

**National Children's Study  
Federal Advisory Committee 29th Meeting  
July 20, 2011  
Natcher Conference Center, National Institutes of Health  
Bethesda, 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**

*Benjamin S. Wilfond, M.D., Acting Chair, National Children's Study Federal Advisory Committee (NCSAC), Director, Treuman Katz Center for Pediatric Bioethics, Professor and Head, Division of Bioethics, Department of Pediatrics, University of Washington School of Medicine, Seattle Children's*

Dr. Wilfond welcomed the meeting participants, who introduced themselves. Dr. Wilfond reviewed the highlights of the April 19, 2011, NCSAC meeting:

- Meeting summary/presentations posted to Study Web site
- Study update
- Principal investigator experiences in the Alternate Recruitment Substudy of the Vanguard Study
  - Hi/Lo recruitment strategy overview
  - Update on provider-based recruitment strategy
  - Enhanced household-based recruitment strategy
- Study sampling strategy: discussion of sampling alternatives—history and current activity
- A new recruitment strategy for the Study
- Discussion on sampling the population for the Study
- Meeting summary.

Dr. Wilfond outlined the advantages of the alternative recruitment strategy that George Rhoads, M.D., proposed at the April 19, 2011, NCSAC meeting:

- 98 percent of births are in the sampling frame
- 80+ percent of women register in the first trimester
- The problem of identifying pregnancies is solved
- Providers are engaged; endorsement is implied
- 80 percent consent rate already demonstrated, so sample probably more representative
- Easier to arrange prenatal and birth biosamples
- Staff not stretched across large number of providers
- Number of birth hospitals reduced in large counties
- Enhances all four key Study features
  - Large sample size
  - Collection of extensive information, including biological and environmental samples, so that hypotheses in many domains can be tested

- Easier to recruit early in pregnancy
- Yields a more representative sample
- Possibly less expensive than other sampling methods.

## **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)*

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 children’s 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. The Study is required to

- Incorporate behavioral, emotional, educational, and contextual consequences to enable a complete assessment of the physical, chemical, biological, and psychosocial environmental influences on children’s well-being
- Gather data on environmental influences and outcomes on diverse populations of children, which may include the consideration of prenatal exposures
- Consider health disparities among children, which may include the consideration of prenatal exposures.

Dr. Hirschfeld briefly reviewed the following topics:

- Study principles
- Exposure areas of interest
- Examples of outcome areas of interest
- Study structure
- Study activities
- Vanguard Study goals
- Alternate Recruitment Substudy
- Study recruitment as of June 2011
- Extent of Study coverage in household-based recruitment.

Dr. Hirschfeld noted the following interpretation of Study recruitment:

- **Household-based recruitment (going door to door with field workers).** About 10 percent of the women contacted are eligible, and of those, between 55 percent and 65 percent of them enroll in the Study.
- **Provider-based recruitment.** About 30 percent of the women are eligible (note that in provider-based recruitment, the women must reside in the preselected geographic segments), and about 85 percent of eligible women enroll. Perhaps the trusted environment improves the consent rate.
- **Direct to public.** About 35 percent of women who voluntarily contact the Study field office are eligible and essentially 100 percent of them enroll.
- **Speed of enrollment.** Household contact is the fastest followed by the direct-to-public approach.

Dr. Hirschfeld also reviewed the following topics:

- **Increased sample size.** Based on field data and reexamination of assumptions and projections, the Program Office proposes to increase the recruitment sample size and the number of locations for the Study. Recent calculations were conducted to estimate a sample size that would yield 100,000 participants after 21 years in order to have sufficient longitudinal data on exposures and conditions with a prevalence of less than 5 percent. Modeling using assumptions about attrition and data collection efficiency for a starting population of 100,000 children estimated a population remaining after 21 years of 39,000–45,000. Modeling of compliance with the projected Study visit schedule estimated that less than 10 percent of initial participants will have all data points. If the Study can retain 90 percent of enrolled women between the initial consent process and birth of the child into the Study (retention is currently 80 percent) and then retain the child cohort with 1–3 percent annual attrition, the Study should enroll 225,000–250,000 women. This larger sample size will provide opportunities to increase the number of locations where women are recruited through increasing the size of current locations and adding additional locations to include sampling units with at least 1,000 live births per year. The current strategy targets sampling units with about 250 live births per year.
- **Protocol development.** The Main Study protocol will emphasize early data collection during pregnancy and childhood because the largest knowledge gaps and potentially influential events occur during these periods. Visit frequency will diminish after 5 years. Recruitment phase is targeted as 2 years for any Study location. Multiple geographic locations can form a sampling unit.
- **Protocol development process.** Inputs for protocol development include Vanguard Study protocol data, the Study’s Scientific Plan and subsequent Institute of Medicine review, and comments from multiple advisors and consultants. A multidisciplinary team contributed to the draft concept document. The NCSAC and the Interagency Coordinating Committee will provide comments on the draft protocol concept document. The next steps are to complete the draft protocol document and submit it to the NIH Office of the Director subcommittee for review. The protocol concept document will be revised based on comments from multiple parties and submitted to the Office of Management and Budget (OMB) in the fourth quarter of calendar year 2011.
- **National Children’s Study Research Day.** The National Children’s Study Research Day will highlight Study scientific progress, invite future collaboration, and focus on scientific accomplishments and opportunities. The goal is to learn, collaborate, and innovate. All NIH Institutes and Centers, other HHS agencies, and other federal departments will be invited. Professional societies and advocacy groups will also be invited. The National Children’s Study Research Day will be open to the public. It will be held on August 24, 2011, at Natcher Conference Center.
- **International collaboration.** Study staff are communicating with multiple international partners, including the World Health Organization, the International Childhood Cancer Cohort Consortium, and several national studies in Europe and Japan regarding opportunities for collaboration across related studies. In this way, the impact of the Study will be broadened as it joins with international partners in the shared goal of improving the health of

all children wherever they may live.

