

**National Children's Study
Federal Advisory Committee 28th Meeting
April 19, 2011
5635 Fishers Lane Conference Center
Rockville, MD**

The National Children's Study (the Study) is led by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) of the National Institutes of Health (NIH) in collaboration with a consortium of federal government partners. Study partners include the National Institute of Environmental Health Sciences (NIEHS) of the NIH, the Centers for Disease Control and Prevention (CDC), and the Environmental Protection Agency (EPA).

Welcome and Introductions

*Carol Henry, Ph.D., Chair, National Children's Study Federal Advisory Committee (NCSAC),
School of Public Health and Health Services, George Washington University*

Dr. Henry welcomed the meeting participants, who introduced themselves. Dr. Henry reviewed the highlights of the January 26, 2010, NCSAC meeting:

- Summary of meeting and presentations posted to Study Web site
- National Children's Study update
- Compensating providers for facilitating recruitment efforts
 - Prenatal care provider recruitment in Montgomery and Schuylkill Counties, PA
 - Explanation of Health Insurance Portability and Accountability Act (HIPAA) Waivers
 - Compensating providers in Wayne County, MI
- Provision of educational materials to potential Study participants
- Vanguard Study recruitment data update and presentation plans for legacy Vanguard data
- Rapporteur's summary of January 2011 meeting:
 - Issues of and data from provider compensation should be further explored.
 - The identification of potential participants not identified through household enumeration and screening should be further investigated.
 - Further discussions are needed on challenges and potential hurdles of HIPAA regulations.
 - Provider compensation is an essential element of provider-based recruitment, but compensation is not necessarily monetary. Study Centers will have to find ways to work with providers and ensure that their engagement is fulfilling.
 - With regard to the provision of educational materials, the Study needs to balance education against confounding or biasing data and maintain the rigors such that the data are interpretable.
 - The NCSAC agreed that the new formats for presenting data are much improved and informative about past activities.
 - The NCSAC agreed that the reasons for participants' refusals, withdrawals, and loss of eligibility due to change in pregnancy status and moving should be further explored.
- Next steps
 - Alternate Recruitment Substudy enrollment and data analysis
 - Continued analysis of data from initial seven Vanguard Study locations including biospecimens and environmental samples

- Gap analysis for formative research opportunities
- New models for visit schedule
- New models for visit assessments
- Introduction of specimen and sample collection across Vanguard Study locations
- Begin construction of framework and architecture for Main Study protocol and infrastructure.

Dr. Henry briefly reviewed the Study’s projected timeline and outlined the agenda for the April 19, 2011, NCSAC meeting.

National Children’s Study Update

Steven Hirschfeld, M.D., Ph.D., Acting Director, National Children’s Study, NICHD, NIH, Department of Health and Human Services (HHS)

- **Overview.** The Study was congressionally mandated by the Children’s Health Act of 2000. It is an integrated system of activities to examine the effects of environmental exposures and genetics on growth, development, and health. The environment is broadly defined to include factors such as air, water, soil, dust, noise, diet, social and cultural settings, access to health care, socioeconomic status, and learning.
- **Study principles.** The Study will be data driven, evidence based, and community and participant informed. Community is broadly defined. Communities will be actively engaged.
- **Exposure areas of interest.** Examples of exposure areas of interest are:
 - Industrial chemicals and byproducts in the air, water, soil, and commercial products
 - Natural products in the air, water, soil, and commercial products
 - Pharmaceuticals used for therapy and in the environment
 - Radiation exposure
 - Proximity to manufacturing, transportation, and processing facilities
 - Living with animals, insects, and plants
 - Media and electronic device exposure, noise
 - Access to routine and specialty health care
 - Structured and unstructured learning opportunities
 - Diet and exercise
 - Family and social network dynamics in cultural and geographic context.
- **Examples of outcome areas of interest.** Examples of outcome areas of interest are:
 - Interpersonal relationships and bonding
 - Inflammatory processes including allergies, asthma, and infections
 - Genetic and epigenetic status
 - Epilepsy and other neurologic disorders
 - Cardiovascular screening and function
 - Childhood cancer
 - Multidisciplinary, multidimensional aspects of sensory input, learning, and behavior
 - Precursors and early signs of chronic diseases such as obesity, asthma, hypertension, and diabetes.
- **Study structure.** The Study is an integrated system of activities. All components and phases together form the Study. Current major components are:

- The Vanguard Study—the pilot phase for methods runs for 21 years. It started in January 2009 with 7 Centers and was expanded in 2010 with 30 additional Centers.
- The Main Study—the exposure-response phase runs for 21 years, about 3 years time-shifted from the Vanguard Study. The Main Study is planned to start in 2012.
- Substudies—studies within studies.
- Formative research—short-term limited studies, focused primarily on methods development, to support and inform the Vanguard and Main Studies. About 270 formative research studies are currently in progress.
- **Vanguard Study goals.** The Vanguard Study is designed to evaluate the feasibility (technical performance), acceptability (impact on participants, Study personnel, and infrastructure), and cost (personnel, time, effort, and money) of Study recruitment, logistics and operations, and Study visits and Study visit assessments. About 500 data items have been developed to examine Study logistics and operations.
- **Vanguard Study protocol development.** The protocol development process follows a hierarchy:
 - Selection of topic areas and items within a topic area
 - Assembly of items into questionnaires known as Study instruments
 - Assembly of instruments and other types of data collection into a visit
 - Assignment of operational data elements to each logistical component and operation of the visit.

The operational data elements form the basis for the primary Vanguard Study analyses.

