCS or "the Study") is a prospective national longitudinal study of the effects of environment and genetics on child health, growth, and development. The Study was mandated by the Children's Health Act of 2000 (Public Law 106-310) and is implemented by the National Institutes of Health (NIH) with consultation from a Federal Consortium that includes the U.S. Centers for Disease Control and Prevention (CDC) and the U.S. Environmental Protection Agency (EPA). Funding is provided by a congressional appropriation to the Office of the Director, National Institutes of Health. Within the NIH, the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) provides the resources including space, personnel, expertise, and additional funding and support for the administration and conduct of the Study; the National Institute of Environmental Health Sciences (NIEHS) has provided some additional scientific advice. Oversight is provided by the Director, National Institutes of Health, and an Independent Study Monitoring and Oversight Committee. Strategic advice is provided by a federally chartered Advisory Committee. The Federal Advisory Committee meets quarterly and is a major venue for the National Children's Study to share with the public current activities and receive input and advice.

The goal of the Study is to explore environmental exposure-health outcome relationships for children, where environment is broadly defined and interpreted in a genetic context. Health outcomes include but are not limited to acute and chronic diseases and conditions, and are intended to be assessed with quantitative and objective criteria for both improvement and decline.

The overall principles of the National Children's Study are:

- Data-driven
- Evidence-based
- Community and participant informed

A Concept of Operations document based on the planned data life cycle is available on the National Children's Study Web site (<http://www.nationalchildrensstudy.gov>) at http://www.nationalchildrensstudy.gov/about/overview/Pages/NCS_concept_of_operations_04_28_11.pdf.

The National Children's Study is run by contracts to provide the federal government with flexibility in deploying resources and to ensure that the data collected are not the property of multiple awardees. Contracts are awarded for periods of performance that end upon contract expiration. The National Children's Study awards contracts for data collection, data analysis, data and specimen archiving, and for multiple support functions. Both data collection and support contracts are subject to full and open competition.

The National Children's Study Program Office

The National Children's Study (NCS) Program Office has a full time staff of 19 people and is organized on the basis of functional teams. The teams are:

The Planning Team collects information and develops recommendations for the National Children's Study programmatic activities, including protocol development needs and analysis plans. The team

conducts gap analyses, identifies risks to Study schedule, and locates additional external resources—projects, programs, and organizations—that can be engaged to support the mission of the Study.

The Operations Team determines tactical and technical implementation and monitors the ongoing activities of the National Children's Study Vanguard Study, including formative research and supplemental methodological studies. The Operations Team also coordinates functions with the other National Children's Study Program Office teams to produce deliverables, such as Study Visit Instruments, and manage specimen and sample collection and repositories.

The Analysis and Evaluation Team evaluates the integrity (data quality), feasibility (scientific robustness), acceptability (burden on participants and Study infrastructure), and cost of Vanguard Study data. The Team recommends changes in study protocol and operations based on these evaluations.

The Communications Team, in liaison with the NICHD Public Information and Communications Branch (PICB) and the NIH Communications Office, is responsible for supporting effective communications regarding National Children's Study activities and plans with participant communities, Study Centers, and the general public.

The Regulatory Team manages oversight and compliance with applicable laws and regulations.

The Administrative Team manages NCS Program Office processes, including personnel and resources.

In addition to the process teams, the National Children's Study has content teams.

Study Visit Development is managed by three coordinated teams—the Study Visit Content Team, the Instrument Development Team, and the Forms Development Team. The Study Visit Development Team also coordinates overall protocol development. Each team works in conjunction with support and field contractors to develop the concepts and then the specific elements of each Study visit.

A summary of the workflow process for instrument and visit development is in Figure 1.

