The information contained in this appendix represents a draft report produced by Battelle on behalf of EPA's National Exposure Research Lab under Task Order 19 of Contract 68-D-99-011.

White Paper on Evaluation of Sampling Design Options for the National Children's Study

Appendix C

Development of Exposure Assessment Study Design for the National Children's Study: Project Overview, Results, and Recommendations

The information contained in this appendix represents a draft report produced by Battelle on behalf of EPA's National Exposure Research Lab under Task Order 19 of Contract 68-D-99-011.

The information contained in this appendix represents a draft report produced by Battelle on behalf of EPA's National Exposure Research Lab under Task Order 19 of Contract 68-D-99-011.

September 30, 2003

Development of Exposure Assessment Study Design for the National Children's Study

Project Overview, Results, and Recommendations

Task 5

Task Order 19
Contract No. 68-D-99-011

Prepared by

Warren Strauss, Jeff Lehman, Michele Morara
Battelle
505 King Avenue
Columbus, OH 43201

Louise Ryan
Harvard School of Public Health - Department of Biostatistics
655 Huntington Avenue
Boston, MA 02115

for

Thomas McCurdy
Task Order Project Officer

Haluk Ozkaynak
Technical Lead

and

Ellen Streib
Project Officer
National Exposure Research Laboratory
Office of Research and Development

The information contained in this appendix represents a draft report produced by Battelle on behalf of EPA's National Exposure Research Lab under Task Order 19 of Contract 68-D-99-011.

DISCLAIMER

Appendix C contains a draft report prepared by Battelle as part of Task Order 19 under Contract 68-D-99-011 for the National Exposure Research Laboratory at the U.S. Environmental Protection Agency (EPA). This draft report has not yet been released by the EPA for public dissemination, and is currently under peer review prior to being released as an EPA report. Therefore, the information contained in this Appendix should not be cited or quoted.

The information contained in this appendix represents a draft report produced by Battelle on behalf of EPA's National Exposure Research Lab under Task Order 19 of Contract 68-D-99-011.

Development of Exposure Assessment Study Design Options for the NCS Project Overview, Results, and Recommendations

C-1.0 INTRODUCTION

This project was performed to develop innovative statistical study design guidance for the acquisition of exposure data over time in a longitudinal study of children who participate in the National Children's Study (NCS), sponsored by a consortium of Federal Agencies[1]. The NCS was authorized by the Children's Health Act of 2000[2], and is intended to investigate environmental influences on children's health and development–including understanding any environmental exposures that may cause or exacerbate health impacts. Due to the large sample size and longitudinal nature of the Study, unique statistical issues arise that must be addressed before a cost-effective sampling design can be developed to gather the environmental and personal exposure data needed in the study. A very important issue that must be addressed in this study is that of obtaining enough samples to provide adequate statistical power to detect health effects attributable to environmental and personal exposures with a minimum amount of burden, while being cost-effective and staying within the study's overall budget. The specific objective of this project was to develop cost-effective statistical sampling strategies and optimal design considerations for the NCS. Relevant specific issues include identification of potential sources of bias and/or uncertainty in the environmental/personal estimates of exposure (in particular, non-responses, subject over-burden, subject drop-out rate, and measurement error); and strategies to address these issues, including sample weighting techniques, and replicate sampling to assess measurement error variance.

The project began with a comprehensive literature review of statistical design and analysis topics relevant to the NCS exposure assessment design[3]. Based on the results of this literature search, research on the development of optimal design strategies and advanced statistical analysis tools was pursued on the following topics:

- ***Optimal Design Considerations for Selecting Subsets of Study Participants for Investigation of Relationship between Exposure and Health Outcome[4]:*** Given the large sample size of the planned NCS cohort and the high costs and burden associated with environmental sampling, it will not be possible to collect detailed longitudinal exposure information across the cohort and at all time periods to support the myriad of hypotheses that relate environmental exposure to potential adverse health outcomes. However, well designed sub-studies can be carried out within the NCS cohort (using only a small fraction of the sample size) to address most hypotheses, with sufficient power to characterize the relationship between exposure and health outcome. Potential methods for selecting a sub-sample include stratified sampling (e.g., specific sampling of high-

exposure and/or high-risk individuals), multistage sampling, two-stage case-control studies and other outcome dependent designs, case cohort designs, and nested case-control studies. Research was conducted to provide guidance on optimal design for selecting an appropriate and efficient sub-sample of the NCS cohort for exploring exposure-related hypotheses as a function of several factors, including the availability of demographic and exposure-related questionnaire information and perhaps ambient monitoring data across the entire cohort, the availability of archived environmental or biological samples, the prevalence of the health outcome of interest, the strength of the relationship between exposure and health outcome, and variability in exposure. Appropriate statistical methods for calculating selection weights and tools for analysis of the resulting data corresponding to each type of design were also developed as part of this research.