- **Additional Vanguard Study activities.** The Study’s Program Office and Study Centers launched an initiative to define child health in positive objective terms across developmental stages. The Program Office, Study Centers, support contractors, and multiple domestic and international partners launched a neonatal terminology harmonization effort as part of a larger effort to harmonize terminology across developmental stages. A related effort will focus on the appropriate, consistent, and harmonized terminology for developmental stages and for various conditions and behaviors in children as they grow up.
- **Vanguard Study methods development.** Vanguard Study methods development includes:
 - Systematic testing of all Study components prior to scale up
 - Distributed operations to leverage local and collaborative expertise within a framework of centrally developed specifications
 - Vanguard Study focus on operational data elements to quantitatively describe Study operations and costs
 - Integrating data standards into an “end-to-end” process across the data life cycle (design, planning, implementation, analysis, archiving, and dissemination)
 - Harmonizing terminology across subspecialties and developmental stages.
- **Alternate Recruitment Substudy.** The Vanguard Study is now at 37 locations across the country with 30 locations engaged in new recruitment using one of three strategies. In the enhanced household-based strategy, participants learn about the Study through field workers walking through neighborhoods. In the provider-based strategy, participants learn about the Study through trusted health care providers—broadly defined to include physicians, public health nurses, midwives, and other licensed health care providers. In the Hi/Lo (direct-to-the-public) strategy, participants learn about the Study directly through media and community outreach. The goal of the Alternate Recruitment Substudy is to compare strategies to

assemble a toolkit for cost-effective directed recruitment for the Main Study launch. Both direct data analysis and predictive modeling will be used.

- **Study recruitment as of April 2011.**

	Enhanced Household	Provider	Hi/Lo	All Alternate Recruitment	Initial Household
Locations	10	10	10	30	7
Locations currently in the field	10	9	10	29	7
Recruitment duration, months	5	5	5	5	18
Women identified for contact	14,850	3,650	3,050	21,550	33,000
Women contacted	11,200	1,150	1,450	13,800	30,000
Women eligible	1,150	300	650*	2,100	2,450
Women consented	600	250	500*	1,350	1,400
Babies	50	50	^	100	600

Person and participant numbers have been rounded to the nearest 50, following the Study rounding policy.

*Numbers are counts of low-intensity participants.

^Number rounds to zero.

- **A learning community.** Except for the focus of the Study remaining on the health and exposures of children, all other aspects of the Study are potentially subject to reevaluation and change. The concurrent deployment of alternative recruitment strategies plus a formative research program provides an exceptional opportunity for launching a learning community with structured and systematic training, feedback, process maps, process improvement, modeling, and simulations. The Study has adapted these approaches both centrally and in the field, particularly in the Hi/Lo recruitment cohort, to build an effective learning community.

NCSAC Discussion and Recommendations

- Joan Y. Reede, M.D., M.P.H., M.B.A, asked what the 650 eligible women in the Hi/Lo recruitment strategy represent. Dr. Hirschfeld explained that this recruitment strategy approaches potential Study participants directly. The recruitment catchment area is expanded compared with the other strategies. The enhanced household-based and provider-based strategies adhere to the geographic segments that were selected for the initial household-based strategy. The direct strategy recruits participants who live in the initial geographic segments but also in surrounding geographic areas. Participants in the surrounding areas receive less intensive data collection compared with the other participants.

Principal Investigator Experiences in the Alternate Recruitment Substudy of the Vanguard Study

Hi/Lo Recruitment Strategy Overview

Patricia McGovern, Ph.D., M.P.H., Principal Investigator, University of Minnesota Study Center, University of Minnesota School of Public Health

The Hi/Lo recruitment strategy approaches the public and potential participants directly—not through households or health care providers. Outreach activities include outdoor advertising, media campaigns, and launch events with key stakeholders. The goal is to create community awareness and support, create a favorable view of the Study, engage key stakeholders, and create a “buzz” by the time potential participants receive the advance letter. The purpose of Hi/Lo recruitment strategy is to inform the Main Study of the optimal size of the secondary sampling unit to yield sufficient numbers of Study-eligible pregnant women to meet Study goals.

The Hi/Lo recruitment strategy involves creating a larger pool of participants and varying the intensity of participation. Intensity refers to the type of data collection and the venue for interactions with participants. Women in the high-intensity group provide environmental samples and biospecimens. Data are collected during home visits. Women in the low-intensity group are contacted only by phone, mail, and online (Internet). All women who live in the high-intensity and low-intensity geographic segments receive advance mailings and are asked to contact the Study. Respondents complete a pregnancy screener by phone, mail, or online. Eligible women who live in the high-intensity geographic areas are invited to “convert” into the high-intensity group. Data collection from these women is the same as other recruitment strategies. Data from women in the low-intensity group are collected less frequently by phone, mail, and online.

The Hi/Lo Study Centers have developed a collaborative improvement network (CoIN), which combines collaborative learning and quality improvement processes. The purpose of CoIN is to improve enrollment outcomes; accelerate improvement through shared learning, problem solving, and idea stealing; optimize implementation using a structured learning system to test and document learning; and test theories of how to make work more effective and efficient.

Implementing the Hi/Lo recruitment strategy has been challenging because it has not been done before. New instruments, scripts, and procedures have to be developed, particularly for the low-intensity cohort. Processes require development and standardization across the 10 participating Study Centers. The processes have required approvals from the Office of Management and Budget and Study Center institutional review boards (IRBs). Community outreach and engagement requires establishing a relationship and trust with neighborhoods. Problems include the inability to name neighborhoods and residents’ identification with municipalities and communities, not with counties.

The many strengths of CoIN have facilitated collaboration across Study Centers and have improved the development of new instruments, scripts, data definitions, procedures, and best practices. Through CoIN, Study Centers are sharing lessons learned from community outreach and engagement, recruitment, and efforts to convert women from the initial low-intensity group to the high-intensity group.