- **Summary.** The Study will be able to collect prospectively data from a broad scope of measures beginning before pregnancy or during pregnancy through 21 years of age on a large sample of children throughout the United States. Ultimately, the Study will be one of the most robust research efforts geared toward studying children's health and development and will contribute to the formation of child health guidance, interventions, and policy for generations to come.

## **NCSAC Discussion and Recommendations**

- Ellen Silbergeld, Ph.D., asked for clarification of the percentage of estimated births in geographic segments in the household-based recruitment. Dr. Hirschfeld explained that the numbers of births in geographic segments over a given period are based on birth records. The percentage of estimated births is back calculated from the number of Study births in a segment over a similar period of time.
- In response to a question from Everett Rhoades, M.D., Dr. Hirschfeld said that births to women 17 years and younger are currently excluded from the Study.
- Steven K. Galson, M.D., M.P.H, asked whether the Study will provide cost estimates to OMB. Dr. Hirschfeld said that cost estimates will be included, but other information relevant to the protocol (for example, compliance with the Paper Reduction Act, the Privacy Act, and good scientific practice) will be submitted to the OMB's Office of Information and Regulatory Affairs.
- Dr. Silbergeld asked whether the NCSAC could review the reports on sample size calculations and attrition estimates that were prepared by the Program Office, Battelle, Booz Allen Hamilton Inc., Research Triangle Institute, Westat, and the National Center for Health Statistics (NCHS). She noted that the 1–3 percent annual attrition rate seemed low. Dr. Hirschfeld explained that the analysts considered (1) annual attrition rates from 1 percent to 10 percent based on reported and simulated data and (2) estimates of the number of data points that participants would complete over 21 years.

## **Provider-based Recruiting Opportunities and Challenges**

*Daniel Hale, M.D., Department of Pediatrics, University of Texas Health Science Center at San Antonio*

The provider-based alternative recruitment strategy is being implemented in 10 Study locations. Each location is typically focused on a subset of the residents predetermined by the home address of potential participants in a single county. The number of county births range from 670 to 25,688.

Two factors influence recruitment efficiency—the number of births per geographic area and the number of providers per geographic area. The counties with the lowest number of births have the highest percentage of births in Study geographic segments. The counties with the highest number

of births have the lowest percentage of births in Study geographic segments. The counties with the fewest providers (“few providers” counties) have the highest calculated segment births per practice. The counties with the most providers (“many providers” counties) have the lowest calculated segment births per practice.

Characteristics of the “few providers” counties are as follows:

- Relatively easy to engage all of the providers
- Easy to engage the hospitals (usually one or two)
- Uncomplicated practice structure
- Limited turnover of providers
- Few or no competing research projects
- Study investigators able to give all providers frequent attention and recognition
- Study staff able to get to know providers’ staff members
- Relatively straightforward to build and sustain trust due to frequent contact
- Relatively easy to have adequate Study staff in the providers’ offices at almost all times.

Characteristics of the “many providers” counties are as follows:

- Practitioner/practice prioritization
- Complex practice structures
- Refusal of a high volume practice to participate can have major “ripple effects”
- Takes considerable time to get to know all providers’ office and support staff
- Takes considerable effort to build trust, especially with the “low volume” offices
- Logistically difficult to adequately “staff” the offices
- Time consuming to engage all hospitals.

Other challenges and opportunities of provider-based recruiting include:

- Many offices/practices still lack an electronic medical records (EMR) system
- Many of the office staff do not know how to use the filter functions of their EMR system
- Because of the limited ability to incentivize provider staff for assistance, Study staff need to be creative to engage practices
- Limits placed on Study recruitment staff by local providers or staff, including time and space.

The common themes for provider-based recruiting are as follows:

- Personal connections are valuable
  - Find a Study champion
  - Find someone in each practice who “gets” the Study
- Every office and practice is different—assume nothing
- Most private practice offices have not previously participated in research activities
  - Be prepared to explain everything again and again
  - Offer suggestions of how things might work
- The more of the recruitment process that the Study team controls, the more effective the recruitment
  - Gaining trust is the key to gaining control.

## **Provider-based Sampling for the National Children's Study: Some Thoughts**

*Michael Elliott, Ph.D., Assistant Professor of Biostatistics, University of Michigan School of Public Health, Michigan Alliance for the National Children's Study (MANCS)*

The current household-based Study design uses a multi-stage area probability sample. From the United States' 3,141 counties, 105 Primary Sampling Units (PSUs) were selected based on probability proportional to size (PPS). Each PSU was divided into 10–15 segments of about equal size. It was estimated that there would be 250 births per year per PSU (16–25 per segment).

The proposed provider-based sampling design uses a first-stage area probability sample with the 105 previously selected PSUs and a second-stage physician practice sample. It was estimated that there would be 25 practices per PSU and 10 births per year per practice (by simple/systematic random sample).

The PPS sample can use birth certificate data to obtain a sampling frame of providers that has an associated size measure. This approach can be used to obtain a PPS sample of providers in the same way that a PPS sample of PSUs (counties) was obtained. Sampling a constant number of births from each provider yields an equal probability of selection sample of births from each PSU.