- ***Optimal Design for Studies Involving Multiple Methods for Characterizing Exposure, with Differing Properties Related to Measurement Error and Bias***[5]***:*** During the implementation of the NCS, it is likely that low-cost, low-burden methods, such as the acquisition of questionnaire information, will be employed across the entire cohort – with smaller subsets of respondents undergoing more extensive environmental exposure assessment using more expensive and more accurate environmental and exposure measurements. Based on the expected properties (cost, burden, bias, error) of the different measurement methods available, this research provides guidance into optimal designs for characterizing the relationship between environmental exposure and health outcomes using a mixture of available measurement methods across the cohort. This work was motivated by comparisons of different methods used to characterize physical activity (questionnaire and accelerometer) of adolescents participating in the Planet Health Study. Based on cost and burden constraints associated with exposure assessment in the NCS, this work provides a theoretical basis for the efficiency gained by employing more expensive and more accurate assessment tools on a small subset of study participants. By employing these more costly and accurate measures across a small validation subset of study participants, we are able to characterize and account for the precision of the less accurate and less expensive assessment tools, in a sense making them more valuable as an informative tool.

- ***Latent Variable Models for Exposure Measurement Error in Environmental Studies***[6]***:*** This research provides flexible statistical analysis tools that correctly assess the relationship between health outcome and environmental exposure in the presence of measurement error. The statistical theory that serves as a basis for these analysis tools was motivated by a sub-study of the Home Allergens and Asthma Cohort Study, which investigated in vivo allergen induced cytokine production as a function of endotoxin measured in household dust. Although the study includes multiple measures of endotoxin in each household, it is assumed that true endotoxin exposure for the subject

child can never be observed, which is likely a realistic assumption if only indoor residential measurements are utilized since these will not include contributions from other environments. A latent variable approach is therefore used to model cytokine production of the child as a function of the observed, yet error-prone assay results. The latent variable approach analytical tools that were developed as part of this research are easily implemented using existing software tools in SAS or S-Plus, and provide efficient and unbiased estimates that are robust to many of the technical challenges anticipated in the NCS, including repeated measures, attrition, missing covariate information, missing visits/observations, and measurement error in both exposure and health outcome.

- ***Optimal Design for Studies Involving Replicate Exposure Measurements Subject to Short and Long Term Variation***[7]***:*** This research provides optimal design considerations regarding how many replicate samples are required to adequately characterize measurement error in the longitudinal environmental exposure measures, and correct for the measurement error bias in models relating exposure to health outcome. This work was motivated by a study of the effects of pesticide exposure on male reproductive health outcomes among a cohort of 750 subjects. In this study, pesticide exposure was assessed using assays of pesticide metabolites in urine, which are costly. Optimal design considerations were developed to determine the number of study subjects requiring replicate sampling to provide an unbiased characterization between the exposure and health outcome, under varying levels of person-to-person variability and within-person short- and long-term variability in pesticide exposure measurements. Cost considerations and expected respondent burden can also be factored into the design, resulting in usable design guidance for the identification of an appropriate sub-sample of respondents from the cohort that will require replicate environmental exposure sampling.

While the results of each of the above statistical design and analysis investigations are highly relevant to the acquisition of exposure information as part of the NCS, they must be integrated with other statistical and exposure assessment concepts to yield appropriate guidance for the NCS design. However, based on the results of this research it quickly became evident that optimal design considerations will be uniquely dependent on the characteristics of the particular research hypothesis of interest[8]. Therefore, instead of attempting to determine global design concepts across all potential hypotheses, we investigated optimal design considerations for the following small sample of example research objectives that may be of interest for the NCS[9] (we envision that these design considerations will be relevant to other research objectives):

1. ***Determination of the relationship between pre-natal exposure to pesticides and current exposure to lead on children's IQ at age three:*** This research objective focuses on a continuous health outcome with two competing sources of exposure that are thought to influence the outcome. It is assumed that current lead exposure can be measured with minimal error, while pre-natal exposure to pesticides – based on assays of archived urine

samples collected during pregnancy – is error prone due to the time-varying nature of pesticide exposure and the limited number of urine assays that would be available for each subject.

1. ***Determination of the relationship between pre-natal exposure to pesticides and miscarriage:*** This research objective focuses on the same error-prone measure of pesticide exposure and a binary health outcome with reasonably high prevalence (approximately 15 percent of women planning pregnancy recruited into the NCS would be expected to experience miscarriage).

2. ***Determination of the relationship between pre-natal exposure to pesticide and autism:*** This research objective focuses on the same error-prone measure of pesticide exposure and a binary health outcome with very low prevalence (approximately 0.34 percent of children (34 per 10,000) recruited into the NCS would be expected to be autistic).

3. ***Characterization of gene-environmental interaction between exposure to particulate matter during the first two years of life and genetic variability in members of the Glutathione S-Transferase (GST) gene family, and the development of childhood asthma later in life (by age 6):*** This research objective focuses on a binary health outcome with reasonably high prevalence (approximately 10 percent of children are expected to develop asthma). It is assumed that the genetic tests are costly, and that there are multiple methods for characterizing exposure to particulate matter. For example low-cost and low-burden methods could be used, such as use of questionnaires to quantify exposure (environmental tobacco smoke, fireplace usage, etc.) and physical activity, and reliance on ambient air monitoring data to quantify particulate matter concentrations. These low-cost methods could be reasonably applied to a large cross section of the NCS cohort in a retrospective manner – however, they are likely not to be sufficient for accurately characterizing true exposure. Alternatively, higher cost methods could be employed such as personal air monitors for quantification of particulate matter concentrations and accelerometers for describing physical activity. However, these sampling methods would require identification of the study subjects prior to the development of asthma which may offset any efficiencies gained in using more precise measurement methods.