NCSAC Discussion and Recommendations

- Benjamin S. Wilfond, M.D., asked whether the impact of disclosing potential participants' geographical eligibility (that is, whether they live in a Study segment) on recruitment could be empirically studied. He also asked whether participants who move in or out of the high- or low-intensity segments would remain in their respective groups.
- Dr. McGovern said the impact of disclosing segment boundaries on recruitment could be studied. Because there are still concerns about privacy and identification of participants, the Study has adhered to laws and regulations to protect participants' privacy and has not disclosed segment boundaries.
- Dr. Hirschfeld noted that the Program Office is seeking input on relaxing the stringency that participants live strictly within segment boundaries. In one possible scenario, women who live outside of segment boundaries could be enrolled and blended with other Study participants. Primary analyses would be conducted only on women living in segments. The women would not know whether they resided in a segment and were part of the primary analyses. The stringency of data protection is an issue.
- Dr. McGovern said participants who move out of a segment but within a 50-mile radius will still be followed. Dr. Hirschfeld noted that although the Study intends to follow all participants, the current Study policy is informed by federal government travel rules, which impose a 50-mile threshold for reimbursement. Study participants who move within the 50-mile radius will continue to be followed by their Study Center, and participants who move to other Study locations will be followed by that Study Center. Although the Study plans to track the migrations of all participants, specific mechanisms are still under development.
- Maria Cancian, Ph.D., asked about the data implications of participants moving in or out of high- and low-intensity groups. She noted that families that are more stressed may be less likely to participate in intense data collection. Dr. McGovern said methods to address this issue have not yet been discussed. Dr. Hirschfeld said the Study would like to gather information on why people choose not to participate in the Study, but instruments to collect this type of data have not yet been developed.
- Edward J. Sondik, Ph.D., M.S.Hyg., asked Dr. McGovern to summarize the amount of time spent respectively in the high- and low-intensity data collections. Dr. McGovern said that the original Vanguard Centers spent several hours for the initial in-home data collection. The initial questionnaire in the Hi/Lo recruitment strategy takes about 30 minutes to complete, and the in-home visit takes about 60 minutes.
- Dr. Sondik asked how frequently data are collected, that is, the visit schedule. Dr. Hirschfeld explained that the Study's emphasis is on early data collection. Ideally, there will be two to three pregnancy visits. For the child, there will be visits at 2, 4, 6, 9, and 12 months and then every 6 months until 5 years of age. The target time for low-intensity in-home visits is 20–30

minutes and 45–90 minutes for high-intensity visits. The low-intensity approach will not have in-home visits.

Update on Provider-based Recruitment Strategy

Stephen L. Buka, Sc.D., Principal Investigator, Rhode Island Study Center, Brown University

The goal of the provider-based strategy is to recruit participants from within the initial Study segments by identifying births through provider offices. One of the challenges to this strategy is the diversity of the 10 Study locations (counties). Dr. Buka highlighted case scenarios from five Study locations: Schuylkill County, PA; Lamar County, TX; Providence County, RI; Durham County, NC; and Wayne County, MI. Across all locations, the responses from providers were generally positive.

The provider-based recruitment strategy offers considerable promise and can contribute to eventual “blended” site-specific strategies for the Main Study. This recruitment strategy requires extensive education and communication, intensive local preparation, maintenance, and “nourishment.” Providers have consistent concerns regarding integration and perceived conflict with clinical duties, patient confidentiality, and space and time constraints. Study Centers must develop creative strategies to address providers’ concerns. There are concerns regarding provider and staff fatigue over time. The strategy is probably not feasible for all providers in most Study locations, but there is value in selecting a probabilistic sample of provider offices, investing heavily in a smaller number of high-volume provider offices, and engaging and compensating providers (for time and office space).

The provider-based recruitment strategy can be successful in diverse settings for several reasons. There are a number of strategies to sample providers and enroll a representative sample of pregnant women. Providers can be added or dropped to refine the sample. The strategy can supplement household sampling. There is a high documented rate of enrollment per week. The strategy offers an optimal approach for identifying, consenting, and enrolling women early in pregnancy or before pregnancy. The strategy provides an excellent opportunity (with provider involvement) to engage participants for long-term retention and foster long-term clinical partnerships critical for the successful conduct of the Study.

NCSAC Discussion and Recommendations

- Dr. Henry asked about the factors contributing to successful retention so far. Dr. Buka said that one of the key factors is establishing a strong relationship between the Study and clinical providers and interacting positively with eligible women. However, data on retention are available from only two Study locations (Providence and Wayne Counties). In these locations, the Study investigators are well known to many of the providers.
- Dr. Henry asked about the differences in data for Providence and Wayne Counties regarding the number of eligible women and the percentage of women who consented. Dr. Buka explained that the data from the two locations were derived using different statistics and had different denominators.

Enhanced Household-based Recruitment Strategy

John Bancroft, M.D., Co-Principal Investigator, Maine Study Center, Maine Medical Center

The 10 enhanced household-based recruitment Study locations are widely dispersed across the country and have diverse populations. All Study Centers and subcontractors are compliant with the Federal Information Security Management Act (FISMA) and have authority to operate. The Study Centers have implemented 20 formative research projects.

The enhanced household-based recruitment strategy includes door-to-door enumeration of all households in sampled neighborhoods—as was done by the original seven Vanguard Centers—with pregnancy screening and consent performed in the home. However, there are variations among the 10 Study Centers in how enumeration is conducted. Enhancements include best practices of original Vanguard Centers; multisector, targeted community outreach; coordinated national and local media; and unique local elements. Media campaigns use radio and television advertisements, with some addressing specific groups (for example, Spanish-speaking audiences). “Cinema spots” are being shown in local movie theaters. Direct mailings of postcards are being used in some locations. Social media such as Facebook, Twitter, and YouTube are being used selectively within tight guidelines. The Study Centers provide incentives to local school districts or participants for enumerations completed.