Instrument Development Flow

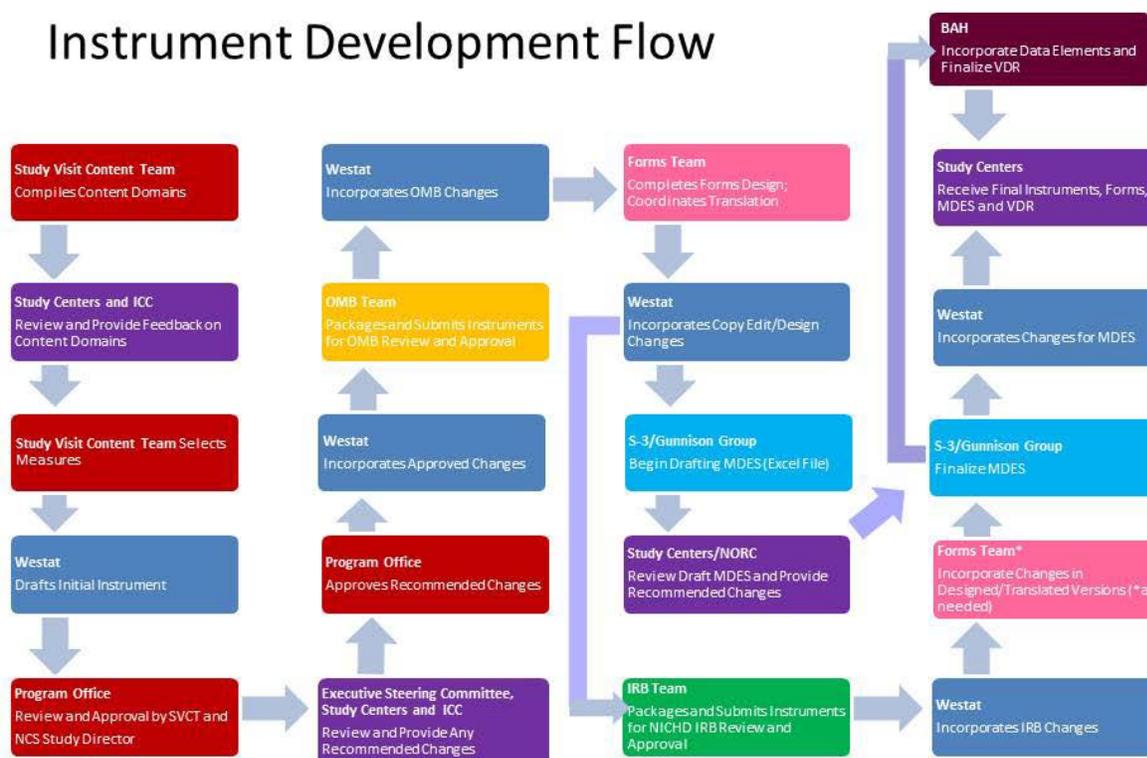


Figure 1: NCS Workflow Process for Instrument and Visit Development with integration of data elements

The Protocol Development Team develops the framework and then the content of the Vanguard Study Protocol and the Main Study Protocol.

The Environmental Team develops formative research projects and contributes to protocol development and analysis.

The Biospecimen Team develops formative research projects and contributes to protocol development and analysis.

The Genetics Team develops formative research projects and contributes to protocol development and analysis, policy, and ethics.

The Data Access and Confidentiality Committee sets policy for data access and confidentiality.

The Publications Committee coordinates the publication of NCS-wide publications through the selection and prioritization of topics and organization of writing and analytic teams.

The Partnership Team develops collaborations and partnerships for the Vanguard Study through Supplemental Methodological Studies and will for the Main Study coordinate Ancillary Studies. The Partnership Team also coordinates the NCS Scholars Program.

Supplemental Methodological Studies (SMS) pertain to focused studies that take place during the Vanguard (pilot) phase of the National Children’s Study. They are geared to inform the Main Study as to

the feasibility, acceptability, and/or cost of items pertaining to recruitment, operational and logistic issues, and Study visit assessments. Supplemental Methodological Studies are initiated from outside of the Program Office and are developed outside the Study protocol planning process. They are funded externally; that is, not with the National Children's Study appropriation. The Principal Investigator will be identified by the applicant. Each SMS will have a Study Co-Investigator from either the Program Office or one of the contracted Study Centers. SMS that will be conducted at more than one location will also have a Study Facilitator for each additional location

Supplemental Methodological Studies are integrated with the Vanguard phase. That is, they involve National Children's Study participants and/or laboratory samples. Requests for just data are not SMS. In contrast, Substudies are a type of formative research involving participants and/or laboratory samples, but initiated and funded by the National Children's Study.

Supplemental Methodological Studies will generally be short-term efforts to support the Vanguard pilot goals. For these studies to inform the design of the Main Study, a prompt turn-around time is pertinent. Please send general inquiries to NCSSupMethStudies@mail.nih.gov

The **NCS Scholars Program** provides federal employees an opportunity to work full time or part time, on site or remotely, on specific projects of mutual interest.

Information Management Systems are coordinated through the Chief Information Officer, NICHD. The Initial Vanguard Study utilized a centralized model of data management, including case management systems and data capture systems. Based on the first year of experience with the centralized model and identification of multiple technical and logistical challenges in planning scale-up, the NCS Program Office implemented a new approach to provide greater flexibility and encourage exploration and innovation to determine preferred methods for case management and data acquisition.

This new approach is termed the "facilitated decentralization" model. In this model, the NCS Program Office develops evaluation questions and plans; data fields, tables and relationships; operational data elements; Study data acquisition instruments; data formatting and transmission standards; a central data archive; and specifications and guidelines for data security, participant confidentiality, and regulatory compliance. This facilitated decentralization model offers distinct advantages over a completely centralized structure: it allows Study Centers under contract with the National Children's Study to select or develop their own case management systems, data acquisition platforms, and as appropriate, data collection modalities to acquire the data. The model builds on local Study Center expertise with existing informatics systems and supports adaptation or development of new systems, with an emphasis on open-source, non-proprietary platforms.

All NCS data systems are certified and accredited per the requirements of the Federal Information Security Management Act of 2002 (FISMA) and related regulations. All NCS data specifications are consistent with international medical research standards, such as those developed by the Clinical Data Interchange Standards Consortium (CDISC).