4. ***Characterization of the combined effects of exposure to ozone and particulate matter and physical activity on wheeze among asthmatic children:*** This research objective combines many of the exposure assessment issues covered in the previous research example, applied to a repeated measure binary response (wheeze).

A flexible simulation tool was created to allow detailed exploration of the above five research objectives. Efforts were made to develop realistic exposure and health outcome

simulated data, based on various parameters available in the scientific literature. Example parameters that served as the basis for the simulations include information regarding the prevalence of the health outcome, distribution of biomarkers (where available) among different subpopulations of interest, measures of chemical concentration in environmental media from aggregate exposure studies, and summaries of exposure-related behavior. Parameters were identified to characterize both average levels of exposure, as well as appropriate trends and sources of variability, such as temporal variability (note that there are many limitations in the available data for characterizing the sources of variability, especially temporal variability in short-term average concentrations). The simulation tool allowed for study subjects to be assigned to different exposure/outcome profiles, in which each different profile would have unique parameters. For example, exposure profiles for lead and pesticides might be different for children living in an urban household with moderate pesticide use compared to children living in a suburban household with high pesticide use. Similarly, particulate matter exposure might differ based on the presence or absence of adult smokers or wood-burning fireplaces. Additionally, the simulation tool allows for the introduction of any number of exposure/outcome profiles, along with specification of the proportion of study subjects in the total NCS cohort that are expected from each profile. Based on the input parameters, a time series of true exposure and health outcome(s) was generated for all study subjects involved in the NCS. Personal and temporal variability was introduced to the simulated time-series data generated among the NCS cohort respondents to provide realistic data to serve as a basis for exploration of different candidate study designs to address each specific research objective.

Based on the unique characteristics of each research objective, candidate sampling design options were developed, refined, and applied to the simulation data. Each candidate design option specified methods to identify study subjects, the number of study subjects participating in the research sub-study, and the manner in which the study subjects would be sampled (frequency, method, duration). Each candidate design option was then applied to the simulated data under a variety of constraints and technical challenges as follows:

- ***Strength of the Relationship between Exposure and Health Outcome:*** For the continuous outcome of IQ, separate simulation models were constructed so that the pesticide and lead exposures accounted for low (5.0%), moderate (7.5%), and high (10%) proportions of the variability in the IQ response. Similarly for the research objectives that focused on binary health outcomes, separate simulation models were constructed so that the corresponding odds ratio that characterizes health outcome as a function of high versus low levels of exposure would range from very weak (odds ratio of 1.1), weak (1.25), moderate (1.5), and strong (2). By varying the strength of the relationship between exposure and health outcome, the power and efficiency of different candidate design options can be compared based on differences in sample size, number of replicate samples per subject, sampling methodology, etc.

- ***Measurement Error in the Exposure and Health Outcome Measures:*** Although in practice the measurement error properties associated with the sampling methodology may be assumed to be known, investigation of different levels of measurement error associated with both exposure and health outcome measurement methods can provide useful insight into the gains and losses in power and efficiency based on the use of different measurement methods with different measurement error properties.

- ***Non-response and Attrition:*** While the candidate design options that were investigated have designed "missingness" – where exposure and/or health outcome are not necessarily measured on every NCS respondent – there is also unplanned missing data that will occur over the course of the study. Two different types of unplanned missing data were introduced into the simulated research outcomes: wave non-response and attrition. Wave non-response occurs when a study subject has missing data for one or more planned sampling events, but remains in the study. There are several potential reasons for wave non-response, including respondent unavailability, difficulty in obtaining a sample, and difficulty in analyzing the sample. Attrition corresponds to when a respondent stops participating in the study, which can be caused by several factors including the burden of the study, when a respondent moves away from the study area, and potential human subjects issues. Since both wave non-response and attrition are anticipated as part of the NCS, the simulated case studies were investigated at varying levels of these two types of unplanned missingness to help determine which sampling designs and corresponding statistical analysis tools are most robust to this anticipated problem.

For each combination of case-study, design option, and technical challenges, approximately 500 NCS simulated outcomes were generated and statistically evaluated in an effort to assess the power, efficiency, and bias associated with the ability to characterize the true relationship between exposure and health outcome. The resulting data from these simulations were analyzed using various statistical methods to help quantify their performance with respect to some of the technical challenges. These methods included different ways of accounting for the selection weights associated with the design, and methods for accounting for measurement error and within subject variability in exposure measurements. Specific results from the statistical evaluation of these design options are provided in this report. A comprehensive review of these results[10] provides the following general guidance for the design of exposure-related research as part of the NCS:

C-2.0 OVER-ARCHING DESIGN FOR LINKING EXPOSURE TO ADVERSE HEALTH EFFECTS

Due to the fact that no single primary hypothesis has been identified for this study, and the fact that there are a multitude of hypotheses that could be formulated for exposure-related research, our recommended design approach centers on hierarchical methods of sampling from the NCS cohort. At its core, we assume that the NCS will collect information related to major health outcomes of interest and limited (e.g., questionnaire-based) risk factor information on all study subjects over time. However, many of the core NCS hypotheses will require much more detailed assessment, which could be addressed using subsampling and other design elements discussed earlier. For each hypothesis relating health outcome to exposure, one approach to this subsampling would entail combining information from two different samples drawn from the NCS cohort as depicted in Figure 1:

- The first sample would be optimally designed to characterize the relationship between the health outcome of interest and some direct or indirect measure of exposure. Example measures of exposure include, but are not limited to, exposure biomarkers, and measures of exposure-related behavior such as activity patterns, diet or consumer product use. The selected measure of exposure may be assessed either prospectively or retrospectively, depending on the design. For example, biomarker concentrations of chemicals may be assessed prospectively from biological samples as they are collected from respondents, or in some cases may be archived for future chemical analysis under a retrospective design. Of course there is the complex issue of when to collect exposure measurements relative to intermittent types of exposure. The optimal statistical design for this first sample would be unique to the specific hypothesis of interest, based on several factors including prevalence of the health outcome, the availability of appropriate measures of exposure at the appropriate time-period(s) of vulnerability for the child (which is typically unknown), and the expected relationship between the health outcome and the measure of exposure. Much of the statistical work conducted as part of this project focused on the selection of appropriate methods, both design and analytical, for characterizing the relationship between health outcome and measures of exposure in this first sample. In particular, much of this work focused on direct measures of biomarker concentrations as an indirect measure of exposure. Note that in general biomarker concentrations are used to estimate a subjects absorbed dose; however, relating biomarker concentrations to external exposure may require considerable additional information about the frequency, duration, and routes of exposure.

- The second sample would be designed to characterize the relationship between the measure(s) of exposure likely to be explored in the first sample (e.g., biomarker concentrations) and measures of aggregate exposure from a representative sub-sample drawn from the total NCS cohort (shown as the box on Figure 1). Aggregate exposure

measures would entail the collection of samples from all potential routes of exposure (air, soil, dust, water, food) as well as specific measures of exposure-related behavior. This sample would be common to all exposure-related hypotheses of interest in the study, and must be carefully designed to provide the linkages between biomarker concentrations and their most likely sources in the environment. Admittedly, this is likely very difficult and careful consideration of appropriate methods for designing and implementing this second sample is fundamental to its success. The specific design of this second sample would include cross-sectional components to capture person-to-person variability across the NCS cohort, as well as longitudinal components to provide estimates of within-person variability in exposure over time. Additionally, this sample should be designed to

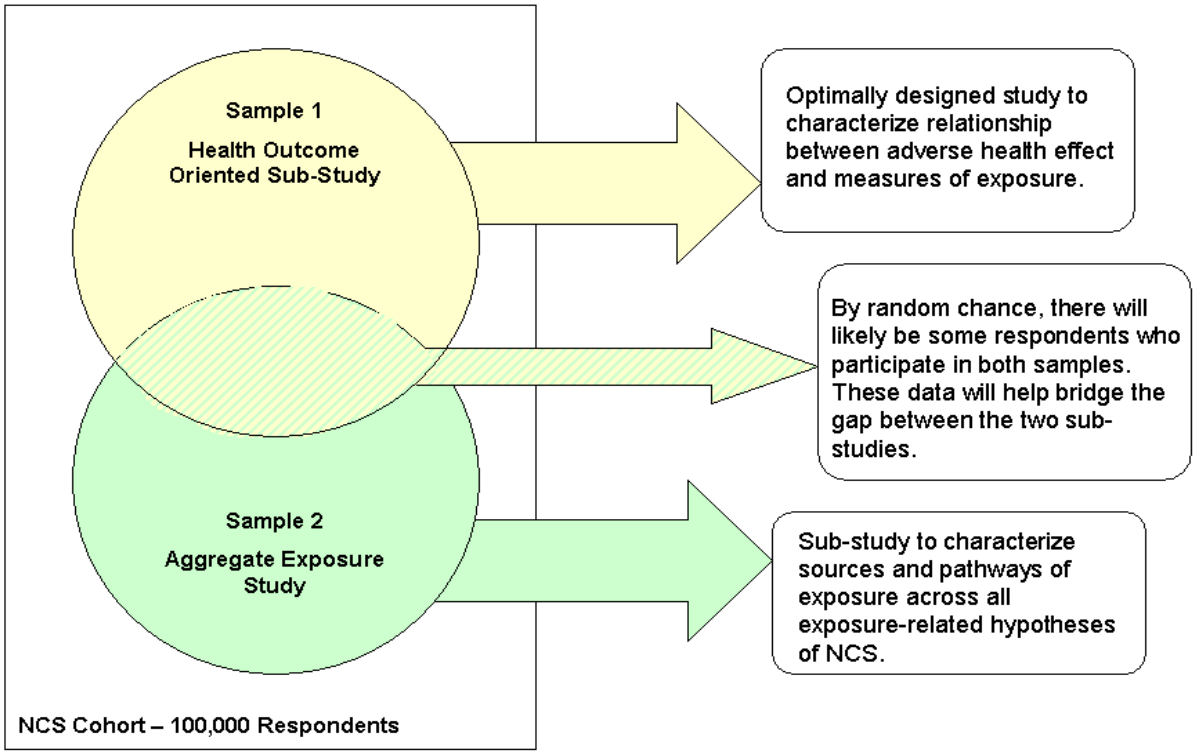

Figure C-1. Overview of the Design for Exploring the Relationship between Adverse Health Effects and Exposure