A variety of community engagement activities are being used across Study locations. For example, within specific segments activities have included giveaway items and food. The Study Centers have engaged community advisory boards and the medical community (for example, birth hospitals and prenatal clinics). Some community events are targeting difficult to reach populations. Enhancements include educational outreach to ineligible community members.

The strengths of the enhanced household-based recruitment strategy are as follows:

- Broad multifaceted community outreach can facilitate household recruitment.
- Local adaptation enhances response.
- Presence of interviewers in community stimulates conversation.
- Enumeration can be accomplished despite elections, holidays, and winter weather.

There are several areas for improvement:

- The FISMA approval process
- Developing open-source information systems and databases within Study Centers
- Enumeration during inclement weather, particularly during the winter
- The frequency and variations in adding protocol elements (about every 90 days) for
 - Federated IRB tiers
 - Security plan revisions
 - Data system modifications
 - Birth hospital memoranda of understanding modifications
 - Staff training
- Transition from enumeration to pregnancy screening (The transition can be abrupt; pregnancy screening involves very personal questions that some people find intrusive when asked by a stranger.).

Discussion Championed by NCSAC Member

Dr. Henry' listed the highlights of the preceding discussion:

- Formalizing the Study Centers' collaborative approach and collaborative learning
- Refining the relationships with providers
- The impact on data of participants switching between high- and low-intensity groups
- Low yield of passive recruitment approaches
- Tracking participants who move out of Study locations
- Challenges to scaling up the number of providers
- Issues of the original birth rate estimates versus a changing or declining birth rate
- Educational outreach to people who are not eligible for the Study
- Cost-effectiveness of alternative recruitment strategies.

National Children's Study Sampling Strategy: Discussion of Alternate Models

Sampling Alternatives: History and Current Activity

L. Randy Curtin, Ph.D., Senior Statistician, National Center for Health Statistics (NCHS), CDC, HHS

The basic consideration for determining the Study's sampling design was using a probability sample versus a convenience sample. In 2002, contractors issued White Papers and reports on the household-, office/provider-, and center-based models and on hybrid options. The Study created a Sample Design Workgroup, and in March 2004, the Study held a Sampling Design Workshop to discuss sampling issues and provide a recommendation, which was a national probability sample using a household-based model. Although the NICHD recognized issues concerning initial response rate, attrition, costs, and operational feasibility, it decided to go forward with the national probability sample. The NICHD also recognized that the national probability sample offered a robust inferential design that could address the Study's multiple objectives, including recruitment of preconception women.

Based on a number of considerations, 110 primary sampling units (PSUs) were chosen. The result was an unbiased sample design with a self-weighting sample. The next phase was to create segments within the PSUs. Based on U.S. birth rates, it was estimated that 12,000 households would have to be screened to yield 1,000 births. To compensate for nonresponders, it was estimated that 16,000 households would have to be screened to yield 1,000 births. Within each PSU, 10–20 segments were selected using stratification criteria. Each segment would have about 1,200 households.

Features of the household-based sampling approach included a well-defined area frame. Every birth has a known probability of being selected to the sample. The demographic and geographic coverage can be controlled. Cluster sizes can be adjusted for cost and data linkage. It was estimated that up to 25 percent of the sample would be preconception women and 90 percent of women would enroll in the first trimester of pregnancy.

The early field results from the Vanguard Study revealed a number of problems with the household-based approach, including high rates of screening and listing, low enumeration and enrollment response rates, high costs, and complexity of field operations. As a result, the Study recognized the need to consider alternative recruitment strategies that would adhere to the overall concepts of data reliability, coverage of the data, cost estimates, and maintaining a nationally “representative” sample. Vanguard sites were surveyed for a preliminary assessment of their ability to conduct alternative recruitment strategies. Criteria were total cost, feasibility, representativeness, and sampling efficiency. Of the 30 new Vanguard sites, 10 were selected to implement an enhanced household-based strategy, 10 sites a provider-based strategy, and 10 a Hi/Lo strategy. The seven original Vanguard sites continued the household model.

The alternative recruitment sample design has potential challenges and analytic issues. A probability sample does not provide valid inference if it is not truly representative. If the sampling frame is not within PSU geographic clustering, it may be challenging to link geographic and contextual variables and collect environmental samples for exposures. There are some limitations in analytic methods. Some of the alternative recruitment strategies may work well in large metropolitan areas, whereas others will work well in small counties. There is a need for consistency in sampling and nonsampling error structure between PSUs.

NCSAC Discussion and Recommendations

- Steven K. Galson, M.D., M.P.H., asked whether there is a precedent for a sampling design that uses different sampling models in different geographical areas. Dr. Curtin noted that there are less expensive models than household surveys (for example, telephone surveys), but these other models also have issues. The recent trend in surveys is to use mixed modes. The Study is unique in several aspects, particularly its size and scope.
- Michelle A. Williams, Sc.D., S.M., M.S., asked for an example of a mixed-mode study. Dr. Curtin cited the National Health Interview Survey, which uses telephone interviews for nonresponders. Other examples are Internet and market-based surveys. The National Health and Nutrition Examination Survey is also a mixed-mode study.
- Dr. Williams commented that these examples are mixed modes of data collection, not sampling or recruitment strategies. Dr. Curtin explained that if the PSUs are selected as a probability sample and the samples within the PSUs are valid, mixed-mode data can be summed as a national sample. However, there will be issues of data comparability across PSUs.
- Patricia O’Campo, Ph.D., said one of the goals of the Study is to examine health disparities, and a representative sample may not be the best approach to support this type of research. Oversampling in certain racial/ethnic or high-risk groups is one way to examine health disparities. She asked whether the alternative recruitment strategies have approaches that will better support the Study’s ability to examine health disparities issues.
- Dr. Hirschfeld noted that examining health disparities is a critical aspect of the Study. The Children’s Health Act of 2000 cited only three criteria for the Study: that it (1) incorporate a

variety of assessments, (2) gather data on environmental influences on outcomes in diverse populations, and (3) consider health disparities.