Several issues need to be considered with this approach:

- **Providers with less than  $n$  PSU births.** A provider can contribute no more births than he or she actually has. In some cases, the sample size will be too small. To obtain the required sample size in the PSU, the number of births can be increased where the number of births allows. Although the epsem design is lost, the representativeness of the sample can be reestablished using sampling weights.
- **Obtaining a subsample when births are greater than  $n$ .** More commonly, there are large providers with far more than  $n$  births in the PSU each year. Therefore, some type of random sampling scheme must be used, such as a sample of women at first visit to provider or a sample based on time (for example, sample 1 week per month, with week chosen randomly).
- **Refusals at the provider and mother level.** The previous two sampling approaches assumed perfect provider response. However, although the Michigan experience with provider-based recruiting has been very positive, response rates have not been perfect. Two options are available: (1) adjust the number of providers and mothers sampled and (2) replace refusing providers and mothers with another sampled provider/mother (PPS in the case of provider), but keep track of refusals so that response rates can be determined.
- **Preference to sample practices, not providers.** The PPS sample can be based on practices instead of providers. The majority of providers will be situated in a practice with other providers, which makes it far easier to work with a sample of practices, not providers. Unfortunately, these data are not contained in birth certificates. In small PSUs, a network of practices might be worked out directly, and provider measures of size can be summed to obtain practice measure of size. In large PSUs, the probability of selection of a given practice can be obtained after sampling the providers only with the names of the providers in the sampled practices. There is no need to work out the full network of providers.

Other issues for provider-based sampling are as follows:

- **“Movers.”** Women who switch practices after being sampled are exactly analogous to movers in the area probability sample. If moving is uncommon, movers can be oversampled. If moving is common, women can be dropped once they leave the sampled practice unless they move to another sampled practice. The Study can try to follow movers but adjust for increased probability of selection at both practices.
- **Stratification.** As in an area probability sample, stratification can be based on factors known for the entire population, such as age of mother, race of mother/child, and geographic region.

### **Discussion Championed by NCSAC Member**

*Michelle A. Williams, Sc.D., S.M., M.S., Co-Director, Center for Perinatal Studies, Swedish Medical Center, Program Director, Multidisciplinary International Research Training, Professor of Epidemiology and Global Health, University of Washington*

- Dr. Williams noted the heterogeneity among practices. For example, the Lamar County, TX, practices are likely to see one or two Study-eligible women per week, whereas the Bexar County, TX, practices are likely to see only two or three Study-eligible women per year. She asked whether (1) the maximum heterogeneity and relative cost-efficiencies among practices had been considered and (2) whether there may be locations where the provider-based practice recruitment approach may not be feasible in terms of trying to maintain a representative sample. Dr. Elliott replied that ideally, there would be about 25 providers per county. However, in some of the smaller PSUs, all of the providers would need to participate. In larger PSUs, the number of providers that participate could be limited to those that have the highest percentage of eligible women.
- Dr. Hale explained that there is a relationship between the number of births in a county and the number of providers in a county. For example, Lamar County, TX, has about 670 births per year and has four practices. Bexar County, TX, has about 25,690 per year and has 130 practices. However, among the 10 provider-based recruiting counties, the number of births per practice is fairly consistent (mean = 215; range = 132–360). Visiting practices on a regular basis is time-consuming, and it is not cost-effective to visit all practices, especially those that see few eligible women per year.
- Dr. Hirschfeld commented that the provider-based recruitment strategy is constricted by the current boundaries on Secondary Sampling Units. The strategy can be changed so that the primary eligibility criterion is where the provider is based. The question is then whether all women seen by the providers are eligible or whether only some women seen by the practice are eligible, depending on certain parameters. The goal is to reduce inefficiencies in screening and recruiting.
- Dr. Hale noted that among the 10 provider-based recruiting counties, there is variability in the number of calculated segment births per practice—84.5 for Lamar County, TX, and 2.7 for Bexar County, TX. Segments are distributed widely across counties.
- Ana V. Diez-Roux, M.D., Ph.D., M.P.H., asked about the challenges of implementing provider-based recruitment, that is, whether each Study location will have its own design,

whether the design will be centralized, or whether information will have to be collected to select the sampling weights. Dr. Elliott said household-based recruitment has the same challenges. Coordination between the Study Centers and Program Office will be critical in addressing these issues. There will need to be some reliance on the Study Centers' ability to implement provider-based recruitment and track appropriate information under the guidance of the Program Office.

- Dr. Silbergeld said the nature of the sample may introduce certain biases at the level of each practice in terms of relative contributions. She asked whether there is information on the demographics of practice-based recruitment versus household-based recruitment and whether the demographic differences can be described. Dr. Elliott said the provider-based recruitment samples have generally been representative at the county level and not different from the neighborhood samples of household-based recruitment. However, the provider-based approach does not include “neighborhood-centric” measures.
- Jennifer Madans, Ph.D., asked whether there are differences across providers as to women's stage of pregnancy when seeking prenatal care, as well as differences in the type of providers sought at different stages of pregnancy. Some women may not seek care until very close to birth, and the provider may be a hospital. There will be differences between these women and women who seek care early in pregnancy, and the information for the women will be different. Given this, the provider-based approach may yield a less representative sample. Dr. Elliott said that so far, the provider-based approach has been recruiting women early in pregnancy. For example, in Wayne County, MI, 60 percent of the participants had second trimester visits.
- Dr. Hale said one challenge with the provider-based strategy is identifying preconception women. Involving clinics that perform pregnancy tests and referrals can help screen and recruit women early in pregnancy. In Bexar County, TX, about half of the women are covered by Medicaid. Of these women, about 60 percent seek prenatal care in the first trimester and about 90 percent seek prenatal care by 20 weeks of gestation.
- Elena Fuentes-Afflick, M.D., M.P.H., asked whether there are differences in how individual providers in the same practice relate to the Study. Dr. Hale replied that there has not been “push back” from individual providers in practices where the head of the practice is enthusiastic about the Study. However, practice managers can be more influential than providers in the relationship with the Study.
- Dr. Williams asked whether current data from the 10 provider-based Study locations could be compared with simulations in order to understand optimal recruitment. Dr. Hale said that there are not enough data to determine whether demographics of babies born in the Study reflect the demographics of babies born in the counties. Because the Study Centers do not have the ability to share data, any data simulations would have to be conducted by the Program Office.