provide statistically valid estimates of the relationship between biomarker concentrations and pathways of exposure for all stages of vulnerability and development during the study, from pre-conception through early adulthood.

- The general hierarchical design for relating health outcomes to exposure described in the previous section provides for enormous flexibility for the pursuit of health outcome and exposure-related research as part of the NCS. Although the design and

implementation of these two samples will likely occur independently over the course of the NCS, the overarching hierarchy of this design for sampling from within the NCS cohort, the application of some reasonably simple design principles, and the availability of statistical analysis tools for appropriately combining information will ensure that NCS researchers will be able to gain insight into the exposure sources and pathways that contribute to adverse health effects with sufficient power and precision, in a resource efficient manner. The following sections provide an overview of the major design and analytical principals associated with each sample, as well as the combination of information across both samples.

The information contained in this appendix represents a draft report produced by Battelle on behalf of EPA's National Exposure Research Lab under Task Order 19 of Contract 68-D-99-011.

C-2.1 SUB-STUDIES RELATING ADVERSE HEALTH OUTCOMES TO MEASURES OF EXPOSURE

There are several key results from this study that are critical to providing optimal design options for characterizing the relationship between an adverse health outcome and measures of exposure. The first, and perhaps most encouraging result, is that there are a variety of statistical designs and associated analytical tools that are available for characterizing these relationships based on relatively small sub-samples from within the NCS cohort. In most cases, these carefully designed sub-samples provide adequate power and precision for characterizing the relationship between health outcomes and measures of exposure using sample sizes ranging between 300 and 1000 respondents – with exceptions typically occurring when the prevalence of the health outcome is very low and the relationship between the health outcome and the measure of exposure is very weak.

A second important result of this study is that there are statistical analysis tools that can correct for the problem of measurement error in the measures of exposure, with very little loss of efficiency or power related to this technical challenge. This result appears to be consistent across all of the potential designs for relating health outcomes to measures of exposure, and has direct positive implications for the design of more cost-efficient research sub-studies. From one perspective, it allows for the use of a single measure of in-body chemical concentration(s) for most respondents in a sub-study, even when there is an anticipated high degree of within-person temporal variation anticipated in these measures. Again, this is an important finding and could have major implications for the NCS in terms of the number of exposure measurements that are necessary in each of the life stages of interest. From a different perspective, it allows for the use of less precise measures of exposure from low-cost questionnaires and diaries of exposure-related behavior over time as the basis for these sub-studies. An important driver, that is required to achieve this result, is the availability of information from within the sample to help characterize the measurement error or uncertainty in the measure(s) of exposure. For example, with the measures of in-body chemical concentration, this requirement could result in a design in which replicate samples are collected from a small validation sample of respondents from the sub-study to help estimate the appropriate within-person variability. Similarly, with respect to the use of low-cost, less precise questionnaire information to serve as measures of exposure across the majority of respondents in the sub-study, this might require the additional use of more expensive and accurate measures from a small validation sample to help characterize the measurement error. Design guidance is also available for the optimal design of these validation samples to help characterize and correct for measurement error in these studies.

Closely linked to the above results on compensating for measurement error in the exposure measures is a result that suggests that for health outcomes that do not alter over time once properly diagnosed (e.g. miscarriage or autism), it is of higher value to measure exposure

and health outcome on a larger number of study subjects – even with error prone measures of exposure – than to measure exposure with high precision on a smaller number of study subjects. Again, this result is dependent on having some information in the sample to properly characterize measurement error in exposure, so that it can be properly accounted for in the statistical analysis phase, and it identifies the competing objectives that are apparent in the NCS paradigm: accurate exposure characterization and powerful assessment of the relationship between exposure and an adverse health outcome.

A third result is that reasonable levels of attrition and wave non-response, ranging from 10 to 30 percent, appear to have minimal effect on the resulting power and efficiency of the sub-study samples. Additionally, higher degrees of attrition and/or wave non-response can be factored into the design of these research sub-studies if they are anticipated due to high expected burden or other factors.

The above three results have practical implications for the collection of exposure-related measurements across the NCS cohort. Since some of the efficient design options for linking health outcomes to exposure metrics are retrospective in nature (outcome dependent), it will be critical to collect exposure-related information from all study subjects in a consistent manner, with a sampling plan that provides coverage across life-stages. Having exposure-related information available for all study subjects at different stages of development for the subject child will also be critical to support health outcome-oriented research in which the biological cause of disease is not well understood and the disease is rare. Using a retrospective study design from within the NCS cohort, researchers could explore the totality of exposure-related information to determine which measures of exposure are most closely associated with disease. The advantage of the NCS cohort for supporting this type of research is that the exposure-related information would be collected from all study subjects in a prospective manner, and therefore would not be prone to the same types of bias and measurement error that many retrospective studies contend with when attempting to recreate exposure information from respondent recall and/or temporally mismatched biological and environmental sampling.