- Jonas H. Ellenberg, Ph.D., commented that the probability sampling model is a step in studying a representative national sample of the U.S. population. The concept of a representative sample should be kept in mind as alternative recruitment strategies are considered. Dr. Ellenberg asked why the number of objectives (that is, one versus multiple) is important when considering a sampling scheme. Dr. Curtin explained that if there is a homogeneous risk across the population (that is, there are no extraneous factors), an exposure-outcome relationship can be measured in any group of people. If there are intervening factors—whether known or unknown—there needs to be a robust method for measuring the effects of these factors. A probability sample with random selection allows protection from extraneous variation mitigating results.
- Dr. Cancian said that having a probability/representative sample does not preclude oversampling groups of interest. Most probability studies in the social sciences have oversamples of relatively small populations. These samples, however, require weighting. Dr. Curtin said an extra screener could be added, but oversampling requires screening a larger pool of potential Study participants. Dr. Cancian noted that a number of surveys have developed methods for oversampling low-income and minority communities in ways that allow data to be analyzed separately as representative of those communities and incorporated nationally.
- Dr. Buka asked whether there was a particular group that recommended that the Study’s sample be self-weighting and when the decision was made. Dr. Curtin said the decision to use a self-weighting sample evolved from the various reports, workgroup discussions, and discussions within the NIH. However, the consensus recommendation was imparted that the Study does not require oversampling minority populations.
- Dr. Hirschfeld acknowledged Dr. Curtin’s decade-long contribution to the Study.
- Dr. Sondik commented that environmental factors were not explicitly considered in selecting the Study’s representative/probability sample. However, because the sample is household-based, the sampling design emphasizes factors related to household environments, and therefore, the Study population is representative of households. Other sampling designs could be used to analyze or simulate the impact of different environmental factors. Sampling designs could be adjusted to collect data on different environmental factors.
- Dr. Curtin explained that the purpose of the Study is not to conduct national exposure assessments and identify distributions of exposures but to analyze the impact of exposures on child health and development. Because of several issues, consensus recommendation was that the Study would not oversample by exposure type or demographics.
- James J. Quackenboss, M.S., noted that Study planners decided not to oversample by exposure type or demographics because oversampling targets known variables. The purpose of the Study is to examine the effects of broadly defined environment variables on child

health and development. Many important variables may not be known, and the relative importance of known and unknown variables cannot be determined a priori. Oversampling may affect the ability to appropriately assess unknown variables.

- Dr. Henry said the probability sample design allows the Study to be data driven, not hypothesis driven.

A New Recruitment Strategy for the National Children's Study

George Rhoads, M.D., M.P.H., Interim Dean and Professor, New York-Northern New Jersey NCS Center, University of Medicine and Dentistry of New Jersey, School of Public Health

The Study is the most expensive single study ever undertaken by the NIH. Its success will be judged by the extent to which it is able to identify and/or rule out potential prenatal and early childhood causes of health problems in children and (ultimately) adults. The Study has four key features that distinguish it from most prior cohort studies:

- Large sample size
- Collection of extensive information, biospecimens, and environmental samples so that hypotheses in many domains can be tested
- Intention to recruit early in pregnancy or even before pregnancy
- A nationally representative sample.

Recruitment by the original seven Vanguard Centers has been slow and expensive, with consent rates around 60 percent. Despite great effort, it appears that fewer than half of pregnancies occurring in segments are enrolled. The proportion of pregnancies captured may drop when intensive recruitment ends. So far, recruitment by the enhanced household-based strategy appears quite similar, with no improvement in enrollment rates.

The Hi/Lo (direct-to-the-public) recruitment strategy is similar to recruiting volunteers. Participation may be related to exposures and perceived vulnerability to disease. This combination of biases invites distortion of Study findings. Many volunteers may appear after the first trimester. The strategy is likely to weaken the analyses that appear in future research papers written from the Study.

The current provider-based recruitment strategy retains eligibility based on geographical segments. Compared with the other recruitment strategies, only 10 percent as many women would need to be contacted to achieve the same number of births. The consent rate for provider-based recruitment is about 80 percent. However, the requirement to retain segments means that (1) the Study Centers must work with nearly all providers in their Study locations, (2) many providers have few eligible patients, and (3) address screening is burdensome.

Dr. Rhoads proposed that the Study abandon geographic segments and use providers as the sampling frame. The proposed sampling frame would be as follows:

- An obstetric group or clinic would be considered a single provider.
- Providers would be stratified by race/ethnicity of in-county births and by size of practice (large, small, or clinic).

- Providers would be sorted into random order within strata, and the first one or two practices from each stratum would be chosen, for a total of six to eight providers. Providers that refused would be replaced by the next provider on the list.
- The Study would recruit the number of women required from each stratum to yield the proportion of 1,000 births that is equal to the proportion of all births in the county from that stratum.
- The quota for strata with two or more providers would be divided in proportion to the number of in-county births per provider.
- Recruitment would be paced to a desired rate from each provider by recruiting every *n*th new prenatal registrant.
- Race/ethnicity and educational status of practices could be easily monitored, so over- or undersampling could be imposed if necessary.
- A little staff time could be reimbursed to providers.
- This sampling frame would probably exclude large outside providers with fewer than 20 county births.
- Providers delivering in unique or uncooperative hospitals could be excluded.
- Providers may be able to identify women who are planning to get pregnant.

Advantages of the proposed sampling frame include:

- About 98 percent of births could be in the sampling frame.
- A little more than 80 percent of women could register in first trimester.
- The challenge of efficiently identifying pregnancies is solved.
- Providers are engaged; endorsement is implied.
- The 80 percent consent rate has already been demonstrated, so the sample probably would be more representative.
- It would be easier to arrange prenatal and birth biosamples.
- Study staff would not be stretched across a large number of providers.
- The number of birth hospitals would be reduced in large counties.