- Dr. Williams said one issue of provider-based recruitment is women who do not receive prenatal care. Dr. Elliott said that there are multiple approaches to identify such women. Dr. Hale said the Study Centers could be surveyed to determine the percentages in each county.
- Dr. Ellenberg asked whether there is a metric to evaluate which Study locations would use the provider-based approach versus the household-based approach. Dr. Hirschfeld explained that Study Centers may need a toolkit of recruitment approaches. However, at this time, the Program Office needs to know the performance characteristics and biases of the three alternative recruitment strategies and then understand settings in which a particular strategy might be most effective. Data will be used to guide recruitment strategy decisions.
- Dr. Ellenberg said that retention is a critical issue in evaluating the recruitment strategies. The way in which a woman is recruited (for example, by a trusted physician) into the Study may affect retention over the first 2–3 years.
- In response to a question from Bruce D. Gelb, M.D., Dr. Hale said there are not enough data to know how many Study participants will complete the protocol (that is, all Study visits). Dr. Madans said it will be important to determine completion rates for high-risk groups.
- Dr. Hale said that pediatricians are one of the keys to retention, and Study Centers have already started to engage them. Pediatricians often make their first contact with mothers in the hospital soon after birth.
- Dr. Hale commented that county population size and number of county births may be an issue if the recruitment period is shortened. For example, Lamar County, TX, has about 670 births per year and would not be able to enroll enough women in a 2-year period. In such a case, the PSU may have to be expanded into multiple counties.
- Dr. Fuentes-Afflick asked whether turnover in practice staff would affect recruitment in the provider-based approach because of the importance of personal relationships with this approach. Dr. Hale said that practice staff does change but not rapidly. Turnover may be one rationale for accelerating recruitment.
- Dr. Diez-Roux said the Study will have to develop guidelines for the selection and implementation of different recruitment approaches in different Study locations. Recruitment strategies may have to be tailored to each Study location. The Study may have to clearly explain why different approaches are being used. Other issues are the differences in data use and attrition. Dr. Hale said the Study Centers are already tailoring approaches. Because of the diversity of Study locations, the Study must have flexible recruitment approaches.
- Dr. Wilfond commented that retention may be linked to the relationship participants have with Study staff.
- Dr. Williams said that retention should be considered in stages. Prenatal retention is particularly important for comparing the alternative recruitment strategies.

- Carol J. Henry, Ph.D., asked whether one recruitment approach would be used in all Study locations in the Main Study or whether different approaches would be used. Dr. Hirschfeld explained that one approach will not be used in all Study locations. Based on data, the Study will evaluate the alternative recruitment strategies and determine which strategy would be most appropriate in different Study locations. Different strategies may be used in one Study location. Dr. Hirschfeld said that an adaptive approach may be necessary. He noted the difference between sampling frame and recruitment approaches. Different recruitment approaches may potentially be used in the same sampling frame.
- Dr. Silbergeld asked whether the Study could evaluate cohort demographic characteristics from stage to stage. Knowing such characteristics is essential to understanding disparities among Study populations. Knowing whether characteristics change from stage to stage may be important in evaluating the alternative recruitment strategies. Dr. Hirschfeld said that there are not enough data yet from the alternative recruitment strategies, but data from the Vanguard Study show that Study participants reflect the county demographics. Dr. Silbergeld said the Study will need data such as income, education, and housing type to more accurately assess disparities.
- John Bancroft, M.D., noted that the person who provided prenatal care may not be the person who delivers the baby. This difference may vary not only among PSUs but within PSUs. Some obstetric hospitals deliver babies but do not participate in prenatal care. So births of providers may be misattributed in the sampling model. Dr. Hale said in university and military hospitals births may be attributed to one or only a few physicians.
- In response to a question from Laura E. Caulfield, Ph.D., Dr. Elliott said women recruited through providers have to live in a PSU but not the geographic segment. Dr. Hale commented that it is much easier to identify and recruit women who live a PSU than in a segment.
- Dr. Hirschfeld clarified that there are two provider-based sampling frames: (1) Eligible women must live in the PSU and receive care only from providers in the PSU, and (2) women are eligible regardless of residence as long as they receive care from providers in the PSU.
- Dr. Silbergeld said it is important to retain the Study's original geographically based sampling frame in order to correlate environmental/exposure data with health outcomes.
- Allen Dearry, Ph.D., said neighborhood characteristics, such as social and built environments, are important to understanding health disparities.
- Dr. Madans said the Study will be able to collect geographic information through home visits, and existing environmental data can be linked with participants' geocodes. Knowing what information is available and what information is missing from each of the alternative recruitment strategies will be important. Another issue is whether the Study is trying to make national estimates or site-specific estimates. Having different sampling schemes will be challenging for the statistical analyses.