The collection and archiving of biological specimens (blood and/or urine) could serve as a foundation for some, yet not all, exposure-related research. For example, while archived bio-specimens can be used for retrospective investigation of heavy metals, some (persistent)pesticides, other persistent pollutant chemicals, and potential genetic links to disease, they cannot be used for most volatile organic and semi-volatile organic (VOC and SVOC) chemical species, non-persistent pesticides, VOCs, allergens, most air pollutants, tobacco smoke, or biological pathogens. To provide coverage across exposures that cannot be assessed retrospectively using archived bio-specimens, the NCS will likely need to employ the prospective collection of less-precise exposure-related information. This could include using questionnaires to capture exposure-related behavior information on activity, diet, and consumer product use; abstraction of medical records and/or diaries during pregnancy to capture fever and

exposure to biological pathogens; and reliance on independent data sources such as ambient air monitoring data obtained from EPA's Aerometric Information Retrieval System (AIRS). To emphasize the utility of both biomarkers of exposure as well as questionnaire information on exposure-related behavior, consider studies that focus on adverse health effects from pesticide exposure. Some health effects might be related to average pesticide exposure, in which case a biomarker concentration (e.g., a urinary pesticide metabolite) might be an appropriate measure of exposure for use in these studies. Alternatively, if an adverse health effect is related to an acute pesticide exposure event, questionnaire information regarding consumer product use and other exposure-related behavior might be better suited to capture these events.

After careful review of all potential NCS hypotheses, the design for routine collection of biological specimens and other exposure-related information across the NCS cohort will need to be carefully planned to provide coverage for all stages of vulnerability in a child's development. In addition to routine data collection that applies to all NCS participants, a series of validation sub-studies should be planned to capture information regarding the uncertainty and variability in exposure information that results from this routine data collection. These validation studies would be conducted on very small, yet representative sub-sets of the NCS cohort, and may include additional replicate sampling for biological specimens to capture temporal variability in biomarker chemical concentrations; concurrent analysis of a subset of bio-specimens to measure VOCs, SVOCs and biological pathogens to characterize measurement error in questionnaire and other methods used to act as surrogates for these types of exposures; and higher technology methods to capture exposure-related behavior (e.g., GPS, accelerometer, or heart-rate monitor to capture physical activity) with a higher degree of precision.

Finally, we would like to note that these results were developed in accordance with a small set of simulated examples that are based on a variety of simplifying assumptions[10] so that a manageable number of factors could be investigated. As noted previously, the optimal sampling design can be highly dependent upon the details of the scenario of interest (in this case simulated scenarios), and, thus, the design considerations identified here are dependent upon the assumptions that were made in developing the simulated scenarios. In defense of the simplifying approaches, we emphasize that this task was to develop general design guidelines and considerations that are appropriate for a study such as the NCS. Certainly, from an intuitive standpoint, the design guidelines identified above would be appropriate for many types of exposure and adverse health outcome scenarios (as well as other types of studies); however, it is important to note that this claim is not rigorously justifiable. In other words, these results should be considered to offer a starting point for efficient design in exposure assessment, and illustrate the need for further research in this area.

C-2.2 DESIGN FOR A GLOBAL AGGREGATE EXPOSURE STUDY

The previous section provides guidance on the collection of surrogate exposure information that can be used to help identify chemicals that are closely associated with adverse health effects. The specific pathways and sources of exposure for many of the chemicals that will be identified in these health outcome-oriented studies are not currently well understood. Furthermore, the sources and pathways of some chemicals that have been characterized in the past may be subject to temporal or spatial variations that limit the application of these past relationships to this study. In order to translate the results of the NCS health outcome research into practical public health solutions, it will be important to first identify and then eliminate the most likely sources and pathways of exposure for chemicals that are shown to have serious adverse health consequences. In this section, we provide some limited design guidance for a representative sub-study of the NCS which characterizes aggregate exposure across all relevant chemicals. This study can be thought of as an extensive validation sub-study of the NCS that provides detailed information on the routes and pathways of exposure for all exposure-related research conducted as part of the NCS. The following are some important design elements for this study:

- The aggregate exposure sub-study must be designed to provide coverage for all potential routes and sources of exposure. This would include sampling air, food, water, dust, and soil for the selected study subject and their appropriate micro-environments, detailed characterization of exposure-related behavior, and measures of corresponding biomarker concentrations (e.g., blood, urine, saliva, hair). These sampling activities would be scheduled to occur in concert with other planned NCS sampling and information gathering activities, to ensure efficiency and consistency in this effort.

- While some of the sampling specified above is both burdensome and invasive, it will likely be necessary to conduct longitudinal aggregate exposure sampling on small subsets of individuals to help determine within-person variability in exposure and how exposure changes as a function of life-stage. Proper incentives will need to be planned to offset the burden placed on NCS respondents that participate in the aggregate exposure study.