Dr. Rhoads concluded that this new recruitment strategy would enhance all four key features of the Study. By conserving resources, it would maintain the sample size, allow the collection of extensive information earlier in pregnancy than the household-based strategy, and probably be less expensive than the other sampling methods. Because of higher response and consent rates, the proposed strategy would probably yield a more representative sample than the household-based strategy.

NCSAC Discussion and Recommendations

- Dr. Ellenberg asked for clarification on why this proposed sampling frame would miss women before conception. Dr. Rhoads said that most women go to an obstetric group or clinic after they are pregnant. Very few will consult an obstetrician before getting pregnant; those who do are generally trying to become pregnant.

Discussion on Sampling the Population for the National Children's Study

Kathleen Belanger, Ph.D., Principal Investigator, Connecticut Study Center at Yale University, Yale University

The Study's purpose is to estimate population parameters and to investigate exposure-response associations to identify potential causal relationships. Probability sampling is necessary to estimate population parameters, and high response rates are necessary to prevent bias in the estimates. Response rates for survey designs (for example, the National Health Interview Survey) cannot be extrapolated to estimate response rates for recruitment for 20-year follow-up. In cohort designs, bias in estimating exposure-response relationships is related to loss to follow-up, not to sampling. However, the results of cohort studies can only be generalized to similar populations; thus, enrolling a representative sample is important.

Problems identified during recruitment in the original seven Vanguard Centers include (1) a large number of households were visited but the yield of current pregnancies identified was low, (2) follow-up of women of reproductive age resulted in small numbers of pregnancies, and (3) the sampling strategy was not cost-effective.

In the provider-based recruitment strategy, a large number of pregnant women can be easily identified, but only women living in the selected segments are eligible. In large counties, only 1 percent to 5 percent of women are eligible. This sampling strategy is not cost-effective. The provider-based recruitment strategy is specifically inefficient in urban, high-density counties. Two other methods could be used to align sampling design with recruitment—sampling practices rather than households (a probability-based sample) and sampling women interested in the Study (a representative sample, rather than a probability-based sample):

- **Sampling practices.** Sampling practices would determine the number of births in each practice in the past 12 months. Using birth certificate data, practices would be stratified by demographic variables and geographic location. The appropriate number of practices would be randomly selected from each stratum to represent the source population. Recruitment and enrollment would be tracked at each practice to ensure a probability-based sample. If a practice refused, it could be replaced by alternate practices in the same strata, or by extending enrollment in selected practices. This approach would focus only on selected practices from each stratum. Time, effort, and resources could be targeted to these practices. The sampling method would be more cost-efficient than household-based sampling or the currently designed provider-based sampling. Pregnant women may be much more willing to learn about the Study, provide information to determine eligibility, and potentially enroll, if they are approached in their prenatal care sites. This approach would preserve probability-based sampling.
- **Sampling women interested in the Study.** The current screening tool collects information about residence location, age, race/ethnicity, and education. This information is entered into the computer at the time of screening. For sampling women interested in the Study, a computer algorithm would determine eligibility by randomly selecting women as eligible, based on their geographic and demographic characteristics. The women interested in the Study would be recruited and enrolled. This approach would ensure that a representative sample could be drawn, distributed geographically across the county and be demographically similar to the births in the county. Sampling women interested in the Study would be more

cost-efficient than household sampling or the currently designed provider-based sampling. The approach preserves the goal of obtaining a representative sample, and it is much easier to advertise the Study to women throughout the county. The algorithm could be revised to select, as eligible, more women from any underrepresented group. This would be an effective strategy even if every prenatal care site did not participate.

Dr. Belanger concluded:

- The current sampling strategy is inefficient for household- and provider-based recruitment.
- A probability-based sample of providers is possible and more efficient.
- Any sampling design must consider retention in the cohort, as well as recruitment.
- A probability-based sample may still have differential loss to follow-up.
- A representative sample (not probability based) will not bias causal inference.

NCSAC Discussion and Recommendations

- Dr. Reede asked how women who do not receive early prenatal care will be recruited using the practice-based recruiting model. Demographic data may not be collected for this population, and issues of health disparities may not be able to be addressed. The Study's sample could become skewed. Dr. Belanger said about 1 percent of women in Connecticut do not receive prenatal care and about 2 percent wait until the third trimester before they seek prenatal care. Women who do not receive care at a provider's office could be recruited at delivery. The women who do not receive prenatal care are overrepresented by women who are drug users and therefore would be difficult to enroll into the Study under any circumstances. Some demographic characteristics could be inferred based on geographic location. Some demographic characteristics could be determined through the provider's patient population and birth certificate data.
- Dr. Rhoads explained that providers would be stratified and that one of the strata would be clinics. Women who seek prenatal care in clinics tend to be of lower socioeconomic status. These women would be included in the practice-based sampling frame.
- Virginia Delaney Black, M.D., M.P.H., of the Michigan Alliance for the National Children's Study noted that this Study Center has successfully identified eligible women at ultrasound visits and through hospital preregistration information. Based on state health department data, the Study Center knows the likelihood of where these women will deliver. Dr. Black explained that other cohort studies have developed mechanisms for recruiting and retaining women who are drug users. Study Centers can develop strategies to retain high-risk groups.
- Dr. Wilfond asked whether the Framingham Study was a convenience sample or a representative sample. Dr. Curtin said the Framingham Study recruited 100 percent of men and women of certain age groups who lived in Framingham, MA. This community-based study was not a sample but a census because all residents were invited to participate. The study also accepted volunteers from the Framingham area.