**Summary.** Dr. Williams summarized the discussion topics as follows:

- The decision-making process for determining the most appropriate recruitment strategy
- Metrics for evaluating recruitment strategies
- Costs and efficiencies of recruitment
- Representativeness of the sample
- Retention up to delivery
- Logistics.

The Study's design and approach to implementation have to be driven by research hypotheses. If the hypotheses are environmentally and community based, then location is important. If the hypotheses involve individual toxic levels and body burden, then location is less important. Once hypotheses are determined, the Study would be optimally designed and recruitment and retention strategies could be tailored to address the hypotheses.

### **Draft Concept of the National Children's Main Study**

*Ruth Brenner, M.D., M.P.H., Associate Director for Science and Protocol Development (Study Visit Measures), National Children's Study, NICHD, NIH, HHS*

The concepts of the Main Study include:

- Assessments from before birth through 21 years of age
- Concentration of visits during (1) the prenatal period and infancy and (2) periods of rapid development and vulnerability, including periods with potentially important knowledge gaps
- Examination of determinants and impact of health disparities
- Use of data and sample repositories to serve as a resource for future studies.

At a Main Study sampling retreat held on May 27, 2011, and during weekly conversations from June 3 through July 1, 2011, the following topics were discussed:

- Physician-based sampling
- Main Study sample—realistic expectations of recruitment rates and attrition
- Additional PSUs
- Alternative SSUs
- Recruitment period.

As a result, a new general sample size goal was established to enroll and retain a sufficient number of women such that there are 100,000 children enrolled in the Study after 21 years. This goal is a departure from previously stated goal of 100,000 children initially enrolled in Study. The justification for the new goal is that for relatively rare outcomes, a larger sample size is needed to detect modest effect sizes, particularly for infrequent exposures.

Assuming an optimistic 2–3 percent yearly attrition rate, about 150,000–200,000 infants would need to be enrolled to have a sample of 100,000 at 21 years. The Vanguard Study experience showed that 80 percent of women enrolled during pregnancy were retained through birth of the child. Given this retention rate, up to 250,000 pregnant women would need to be enrolled to yield a sample of 100,000 at 21 years. Sample size can be increased in several ways, for

example, by improving recruitment rates, increasing the number of PSUs, and increasing the number or size of SSUs.

Inclusion criteria for Study eligibility are as follows:

- Reside in Study segment (for the geographic approach)
- Pregnant adult women
- Pregnant emancipated minors—emancipated minors per laws of jurisdiction
- Pregnant minors—14 years or older, not emancipated, with permission for enrollment from legally authorized representative
- Nonpregnant adult women—49 years or younger with a high probability of pregnancy
- Children born to enrolled women
- Fathers of enrolled children as identified by enrolled women
- New adult caregivers or adult guardians of enrolled children.

Exclusion criteria for Study eligibility are as follows:

- Women who are surgically, medically, or genetically infertile and not responsive to interventions
- Women who are unable to understand Study participation and grant informed consent.

The following data collection schedule has been proposed:

- Pregnancy—three in-person data collections
- Birth—one or two data collections at place of delivery
- 2 months and 4 months—remote contact (for example, by phone)
- 6 months—in-person visit
- 9 months—remote contact
- 12 months—in-person visit
- Data collection every 6 months through age 5 years—in-person visit on the year, remote contact on the half year.

Types of data collection may include:

- Questionnaires and diaries
- Select medical record abstractions
- Observations
- Physical measures
- Photographs and videos
- Biologic specimens
- Environmental samples
- Revisiting a matrix sampling approach.

In the initial seven Vanguard Centers, women with a high probability of becoming pregnant received an in-person data collection, which included questionnaires, environmental samples, biologic specimens, and physical measures. The purpose was to collect data before, but close to, conception and facilitate data capture earlier in pregnancy. From January 2010 to September 2010, 150 women had completed a preconception visit and 3–9 months of follow-up. For women in the high probability of pregnancy cohort who had at least 3 months of follow-up after the preconception data collection, about 14 percent became pregnant. For those who became

pregnant in the high-probability preconception cohort, the first trimester data collection occurred about 30 days earlier compared with women who were not in the high-probability preconception cohort. In summary, although a relatively small percentage of women enrolled into the high probability of pregnancy cohort became pregnant within 3 months of the preconception data collection, those who did become pregnant completed the first trimester visit earlier in pregnancy than did the rest of the cohort.

### **Discussion Championed by NCSAC Member**

*Bruce D. Gelb, M.D., Gogel Family Professor of Child Health and Development, Professor of Pediatrics, Professor of Genetics and Genomic Sciences, Director, Child Health and Development Institute, Mount Sinai School of Medicine*