- The sampling methodology and chemical analysis of resulting samples should be designed to provide the results across as many chemical species as possible, subject of course to the logistics problems associated with collecting samples for multiple analyte classes, and the degree of burden associated with that collection. In addition, when possible (including consideration of sample/extract stability), these samples (or extracted materials) should be archived for future chemical analysis to cover chemical species that cannot be easily or cost-effectively assessed with current chemical analysis methods.

- The aggregate exposure study must be conducted on a representative sample of NCS respondents, so that the results can be generalized back to the NCS cohort population. This requirement also extends to temporal representation within the NCS, so that the aggregate exposure study is not conducted only on respondents who enroll into the study at an early or late date.

- This study must provide coverage across all stages of vulnerability, including pre-conception exposures among women planning pregnancy, exposures during pregnancy, and exposures during infancy through early adulthood.

- The aggregate exposure study must be planned so that sources and pathways of exposure can be characterized within each stage of vulnerability with known precision and accuracy.

Given the above constraints, one possible example sampling plan for the aggregate exposure study might be to randomly select and recruit a sub-sample of participants upon enrollment in the larger NCS. For example, upon enrollment in the NCS, 1 percent of respondents could be selected at random for initial inclusion in the aggregate exposure study, for a total of 1,000 participants among women planning pregnancy or in early stages of pregnancy. Of these 1,000 women who participate in the aggregate exposure study during this first stage, 40 percent (or 400) women would be selected at random to participate in the aggregate exposure study during the first two stages of vulnerability, and 16 percent (or 160 women) would be encouraged to participate in the aggregate exposure study for the first three stages of vulnerability. At each subsequent stage of vulnerability covered by the NCS, the aggregate exposure study would be replenished to achieve a total sample size of 1,000 study subjects – by enrolling 600 study subjects for the aggregate exposure study from a pool of available NCS study participants who previously had not participated in the aggregate exposure study. Of the 600 study subjects who are chosen for participation in each subsequent phase, 240 would participate in two consecutive phases, of which 160 would participate in three consecutive phases.

The above sampling plan for aggregate exposure assessment is one simple example of how NCS study participants could be identified and enrolled for this important sub-study. The actual sample size necessary to provide accurate exposure assessment information to the NCS to serve as a basis for adjusting relationships for measurement error in less precise measures of exposure may, in fact, be smaller or greater than the 1,000 subjects chosen here as an example at each stage of life. This will certainly be a topic of future research if the above recommended approach for an aggregate exposure study is adopted as part of the overall NCS strategy for exposure assessment. Clearly, additional design considerations for the aggregate exposure study could be integrated, such as stratified sampling approaches to over-sample important subpopulations. In addition, the above design requirements, constraints, and options leave much

of the sampling plan to be determined with future collaborative design work among chemists, statisticians, survey methodologists, and exposure assessment experts. However, compliance with the above stated requirements for the aggregate exposure sub-study will ensure that the information gathered in this effort can be appropriately combined with the results of the health outcome research sub-studies to help identify the pathways and sources of exposures that are statistically associated with adverse health effects as detailed in the following section.

C-2.3 COMBINING INFORMATION ACROSS THE TWO LINKED SAMPLES

The hierarchy of the planned study design facilitates combining data across two sub-studies of the NCS to help comment on the exposure-related causes of adverse health effects. The first sub-study would be optimally designed to characterize the relationship between an adverse health effect and some surrogate measure of exposure. From a public health perspective, and certainly as a first stage analysis, it is important to first understand what substances or exposure-related behaviors are related to an adverse health outcome. If the substance or behavior is determined to be a significant risk factor for the adverse health outcome, a second stage analysis that determines the external sources and pathways that can serve as the biological basis for the risk factor is an appropriate follow-on activity. From an environmental policy perspective, internal doses of a chemical cannot necessarily be regulated – however, external exposure sources that are linked with high internal doses may be regulated. The result of combining the data across the two sub-studies would then be the identification of substances or behaviors that are significant risk factors for causing adverse health effects, and the identification of external exposures and/or exposure pathways which can be potentially eliminated through various types of intervention.

The combining of information across the two samples depends on an assumption of transportability – that the exposure pathways that are found to be highly linked with increased levels of biomarker concentrations (or absorbed doses) in the aggregate exposure study are likely the same pathways that lead to an adverse health effect in the first stage analysis. One advantage to the hierarchical sampling strategy that is being employed is that both samples are being selected from the same finite study population. Therefore, there will be a small number of study participants that occur in both samples that can provide data to assess this assumption of transportability. Furthermore, the availability of detailed aggregate exposure information could be used as a stratification variable for over-sampling in the first sample that relates health outcome to a measure of exposure.

Although we discuss the combining of information across both samples as a two-stage process, there are appropriate statistical tools available for combining the data under a single statistical model. Bayesian hierarchical modeling, while typically complex in implementation, appears especially well suited for the structure of the anticipated data given the above paradigm.