The new approach to informatics for the National Children's Study is informed by several trends in informatics, including modular architecture, use of standardized terminology with curation, semantic awareness, scalability, defined transmission standards, open source platforms with development communities, vertical and horizontal integration of process, and interoperability. The NCS emphasis on interoperable modular architecture means that any component of a data system can accurately and

efficiently communicate with other data systems, while adhering to international data standards. The approach is flexible to support innovation, accommodate evolving technology, and extend functionality. In addition, its components can be reused or adapted for other studies.

Major Initiatives Coordinated by the National Children's Study Program Office

Examples of trans NCS activities that have implications for other research efforts are:

Health Measurements Network – The concept of health is complex and multi-dimensional. Precise quantitative objective age and developmental stage measures for different health dimensions are not available. The National Children's Study, in partnership with other NIH initiatives, particularly the Patient Reported Outcomes Measurement Information System (PROMIS) and the NIH Toolbox for assessment of neurologic and behavioral function, began a formal initiative to develop relevant assessments for the ages and stages of child development. The resulting tools will be tested and validated in the Vanguard Study as well as in other venues.

Terminology – In conjunction with the NICHD and the National Cancer Institute, the National Children's Study began a systematic effort to develop relevant terminology for all ages and stages of development due to important gaps in all the major terminology systems.

Metadata Tagging of Operational and Longitudinal Data – the National Children's Study is required to integrate data collection and data analysis from multiple domains. Due to different data types, structures and formats, the integration has multiple technical challenges. In addition, in order to develop new instruments, consistency and efficiency in the workflow process is essential. The technical solution is to use metadata to describe both structure and content of the data and append or tag data elements with the structured metadata.

Some business practices adopted by the NCS Program Office are described in the following table.

EXAMPLES OF BUSINESS PRACTICES INITIATED FOR THE NATIONAL CHILDREN'S STUDY

Business Practice		Description
1.	Business Plans and Process Mapping	Developing business plans and process maps that outline the comprehensive set of activities along with timeline and resource requirements for the program.
2.	Risk Assessment and Management	Identifying risks and developing mitigation strategies to address them proactively; formalized in a Risk Management Matrix.
3.	Critical Path Plan Development	A plan that sequences a set of tasks (or workflows) with the shortest time required to complete those tasks towards achieving a goal. Any delays in the critical tasks can jeopardize the project if it is not compensated by acceleration of a later task.
4.	Program/Study Retreats	Monthly retreats designed as an open forum where all of the program/Study staff convenes to review program goals, develop plans, discuss operational issues, address and manage risks, obtain input from external experts, and identify implementable best practices from other studies.
5.	Concept of Operations (CONOPS) Development	A manual of operations designed to provide an overall view of the organizational structure, responsibilities, and interactions between the various components of the program.
6.	Decentralized Model for Operations	An operational model that implements field expertise and control at the local level with general programmatic guidance and oversight at the central level.
7.	Continual Program Monitoring	Using project trackers that list the various activities with target dates to periodically monitor progress of the program.
8.	Federated Institutional Review Board (IRB) Model	A model that establishes a shared set of IRB principles and processes for reviewing Study protocols. It permits information sharing across all IRBs and provides an opportunity to facilitate local IRB review by allowing reliance on the NICHD intramural IRB.
9.	Functional Domain-based Organizational Structure	Organizing the program/administrative office based on functional domains of the program/Study and aligning staff based on their expertise into these domains (For example—planning, operations, analysis and evaluation, communication and administration).
10.	Continual Resource Optimization	Assessing and optimizing staff and their roles to meet new or changing needs of the program. This also ensures that the program staff is aligned to activities based on their expertise and interests.
11.	Project Management Training	Training clinical research scientists in the principles and practices of project management to assist them in managing clinical research studies more effectively.
12.	Scholars Program	A program where interested federal employees from various fields, backgrounds, and training join the Study for 6 months to 1 year and contribute <i>in-kind</i> to the development of the clinical research/study, while also enhancing their own career goals.

13.	<p>Internal Communication Enhancement</p> <ul style="list-style-type: none"> • Daily Muster and Activities Dashboard 	<p><i>Daily Muster</i> provides up-to-the-minute updates, information, requests, and status of various activities or tasks.</p> <p><i>Activities Dashboard</i> provides a weekly status update of all activities within the program categorized by functional areas (For example—planning, operations, analysis, communications, etc.).</p>
14.	NIH Plain Language Initiative Training	<p>The Plain Language Initiative requires the use of plain language in all new documents written for the public, other government entities, and fellow workers. The Plain Language Initiative training focuses on writing that is clear and to the point so as to improve communication between the government and the public.</p>

In sum, the National Children’s Study Vanguard Study reached a critical milestone with preliminary analysis of the recruitment phase that will guide immediate Vanguard activities and Main Study design. The National Children’s Study Program Office is continuing improvements in the business model. All operations continue to undergo evaluation with a commitment to adjust based on performance metrics.