- Dr. Gelb listed three issues for discussion:
  - The increased sample size (250,000 women enrolled instead of 100,000)
  - The recruitment approaches to achieve the increased sample size
  - The flexible approach to recruitment.
- Dr. Wilfond commented that the sampling approach has not been determined, that is, whether it will be provider-based or population-based.
- Dr. Gelb asked for clarification on what a realistic attrition rate would be. He also asked how front-loading data collection would affect later attrition. Dr. Hirschfeld explained that attrition rate estimates from similar longitudinal cohort birth studies in the United States and other countries were examined. Cultural factors and health care delivery factors in these studies were considered. Program Office staff thought that a 5 percent annual attrition rate might be too high and that 1 percent might be too optimistic. A more realistic estimate for the Study would be about 3 percent. Data from other large birth cohort show that most attrition occurs in the first several months to years, which could be up to 20 months. After this period, attrition rates tend to plateau. Some studies have been able to recapture participants that have dropped out, and as a result, there may be some increases from year to year. However, there are then missing data points for these participants. The Program Office will continue to gather attrition data from the Vanguard Study to determine whether the 3 percent attrition rate is realistic. The Study will gather data to project demographic trends. Data will also be gathered on participants who move, and ways to keep these participants in the Study will be explored.
- Jeffrey Krischer, Ph.D., noted that the TEDDY Study—a prospective birth cohort study of the causes of type 1 diabetes mellitus—has enrolled more than 8,000 babies, who will be followed for 15 years. Data from this study show that the attrition rate is high in the early phase but then tapers off quickly. The peak attrition rate is around 18–24 months. The attrition rate over the first 6 years is approaching 20 percent, which is annual rate of about 3 percent.
- Dr. Fuentes-Afflick said it may be necessary to have some worst-case planning, given the diversity of the Study’s participants. Because there may be different retention rates for different populations, oversampling of certain populations may be needed.

- Dr. Ellenberg said that the original Study had considered assessments of rare outcomes such as schizophrenia. It was understood that after 21 years, the sample size would be less than 100,000, which would probably not be a sufficient number to assess rare outcomes. Dr. Ellenberg asked what the rationale was to increase the sample size in order to yield a sample of 100,000 after 21 years. Because of the greater costs of the increased sample size, a shorter study could be considered. Dr. Hirschfeld said the initial assumptions about recruitment rates, birth rates, and efficiencies that were used to construct the original sample size and the estimated retention rates were not supported by data. Study data have revealed that other assumptions are not as robust as needed. It was determined that the original sample size would not yield a sufficient number of participants after 21 years to answer proposed Study questions. The increased sample size of 250,000 would yield a sufficient number for data analyses of informative events and a sufficient data set to answer previous questions, new questions, and future questions not yet contemplated; in addition, the larger sample is more likely to address Study mandates.
- Dr. Hirschfeld further noted that based on the original Study assumptions about attrition, it was estimated that after 21 years, there would remain 70,000–75,000 participants. Current data do not support this estimate.
- Dr. Galson asked whether there are examples of the original and new attrition rate calculations. Lester R. Curtin, Ph.D., described the assumptions and methodologies for determining the Study’s sample size. Based on standard equations, it was determined that analysis of rare outcomes with 2 percent prevalence would require a sample size of 100,000. He noted attrition rates have been continuously discussed since planning Study design began. Many of the assumptions were debatable and considered optimistic. Vanguard Study data have shown that the assumptions were very optimistic. Because of this, the sample size has been reevaluated. The Study’s 4-year enrollment period has also been reevaluated. The enrollment period might be too long for a number of reasons. PSUs would have to be added to enroll 250,000 women in a shorter enrollment period (for example, 2 years). Historical data of attrition rates from other studies have been examined to determine realistic assumptions and calculate an initial sample size that would yield 100,000 participants after 21 years.
- Juergen A. Klenk, Ph.D., explained that compliance with completing Study visits was another key issue in calculating the increased sample size. The goal is to have as many complete data sets for children after 21 years as possible in order to analyze rare outcomes.
- Dr. Ellenberg commented that there are higher costs with the increased sample size. Analysis of rare outcomes may have to be sacrificed in order to keep the Study affordable. Dr. Hirschfeld responded that the Study is for the moment only considering what the best scientific design and outcomes should be. Cost issues are not an initial part of the discussion of determining the best science, but will have to follow.
- Dr. Silbergeld said the Study needs to define its goals. It is difficult to design a study without having goals. The Study’s original goals were not to study the origins and risk factors for rare

outcomes but to understand the trajectory of normal development and deviations from normal development. Certain outcomes may be difficult to detect with the Study's design. Dr. Silbergeld said the Study's goals appear to be moving. Dr. Hirschfeld explained that there are no data to inform normal development. The Study design will have to be able to address what is not known about exposures and outcomes. The Study design is being geared toward answering questions about a range of outcomes with a range of prevalence. The Study needs to be sufficiently powered to answer questions about rare outcomes, both known and unknown, and should be able to answer questions about disparities.

- Dr. Wilfond noted there are three issues to consider: sample size as it relates to the incidence of disease, differential attrition, and maintaining the Study's representativeness if the sample size is increased.
- Dr. Diez-Roux asked whether increasing the sample size is justified. There are trade-offs involved with the increased sample size, and the implications need to be considered irrespective of cost. A sample size of 100,000 can be used to study a number of effects with different strengths, both large and small. Repeated measures over time can provide sufficient power for a variety of analyses.
- Patricia O'Campo, Ph.D., noted the importance of keeping a representative sample. Increasing the number of PSUs provides an opportunity to rethink issues about disparities. The types of disparities and populations should be identified and discussed. Whether a representative sample is the best design to answer questions about disparities should also be discussed. Dr. Hirschfeld said the increased sample size allows a reexamination of the types of analyses that can be done. The Study's domains all focus on children's health, but powering the Study for certain conditions may limit analyses or introduce biases. A probability sample can help reduce biases.
- Dr. Curtin commented that minority groups and disadvantaged populations can be oversampled without moving away from a probability design as long as sampling weights are used in the analyses.
- Dr. Fuentes-Afflick cautioned against causing participant fatigue by coinciding Study visits with pediatric visits. Because Study visits are time-consuming, participants may push back if Study visits occur too close to pediatric visits.
- Dr. Wilfond said the issue of sampling frame (that is, population sampling versus provider-based sampling) needs to be addressed before increasing the number of PSUs can be discussed.
- Dr. Gelb asked whether there are data on the percentage of providers that are outside PSUs in the provider-based approach and whether there are any potential biases in using providers outside PSUs. Dr. Hirschfeld replied that there are no data at this time.
- Graham Kalton, Ph.D., distinguished the two provider-based approaches. In the first, only births in a PSU would be eligible for Study, regardless of provider location. In the second,

only providers in a PSU would be considered, and all the births for those providers would be eligible, regardless of where a mother resides. Both approaches will yield a probability sample. The challenge is comparing and analyzing the two approaches when used in different PSUs.