- Dr. Curtin explained that probability sampling is well defined. However, representative sampling is difficult to define, and there is a spectrum of interpretations on what it means to be “representative.”
- Dr. Wilfond asked whether the proposed provider-based sampling frame is a probability sample or a representative sample. Dr. Rhoads said his proposed sampling frame is not rigorously defined as a probability sample like the alternative provider-based recruitment strategy, but because the response rates could be better, the proposed sampling frame could function as a probability sample.
- Dr. Belanger said the Study can draw a sample of providers that is probability based.
- Dr. Cancian commented that a probability sample is used because differences in a population are not known and there is an equal probability of sampling all individuals in the population. However, a low response rate affects the equal probability of the sample. A random sample should be able to replicate a probability study if the characteristics of a population are known. Issues arise in sampling when many of the characteristics are not observable and a representative sample cannot explicitly be chosen. The concerns about using a provider-based sampling frame would be less if women are recruited at any time during their pregnancy and at birth. Sampling fractions could be adjusted to examine women who seek prenatal care in different trimesters.
- Dr. Ellenberg asked whether the hospitals that deliver babies in the Study location essentially become Study Centers. The participants would be recruited based on where they deliver. The sampling frame would not be a random sample of providers.
- Ana Diez-Roux, M.D., Ph.D., M.P.H., noted that the proposed provider-based sampling frame should be able to address two issues of concern: sufficient variability in key exposures of interest and minimizing the likelihood of biasing exposure-outcome associations. Although the sampling frame may not be able to recruit as many women in early pregnancy, the approach is reasonable.
- David L. Hubble, Ph.D., asked whether any previous study has used a provider-based sampling frame. One concern is how the approach would deal with women who go to multiple providers or could be recruited into a study in multiple ways. Dr. Belanger said the probability that a women would go to more than one provider for a single pregnancy is low. Providers include group practices and clinics and are not necessarily individual physicians.
- Dr. Rhoads commented that women who move out of segments versus those who do not may have different risks. He suggested that women who move out of segments should not be excluded from the Study.
- Dr. Sondik said the provider-based sampling would be an efficient recruiting approach because of the high probability of pregnant women receiving prenatal care. This approach may not necessarily overlook preconception women.

- Dr. O'Campo asked whether there are data on participating providers (for example, the extent to which they will comply with procedures and provide information). Dr. Belanger said such data from the alternative provider-based recruitment strategy are forthcoming. If a provider refuses to participate or underperforms, it would be replaced by another provider within the stratum. A provider's willingness to participate may be more a matter of how many patients are eligible (that is, a high proportion or a low proportion) and the relative amount of effort required to recruit a participant.
- Mr. Quackenboss explained that a body of evidence is accumulating on the household-based approach in terms of its advantages and limitations. The original goal of this approach was to be able to recruit women before conception and collect early exposure data. It is becoming clearer that alternative approaches can still have a probability sample but use fewer resources by not collecting early exposure data.
- Dr. Reede noted that women may receive prenatal care from providers who are not obstetricians/gynecologists. The provider-based sampling frame should consider its concept of what constitutes a provider. Dr. Belanger said birth certificates state who delivered the baby, which is a good indication of where a woman received her prenatal care. Birth certificates can provide some information about the woman, and where a woman receives prenatal care (for example, at a federally funded clinic) may provide additional information.

Discussion Championed by NCSAC Member

Dr. Wilfond

- Dr. Wilfond described the discussion as stimulating, provocative, and thoughtful. Although much work needs to be done, exploring new recruiting strategies is a starting point in asking important questions about their applicability across Study locations. Issues of health disparities are important and should be considered in the recruiting strategies. The issues and challenges of retention should not be overlooked.
- Dr. Hirschfeld commented that the Study is defining providers broadly. The concept of geographically based PSUs (that is, the 105 Study locations) will be retained, but the reference frame may be the women's address or the health care provider's address.
- Alan E. Guttmacher, M.D., said the discussions have been important, informative, and collaborative and will help him translate the discussions to various audiences.

Meeting Summary by NCSAC Member

Dr. Reede

Dr. Reede listed the meeting highlights as follows:

- The three criteria of the Study's legislative mandate should be revisited.
- Sampling frames and models should be able to address these criteria.
- The Study is data driven, not hypothesis driven.
- The Study's four key features provide the backdrop for assessing feasibility, acceptability, and cost.

- Harmonizing terminology is essential to effective Study communication.
- Implementation of the alternative recruitment strategies has facilitated collaborative learning among the Study Centers.
- The Hi/Lo recruitment strategy outreach and engagement metrics can be applied to other recruitment activities.
- Environmental factors such as stress may play a role in participants moving in or out of high- and low-intensity groups, which may affect data.
- Establishing working relationships with providers in the provider-based recruitment strategy is important.
- There may be selection issues in the proposed provider-based sampling frame.
- In the household-based recruitment strategy, field staff issues need to be better understood.
- Advantages for each alternative recruitment strategy should be identified, and hybrid options should be considered.
- Learning communities may provide opportunities for training.
- Educational outreach is important in engaging providers and participants.
- Issues of participants moving out of Study locations and how these participants will be tracked remain.
- Declining birth rates is an issue that needs to be considered.
- The Study's sampling frame should ensure that health disparities and diverse populations can be examined.
- Retention factors and issues need to be considered during recruitment.
- Difficult-to-reach populations need to be included in the Study.
- The Study needs to be as inclusive as possible.