This approach allows a high degree of flexibility in the modeling process, allows for the combination of data from different sources (or samples, in this case), and allows flexibility in the distributional assumptions of the data being modeled. While this approach provides a convenient structure for combining data across multiple samples, additional research would be necessary for integrating the statistical concepts of designed missingness and latent variable methods to correct for measurement error bias with the hierarchical Bayesian analysis approach.

One final important design item to consider is that it is possible to (1) demonstrate a highly statistically significant relationship between an adverse health effect and a biomarker or absorbed dose chemical concentration in the first sample, (2) demonstrate highly statistically significant relationships between that same biomarker chemical concentration and measures of exposure and exposure-related behavior in the second aggregate exposure sample, and (3) combine the data from these two samples and not have sufficient power to show a statistically significant relationship between the adverse health effect assessed in the first sample and the most significant source or pathway of exposure from the aggregate exposure study. From an intuitive perspective, this result makes sense: if a measure of exposure (biomarker concentration) explains only part of the variability in an adverse health effect response, and the pathways of exposure assessed in the aggregate exposure study explain only a fraction of the variability of that measure of exposure, then when combining both sources of information we would likely find that the exposure pathways will explain an even smaller amount of variability in the adverse health effect response. Given this important result, it will be necessary to ensure higher degrees of statistical precision and power in both samples if the ultimate goal is to yield statistically significant relationships between aggregate measures of exposure and adverse health effects in children.

Finally, we hope that the design guidelines that resulted from this project foster more discussion and collaborative work in developing powerful and efficient designs for the NCS project. There are most certainly other design guidelines that should be considered in the exposure assessment design for the NCS study, and there are most certainly necessary and advantageous refinements of the guidelines presented. The need to examine these avenues more thoroughly and to consider and investigate the many technical challenges involved in a study as large and complex as the NCS is fundamental to its ultimate success. Scientific hypothesis generation, methods development for health outcome and exposure assessment and characterization, appropriate statistical analysis tools, and appropriate, powerful and efficient statistical designs, are just a few of the core areas that will need to be carefully planned and clearly specified to ensure the success of this groundbreaking public health study. Solutions to these multidisciplinary issues call for collaboration among scientists and researchers spanning a wide range of expertise, and can lead to a successful, powerful, and resource efficient National Children's Study.

C-3.0 REFERENCES

1 Branum, A.M., Collman, G.W., Correa, A., Keim, S.A., Kessel, W., Kimmel, C.A., Klebanoff, M.A., Longnecker, M.P., Mendola, P., Rigas, M., Selevan, S.G., Scheidt, P.C., Schoendorf, K., Smith-Khuri, E., Yeargin-Alsopp, M. (2003) "The National Children's Study of Environmental Effects on Child Health and Development." Environmental Health Perspectives 111:642-646.

2 The Children's Health Act of 200. Public Law 106-310, 106th Congress, Section 1004

3 Strauss, W.J., Ryan, L., et. al. "Development of Exposure Assessment Study Design for the National Children's Study: Literature Search Report." Technical Report submitted by Battelle to U.S. Environmental Protection Agency, National Exposure Research Laboratory under Task 2 of Task Order 19 on Contract 68-D-99-011, May 2003

4 Loecke, D (2003) "Mixed Effects Mean Score Method for Auxiliary and Missing Covariate Data in Longitudinal Regression Models." Ph.D. Thesis, Harvard School of Public Health, Department of Biostatistics.

5 Arroyo, C. (2003) "Mixed Effects Models for Multiple Outcomes: Application for Physical Activity Data." Ph.D. Thesis, Harvard School of Public Health, Department of Biostatistics.

6 Roberts, K (2003) "Latent Variable Models for Exposure Measurement Error in Environmental Studies." Ph.D. Thesis, Harvard School of Public Health, Department of Biostatistics.

7 Park, S. and Ryan, L. (2003) "Optimal Design for Studies Involving Replicate Exposure Measurements Subject to Short and Long Term Variation." Technical Report, Harvard School of Public Health, Department of Biostatistics.

8 Ryan, L., Loecke, D., Roberts, K., Park, S. Arroyo, C., Strauss, W.J. "Development of Exposure Assessment Study Design for the National Children's Study: Optimal Design Methods and Statistical Analysis Methods." Technical Report submitted by Battelle to U.S. Environmental Protection Agency, National Exposure Research Laboratory under Task 3 of Task Order 19 on Contract 68-D-99-011, June 2003

9 Strauss, W.J., Lehman J., Morara, M., Ryan, L., "Development of Exposure Assessment Study Design for the National Children's Study: Simulation of Exposure and Health Outcome Research Sub-Samples for Guiding Environmental Sampling Design Options." Technical Report submitted by Battelle to U.S. Environmental Protection Agency,

National Exposure Research Laboratory under Task 4 of Task Order 19 on Contract 68-D-99-011, May 2003

10 Strauss, W.J., Lehman, J., Ryan, L., Morara, M., "Development of Exposure Assessment Study Design for the National Children's Study: Power and Efficiency Analysis." Technical Report submitted by Battelle to U.S. Environmental Protection Agency, National Exposure Research Laboratory under Task 5 of Task Order 19 on Contract 68-D-99-011, June 2003