- Dr. Madans noted that there needs to be sufficient information, particularly from a study's early phases, to account for biases in probability samples. She also noted that even with a large representative sample, there still may not be sufficient data to allow analyses of a broad range of outcomes, particularly those with low prevalence.
- Dr. Williams asked whether there is an explanation for the low percentage of pregnancy (14 percent) among the women in the high probability of pregnancy cohort. Dr. Brenner said the percentage was lower than expected and cannot be explained with current data. Additional follow-up data will be analyzed to see whether the pattern continues. She noted that the 14 percent is based on data from early recruitment.

### **Report of the NCSAC Data Presentation Working Group**

*Jonas H. Ellenberg, Ph.D., Professor of Biostatistics, Department of Biostatistics and Epidemiology, Associate Dean of Research Program Development, University of Pennsylvania School of Medicine*

The NCSAC Data Presentation Working Group is composed of Dr. Ellenberg; Dr. Henry; Joan Y. Reede, M.D., M.P.H., M.B.A.; Dr. Silbergeld; and Dr. Williams. The working group was formed to provide advice on the table shells for presenting Vanguard data. The working group met three times to discuss the data shells. The working group's recommendations were given to the Program Office and are reflected in the following presentation.

### **Table Shells for Presenting Vanguard Study Data from the Alternate Recruitment Substudy**

*Brian Haugen, Ph.D., Senior Scientist (Analysis and Evaluation), National Children's Study, NICHD, NIH, HHS*

Dr. Haugen explained that, as recommended by the Data Presentation Working Group, a detailed glossary has been developed for table headings, terms, and categories (for example, outcome codes, response codes, and disposition codes). The glossary will be included with all data presentations.

Dr. Haugen reviewed six table shells:

- Overall summary
- Pregnancy screening completion rates
- Consent rate by recruitment strategy
- Pregnancy characteristics of enrollees
- Demographic characteristics of enrollees (race and ethnicity only)
- Additional measures of recruitment and retention.

The following comments about the tables and data were made:

- The same label should be used when referring to the same number in multiple tables.
- Line 4 in Table 1 should be “Women eligible for consent” (not “Women asked for consent”).
- Study policy on rounding numbers to the nearest 50 follows federal standards and guidelines for statistical surveys.
- In Table 3, line c, “Ineligible – pregnancy loss” refers to women who were eligible because they were pregnant but became ineligible because of pregnancy loss before consent.
- The tables should present more explicit calculations (for example, statistics, ratios, and percentages) to show changes/trends from stage to stage and overall success rates.
- In order to be geographically eligible, women must reside in a segment (SSU) at the time of consent.
- Line b in Table 4 (“Pregnancy < 14 weeks gestational age at consent”) is included to understand operational reasons for why women are not consented early in pregnancy.
- Table 5 uses race and ethnicity for demographic characteristics of enrollees. The table shell is an example for presenting demographic characteristics.
- Pre- and postconsent demographic data—including race/ethnicity, age, household income, and education level—are being collected from the Alternate Recruitment Substudy and will be presented at future NCSAC meetings.
- Demographic data of fathers were collected by the original seven Vanguard Centers.
- Demographic data of fathers and babies in the Alternate Recruitment Substudy have not yet been collected.
- Race/ethnicity data collection instruments allow for multiple responses and record the order of responses.
- Collecting data on immigration status has been discussed. The Program Office welcomes input on this topic.

The NCSAC made the following recommendations for data tables:

- Tables should be provided on paper at meetings.
- Tables provided before meetings could be in Word or Excel format.
- Tables should provide more detailed information.
- Each table should be interpretable by itself.
- Tables should include the numbers used to determine percentages/proportions.
- Row headings could be abbreviated or simplified and explained in detail in the glossary.
- Glossary definitions should clearly denote nominators and denominators.

Dr. Haugen reviewed two figures showing cumulative enrollment by weeks since fieldwork began and cumulative births by weeks since fieldwork began for the alternate recruitment strategies.

### **Meeting Summary by NCSAC Member**

*Ana V. Diez-Roux, M.D., Ph.D., M.P.H., Professor of Epidemiology, Director, Center for Social Epidemiology and Population Health, University of Michigan*

Dr. Diez-Roux summarized the discussion topics and key issues:

- Reconsidering the Study design
  - Sampling and recruitment
  - Sample size

- Evaluating different sampling and recruitment approaches
- Establishing evaluation criteria including
  - Cost
  - Representativeness
  - Logistics/feasibility
  - Potential yield of follow-up and retention
- Establishing criteria for assessing follow-up and retention
- The need to be explicit about the evaluation and decision-making criteria
- The possibility that different recruitment and sampling approaches may be used at different Study locations
- Establishing criteria for selecting different recruitment and sampling approaches
- Understanding the implications of using different approaches at different Study locations
- Whether the provider-based approach will be geographically based and, if not, the effect on geographic data
- Increasing the sample size based on projected attrition rates
- The need to scientifically justify the increased sample size
- Maximizing follow-up and retention
- Scheduling visits
- Data presentation
- Providing data before NCSAC meetings.