NCSAC Members

*Wilma Brakefield-Caldwell, R.N., Public Health Nurse Administrator

Maria Cancian, Ph.D., University of Wisconsin-Madison

*José F. Cordero, M.D., M.P.H., University of Puerto Rico

Ana V. Diez-Roux, M.D., Ph.D., M.P.H., University of Michigan

Jonas H. Ellenberg, Ph.D., University of Pennsylvania Medical School

*Elena Fuentes-Afflick, M.D., M.P.H., University of California, San Francisco

Steven K. Galson, M.D., M.P.H., Amgen

*Bruce D. Gelb, M.D., Mount Sinai School of Medicine

Carol J. Henry, Ph.D., NCSAC Chair, George Washington University School of Public Health and Health Services

Jeffrey Krischer, Ph.D., University of South Florida

*Alma M. Kuby, M.A., M.B.A., Survey Methodologist

Patricia O'Campo, Ph.D., University of Toronto

Joan Y. Reede, M.D., M.P.H., M.B.A., Harvard Medical School

Everett Rhoades, M.D., University of Oklahoma Health Sciences Center

*Ellen Silbergeld, Ph.D., Johns Hopkins University Bloomberg School of Public Health

Benjamin S. Wilfond, M.D., University of Washington School of Medicine

Michelle A. Williams, Sc.D., S.M., M.S., University of Washington School of Public Health

**Did not participate*

Ex Officio Members

Allen Dearry, Ph.D., NIEHS, NIH, HHS

Alan E. Guttmacher, M.D., NICHD, NIH, HHS

Edward J. Sondik, Ph.D., M.S.Hyg., CDC, HHS

*Kevin Y. Teichman, Ph.D., Office of Research and Development (ORD), EPA

**Did not participate*

Designated Federal Official/Executive Secretary

Kate Winseck, M.S.W., NICHD, NIH, HHS

ICC Members

Amy Branum, M.S.P.H., NCHS, CDC, HHS

*Adolfo Correa, M.D., Ph.D., National Center on Birth Defects and Developmental Disabilities, CDC, HHS

Sally P. Darney, Ph.D., ORD, EPA

*Michael Firestone, Ph.D., Office of Children's Health Protection, EPA

Kimberly Gray, Ph.D., NIEHS, NIH, HHS

Steven Hirschfeld, M.D., Ph.D., NICHD, NIH, HHS

Danelle Lobdell, Ph.D., ORD, EPA

Mary E. Mortensen, M.D., M.S., National Center for Environmental Health, CDC, HHS

Sheila A. Newton, Ph.D. (chair), NIEHS, NIH, HHS

**James J. Quackenboss, M.S., ORD, EPA

Marshalyn Yeargin-Allsopp, M.D., National Center on Birth Defects and Developmental Disabilities, CDC, HHS

**Did not participate*

***Represented agency's Ex Officio member at this meeting*

Program Office Members

Marion J. Balsam, M.D., NICHD, NIH, HHS

Debowanna Blackshear, NICHD, NIH, HHS

Ruth A. Brenner, M.D., M.P.H., NICHD, NIH, HHS

Elizabeth L. Cho-Fertikh, Ph.D., NICHD, NIH, HHS

Danielle Cloutier (contractor), NICHD, NIH, HHS

Jessica E. Graber, Ph.D., NICHD, NIH, HHS

Brian J. Haugen, Ph.D., NICHD, NIH, HHS

Carl V. Hill, Ph.D., M.P.H., NICHD, NIH, HHS

JoEllen Jay (contractor), NICHD, NIH, HHS

Carol H. Kasten, M.D., NICHD, NIH, HHS

Nicole Kramer (contractor), NICHD, NIH, HHS

Scott Lewis, NICHD, NIH, HHS

Maria Lopez-Class, Ph.D., M.P.H., NICHD, NIH, HHS

John Lumpkin, M.S., NICHD, NIH, HHS

John Moye, Jr., M.D., NICHD, NIH, HHS
Nancy Parfitt Hondros, NICHD, NIH, HHS
Christina H. Park, Ph.D., NICHD, NIH, HHS
Nicole Pultar (contractor), NICHD, NIH, HHS
Gitanjali S. Taneja, Ph.D., NICHD, NIH, HHS
Maureen R. Wildman (contractor), NICHD, NIH, HHS
Nancy D. Wirth, NICHD, NIH, HHS

Observers and Other Participants

John Bancroft, M.D., Maine Medical Center
James Baumberger, American Academy of Pediatrics
Kathleen Belanger, Ph.D., Yale University School of Public Health
Art Bennett, B.E.E., M.E.A., Consultant, NICHD, NIH, HHS
Stephen L. Buka, Sc.D., Brown University
Richard A. Chestek, Ph.D., Booz Allen Hamilton Inc.
Marguerite Clarkson, Ph.D., PricewaterhouseCoopers LLP
Lester R. Curtin, Ph.D., NCHS, CDC
Angela DeBello, M.A., National Opinion Research Center (NORC) at the University of Chicago
Virginia Delaney-Black, M.D., M.P.H., Wayne State University
Hannah F. Elson, Ph.D., Fisher BioServices
Ned English, NORC at the University of Chicago
Triesta Fowler-Lee, M.D., NICHD, NIH, HHS
John K. Gohagan, Ph.D., Office of the Director, NIH, HHS
Jennifer E. Gorman, M.P.A., Office of the Director, NIH, HHS
David L. Hubble, Ph.D., Westat
Marina Kleinhapel, M.D., M.P.H., Michigan Department of Community Health
Juergen A. Klenk, Ph.D., Booz Allen Hamilton Inc.
Patricia McGovern, Ph.D., M.P.H., University of Minnesota
John McGrath, Ph.D., NICHD, NIH, HHS
Jin-Young K. Park, Ph.D., Office of Food Additive Safety, Food and Drug Administration, HHS
George Rhoads, M.D., M.P.H., University of Medicine and Dentistry of New Jersey
Susan Schechter, M.A., NORC at the University of Chicago
Branka Sekis, Social and Scientific Systems, Inc.
James P. Shannon, P.E., RTI International
James Ter Maat, Social and Scientific Systems, Inc.

I hereby certify that, to the best of my knowledge, the foregoing minutes are accurate and complete.



June 3, 2011

Date

Carol J. Henry, Ph.D.
Chair
National Children's Study Federal Advisory Committee