### **Next Steps**

*Steven Hirschfeld, M.D., Ph.D., Acting Director, National Children's Study, NICHD, NIH, HHS*

Dr. Hirschfeld listed the following next steps:

- Continue work on the draft protocol document
- Submit the draft protocol for review
- Revise the draft protocol document based on reviewer comments
- Present the status of the draft protocol document at the October 19, 2011, NCSAC meeting
- Hold the Study Research Day on August 24, 2011 in Bethesda, MD.

### **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  
\*\*Jennifer Madans, Ph.D., NCHS, CDC, HHS  
\*Edward J. Sondik, Ph.D., M.S.Hyg., NCHS, CDC, HHS  
Kevin Y. Teichman, Ph.D., Office of Research and Development (ORD), EPA  
*\*Did not participate*  
*\*\*Represented agency's Ex Officio member at this meeting*

### **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*

### **Program Office Members**

Debowanna Blackshear, NICHD, NIH, HHS  
Ruth A. Brenner, M.D., M.P.H., NICHD, NIH, HHS  
Andrew M. Briggs (contractor), NICHD, NIH, HHS  
Danielle Cloutier (contractor), NICHD, NIH, HHS  
Michael J. Dellarco, Dr.P.H., NICHD, NIH, HHS  
Jessica E. Graber, Ph.D., NICHD, NIH, HHS

Jihane Guidy (contractor), NICHD, NIH, HHS  
Brian J. Haugen, Ph.D., NICHD, NIH, HHS  
Carol H. Kasten, M.D., NICHD, NIH, HHS  
Nicole Kramer (contractor), NICHD, NIH, HHS  
Eric Lorenzo, Ph.D., NICHD, NIH, HHS  
John Lumpkin, M.S., M.B.A., NICHD, NIH, HHS  
John Moye, Jr., M.D., NICHD, NIH, HHS  
Nancy Parfitt Hondros, NICHD, NIH, HHS  
Christina H. Park, Ph.D., MHS, NICHD, NIH, HHS  
Jennifer E. Park, Ph.D., NICHD, NIH, HHS  
Nicole Pultar (contractor), NICHD, NIH, HHS  
Julia Slutsman, Ph.D., NICHD, NIH, HHS  
Gitanjali S. Taneja, Ph.D., NICHD, NIH, HHS

### **Observers and Other Participants**

Anjene M. Addington, Ph.D., M.P.H., Booz Allen Hamilton Inc.  
Abdul H. Ally, Fisher BioServices  
Rena Atayeva, NICHD, NIH, HHS  
John Bancroft, M.D., Maine Study Center, Maine Medical Center  
Kevin Beverly, Social and Scientific Systems, Inc.  
Laura E. Caulfield, Ph.D., Johns Hopkins University Bloomberg School of Public Health  
Richard A. Chestek, Ph.D., Booz Allen Hamilton Inc.  
Marguerite Clarkson, Ph.D., PricewaterhouseCoopers LLP  
Lester R. Curtin, Ph.D., NCHS, CDC, HHS  
Deborah W. Davis, Ph.D., University of Louisville Study Center, University of Louisville  
Rebecca A. Davison, American Academy of Pediatrics  
Elaine Eaker, Sc.D., Westat  
Michael R. Elliott, Ph.D., M.S., MANCS, University of Michigan School of Public Health  
Hannah F. Elson, Ph.D., Fisher BioServices  
Ned English, M.Sc., National Opinion Research Center (NORC) at the University of Chicago  
Mischka Garel, M.P.H., Johns Hopkins University Study Center, Johns Hopkins University  
Bloomberg School of Public Health,  
David A. Gordon, M.S.E.E., Booz Allen Hamilton Inc.  
Daniel E. Hale, M.D., University of Texas Health Science Center at San Antonio, University of  
Texas Health Science Study Center  
David L. Hubble, Westat  
JoEllen Jay, Booz Allen Hamilton Inc.  
Lisa Kaeser, J.D., NICHD, NIH, HHS  
Graham Kalton, Ph.D., Westat  
Michael Keller, Ph.D., Booz Allen Hamilton Inc.  
Juergen A. Klenk, Ph.D., Booz Allen Hamilton Inc.  
Colleen Lee, M.S., National Cancer Institute, NIH, HHS  
Yvonne T. Maddox, Ph.D., NICHD, NIH, HHS  
Stephen Marcus, Ph.D., National Institute of General Medical Sciences, NIH, HHS  
Ryan Morrell, Booz Allen Hamilton Inc.

Barbara O'Brien, M.T., M.P.H., Westat  
Kristina Rother, M.D., M.H.Sc., National Institute of Diabetes and Digestive and Kidney  
Diseases, NIH, HHS  
Mona Rowe, M.C.P., NICHD, NIH, HHS  
Rajni Samavedam, M.P.H., Booz Allen Hamilton Inc.  
Terry Savage, Westat  
Susan Schechter, M.A., NORC at the University of Chicago  
Daniela Smith, Booz Allen Hamilton Inc.  
Megan Stacy, NICHD, NIH, HHS  
Andrew Westdorp, Social and Scientific Systems, Inc.  
Tisha Wiley, Ph.D., Office of the Director, NIH, HHS  
Wei-Wen Wu, Ph.D., R.N., National Taipei University of Nursing and Health Sciences

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



October 11, 2011

\_\_\_\_\_  
Date

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